Spatial Statistics 54 (2023) 100738

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

A spatio-temporal multinomial model of firearm death in Ecuador

Jorge Sosa [a], Álvaro Briz-Redón [b,*], Miguel Flores [a], Mauricio Abril [c], Jorge Mateu [d]

[a] *Department of Mathematics, Escuela Politécnica Nacional, Quito, Ecuador*
[b] *Department of Statistics and Operations Research, University of Valencia, Valencia, Spain*
[c] *Faculty of Engineering and Applied Sciences, Universidad Central, Quito, Ecuador*
[d] *Department of Mathematics, University Jaume I, Castellon, Spain*

ARTICLE INFO

Article history:
Received 19 January 2023
Received in revised form 6 March 2023
Accepted 7 March 2023
Available online 15 March 2023

Keywords:
Bayesian inference
Crime
Firearms
Multinomial distribution
Spatial effects
Temporal effects

ABSTRACT

This paper presents a statistical model based on a multinomial distribution with fixed and random effects, the latter effects being structured and non-structured in space and time. Inference is performed through a Bayesian framework. We are interested in analyzing violent deaths at the level of parroquia in Ecuador. Noting that most of the deaths are linked to firearms, and much less with knifes, we build a multinomial model to predict the probability of three different types of deaths as a close proxy to a violent death having had occurred. We provide a practical and realistic interpretation of the model putting this in the real crime context and scenario in Ecuador.

1. Introduction

Latin America is one of the most violent regions in the world. However, this situation is concentrated in certain countries such as Brazil, Colombia, El Salvador, Mexico, and Venezuela, among others (Bergman, 2018; Dávila-Cervantes and Pardo-Montaño, 2018), where inequality and informal labor are exploited by organized crime, leading to an increase in criminality (Hernández Bringas, 2022). Firearms are therefore a vehicle for the perpetration of various crimes such as robbery, kidnapping, homicide, etc. Several studies have been developed in Latin America, analyzing factors

* Corresponding author.
E-mail address: alvaro.briz@uv.es (Á. Briz-Redón).

https://doi.org/10.1016/j.spasta.2023.100738

associated with crime such as social aspects, gender, age (Neapolitan, 1994; Briceño-León et al., 2008; Rivera, 2016), as well as social and human behaviors related to firearms (Ceccato et al., 2007; Vilalta and Muggah, 2014; Dare et al., 2019). However, studies in Ecuador that examine homicides from a spatial perspective or that relate the type of weapon are scarce and are concentrated in specific cities of the country (Arvelo et al., 2019; Garc, 2021; Ochoa Veloz, 2021). In this sense, this study aims to analyze the evolution of the homicide rate in general and by firearm in Ecuador, but above all, to examine its spatial behavior at the cantonal level for the period 2018–2022. Since 2018, Ecuador has maintained a significant increase in the death rate, which is attributed to the presence of drug trafficking cartels as the first response from the official entity; if this behavior continues, it is expected that in 2022 it will reach the highest homicide rate in the history of Ecuador, surpassing the average rate in Latin America (Primicias, 2022a). This type of violence is more frequent in provinces of the Ecuadorian Coast, predominantly in Guayas, Manabí, and Esmeraldas (GK, 2022), in addition to events of extreme violence in the Social Rehabilitation Centers (SRC), which shows a clear power struggle and conflicts to maintain control of these centers; we note the recent amount of eleven prison massacres and 413 prisoners that have been killed in 21 months (Primicias, 2022b). The Central Government has responses and actions that are poorly supported by information, so it is necessary to generate knowledge of the phenomenon, which will improve decision-making at different levels of care. To this end, in this paper we focus on analyzing violent deaths, together with the type of weapon used to commit the crime (firearm, knife, others), and also with information on the victim (sex, age, marital status, etc.). We also take into account the date and time given by the forensic examination, as well as the position of the event by longitude and latitude, which makes it possible to disaggregate down to the level of canton so that a Public Policy proposal can be generated according to each locality according to the reality of the crime.

Fig. 1 depicts the evolution of the violent death rate in Ecuador from 1980 to 2022. It shows a behavior that requires careful study since between 1980 and 2009 there is a growing trend with slight increases and decreases, that overall are not very pronounced, and where the mortality rate was moderate for Ecuador (Briceño-León et al., 2008). Between 2010 and 2017, there is a marked and sustained decrease in the violent death rate, which is attributed to the implementation of social public policies, aimed at improving the welfare of the population, improving social investment in education and health, and reducing the social inequality gap. What is worrying and requires an in-depth study is the change in trend between 2018 and 2022, with a pronounced growth, with the violent death rate exceeding the average of the entire series in 2021, and almost doubling it in 2022 according to the projection made on the basis of the data available at the time of conducting this study (data up to July 2022), by multiplying the number of events recorded at that moment of the year by the inverse of the fraction of the year elapsed.

Although according to reports of seizures of controlled substances, large volumes have been seized in the last three years, the drug business is the main cause for this large number of deaths, considering that it is strongly associated with violence, as different groups compete to control the markets (Azaola, 2008); these criminal organizations are creditors of high caliber weapons such as explosives, sub fuses, among others through arms trafficking (Valdivieso, 2015). In the previous paragraphs, possible factors associated with violent deaths were presented, however, from the legal point of view, the Organic Criminal Code typifies four types of crimes: Murder, Homicide, Femicide, and Sicariato, which are due to specific criminal behaviors. We note that we use data in different supports, geographical coordinates in form of (x, y, t), marks associated to each crime, and covariates associated to the canton level, and we also have the parroquias, which is a more disaggregated level compared to canton. Our proposed model is flexible enough to account for these support differences and uses all types of information, as described later in Sections 2 and 3.

The paper is structured as follows. Section 2 gives a detailed description of the data, a spatio-temporal multinomial model is used to determine association between the type of weapon as the response variable and several victim-level and environmental-level characteristics in Section 3, and the results are presented in Section 4. The paper ends with some final discussion and conclusions in Section 5.

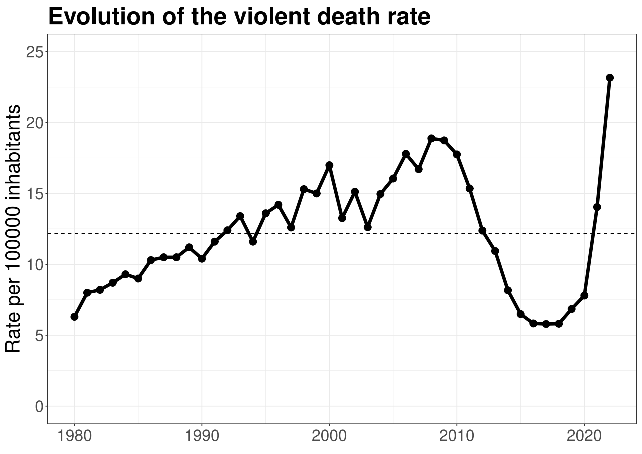


Fig. 1. Evolution of the violent death rate (per 100 000 inhabitants) in Ecuador during the period 1980–2022. The value of the time series for 2022 is a projected value considering the data currently available for the present study. The dashed line represents the average violent death rate during the period 1980–2022 (12.18).

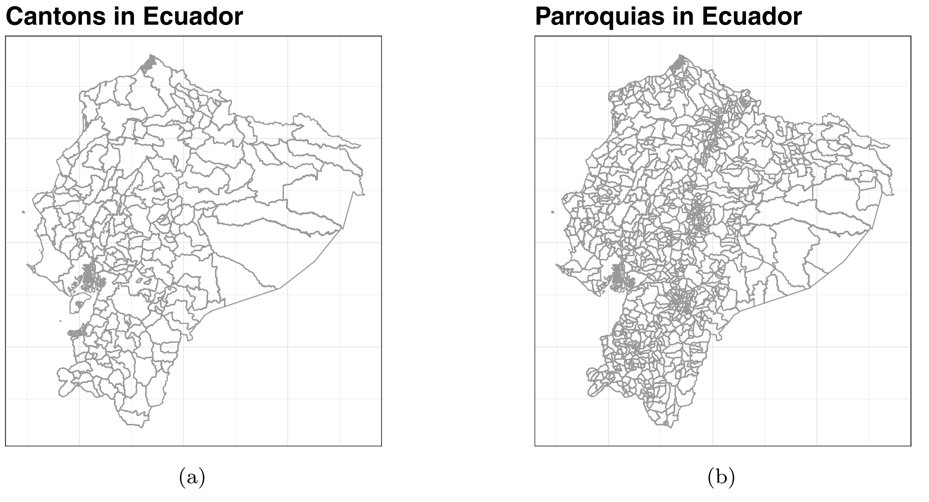


Fig. 2. Spatial configuration of cantons (a) and parroquias (b) across the Republic of Ecuador.

2. Data

The present study uses the violent death database with a total of 8496 observations, where each one of them has longitude and latitude coordinates, making it possible to have information at both the canton and the parroquia level, which are two (small-area) political-administrative divisions of the Republic of Ecuador. Cantons are a geographical unit that can be regarded as equivalent to the concept of a city, in general. Parroquias therefore correspond to intra-city geographical units, such as the municipal district or the borough in a different context. Fig. 2 shows the spatial configuration of cantons and parroquias across the country.

The homicide data are collected by the National Police of Ecuador through the National Directorate of Crimes against Life, Violent Deaths, Disappearances, Extortion and Kidnappings (DINASED), which is the police agency with expertise in Violent Deaths, and after the body is removed with

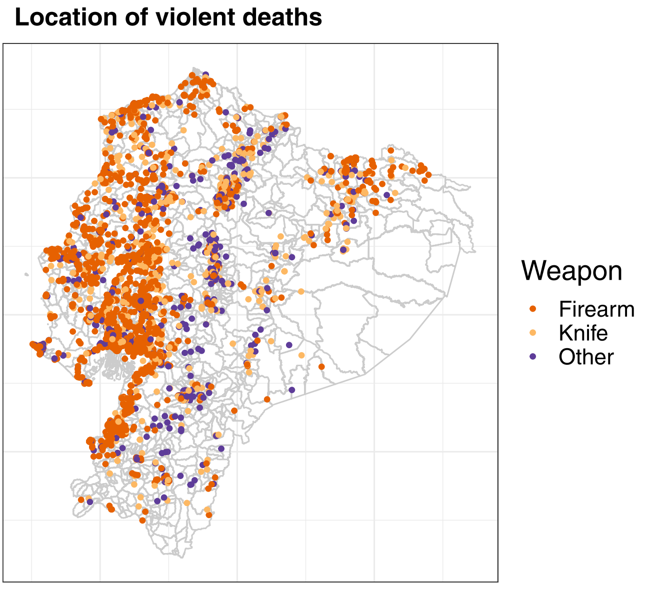


Fig. 3. Location (xy-coordinates) of the violent deaths recorded during the study period by type of weapon.

the knowledge and direction of the Public Prosecutor on duty, the autopsy protocol is carried out in Forensic Medicine. The geographic representation is associated with the shapefiles of cantons provided by the National Institute of Statistics and Census (INEC), a cartography that is updated for the execution of the Population Census. The temporal component focuses from January 2018 to July 2022 to analyze the latest years with a steady increase in the violent death rate. Fig. 3 shows the distribution of intentional homicides committed by different weapons (firearms, white arms/knives, or others), where a high percentage is concentrated in the coastal area of Ecuador, making it the most violent region of continental Ecuador in the last 5 years. According to the national police, this increase in violence is attributed to the dispute over territorial space for the sale of drugs. In addition, the northern part of the country is a conflict zone as it is the border with neighboring Colombia, where irregular warfare is present, as well as several coca leaf plantations.

Interestingly, the vast majority of these crimes were committed with firearms, with 5800 cases of this type, while there were 1752 cases with bladed weapons (knife type), and only 944 cases were attributed to another type of weapon. Although there are more cases involving firearms throughout the period considered, there is an overwhelming increase from April 2021 onwards with respect to other types of weapons (see Fig. 4). This fact makes a close link relationship between the use of firearms and the violent death rate (confirmed by Figs. 3 and 4), and motivates the fact that the use of firearms can be considered a proxy for the death rate. Prompted by this fact, we will focus on the type of weapon as a response variable for our statistical model.

Our data also has additional information, such as the age and sex of the victim, the type of crime, the place where the crime was committed, which can be at home, on a public road, or other, and the time at which the crime occurred, provided by the forensic examination. Regarding the type of crime, the events in the database have been classified into four classes according to the Organic Penal Code: *contract killing* (or *sicariato*, a killing carried out on agreement with a hired killer), *femicide* (a killing of a woman due to the fact of being a woman or due to her gender condition), *homicide* (a killing resulting from an accidental, reckless, or negligent act), and *murder* (a killing with malice aforethought). In addition, this study considers some socio-demographic spatial covariates at the cantonal level such as employment rate, complaint rate (related to crime reporting), robbery rate, health services rate, the rate of high school education institutions, as well as the accrued

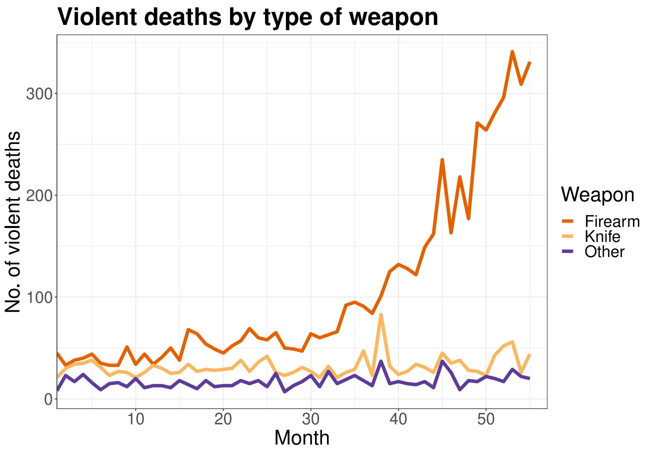


Fig. 4. Time series at the month level of the number of violent deaths by type of weapon during the study period.

Table 1
Summary and description of the spatial covariates and marks considered for the analysis. In the case of the covariates, each rate is computed dividing the quantity in question by the population size of the canton (rate per person). In the case of the marks, the levels that each one presents are provided within brackets. For mark *Time*, the following division of the day into time periods has been considered: from 0 h to 6 h (*Night*), from 6 h to 12 h (*Morning*), from 12 h to 18 h (*Afternoon*), and from 18 h to 24 h (*Evening*).

	Name	Description
Covariates	*Budget*	Accrued budget of the human development bond at the canton level.
	Complaint	Complaint rate at the canton level.
	Employment	Employment rate at the canton level.
	Health	Health services rate at the canton level.
	High education	High-school education institutions rate at the canton level.
	Robbery	Robbery rate at the canton level.
Marks	*Age*	Age of the victim (*<24*, *24–40*, *40–60*, or *>60*)
	Crime	Type of crime (*Contract killing*, *Femicide*, *Homicide*, or *Murder*)
	Gender	Gender of the victim (*Male* or *Female*)
	Place	Place where the crime has occurred (*Home*, *Other*, or *Public street*)
	Time	Time of occurrence of the crime (*Night*, *Morning*, *Afternoon*, or *Evening*)

budget of the human development bond (BDH for its acronym in Spanish). The latter aims to provide families benefiting from this voucher, a minimum level of consumption in addition to providing the opportunity to invest in the education and health of people under 18 years of age. These variables were obtained through the Ministry of Social and Economic Inclusion, the Ministry of Public Health, the Ministry of Education, and the Social Registry. A summary of the spatial covariates and marks considered for the study is provided in Table 1. All the covariates were standardized to be included in the modeling framework to be described in the subsequent section. Recall that a mark is a variable attached to each particular crime location, and a covariate is defined for a region and not only at particular locations where a crime happened. Finally, note that the covariate information is given at the canton level, since parroquia-level covariate information is not available. Nevertheless, as will be described in Section 3, the possible existence of relevant unobserved spatial covariates and spatial heterogeneity in the data will be accounted for through spatial random effects defined at the parroquia level. In this way, model estimates will be spatially smoothed at the most disaggregated level possible, allowing us to delineate the occurrence of the crimes of interest more accurately. Fig. 5 shows the choropleth maps corresponding to the spatial covariates considered for the analysis.

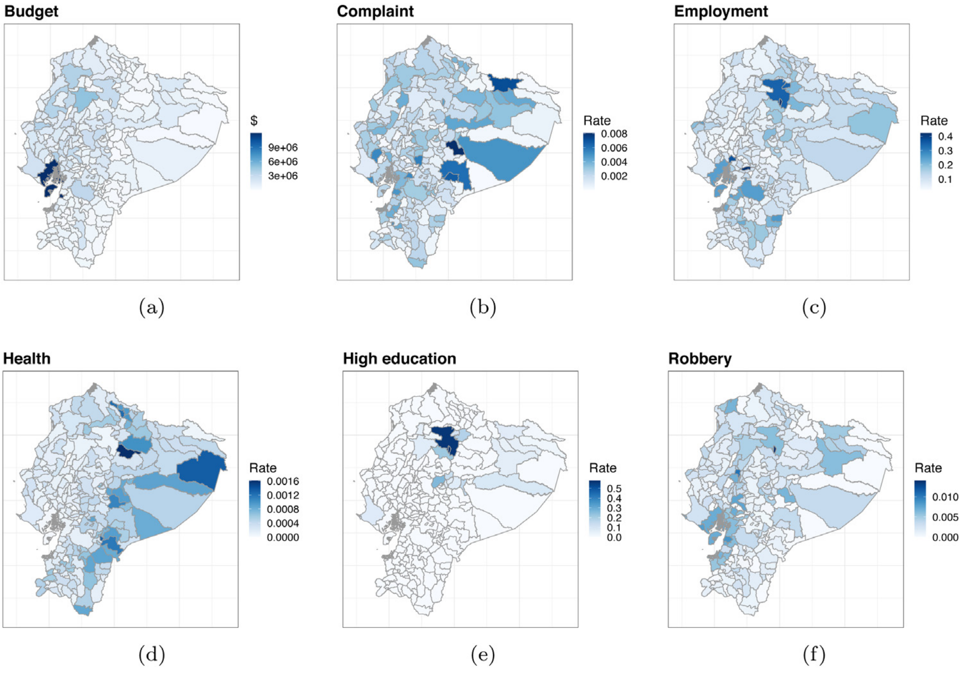


Fig. 5. Choropleth maps at the canton level corresponding to spatial covariates *Budget* (a), *Complaint* (b), *Employment* (c), *Health* (d), *High education* (e), *Robbery* (f).

3. Methodology

3.1. Model description

The response variable considered for the present study corresponds to the type of weapon employed for carrying out the criminal act. We consider three possible values for this random variable: *Firearm*, *Knife*, and *Other*. The category *Other* contains other types of weapons such as sticks, stones, ropes, or poison, among others. However, none of these is frequent enough to consider them alone as a category itself. Then, the database has been constructed with this *Other* category that includes all these kinds of weapons. To simplify the notation, in what follows we will consider these outcomes as numerical values 1 (*Firearm*), 2 (*Knife*), and 3 (*Other*). The main goal of the study is to find associations between the type of outcome and several victim-level and environmental-level characteristics. Specifically, victim-level characteristics can be regarded as marks of the observed pattern of crimes, whereas environmental-level characteristics refer to canton-level covariates and spatio-temporal effects. Therefore, we can assume that the type of weapon employed for crime i, Y_i, follows a multinomial distribution with probabilities $\{\pi_k(\boldsymbol{X}_i, \boldsymbol{R}_i)\}_{k=1}^{3}$, where $\pi_{k_0}(\boldsymbol{X}_i, \boldsymbol{R}_i)$ represents the probability that event i is of type $k_0 \in \{1, 2, 3\}$, given the victim-level and environmental-level characteristics corresponding to event i, $\boldsymbol{X}_i$, and some random effects specified for event i, $\boldsymbol{R}_i$. Again, to make the notation simpler, we write π_{ki} instead of $\pi_k(\boldsymbol{X}_i, \boldsymbol{R}_i)$, implicitly assuming the dependence between this probability value and the spatial covariates, marks, and random effects corresponding to event i.

In order to model the π_{ki}'s, we consider a log-linear modeling structure (Fienberg and Rinaldo, 2012). Specifically, we adopt the following log-linear relationship for $\pi_{k_0 i}$, with $k_0 \in \{1, 2, 3\}$,

$$\log \pi_{k_0 i} = \log P(Y_i = k_0) = f_{k_0}(\boldsymbol{X}_i, \boldsymbol{R}_i) - \log Z_i,$$

where $f_{k_0}(\boldsymbol{X}_i, \boldsymbol{R}_i)$ represents a linear combination (dependent on $k_0 \in \{1, 2, 3\}$) of fixed and random effects and Z_i is a partition function defined as

$$Z_i = \sum_{k=1}^{3} e^{f_k(\boldsymbol{X}_i, \boldsymbol{R}_i)}.$$

The inclusion of Z_i is necessary to ensure that the modeled π_{ki}'s result in a well-defined probability function, that is, with $\sum_{k=1}^{3} \pi_{ki} = 1, \forall i$. This formulation of the model can be also rewritten as follows to ease its implementation

$$\begin{aligned} \log \phi_{k_0 i} &= f_{k_0}(\boldsymbol{X}_i, \boldsymbol{R}_i) \\ \pi_{k_0 i} &= P(Y_i = k_0) = \frac{\phi_{k_0 i}}{\sum_{k=1}^{3} \phi_{ki}}, \end{aligned} \tag{1}$$

where $\sum_{k=1}^{3} \phi_{ki}$ corresponds to the partition function Z_i previously introduced.

Hence, following (1), it is necessary to specify the model for the $\log \phi_{ki}$'s. As indicated, we allow each of these terms, for $k \in \{1, 2, 3\}$, to follow a specific linear combination of fixed and random effects, $f_k(\boldsymbol{X}_i, \boldsymbol{R}_i)$. In particular, we consider type-specific fixed and spatio-temporal random effects for *Firearm* and *Knife*, which are the types of interest, and the simplest form for $f_3(\boldsymbol{X}_i, \boldsymbol{R}_i)$, which corresponds to type *Other*. This leads to the following full specification of the $\log \phi_{ki}$'s

$$\begin{aligned} \log \phi_{1i} &= \alpha_1 + \sum_{j=1}^{p} \beta_{1j} X_{ji} + u_{1p(i)} + v_{1p(i)} + \delta_{1m(i)} + \varepsilon_{1m(i)} \\ \log \phi_{2i} &= \alpha_2 + \sum_{j=1}^{p} \beta_{2j} X_{ji} + u_{2p(i)} + v_{2p(i)} + \delta_{2m(i)} + \varepsilon_{2m(i)} \\ \log \phi_{3i} &= \alpha_3 \end{aligned} \tag{2}$$

In (2), the α_k's (for $k \in \{1, 2, 3\}$) are intercept parameters that represent the mean value for the $\log \phi_{ki}$'s, whereas the β_{kj}'s (for $k \in \{1, 2\}$) represent the (fixed) effect that covariate/mark j has on $\log \phi_{ki}$. A larger (smaller) estimate of β_{kj} means that covariate/mark j is associated with a larger (smaller) probability of occurrence for type k. We note that both the spatial covariates and the marks participate in the model in the same way. For this reason, we use a single index, $j \in \{1, \ldots, p\}$, that refers to the corresponding covariate/mark. In other words, we can write $\{1, \ldots, p\} = \mathscr{C} \cup \mathscr{M}$, where $\mathscr{C}$ and $\mathscr{M}$ refer to the indices corresponding to the spatial covariates and marks, respectively (this notation will be helpful in subsequent sections).

The remaining terms in (2) correspond to spatial and temporal random effects that we introduce for types 1 and 2. The terms $u_{kp(i)}$ correspond to the spatial random effects given by the Besag–York–Mollié (BYM) model (Besag et al., 1991), where $p(i)$ denotes the parroquia where event i has taken place. Under the BYM model, the conditional distribution of the spatially-structured effect on parroquia p is

$$u_{kp} | u_{kp' \neq kp} \sim N\left(\sum_{p' \neq p = 1} w_{pp'} u_{kp'}, \frac{\sigma_{ku}^2}{N_p} \right)$$

where N_p is the number of neighbors for parroquia p, $w_{pp'}$ is the (p, p') element of the row-normalized neighborhood matrix, and σ_{ku}^2 represents the variance of this random effect. The spatially-unstructured effect over the parroquias, v_{kp}, follows a Gaussian distribution, $v_{kp} \sim N(0, \sigma_{kv}^2)$, where σ_{kv}^2 is the variance of this random effect. The temporal random effects are defined at the month level, with $m(i)$ representing the month of occurrence of event i ($m(i) \in \{1, \ldots, 55\}$). The temporally-structured random effect, δ_{km}, is specified through a second-order random walk, $\delta_{km} | \delta_{km-1}, \delta_{km-2} \sim N(2\delta_{km-1} + \delta_{km-2}, \sigma_{k\delta}^2)$, with $\sigma_{k\delta}^2$ the variance component. Finally, the temporally-unstructured random effect, ε_{km}, follows an identically distributed Gaussian prior, $\varepsilon_{km} \sim N(0, \sigma_{k\varepsilon}^2)$.

Therefore, the structure chosen in (2) constitutes a multinomial regression model with spatio-temporal random effects. The regressors and random effects are used to describe the occurrence of the types of interest, *Firearm* and *Knife*. The consideration of multiple regressors allows us to detect associations between the typology of the victim and the outcome, as well as to find environmental characteristics that correlate with certain crime type. The spatial and temporal random effects included in the model allow smoothing model outcomes in space and time, producing more reliable estimates of how each type of crime is distributed over the study region during the study period. Regarding temporal random effects, a smaller time unit, such as a week, could have been chosen. However, working at the month level seems sufficient to establish an overall time trend for each crime type that can then be subsequently analyzed by the competent authorities. Modeling at the week level would complicate the model with little benefit in this regard. Similarly, the model could also include spatio-temporal interaction terms such as those described in Knorr-Held (2000), allowing us to detect variations in these trends between parroquias. However, this would increase the computational cost of the model considerably and has been discarded.

The spatio-temporal multinomial model described in previous lines has been implemented in the NIMBLE system for Bayesian inference (de Valpine et al., 2017), which relies on Monte Carlo Markov Chain (MCMC) procedures. A vaguely informative Gaussian prior, $N(0, 1000)$, has been assigned to the α_k's and the β_{kj}'s. The variance parameters of the spatial and temporal random effects have been assigned a Gamma-distributed prior, $Ga(1, 0.5)$.

3.2. Model selection and validation

In order to justify the need to include the different components of the multinomial model previously described, a comparative study has been carried out considering some submodels of the full model. In particular, we consider a model without spatial covariates and space–time random effects that only accounts for mark effects (mark-effects model), a model without spatial covariates that accounts for mark and space–time random effects (mark-ST-effects model), and the model previously described that accounts for mark, covariate, and space–time effects (full model). Besides, we implement a Bayesian variable selection procedure (Kuo and Mallick, 1998; Dellaportas et al., 2002) for the spatial covariates considered in the analysis in the context of fitting the full model. Basically, this procedure consists of replacing the β_{kj} terms in (2) by $I_{kj}\beta_{kj}$ (if $j \in \mathscr{C}$), where I_{kj} is a binary variable that is assigned a Bernoulli $Ber(0.5)$ prior. We decided to apply this procedure only on the spatial covariates and not on the marks in order to alleviate the computational cost of the procedure and because of the intrinsic relevance of the marks considered. This model is referred to as the full model with BVS (Bayesian variable selection) in the remainder of the paper.

The Watanabe–Akaike Information Criterion (WAIC) proposed by Watanabe and Opper (2010) has been used for model comparison. A reduction in the value of the WAIC indicates greater model performance. In addition to this, we have followed Nakagawa and Schielzeth (2013) in order to quantify the level of variability captured by each model component. In particular, we have estimated the variability corresponding to the fixed effects of the model as $\sigma^2_{kf} = Var(\sum_{j=1}^{p} \beta_{kj}X_{ji})$ by using the posterior estimates of the β_{kj}'s for $k \in \{1, 2\}$. Then, we are interested in computing the fraction $\omega_k = \frac{\sigma^2_{kf}}{\sigma^2_{ku}+\sigma^2_{kv}+\sigma^2_{k\delta}+\sigma^2_{k\varepsilon}}$, for $k \in \{1, 2\}$, considering empirical estimates of these variance parameters. The lower the value of ω_k, the greater is the contribution of space–time random effects.

Model validation has been carried out through several measures of in-sample predictive performance. Specifically, the posterior mean of the π_{1i}'s and the π_{2i}'s depending on the outcome actually observed has been employed as a tool of model criticism that also allows us to guess the classification ability of the model based on the estimation of these probabilities. Besides, the receiver operating characteristic (ROC) curves have been also computed as another measure of predictive quality. We note again that all these correspond to in-sample metrics which somehow act as goodness-of-fit measures for a categorical response variable. Performing an out-of-sample forecasting assessment of the model fitted was beyond the scope of the paper.

Table 2
Comparison of the multinomial models fitted in terms of their WAIC. The structure of the model corresponds to the specification of the term $\log \phi_{ki}$ (for $k \in \{1, 2\}$) in (2). The sets $\mathscr{C}$ and $\mathscr{M}$ refer to the indices of the spatial covariates and the marks, respectively, as described in Section 3.

Model	Structure	WAIC
Mark-effects	$\alpha_k + \sum_{j\in\mathscr{M}} \beta_{kj}X_{ji}$	11 095.15
Mark-ST-effects	$\alpha_k + \sum_{j\in\mathscr{M}} \beta_{kj}X_{ji} + u_{kp(i)} + v_{kp(i)} + \delta_{km(i)} + \varepsilon_{km(i)}$	9951.89
Full	$\alpha_k + \sum_{j\in\mathscr{C}\cup\mathscr{M}} \beta_{kj}X_{ji} + u_{kp(i)} + v_{kp(i)} + \delta_{km(i)} + \varepsilon_{km(i)}$	9850.35
Full with BVS	$\alpha_k + \sum_{j\in\mathscr{C}} I_{kj}\beta_{kj}X_{ji} + \sum_{j\in\mathscr{M}} \beta_{kj}X_{ji} + u_{kp(i)} + v_{kp(i)} + \delta_{km(i)} + \varepsilon_{km(i)}$	9855.60

3.3. Mapping model outcomes

The spatial random effects included in the multinomial model provide insight into the influence of the parroquias on the probability that a violent death has been caused by a firearm. That is, by mapping the estimates of the spatial random effects we can establish areas of elevated risk of firearm-related events. In particular, mapping the spatially-structured effects, u_{kp}, is of special interest, as they provide a smoothed picture of the risk, which facilitates the identification of patterns. Under the BYM model chosen for the analysis, the spatial random effects u_{kp} and v_{kp} are not uniquely identifiable (only their sum is identifiable), as discussed by Eberly and Carlin (2000). Thus, we corrected the posterior estimates of u_{kp} and v_{kp} following Duncan et al. (2019). Specifically, we compute the spatial fraction $\psi_k = \frac{s_{u_k}}{s_{u_k}+s_{v_k}}$, where s_{u_k} and s_{v_k} are empirical estimates of the posterior marginal standard deviation of the spatial random effects. The value of ψ_k represents the proportion of the spatial variability that has a geographical structure. Thus, values of ψ_k close to 1 mean that the spatially-structured effects, u_{kp}, have a greater impact than the spatially-unstructured ones, v_{kp} (a value of ψ_k close to 0 means the opposite). Then, the corrected estimates of u_{kp} and v_{kp} are, respectively, $\hat{u}^c_{kp} = \hat{u}_{kp} - \psi_k\hat{v}_{kp}$ and $\hat{v}^c_{kp} = \hat{v}_{kp} + \psi_k\hat{v}_{kp}$, where $\hat{u}_{kp}$ and $\hat{v}_{kp}$ denote the posterior mean of the corresponding parroquia-level random effects.

3.4. Software and computation

The R programming language (R Core Team, 2022) has been used for the analysis. In particular, the R packages ggplot2 (Wickham, 2016), nimble (de Valpine et al., 2017), rgdal (Bivand et al., 2019), rgeos (Bivand and Rundel, 2020), and spdep (Bivand et al., 2008) have been used. A remote computer with 150 GB of RAM has been used to fit all the multinomial models described in the paper. Fitting the full model presented in the paper without including the Bayesian variable selection procedure takes approximately 60 h in this computer. Incorporating the Bayesian variable selection procedure for the spatial covariates increases computation time to nearly 80 h.

4. Results

We begin the analysis by comparing the four specifications considered for the multinomial model in terms of performance. Table 2 shows the structure of the four models and the corresponding WAIC values obtained. It is clear that the inclusion of spatio-temporal effects in the model results in a significant improvement of the model fit, while the addition of the spatial covariates only provides a modest gain. In addition, the BVS procedure yields a slightly inferior model than the full model. For this reason, we will focus on the full model (without BVS) to analyze the spatio-temporal dependence of the observations through the estimates of the corresponding random effects in space and time. Subsequently, we analyze the estimated fixed effects, considering the outputs of both the full model and the full model with VBS. On the other hand, the computation of the ω_k's enables us to see that the spatio-temporal random effects explain a large proportion of the variability of the data, especially in the case of firearm-related events. Specifically, we have obtained $\omega_1 = 0.64$ and $\omega_2 = 0.95$ for the full model, which tells us the proportion of the variability of the $\log \phi_{ki}$'s

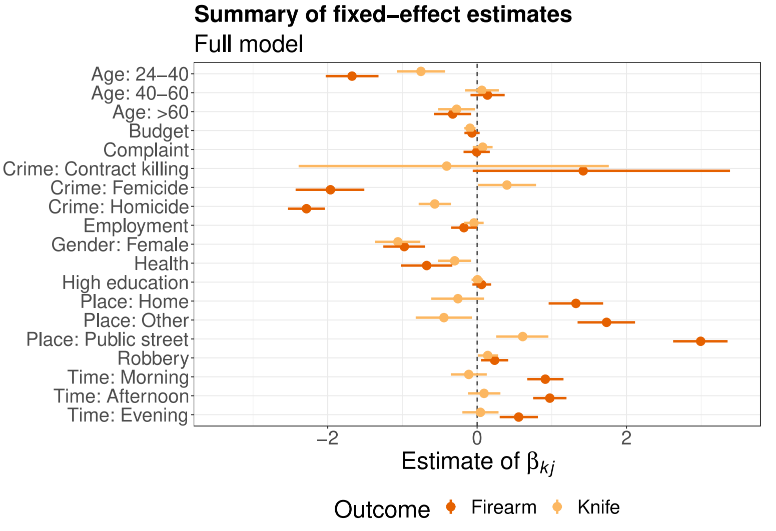
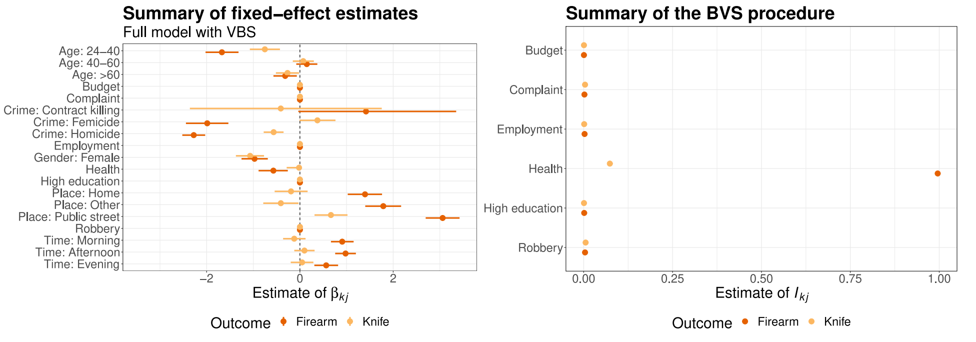


Fig. 6. Summary of the covariate/mark effects estimated through the full model in terms of the posterior mean and 95% credible interval corresponding to each of the β_{kj} parameters in (2).

that is captured by the fixed effects of the model. Hence, although the contribution of space–time effects for knife-related events is low, the benefit from the inclusion of such effects for modeling firearm-related events seems substantial.

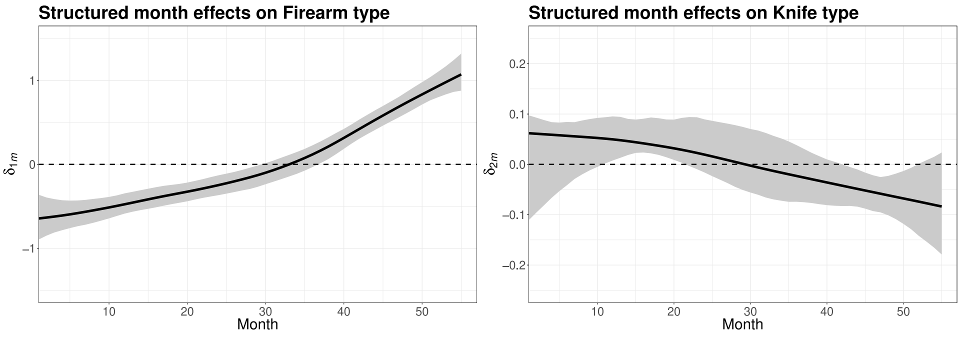
Fig. 6 summarizes the results yielded by the full model for the fixed-effects estimates corresponding to the set of victim-level and environmental-level characteristics considered in the analysis. For those characteristics that are encoded as categorical variables, we must keep in mind that the estimate has to be interpreted in relation to the reference level (absent in the graph). We can observe that the fact that the event has occurred at home, or in public streets is strongly associated with a greater probability that the crime has been committed with a firearm. At the same time, crimes that have not occurred in the night hours (which acts as the reference level) are more likely to be conducted with the aid of a firearm rather than a knife. Besides, it is also worth noting that belonging to the 24–40 age group, as well as being a female, are two characteristics that are associated with a lower probability that an observed crime event has been committed with a firearm, even though a similar result is obtained for the knife type. Therefore, other types of weapons or *modus operandi* might be more frequent in victims with these characteristics. Regarding the covariates available at the canton level, the rate of health services displays a strong negative association with both the firearm and the knife types. Less notably, a higher employment rate at the canton level is also associated with a lower probability of firearm, while a higher robbery rate exhibits the opposite behavior. The little effect of most of the spatial covariates is confirmed through the employment of the VBS procedure. Indeed, Fig. 7(a) shows the fixed-effect estimates provided by this model according to the probabilities of inclusion given by the posterior estimate of the I_{kj} parameters, which are also shown in Fig. 7(b). Thus, the only covariate effect that seems relevant is the one that *Health* has on firearm-related events. We also note that the incorporation of the BVS procedure barely alters the estimates of the effects corresponding to the marks included in the analysis, as can be observed by comparing Figs. 6 and 7(a).

Fig. 8 shows the estimates of the temporally-structured month-level random effects corresponding to types firearm (Fig. 8(a)) and knife (Fig. 8(b)). It can be observed that the behavior is the opposite: the probability that a crime event of this kind has been committed with a firearm shows an increasing trend during the study period, while in the case of a knife, the tendency is decreasing. However, we should note that the magnitude of the estimation of these temporal random effects is



(a) (b)

Fig. 7. Summary of the covariate/mark effects estimated through the full model with BVS: (a) in terms of the posterior mean and 95% credible interval corresponding to each of the β_{kj} parameters and summary of the BVS procedure, (b) in terms of the posterior mean of the corresponding I_{kj} parameters.

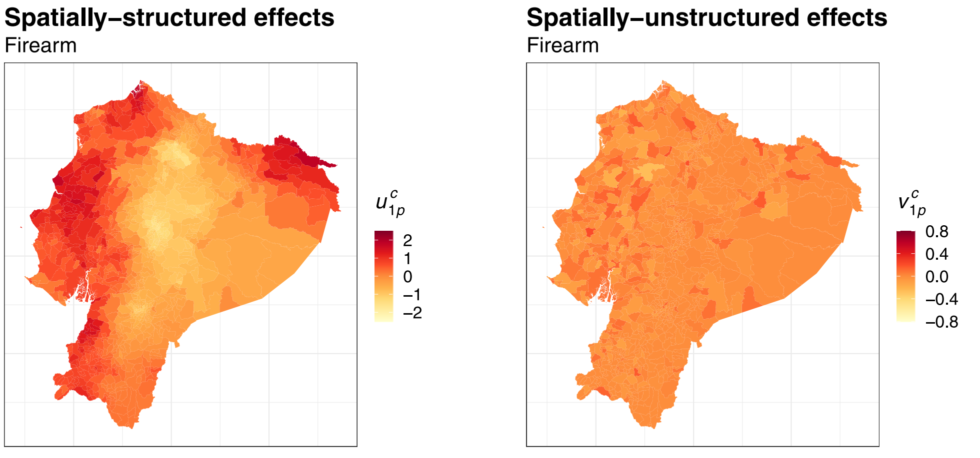


(a) (b)

Fig. 8. Temporally-structured random effects estimated for types *Firearm* (a) and *Knife* (b) with the full model. These correspond to δ_{1m} and δ_{2m}, respectively, in (2). The solid line represents the posterior mean of δ_{km} (for $k \in \{1, 2\}$) at each month, m, within the study period ($m \in \{1, \dots, 55\}$), whereas the gray areas represent 95% credible bands.

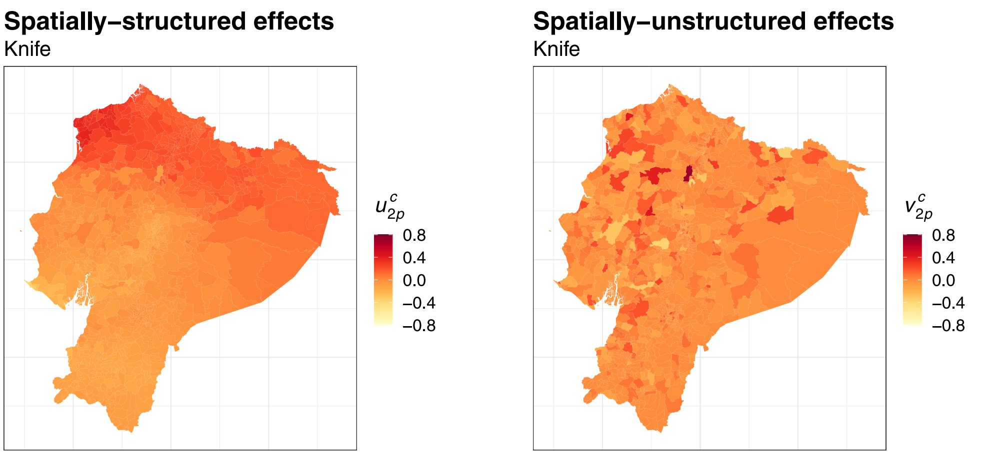
markedly different. Thus, the estimate of this effect in the case of knife is much lower, suggesting that despite the presence of a decreasing trend, the time factor has less weight as an explanatory factor of the probability that an observed crime is caused by a knife. In addition, the credibility bands include 0 for much of the study period, indicating that the temporal effect may not be relevant in this case.

Figs. 9 and 10 display the estimates of the spatial random effects at the parroquia level, after applying the correction of Duncan et al. (2019) to alleviate identifiability issues. Note that the scale used in the maps is not always the same in order to be able to observe the existing variations for each estimated spatial random effect. A greater value of either the spatially-structured (u_{1p} and u_{2p}) or spatially-unstructured (v_{1p} and v_{2p}) random effects for a certain parroquia is associated with a higher probability that a crime that takes place in that parroquia is carried out with the weapon in question. Thus, it can be observed that the distribution of both firearm-related and knife-related crimes is markedly spatial, according to the estimates corresponding to the spatially-structured random effects. On the one hand, the choice of a firearm is more common along the east coast of Ecuador and the inland area to the east, and also at the northwest of the country (Fig. 9(a)). On the other hand, knife-related crimes show a higher probability of occurrence in the north of Ecuador, especially in the northeast area (Fig. 10(a)). In addition to this, spatially-unstructured



(a) (b)

Fig. 9. Spatially-structured (a) and spatially-unstructured (b) random effects at the parroquia level estimated for type *Firearm* with the full model. These correspond to u_{1p} and v_{1p}, respectively, in (2), after applying the correction procedure on the spatial random effects to alleviate the identifiability issues between u_{1p} and v_{1p} (for this reason, we use u^c_{1p} and v^c_{1p} in the legend).



(a) (b)

Fig. 10. Spatially-structured (a) and spatially-unstructured (b) random effects at the parroquia level estimated for type *Knife* with the full model. These correspond to u_{2p} and v_{2p}, respectively, in (2), after applying the correction procedure on the spatial random effects to alleviate the identifiability issues between u_{2p} and v_{2p} (for this reason, we use u^c_{2p} and v^c_{2p} in the legend).

effects capture the remaining spatial heterogeneity existing in the data, mainly the presence of parroquias that behave differently from their neighboring parroquias. In the case of firearm, the spatially-unstructured effects present a relatively flat distribution across the entire country, which allows discarding the presence of spatial outliers. In contrast, for knife-related events we find moderate spatial heterogeneity, especially in the eastern half of Ecuador. In particular, it could be useful to study the parroquias that present a very high estimate for this random effect, since they

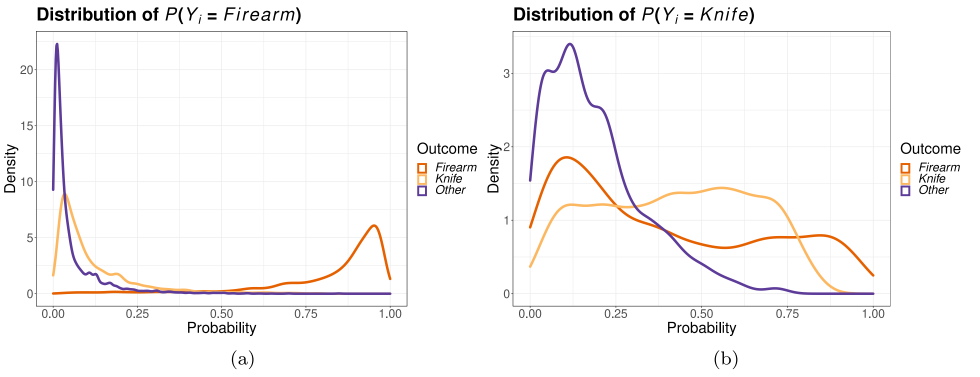


Fig. 11. Distribution of $\pi_{1i} = P(Y_i = Firearm)$ (a) and $\pi_{2i} = P(Y_i = Knife)$ (b) for the full model depending on the outcome observed.

might deserve further study. Furthermore, we have obtained spatial fraction values of $\psi_1 = 0.97$ and $\psi_2 = 0.73$, which allow us to conclude that the pattern of firearm-related events is markedly spatial.

Regarding model validation, Fig. 11 shows the distribution of the posterior mean of the π_{1i}'s and the π_{2i}'s depending on the outcome actually observed. The main conclusion that can be drawn from this analysis is that events caused by a firearm are better characterized by the model than those caused by a knife. Indeed, the distribution of the posterior mean of $\pi_{1i} = P(Y_i = Firearm)$ is located close to 1 for those events that have been actually committed with a firearm. At the same time, the distribution of the posterior mean of the π_{1i}'s is close to 0 for the other possible outcomes. In contrast, in the case of $\pi_{2i} = P(Y_i = Knife)$, the distribution of the posterior mean does not differ substantially by observed outcome, although higher values tend to be observed for events that have actually been carried out with a knife. Finally, Fig. 12 shows the ROC curves obtained for each crime type. We note that for the construction of this curve, the predicted type for the observation i is obtained as the category $k \in \{1, 2, 3\}$ that maximizes the set of posterior probabilities of the π_{ki}'s. The results allow us to observe again that the model performs better for firearm-related events.

5. Discussion and conclusions

The fitted multinomial model accurately depicts violent deaths with a firearm in Ecuador, which widely affect young men on public roads. This finding may have anything to do with the selective killings carried out by organized criminals to capture specific territories. Estimating the temporal component of firearm death in Ecuador throughout the study period reveals a significant and steady upward trend. Only as a result of agreements between the government and such organized groups have other nations seen significant declines in criminality (Cano and Rojido, 2017). In addition, the commission of other crime types, including drug trafficking, kidnapping, and human trafficking is frequently linked to the incidence of this kind of organized violence. The existence of firearms, which are illegal in Ecuador, is also problematic since it breeds more aggressive communities and unregistered, uncontrolled people that threaten public safety and interpersonal coexistence.

Although femicides have been also included in the analysis, this is an issue that might deserve individual study. The fact that the occurrence of knife-related crimes displays a decreasing trend since April 2021 might give the perception that the incidence of this phenomenon is reducing. However, domestic violence is not an isolated act of violence but responds to a patriarchal social construction (Cespedes et al., 2018). From a previous exploratory analysis of the available data, it has been observed that femicide typically occurs in the victim's home and is committed by her partner or ex-partner. This pattern has been observed as well in other Latin American countries, in which bladed weapons tend to be the most commonly employed (Carcedo and Ordóñez Laclé, 2011).

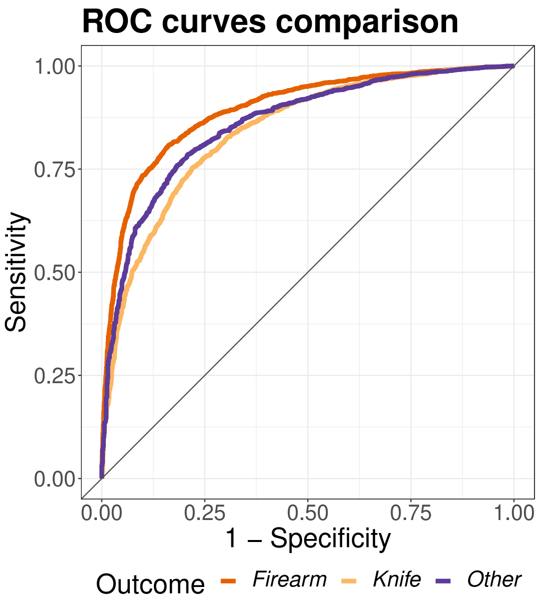


Fig. 12. Comparison of the ROC curves yielded by the full model considering the three possible values of the response variable.

This study demonstrates the need to continue deepening the analysis of firearm deaths in the cantons and provinces of the coast, to determine the type or types of violence which they are associated with, allowing authorities and researchers to have much more efficient tools to mitigate the existing situation. In the case of deaths with a knife, it is also necessary to deepen the analysis from the behavioral criminological field. Deaths with a knife are usually associated with fights, femicides, and domestic violence. The present study also allows us to approach the study of this kind of interpersonal violence, but the development of more focused studies seems advisable to better characterize this type of crime. Having socio-demographic covariates at a more accurate spatial level would be useful as well to gain insights about the influence of the economic context on the incidence of these offenses. Indeed, the high levels of violence observed in highland territories are more associated with inequality rather than with poverty itself (Hernández Bringas, 2022). An example of this is the level of violence in Quito, the capital of the country, where wealthy and deprived areas tend to coexist.

Finally, we should note some limitations of the study we have carried out. From a proactive surveillance perspective, the study of crime in Ecuador at the parroquia level may be insufficient for forecasting purposes. As previous literature has shown, it is convenient to use fine spatio-temporal units that allow us to capture the existing interaction between events (Mahfoud et al., 2021). In addition, despite the availability of some marks for the events under study, information on the victim is still quite scarce. It would be of particular benefit to collect certain socio-demographic characteristics about the victims. We have attempted to alleviate this deficiency by including socio-demographic covariates at the canton level, but they have proved to be insufficient in this regard. From a methodological point of view, it is also noteworthy that other modeling choices could be explored in order to capture the spatial variability of the observations. In particular, the consideration of multiple neighborhood relationships (Briz-Redón et al., 2022) or the employment of alternative spatial models that avoid the identifiability issues mentioned (Riebler et al., 2016) could be of interest.

Acknowledgments

We acknowledge the National Police of Ecuador for providing the data and giving us access to the associated information. Profs. Sosa and Flores also acknowledge the Escuela Politécnica Nacional for providing the necessary resources for a visit to University Jaume I. Profs. Mateu and Briz-Redón were partially funded by grant PID2019-107392RB-I00, from the Spanish Ministry of Science.

References

Arvelo, P.M.M., Zambrano, J.C.A., Robles, G.K., Zambrano, J.E.C.P., Pita, G.F.V., Carolina, D., Nolivos, A., Paucar, C.E.P., 2019. Neutrosophic model for the analysis of criminal behaviour in quevedo, ecuador, from a spatial econometric analysis. In: Neutrosophic Sets and Systems, Book Series, Vol. 26, 2019: An International Book Series in Information Science and Engineering. Infinite Study, p. 48.

Azaola, E., 2008. Crimen, Castigo Y Violencias En MÉXico. FLACSO Ecuador.

Bergman, M., 2018. More Money, More Crime: Prosperity and Rising Crime in Latin America. Oxford University Press.

Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Ann. Inst. Statist. Math. 43 (1), 1–20.

Bivand, R., Keitt, T., Rowlingson, B., 2019. Rgdal: Bindings for the 'geospatial' data abstraction library. R package version 1.4-6.

Bivand, R.S., Pebesma, E.J., Gomez-Rubio, V., 2008. Applied Spatial Data Analysis with R, Vol. 747248717. Springer.

Bivand, R., Rundel, C., 2020. Rgeos: Interface to geometry engine - open source ('GEOS'). R package version 0.5-3.

Briceño-León, R., Villaveces, A., Concha-Eastman, A., 2008. Understanding the uneven distribution of the incidence of homicide in latin america. Int. J. Epidemiol. 37 (4), 751–757.

Briz-Redón, Á., Iftimi, A., Correcher, J.F., De Andrés, J., Lozano, M., Romero-García, C., 2022. A comparison of multiple neighborhood matrix specifications for spatio-temporal model fitting: a case study on COVID-19 data. Stoch. Environ. Res. Risk Assess. 36 (1), 271–282.

Cano, I., Rojido, E., 2017. Introducción: La singularidad de la violencia letal en América Latina. Revista CIDOB D'Afers Internacionals 7–24.

Carcedo, A., Ordóñez Laclé, C., 2011. Femicidio En Ecuador. Manthra Editores, Ecuador.

Ceccato, V., Haining, R., Kahn, T., 2007. The geography of homicide in São Paulo, Brazil. Environ. Plan. A 39 (7), 1632–1653.

Cespedes, L.G., Noles, S.P.Z., Espino, I.C., 2018. Violencia contra la mujer y regulación jurídica del femicidio en ecuador. DÍKÊ. Revista Investig. Derecho, Criminol. Consult. Juríd. 12 (23), 129–150.

Dare, A.J., Irving, H., Guerrero-López, C.M., Watson, L.K., Kolpak, P., Shigematsu, L.M.R., Sanches, M., Gomez, D., Gelband, H., Jha, P., 2019. Geospatial, racial, and educational variation in firearm mortality in the USA, Mexico, Brazil, and Colombia, 1990–2015: a comparative analysis of vital statistics data. Lancet Public Health 4 (6), e281–e290.

Dávila-Cervantes, C.A., Pardo-Montaño, A., 2018. Violence in Colombia and Mexico: trend and impact on life expectancy of homicide mortality between 1998 and 2015. Public Health 163, 1–8.

Dellaportas, P., Forster, J.J., Ntzoufras, I., 2002. On Bayesian model and variable selection using MCMC. Stat. Comput. 12 (1), 27–36.

Duncan, E., Cramb, S., Baade, P., Mengersen, K., Saunders, T., Aitken, J., 2019. Developing a Cancer Atlas Using Bayesian Methods: A Practical Guide for Application and Interpretation. Queensland University of Technology (QUT) and Cancer Council Queensland, Brisbane, Australia.

Eberly, L.E., Carlin, B.P., 2000. Identifiability and convergence issues for Markov chain Monte Carlo fitting of spatial models. Stat. Med. 19 (17–18), 2279–2294.

Fienberg, S.E., Rinaldo, A., 2012. Maximum likelihood estimation in log-linear models. Ann. Statist. 996–1023.

Garc, K.J.T., 2021. Calculating Years of Life Lost to Estimate Premature Mortality Due to Firearms in Ecuador 1998–2018 (Ph.D. thesis). Yonsei University.

GK, 2022. Homicidios en ecuador se han duplicado en 2022. https://gk.city/2022/08/10/homicidios-en-ecuador-se-han-duplicado-en-2022/. Last accessed on 2023-01-13.

Hernández Bringas, H., 2022. Homicidios en América Latina y el caribe: magnitud y factores asociados. Notas Poblac. 113, 119–144.

Knorr-Held, L., 2000. Bayesian modelling of inseparable space-time variation in disease risk. Stat. Med. 19 (17–18), 2555–2567.

Kuo, L., Mallick, B., 1998. Variable selection for regression models. Sankhyā B 65–81.

Mahfoud, M., Bernasco, W., Bhulai, S., van der Mei, R., 2021. Forecasting spatio-temporal variation in residential burglary with the integrated Laplace approximation framework: Effects of crime generators, street networks, and prior crimes. J. Quant. Criminol. 37 (4), 835–862.

Nakagawa, S., Schielzeth, H., 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods Ecol. Evol. 4 (2), 133–142.

Neapolitan, J.L., 1994. Cross-national variation in homicides: The case of Latin America. Int. Crim. Justice Rev. 4 (1), 4–22.

Ochoa Veloz, P.E., 2021. Spatial Patterns of Crime in District 05, Quito-Ecuador, 2016–2020 (Master's thesis). Minnesota State University, Mankato.

Primicias, 2022a. Ecuador alcanza la tasa más alta de muertes violentas de la última década. https://www.primicias.ec/noticias/en-exclusiva/ecuador-tasa-muertes-violentas-ultima-decada/. Last accessed on 2023-01-13.

Primicias, 2022b. Once masacres carcelarias y 413 presos asesinados en 21 meses. https://www.primicias.ec/noticias/en-exclusiva/carceles-nueve-masacres-victimas-ecuador/. Last accessed on 2023-01-13.

R Core Team, 2022. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Riebler, A., Sørbye, S.H., Simpson, D., Rue, H., 2016. An intuitive Bayesian spatial model for disease mapping that accounts for scaling. Stat. Methods Med. Res. 25 (4), 1145–1165.

Rivera, M., 2016. The sources of social violence in latin america: An empirical analysis of homicide rates, 1980–2010. J. Peace Res. 53 (1), 84–99.

Valdivieso, C., 2015. Armas de Fuego En Ecuador. FLACSO Ecuador.

de Valpine, P., Turek, D., Paciorek, C.J., Anderson-Bergman, C., Lang, D.T., Bodik, R., 2017. Programming with models: writing statistical algorithms for general model structures with NIMBLE. J. Comput. Graph. Statist. 26 (2), 403–413.

Vilalta, C., Muggah, R., 2014. Violent disorder in ciudad juarez: a spatial analysis of homicide. Trends Organ. Crime 17 (3), 161–180.

Watanabe, S., Opper, M., 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. J. Mach. Learn. Res. 11 (12).

Wickham, H., 2016. Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.