Fire Safety Journal 138 (2023) 103802

Contents lists available at ScienceDirect

Fire Safety Journal

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

Spatial modelling and mapping of urban fire occurrence in Portugal

Regina Bispo [a,b,*], Francisca G. Vieira [a], Nádia Bachir [a], Pedro Espadinha-Cruz [c], José Pedro Lopes [d], Alexandre Penha [e], Filipe J. Marques [a,b], António Grilo [c]

[a] NOVAMath Center for Mathematics and Applications, NOVA School of Science and Technology, Universidade NOVA da Lisboa, Caparica, Portugal
[b] Department of Mathematics, NOVA School of Science and Technology, Universidade NOVA da Lisboa, Caparica, Portugal
[c] UNIDEMI, Department of Mechanical and Industrial Engineering, NOVA School of Science and Technology, Universidade NOVA da Lisboa, Caparica, Portugal
[d] Escola Nacional de Bombeiros and Instituto Superior de Educação e Ciências, Coimbra, Portugal
[e] Comando Nacional de Emergência e Proteção Civil, ANEPC Autoridade Nacional de Emergência e Proteção Civil, Carnaxide, Portugal

ARTICLE INFO

Keywords:
Hotspot analysis
Spatial autocorrelation
Spatial modelling
Urban fires

ABSTRACT

Fires in urban areas typically carry severe consequences. High population density together with the complexity of urban network potentially imply significant impacts in property loss, physical damage and life losses. However, despite the impact that fires may have in urban areas, research in urban fire prediction remains limited. In this study, we modelled urban fire occurrences while making a comparative analysis of different strategies to account for spatial autocorrelation. Considering space dependence in addition to a range of social-economic explanatory variables has proven to strengthen the validity of the fitted models.

The spatial Durbin error model, including population density, degraded buildings density and buying power, was selected as having the best fit. This model allowed to map the estimated probability of fire occurrence across Portugal, revealing a spatial pattern with clusters centred on the two main Portuguese city districts (Lisboa and Porto). Ultimately, the analysis of the relation between the observed urban fire incidence and the actual number of fire stations in each municipality allowed to underline the need for planning the spatial configuration of fire stations, both in number and location, at a regional scale.

1. Introduction

Urban fires, defined as any fires occurring within an urbanized area, represent a considerable proportion of the fire stations (FS) service and have significant financial and social impacts [1]. As a disaster that occurs with high frequency it affects ecosystems and human safety since ancient times [2]. Urban fires result in more than 300,000 deaths annually worldwide [3]. A total of 3,704 deaths and $14,8 billion direct property losses were registered due to urban fires in the United States in 2019 [2]. Studies in Europe reported, in the United Kingdom, around 212,500 fires during 2013 and 2014, involving 322 deaths and more than 9,700 non-fatal casualties [4]. A comprehensive study regarding urban fire events in Greece analysed fire incidents from 2000 to 2019 and concluded that residential fires represent 25,6% of the total incidents in Greece, with an average of 3,989 incidents per year [5]. In Australia, direct urban fire losses were estimated at $885 million per annum with an average of 0.07 per cent of GDP per annum [6]. In 2020, China registered 252,000 fire incidents resulting in 1,183 deaths and $621 million direct property losses [7]. Himoto et al. [8] studied fire events in Japan, emphasizing the great impact of fire events occurring at Sakata, in 1976, and at Hanshin-Awaji in 1995, being the last a consequence of a large earthquake. In 2019, a total of 31,061 urban fires occurred in Japan [9]. In Portugal, according to data supplied by the Portuguese National Emergency and Civil Protection Authority (ANEPC, *Autoridade Nacional de Emergência e Proteção Civil*) the average of urban fires, between 2012 and 2020, was around 8,841 events/year. Since 43% of the Portuguese population lives in urban areas, these numbers potentially carry severe consequences. High population density together with the complexity of the urban network imply that urban fires may have a significant impact in firefight operational costs, property damage and the potential loss of life. In addition, the proximity to industry or commercial areas typically increases urban fire probability [10]. Property impacts encompass the loss of buildings (*e.g.* residential, commercial, and industrial), structures, equipment and other types of properties caused by burning, smoking, radiation, demolition, collision, water stains and pollution during the firefighting [3]. Densely populated areas are subject to more harmful consequences in the sense that the potential loss of lives or other socio-economic primordial consequences, such as, *e.g.*, job losses, are more likely to

* Corresponding author at: NOVAMath Center for Mathematics and Applications, NOVA School of Science and Technology, Universidade NOVA da Lisboa, Caparica, Portugal.

E-mail address: r.bispo@fct.unl.pt (R. Bispo).

https://doi.org/10.1016/j.firesaf.2023.103802
Received 14 December 2022; Received in revised form 27 March 2023; Accepted 27 April 2023
Available online 6 May 2023

occur [2]. Ultimately, the combination of all these forms of losses may lead to substantial costs in firefighting operations and severely impact economies both at regional and national scales.

Despite the impact of fires in urban areas, research in urban fire prediction remains limited. According to Jin et al. [10], capturing the urban fire dynamics is an indispensable tool for urban security planning and fire emergency decision-making. Moreover, since public resources are limited, national and regional governments need research-based information to adequately plan how many and where firefighting facilities should be established [11] and decide how to distribute human, material, and financial resources [12].

In literature, several approaches are used to predict and model fire occurrences. Regardless of the modelling approach, most studies emphasize the need to consider spatial techniques. For instance, Corcoran et al. [13] recognize the importance of using spatial analytical techniques to model fire incident data and to explore their spatial dynamics. Song et al. [14] refer to the existence of spatial autocorrelation in the occurrence of fires in urban areas and that this correlation should be considered to model these occurrences. Moreover, Corcoran and Higgs [1] mention that the dynamics of urban fires is related to the neighbouring structure and socio-demographic characteristics. Clark et al. [15] present a review of social science literature and the implications on urban fire research. These authors point out that social and economic factors may influence the understanding of fire risk as they can strengthen spatial correlation. Thus, the use of statistics techniques which consider spatial dependence, arise as a natural tool to understand and predict the occurrence of urban fires.

In this context, several spatial techniques have been used. Spatial kernel density analysis, co-maps and other exploratory spatial data methods are commonly used to examine the spatial, temporal and spatial–temporal variations of fire events [16]. Song et al. [14] considered geographically weighted regression, geographically and temporally weighted regression, and global linear regression to model occurrences of urban fires in a city scale. Oliveira et al. [17] also considered geographically weighted regression to explore spatial patterns of fire density in Southern Europe. Spatial econometric models, as the ones used in this study, appear in literature potentially offering new insights into the modelling of fire occurrences as these allow to consider spatial autocorrelation in a flexible way, either in the response variable, the explanatory variables and/or the random error term [18,19].

The explanatory variables of fire occurrence included in models may differ substantially depending on considering forest or urban events. Unlike wildfires, urban fires are mainly linked to a (complex) set of technical, social, and economic factors. In fact, [20] defined urban fires as a social product which thus is, inherently, affected by various socio-economic factors.

Urban fire predictors usually encompass population density, residential density and industrial agglomeration, being these positively associated with the events occurrence, i.e, higher/lower densities are, on average, associated with higher/lower incidence fire rates [2,4]. The socio-economic predictors are typically organized into six broad categories: (1) Housing quality and condition; (2) Household structure and demographics; (3) Economic conditions; (4) Education; (5) Ethnicity, and (6) Crime.

Housing quality and condition encompass variables linked to the house usage and building characteristics. Jennings [21] did an extensive literature review on the relationship between social-economic characteristics and urban (residential) fire risk. This author reports that the incidence of fires in residential buildings is greatly linked to housing quality, social structure, household income, household overcrowding and general social conditions. Variables such as renter-occupied housing [13,22–24], vacant housing [25,26], social landlords presence [27] and housing age [13,23,25–27] are mentioned in literature as being, in general, positively associated with fire occurrence. However, Duncanson et al. [28] concluded that owner-occupied homes are less prone to fires occurrence. Quality construction materials are also included in this category of predictors. Low quality is reported as generically increasing the chance of fire occurrence [29].

Household structure and demographics cover factors as overcrowding [27,28], parental presence [13,28] and single parent household [30]. Overcrowding and the household structures including parents with children above 15 years old were found to be associated to higher levels of fire occurrence in studies conducted by Hastie and Searle [27], Duncanson et al. [28] and Chhetri et al. [30]. In contrast, parental presence and single parent households type seem to be associated with lower chance of fire occurrence [28,30].

Economic conditions encompass variables that reflect socio-economic disadvantages [26,31] as the income level [3,25,28], employment status [24,25,27,30], and poverty level [28]. Higher income levels are reported as being associated with lower fire incidence rates [28]. Accordingly, Shai [25] and Hu et al. [3] describe that low income levels are linked to a higher fire incidence. The same type of effect is described for poverty level [28]. Anderson-Bell et al. [24] state that employment status does not have a significant influence in fire occurrence. However, Chhetri et al. [30], Shai [25] and Hastie and Searle [27] claim that unemployment or long term-sick leaves increase the chance of fire occurrence.

The education category covers the schooling level effect. According to Duncanson et al. [28], Corcoran et al. [13], Hu et al. [3] and Anderson-Bell et al. [24] higher school or educational levels are, on average, associated with lower fire incidence levels.

Chhetri et al. [30] studied the ethnicity effect using the proportion of indigenous population as a predictor. These authors concluded that the presence of high proportions of indigenous people are associated to higher levels of fire incidents. Morgner and Patel [32] report that the co-existence of different ethnic groups in the same household is associated with a comparatively lower fire incidence. However, Anderson-Bell et al. [24] discarded any influence of ethnicity on fire occurrence.

The last category of socio-economic predictors includes crime related variables. Anderson-Bell et al. [24] categorized crimes in violent and burglary crimes. These authors report that the presence of violent crimes is positively associated with fire occurrence, while burglary crimes seem to be negatively associated with fire occurrence.

Given the above description, the objectives of this research are twofold. First, it aims at modelling urban fire occurrences while making a comparative analysis of different strategies to account for spatial autocorrelation. Aligned with the first goal, the second objective is to identify factors that contribute to explain the relationship between fire events and the urban pattern. Ultimately, this study aims at mapping the probability of urban fires occurrence in mainland Portugal.

In the subsequent sections we describe the data collected (Section 2), detail the methods (Section 3) and present the findings of our research (Section 4). In the last section (Section 5), we highlight the main conclusions of the work developed.

2. Data description

Urban fires occurrences in mainland Portugal, between 2012 and 2020, were supplied by ANEPC. Each record contained the date, time, and incident location (latitude and longitude). These data are regularly recorded for each municipality and stored automatically in a central national platform. The available records were averaged across the years, aggregated at a municipality level, and normalized to the area of the administrative unit. The obtained annual average of urban fires per municipality (target variable) ranged from 2.1×10^{-3} fires/km^2 to 6 fires/km^2, with 94% of the municipalities having an annual average under 1 urban fire/km^2.

The selection of explanatory variables to include in the models was based on factors that were simultaneously described in the literature as influencing urban fires occurrence (see Section 1) and publicly

Table 1
Description of the explanatory variables.

Variable name	Description
Population density	Number of inhabitants per km^2
Buildings density	Number of buildings per km^2
Degraded buildings density	Number of degraded buildings or in need of repair per km^2
Age of buildings	Average age of buildings (years)
Income	Average income per person (€)
Buying power	Consumer buying power (%)
Education	Secondary school completion rate (%)
Crime	Criminality rate (‰)

available at Statistics Portugal website (www.ine.pt) for extraction at a municipality level.

Variables characterizing the number of inhabitants, building structure and the socio-economic traits by municipality, were identified from the last available 2021 Portuguese Census and downloaded from Statistics Portugal website (www.ine.pt). The number of buildings, the number of degraded buildings or in need of repair and the age of buildings were selected to represent housing quality. Average income per person (€) and buying power (%) were chosen to express economic conditions. Finally, the secondary school completion rate (%) and criminality rate (‰) were elected to proxy education and crime levels, respectively.

Count variables (i.e., number of inhabitants, number of buildings and number of degraded buildings or that needs to be repaired) were normalized to the area of the administrative unit, yielding densities.

The final set of explanatory variables is summarized in Table 1.

It is worthwhile noting that as the target variable is the annual average urban fires per municipality, census variables are being used to proxy the annual average of the explanatory variables. The use of census-based proxies to augment the dataset by appending socio-economic variables is a common procedure to overcome data limitations [33]. In addition, there is evidence that census proxies can give reasonable approximations for models parameters, particularly in terms of sign and significance [34], which legitimates the procedure.

3. Methods

In this section we briefly introduce the main concepts and methods used to analyse and model the fire data described in Section 2.

3.1. Spatial autocorrelation

A key insight when analysing spatial data is that observations that are close to each other are likely to be related and, thus, cannot be assumed to be mutually independent, as commonly done in classic statistical analysis [35]. In this context, spatial autocorrelation arises when units that are close to one another present values more associated than units that are far apart. The connectivity between two spatial units i and j $(i, j = 1, \dots, n)$ can be formally defined by a matrix $\mathbf{W}$ with elements $w_{ij} = 1$ if $j \in N(i)$ and $w_{ij} = 0$, otherwise, with $N(i)$ being the set of neighbours of location i and, by definition, $w_{ii} = 0$. In this study we assumed the most straightforward definition of $\mathbf{W}$ provided by the binary matrix where $w_{ij} = 1$ if regions i and j share some common boundary, perhaps a vertex, and $w_{ij} = 0$, otherwise.

With the neighbours defined, spatial autocorrelation may be calculated globally or locally. Global measures create a single value that represents the entire data whilst local indicators give a measure for every single unit in the study area. Popular global measures for area level data include Moran's I [36] and Geary's C [37]. Local Indicators of Spatial Association (LISA) [38] include local Moran's I and Getis–Ord's G and G* [39]. These local indicators allow to define high and low spatial associations between neighbours. Regions with high/low association are commonly named as hot/cold spots, respectively.

3.2. Statistical modelling

Classic (non-spatial) linear regression can be generally represented by equation $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \epsilon$ where $\mathbf{y}$ is the vector $(y_1, \dots, y_n)$ (with y_i $(i = 1, \dots, n)$ being the realizations of the response variable and n the number of observations), $\mathbf{X}$ denotes an $n \times k$ matrix of explanatory variables, $\boldsymbol{\beta}$ represents the $k \times 1$ vector of (unknown scalar) parameters that quantify the contribution of the explanatory variables to the response variable, and $\epsilon_{n\times 1}$ is the error vector (composed by independent and identically normal distributed random variables) with zero mean and constant variance σ^2. As, in this approach, the parameters are commonly estimated by ordinary least squares (OLS), this model is often referred to as the OLS model. By assuming that observations are realizations of independent random variables, OLS models do not account for potential spatial dependence between the observations. However, when dealing with spatial data, ignoring spatial autocorrelation may produce biased and inconsistent estimates [40].

Spatial dependence can be incorporated in OLS models using several strategies, namely, by including an additional regressor in the form of (*i*) spatially lagged dependent variable, (*ii*) spatially lagged explanatory variables, and/or (*iii*) a spatially lagged error term.

The *spatial autoregressive model* (SAR) is easily obtained by combining the conventional regression OLS model with the spatially lagged dependent term, producing a spatial extension of the standard regression defined by

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \epsilon \quad |\rho| < 1 \tag{1}$$

where ϵ is the vector of random errors with independent elements $\epsilon_i \sim N(0, \sigma^2)$ $(i = 1, \dots, n)$ and σ^2 is the error variance. In this model, ρ represents a spatial autoregressive coefficient that takes the value zero if there is no spatial dependence in the vector of cross-sectional observations $\mathbf{y}$. The matrix $\mathbf{W}$, as previously defined, formally describes the geographical relationship between spatial units. Hence, $\mathbf{W}\mathbf{y}$ transforms the vector $\mathbf{y}$ through the spatial relationships defined in $\mathbf{W}$ and defines a spatial lag of $\mathbf{y}$.

Spatially lagged explanatory variables are important tools to use for spatial regression modelling in sharp contrast to the more widely used form of spatial regression modelling, where only the dependent variable is lagged. The model that only includes a *spatial lag of* $\mathbf{X}$ (SLX) may be described by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\gamma} + \epsilon \tag{2}$$

which includes two core components: the direct effects of the covariates represented by $\boldsymbol{\beta}$ and the indirect effects of the covariates in neighbours represented by the $k \times 1$ vector $\boldsymbol{\gamma}$. The model including both the spatially lagged dependent and independent variables is labelled the *spatial Durbin model* (SDM) [41] and defined by

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\gamma} + \epsilon. \tag{3}$$

The *spatial error model* (SEM) defined by

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{4a}$$

$$\mathbf{u} = \lambda \mathbf{W}\mathbf{u} + \epsilon \quad |\lambda| < 1 \tag{4b}$$

only includes the spatial dependence in the disturbance term, with λ being a spatial autocorrelation parameter and, from ((4)b), $\mathbf{u} = (\mathbf{I} - \rho\mathbf{W})^{-1}\epsilon$. Other spatial models may be defined by combining the different spatial repressors. In particular, the expressions

$$\mathbf{y} = \rho \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{5a}$$

$$\mathbf{u} = \lambda \mathbf{W}\mathbf{u} + \epsilon \tag{5b}$$

combine the spatially lagged dependent variable with a spatially lagged error term and define the so called *Kelejian–Prucha model* or SAC model [42,43]. This model is also mentioned in the literature as the SARAR model or the *Cliff-Ord model* [44]. It is also possible to combine

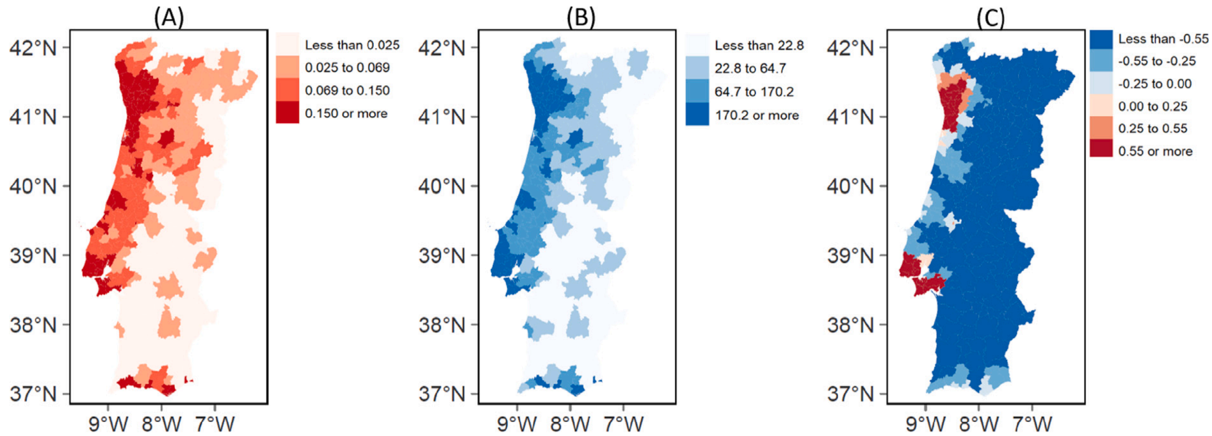

Fig. 1. (A) Observed urban fire incidence (annual average of urban fires per km^2), (B) Population density (number of inhabitants per km^2), and (C) Getis–Ord's Gi* indicator of local spatial association.

the spatially lagged independent variables with a spatially lagged error term which leads to the *spatial Durbin error model* (SDEM)

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{W}\mathbf{X}\gamma + \mathbf{u} \tag{6a}$$

$$\mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \epsilon. \tag{6b}$$

The most general form of a spatial linear regression model is called *General Nesting Spatial model* (GNS) (also mentioned in literature as the *Manski model*), includes all types of spatial effects and can be represented by [42,44]

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\gamma + \mathbf{u} \tag{7a}$$

$$\mathbf{u} = \lambda\mathbf{W}\mathbf{u} + \epsilon. \tag{7b}$$

In the above spatial models it is important to distinguish direct from indirect impacts (frequently mentioned as *spillover effects*). Direct impacts correspond to the effects that changes in one location may have in the response variable within that same region. The indirect effects are linked to the potential effect that changes in a certain location may have on the response variable of neighbouring regions. Note that indirect impact measures are only valid for models that include a spatially lagged variable. Thus, OLS model and SEM do not provide information regarding indirect effects.

To evaluate the importance of considering spatial dependence, several statistics may be used to formally test for spatial autocorrelation and decided on which model to fit. Moran's I and Geary's C statistics may be used to test for spatial autocorrelation. In addition, Moran's I test applied to the residuals of OLS model may be used to test for spatial dependence. However, none of these provide information on the type of spatial dependence in linear models. For this purpose, likelihood ratio (LR) tests may be used to guide model selection [40]. In this study, we adopt a *bottom-up* approach (*specific-to-general*) for model specification starting with the basic non-spatial OLS model, followed by tests for possible misspecification due to omitted spatial effects [45]. For this purpose, the likelihood ratio tests were used to guide model selection [40]. The values for Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and both the standard adjusted R^2 and the Nagelkerke pseudo R^2 [46] were used to compare the models relative fit. Models diagnosis was also based on the spatial predicted and residuals plots for all the models. To evaluate collinearity we examined the bivariate linear correlations matrix of explanatory variables and calculated the Variance Inflation Factor (VIF), which is a common tool used to measure multicollinearity. The values of VIF indicate the magnitude of the inflation in the standard errors (decrease in precision) due to multicollinearity. High values VIF (as a rule of thumb, above 5) indicate the existence of collinearity. Ultimately, the estimated probability of urban fires occurrence was determined and mapped across the country.

Table 2
Summary of global spatial autocorrelation measures.

Test	Statistic value	Std. Error	p-value
Moran's I	0.5535	0.0355	<0.001
Geary's C	0.3753	0.0750	<0.001

All statistical analyses were carried out in R, an open-source software environment for statistical computing [47]. Package `sf` [48] was used to read shapefiles and deal with geometry indexed objects. Spatial autocorrelation measures were calculated using package `spdep` [49]. Package `spatialreg` [50] was used to model data accounting for spatial dependence. All conclusions considered a 0.05 level of significance.

4. Results and discussion

This section presents the results of fitting the spatial models described in Section 3 and the main findings of our research.

4.1. Exploratory data analysis

Fig. 1(A) shows the observed annual average of urban fires per km^2 in mainland Portugal. The fire occurrence across the country is clearly asymmetrical, having a generic greater incidence along the coast and around main central cities (as Lisboa, Porto and Faro).

Fig. 1(B) depicts population density variation across the territory. There is a clear resemblance in both spatial patterns clearly identifying the most populated districts generally matching the ones with higher urban fire incidence.

To quantify the level of spatial dependence and identify clusters across space we have computed the global measures of spatial autocorrelation Moran's I and Geary's C (Table 2).

Both measures indicate that spatial clustering is significantly higher than what would be expected to find if the underlying spatial process would be at random ($p < 0.001$, Table 2). Getis–Ord's Gi* was further used as a local measure of spatial association to identify hot and cold spots (Fig. 1(C)). Locations associated with positive values of Gi* (Fig. 1(C), red clusters) point out administrative units that are highly associated with the neighbour municipalities, being designated as hot spots. These hot spots, mainly localized in Lisboa, Setúbal and Porto districts, reveal a citycentric spatial pattern. In contrast, locations with

Table 3
Summary of goodness-of-fit measures for all the fitted models.

	OLS	SDM	SLX	SAR	SAC	SEM	SDEM	GNS
Deviance	5.487	4.202	4.347	5.110	4.978	4.962	4.083	4.074
AIC	−292.293	−343.271	−339.005	−302.850	−305.199	−302.494	−348.164	−346.182
BIC	−274.155	−288.857	−288.219	−270.202	−268.923	−269.845	−293.749	−288.140
Log-likelihood	151.146	186.636	183.503	160.425	162.600	160.247	189.082	189.091
[a]R^2	0.959	0.969	0.967	0.962	0.963	0.962	0.969	0.969

[a]Adjusted R^2 for models OLS and SLX and Nagelkerke Pseudo R^2 for models SDM, SAR, SAC, SEM, SDEM and GNS.

Table 4
Summary of SDEM model.

		Estimate	Std. Error	z	p
Direct effects	Population density	0.00078	0.00003	25.63895	<0.00001
	Degraded buildings density	0.01193	0.00280	4.26222	0.00002
	Buying power	−0.15525	0.02124	−7.30998	<0.00001
Indirect effects	Population density	−0.00028	0.00007	−4.16253	0.00003
	Degraded buildings density	0.01669	0.00503	3.31827	0.00091
	Buying power	0.28089	0.04530	6.20031	<0.00001
	Intercept	0.04010	0.06889		

negative Gi* values (Fig. 1(C), blue clusters), point out administrative units that are weakly associated (or have no association) with the neighbours, being designated as cold spots.

4.2. Spatial models

The non-spatial OLS and the spatial SAR, SLX, SDM, SEM, SAC, SDEM and GNS models (Eqs. (1) to (7), Section 3), were fitted to data, considering the explanatory variables described in Table 1. Before including spatial dependency in the models, we have examined collinearity between these variables. The bivariate linear correlations between the buildings density and both the population and the degraded buildings densities were above 0.85. In addition, VIF values for these three variables were, respectively, equal to 10.2, 6.7 and 6.1. Removing the buildings density from the linear predictor kept all the VIF values under 5. Thus, this variable was dropped and the models with the remaining variables were fitted to the data. A summary of the obtained goodness-of-fit measures is shown in Table 3.

As mentioned, to select the best model we followed a *bottom-up* approach (see Section 3), starting by fitting the OLS model. This model presents the highest AIC and deviance values and the lowest log-likelihood and R^2 values (Table 3), clearly indicating this model as the worst possible option among the eight fitted models.

Fig. 2 represents schematically the model selection approach. Likelihood ratio (LR) test statistics supported the selection of SAR, SLX and SEM models over OLS model ($p < 0.0001$). SAR model was further compared with SAC model ($p = 0.0370$) and SDM ($p < 0.0001$) leading to SAR model rejection. SLX was rejected while compared with SDM ($p = 0.0123$) and SDEM ($p = 0.0008$) models. SEM was also rejected when compared with these same models ($p < 0.0001$) and with SAC model ($p = 0.0301$). SAC and SDM models were contrasted with GNS model, favouring the selection of the former ($p < 0.0001$ and $p = 0.0267$, respectively). Finally, SDEM was compared to GNS model. SDEM has the lowest deviance, AIC and BIC values (Table 3). On the other hand, considering the remaining goodness-of-fit measures, there is a mild preference for GNS over SDEM. However, we found no significant differences comparing GNS vs. SDEM ($p = 0.8935$). Thus, following the principle of parsimony, SDEM was preferred over GNS and ultimately selected as the most adequate model.

The estimated average number of annual urban fire events given by each fitted model was mapped for each municipality (Fig. 3). The estimates from SLX and SDM (Figs. 3(B) and (E)), generally underestimate the fire occurrences in the interior of the country and seem to overestimate it along the Portuguese coast. Contrarily, SAR, SEM and SAC overestimate urban fire occurrences in the interior of the country. The estimates from SDEM and GNS models (Fig. 3(G) and (H)) show the best approximations to the raw observations (Fig. 3(A)), when compared to the other models (Figs. 3(B) to 3(F)). As mentioned, the fit of these two models is quite similar and SDEM was chosen as the best model based on the parsimony principle.

Table 4 summarizes the results for the fitted SDEM model.

Direct effects show, as expected, a positive relationship between fire occurrences and the population density. This same effect in mentioned in literature by, *e.g*, Liu et al. [2] and Zhang et al. [4]. The degraded buildings density also has a direct positive impact on fire occurrences. This effect is in accordance with the effects described in the literature. In fact, as mentioned, poor housing conditions are typically associated with higher fire incidence (*e.g*, [23], Shai [25], Corcoran et al. [13], Hastie and Searle [27], Anderson and Ezekoye [26]). On other hand, consumer buying power is negatively related to the occurrence of fire events which is also in accordance with previous results [28, e.g].

These explanatory variables reveal significant indirect impacts, meaning that their variation in a given unit will impact the fire occurrence in the neighbouring units (spillover effect). Lower population densities in a particular municipality emerge associated to higher fire occurrences in the neighbouring regions. In fact, a lower level in the population density in a particular region is typically associated with more densely populated neighbouring regions, therefore increasing, on average, the number of fire events in those areas.

Overall, population density is the most relevant predictor. It is also worth noting that consumer buying power stands out for its considerable contribution both direct- and indirectly. The positive indirect effect might be a reflect of the social pattern predominantly found in central metropolitan areas where different municipalities within the same region typically present asymmetric economic and/or education levels [43]. Indeed, the central metropolitan areas, Lisboa and Porto, showed to have the highest association with neighbours (Fig. 1(B)).

Fig. 4(A) depicts the spatial distribution of SDEM residuals. These are generically low with the vast majority within the close vicinity of 0, confirming a global good model fit. Expectedly, relatively higher residuals appear within the hotspot regions where more extreme values, harder to accommodate by the model, were observed.

Finally, based on SDEM, we present the estimated probability of urban fires occurrence mapped across mainland Portugal (Fig. 4(B)).

In general, municipalities with higher estimated probabilities of urban fire occurrence are centred in the main districts, Lisboa and Porto. Within these districts, the municipalities with higher predicted incidence of urban fires per km^2 (over 4 events/year) are Amadora, Porto and Odivelas. These municipalities are the ones characterized by

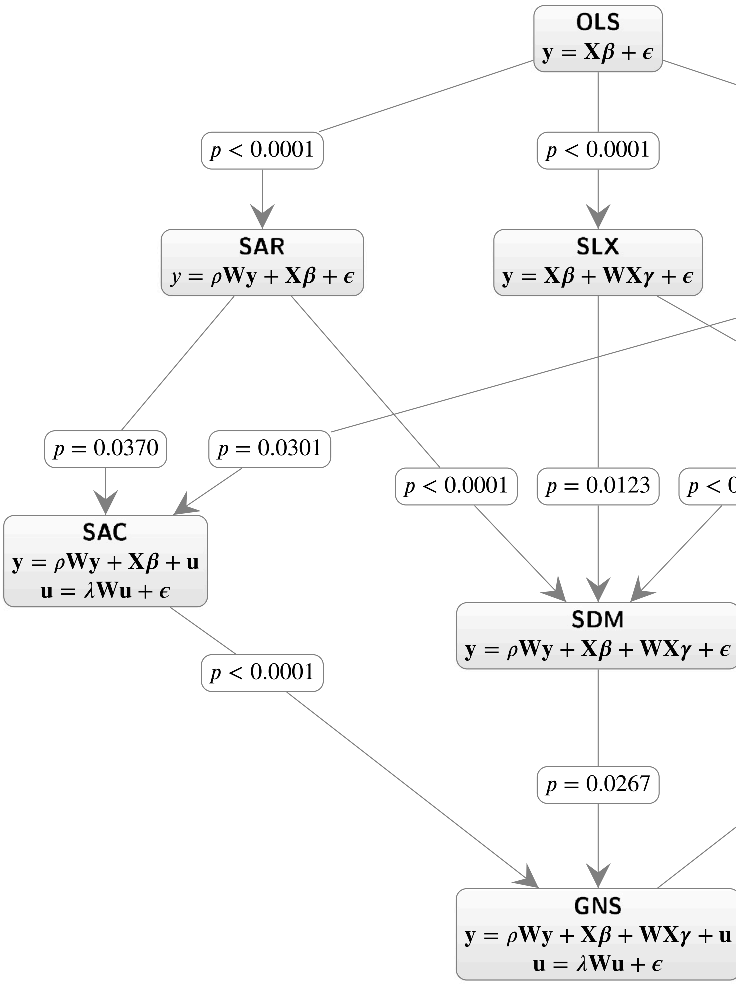

Fig. 2. Model specification: *Specific-to-general* approach. Arrows link the models being compared and point towards the selected model according to likelihood ratio tests (respective p-values given in the white background rectangles). Final selected model (SDEM) in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the highest population density in Portugal (> 5000 inhabitant/km^2). Porto municipality has also the highest building obsolescence (> 70 buildings/km^2, in sharp contrast with the Portuguese mean around 3 buildings/km^2). As mentioned, socio-economic variables have shown to be relevant while modelling urban fire events, namely consumer buying power. In this regard, the higher predicted fire risk in Amadora and Odivelas is likely associated with characteristic asymmetric economic context within Lisboa district. In fact, Lisboa municipality, which is a neighbour of both Amadora and Odivelas, presents the highest proportion of consumer buying power within the district and this variable has a strong positive effect in the neighbouring regions (Table 4). As such, this indirect impact explains the increased predicted fire risk in these two neighbour municipalities.

As underlined previously, one of the most important questions regarding regional and national firefighting planning is related to find the optimal FS spatial configuration both in terms of number and location [11]. Having this issue in mind, we related the observed urban fire incidence based with the actual number of FD in each municipality. Fig. 5 depicts graphically this relationship showing a linear relation between the two quantities ($r^2 = 0.65$, $p < 0.001$).

The municipalities with the highest fire incidence – Amadora and Porto – are clearly under the regression line, having less FS than what would be expected given their fire incidence. In contrast, Oeiras, Odivelas and São João da Madeira are distinctly above the line, thus, having more FS than what would be expected based on the regression. This simple linear model allowed us to find the expected number of FS for each municipality as a function of the annual average of urban fires per km^2. For Amadora one would expect to find 3 FS and yet it has only 1. In turn, the expected number of FS in Porto is 5, but it has only 3. In contrast, for instance, Oeiras has 7 FS but it should be expected to have only 4 given the observed fire incidence. However, having noticed that, it is important at this point to stress that the above comment, although seems note worth, elapses from a rather simplistic analysis lacking in considering fundamental factors such as, *e.g.*, both the occurrence of rural fires and the FS dimensions and resources.

5. Conclusions

In this work we studied comparatively seven different spatial modelling strategies to model urban fire occurrence across mainland Portugal. Considering space dependence in addition to a range of social-economic explanatory variables has proven to strengthen the validity of the fitted model in opposition to adopting the standard classical regression approach.

The spatial Durbin error model (SDEM) was selected among the several modelling approaches as having the best fit, explaining around 97% of the variability in data. Based on this model we found evidence that variables such as population density, degrade buildings density and buying power have both direct and indirect significant impacts. On average, as expected, the increase of the population density (associated with the increase of building density) and the number of degraded buildings or in need of repair have shown to increase the average

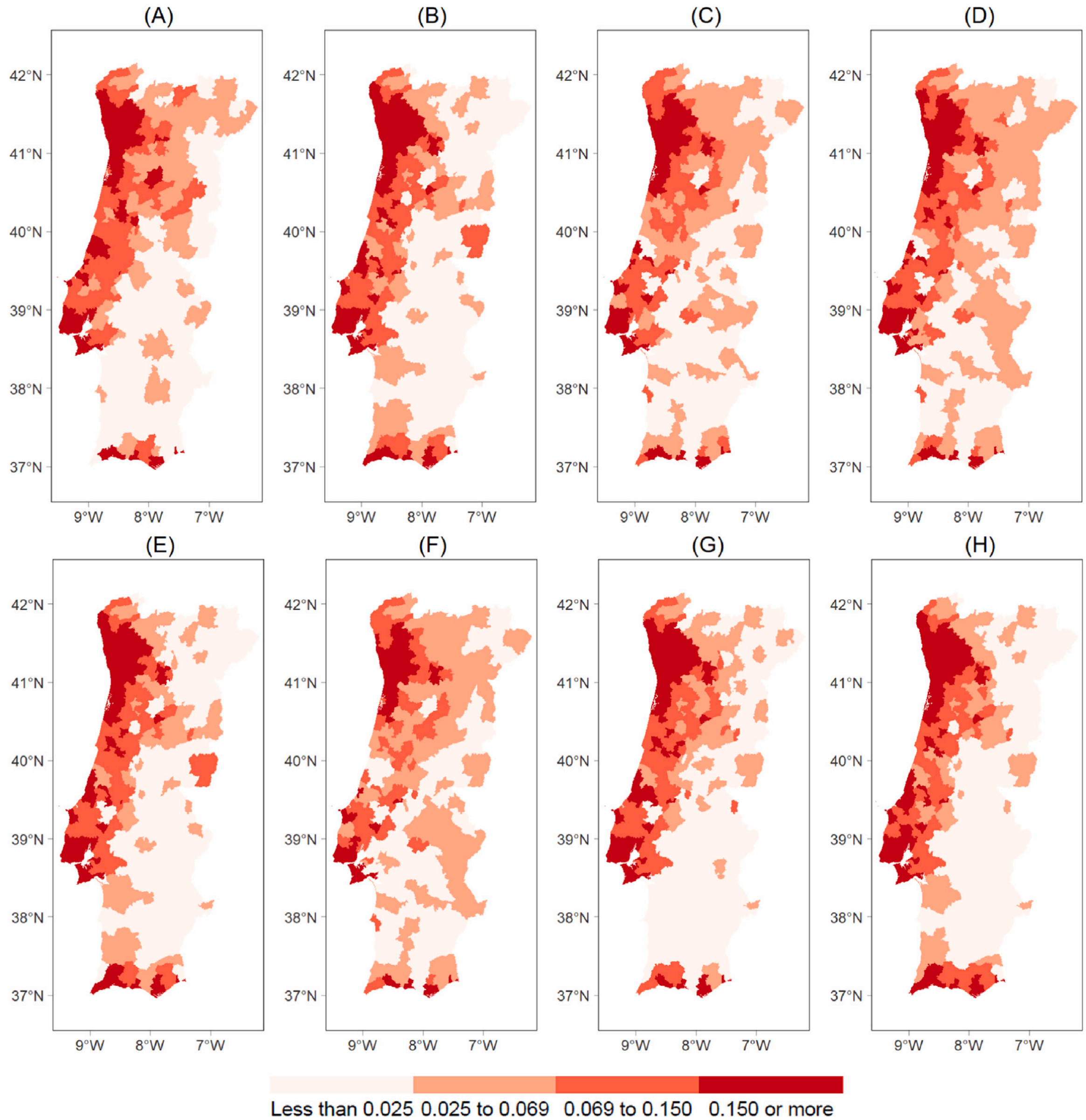

Fig. 3. (A) Observed urban fire incidence (annual average of urban fires per km^2) and predicted urban fire incidence by models (B) SLX, (C) SAR, (D) SEM, (E) SDM, (F) SAC, (G) SDEM and (H) GNS.

number of urban fires both in own- and neighbouring municipalities. In contrast, on average, the increase of the buying power in a particular municipality decreases directly the number of urban fires and, by a spillover effect, increases indirectly the number of urban fires in the neighbour municipalities. These effects are in line with the relations described in literature. In fact, high population density is one of the most frequent predictors associated with the increase of urban fire probability [10, e.g]. Factors associated with socio-economic disadvantages, such as building degradation and low buying power, are often mentioned in literature as increasing the fire risk [31, e.g.]. Negative spatial effects as the ones found in relation to the buying power are typically associated with the urban social patterns that commonly characterize central metropolitan areas where different municipalities, within the same area, frequently present sharp asymmetries in both economic and education domains [43, e.g.].

The fitted spatial model allowed to map the estimated probability of fire occurrence across Portugal. By allowing to easily visualize the region's most prone to fire incidence this map may be a valuable tool namely for planning and implementing education programs to promote a more informed population about fire risks and safety measures. In addition, by exposing a clear spatial pattern with clusters centred on the two main Portuguese city districts (Lisboa and Porto) it also highlights the impact that an adequate territorial planning, focused on areas with high population/building densities, could have on preventing urban fires.

The relation between the observed urban fire incidence and the actual number of FS in each municipality highlights the need for planning the spatial configuration of fire stations, both in number and location, at a regional scale. Since this configuration may considerably influence the effectiveness of the provided services, national and regional governments need research-based advice on how many and where firefighting facilities should be established. Hence, we suggest further research aiming at optimizing the number and locations of FS to ensure the best possible system performance. Moreover, current Portuguese legislation [51] defines the items that must be deployed when considering the degree of readiness for urban firefighting. However, these legal requirements address risk assessment and risk treatment considering only the notion of "special buildings" (*e.g.*, schools, hospitals, shopping

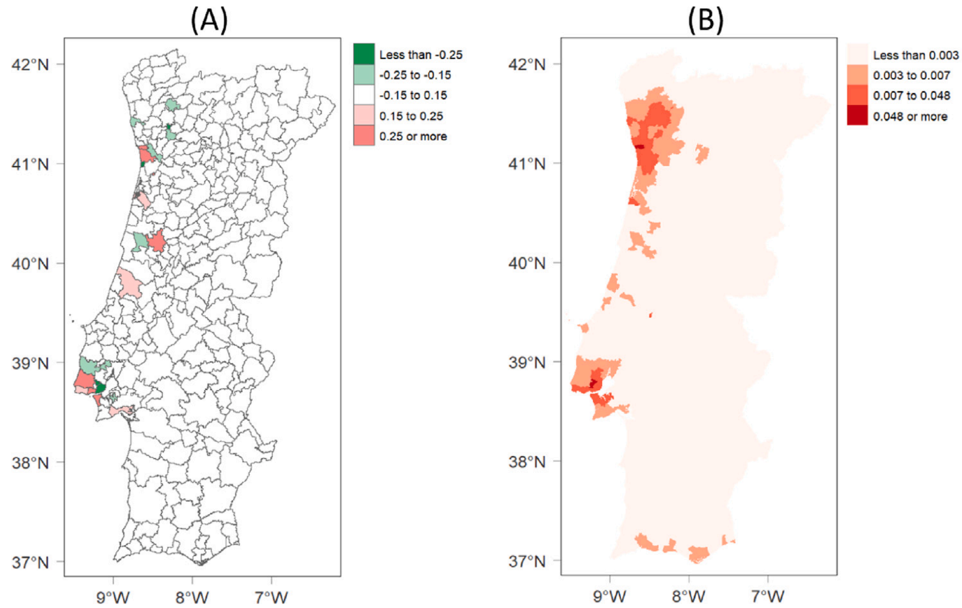

Fig. 4. (A) Spatial distribution of SDEM residuals and (B) Estimated probability of fire occurrence based on SDEM model predictions.

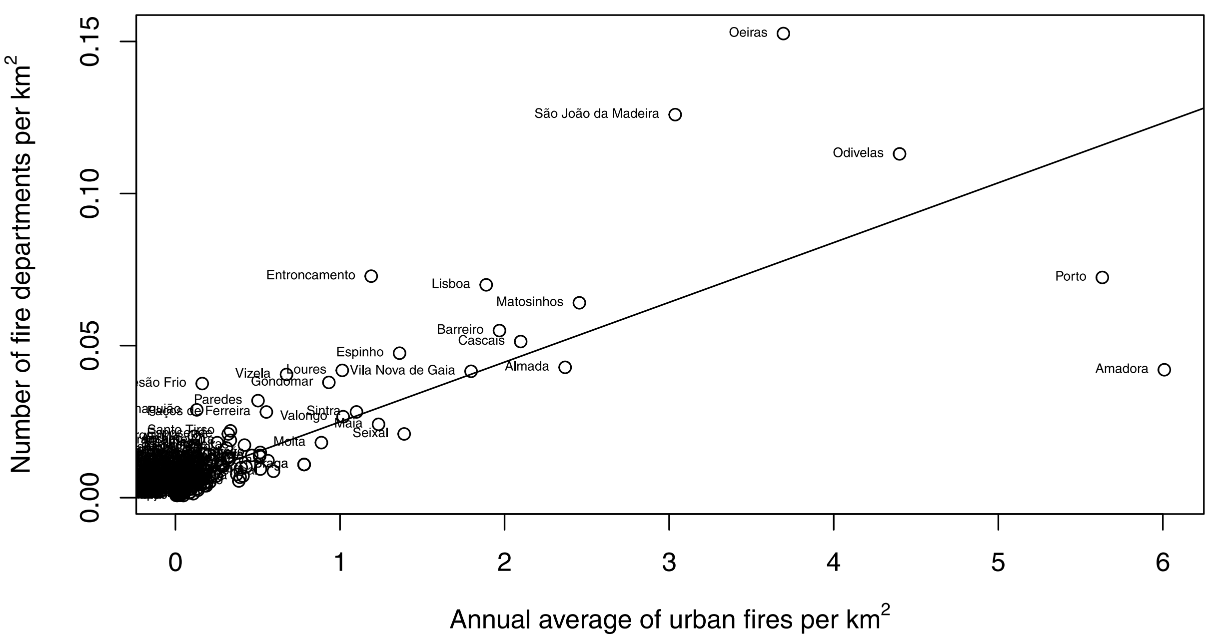

Fig. 5. Relation between the number of fire departments and the annual average of urban fires per km^2.

malls). Our results suggest that considering the probability of occurrence of urban fires, while defining risk assessment and risk treatment, could potentially contribute towards a more accurate definition of these concepts.

To the best of our knowledge, this is the first study that models the urban fire incidence using spatial modelling techniques in relation to socio-economic characteristics on a global scale in mainland Portugal. We conclude suggesting that spatial analytical techniques should be further applied in main districts to explore local dynamics and model the relationship with social-economic and -demographic features on a micro-level urban fire incident data.

Funding

This work is funded by national funds through the FCT - Fundação para a Ciência e a Tecnologia, I.P., under the scope of the project DSAIPA/DS/0088/2019 and research and development units UNIDEMI (project UIDB/00667/2020) and NOVAMath (projects UIDB/00297/2020 and UIDP/00297/2020).

CRediT authorship contribution statement

Regina Bispo: Conceptualization, Methodology, Data analysis. **Francisca G. Vieira:** Data analysis, Writing – original draft. **Filipe J. Marques:** Methodology, Writing – original draft. **António Grilo:** Funding.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We intend to publish a data descriptor on the used data.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.firesaf.2023.103802.

References

[1] J. Corcoran, G. Higgs, Fire Saf. J. 62 (2013) 1–2, http://dx.doi.org/10.1016/j.firesaf.2013.11.001.
[2] D. Liu, Z. Xu, C. Fan, Nat. Hazards 97 (3) (2019) 1175–1189, http://dx.doi.org/10.1007/s11069-019-03694-1.
[3] J. Hu, X. Shu, S. Xie, S. Tang, J. Wu, B. Deng, Fire Saf. J. 110 (2019) 102890, http://dx.doi.org/10.1016/j.firesaf.2019.102890.
[4] X. Zhang, J. Yao, K. Sila-Nowicka, ISPRS Int. J. Geo-Inform. 7 (1) (2018) 7, http://dx.doi.org/10.3390/ijgi7010007.
[5] K. Georgiadis-Filikas, I. Bakas, K. Kontoleon, Fire Technol. 58 (5) (2022) 3191–3233, http://dx.doi.org/10.1007/s10694-022-01308-0, URL https://link.springer.com/10.1007/s10694-022-01308-0.
[6] P. Chhetri, J. Corcoran, R.J. Stimson, R. Inbakaran, Geograph. Res. 48 (1) (2010) 75–85, http://dx.doi.org/10.1111/j.1745-5871.2009.00587.x.
[7] Z.-G. Liu, X.-Y. Li, G. Jomaas, Int. J. Disaster Risk Reduct. 78 (2022) 103138, http://dx.doi.org/10.1016/j.ijdrr.2022.103138, URL https://linkinghub.elsevier.com/retrieve/pii/S2212420922003570.
[8] K. Himoto, M. Shinohara, A. Sekizawa, K.-i. Takanashi, H. Saiki, Fire Saf. J. 96 (2018) 105–114, http://dx.doi.org/10.1016/j.firesaf.2018.01.003, URL https://linkinghub.elsevier.com/retrieve/pii/S0379711217303697.
[9] H. Yoshioka, K. Himoto, K. Kagiya, Fire Technol. 56 (5) (2020) 1885–1901, http://dx.doi.org/10.1007/s10694-020-00960-8, URL http://link.springer.com/10.1007/s10694-020-00960-8.
[10] G. Jin, C. Zhu, X. Chen, H. Sha, X. Hu, J. Huang, IOP Conf. Ser. Mater. Sci. Eng. 853 (1) (2020) 012050, http://dx.doi.org/10.1088/1757-899x/853/1/012050.
[11] S.B. Agbola, O.J. Falola, Int. J. Disaster Risk Reduct. 54 (2021) 102035, http://dx.doi.org/10.1016/j.ijdrr.2021.102035.
[12] D. Rohde, J. Corcoran, P. Chhetri, Comput. Environ. Urban Syst. 34 (1) (2010) 58–69, http://dx.doi.org/10.1016/j.compenvurbsys.2009.09.001.
[13] J. Corcoran, G. Higgs, C. Brunsdon, A. Ware, P. Norman, Comput. Environ. Urban Syst. 31 (6) (2007) 623–647, http://dx.doi.org/10.1016/j.compenvurbsys.2007.01.002.
[14] C. Song, M.-P. Kwan, J. Zhu, Int. J. Environ. Res. Public Health 14 (4) (2017) 396, http://dx.doi.org/10.3390/ijerph14040396.
[15] A. Clark, J. Smith, C. Conroy, J. Risk Res. 18 (9) (2014) 1113–1129, http://dx.doi.org/10.1080/13669877.2014.913660.
[16] J. Yao, X. Zhang, ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inform. Sci. XLI-B2 (2016) 63–69, http://dx.doi.org/10.5194/isprsarchives-xli-b2-63-2016.
[17] S. Oliveira, J.M. Pereira, J. San-Miguel-Ayanz, L. Lourenço, Appl. Geogr. 51 (2014) 143–157.
[18] C. Song, M.-P. Kwan, W. Song, J. Zhu, Sustainability 9 (5) (2017) 819, http://dx.doi.org/10.3390/su9050819.
[19] M.L. Loureiro, J. Barreal, Forest Syst. 24 (2) (2015) e022, http://dx.doi.org/10.5424/fs/2015242-05713.
[20] J.H. Wang, J.H. Sun, S.M. Lo, L. Gao, R. Yuen, Procedia Eng. 11 (2011) 437–444, http://dx.doi.org/10.1016/j.proeng.2011.04.680, URL https://www.sciencedirect.com/science/article/pii/S187770581100868X The 5th Conference on Performance-based Fire and Fire Protection Engineering.
[21] C.R. Jennings, Fire Saf. J. 62 (2013) 13–19, http://dx.doi.org/10.1016/j.firesaf.2013.07.002.
[22] R. Ardianto, P. Chhetri, S. Dunstall, Weber, T., McPhee, M.J. and Anderssen, R.S. (Eds) MODSIM2015, 21st International Congress on Modelling and Simulation, Modelling and Simulation Society of Australia and New Zealand, 2015, http://dx.doi.org/10.36334/MODSIM.2015.M4.ardianto, URL http://www.mssanz.org.au/modsim2015/M4/ardianto.pdf.
[23] O. Špatenková, K. Virrantaus, Fire Saf. J. 62 (2013) 49–63, http://dx.doi.org/10.1016/j.firesaf.2013.07.001.
[24] J. Anderson-Bell, C. Schillaci, A. Lipani, Remote Sens. Appl. Soc. Environ. 21 (2021) 100470, http://dx.doi.org/10.1016/j.rsase.2021.100470, URL https://linkinghub.elsevier.com/retrieve/pii/S2352938521000069.
[25] D. Shai, Public Health Rep. 121 (2) (2006) 149–154, http://dx.doi.org/10.1177/003335490612100208, URL http://journals.sagepub.com/doi/10.1177/003335490612100208.
[26] A. Anderson, O.A. Ezekoye, J. Fire Protect. Eng. 23 (1) (2013) 51–75, http://dx.doi.org/10.1177/1042391512469521.
[27] C. Hastie, R. Searle, Fire Saf. J. 84 (2016) 50–56, http://dx.doi.org/10.1016/j.firesaf.2016.07.002, URL https://linkinghub.elsevier.com/retrieve/pii/S0379711216300832.
[28] M. Duncanson, A. Woodward, P. Reid, Fire Saf. J. 37 (2) (2002) 165–179, http://dx.doi.org/10.1016/S0379-7112(01)00033-9, URL https://linkinghub.elsevier.com/retrieve/pii/S0379711201000339.
[29] M. Shokouhi, K. Nasiriani, H. Khankeh, H. Fallahzadeh, D. Khorasani-Zavareh, J. Injury Violence Res. 11 (1) (2019) http://dx.doi.org/10.5249/jivr.v11i1.1059.
[30] P. Chhetri, J. Corcoran, R.J. Stimson, R. Inbakaran, Geograph. Res. 48 (1) (2010) 75–85, http://dx.doi.org/10.1111/j.1745-5871.2009.00587.x, URL https://onlinelibrary.wiley.com/doi/10.1111/j.1745-5871.2009.00587.x.
[31] T. Buffington, J.G. Scott, O.A. Ezekoye, Comput. Environ. Urban Syst. 88 (2021) 101633.
[32] C. Morgner, H. Patel, Fire Saf. J. 125 (2021) 103384, http://dx.doi.org/10.1016/j.firesaf.2021.103384, URL https://linkinghub.elsevier.com/retrieve/pii/S0379711221001259.
[33] A. Geronimus, J. Bound, L. Neidert, On the Validity of Using Census Geocode Characteristics to Proxy Individual Socioeconomic Characteristics, Technical Report, National Bureau of Economic Research, 1995, http://dx.doi.org/10.3386/t0189.
[34] M. Adjemian, J. Williams, Transportation 36 (2) (2009) 223–241, http://dx.doi.org/10.1007/s11116-009-9191-2.
[35] R. Bivand, G. Piras, J. Stat. Softw. 63 (18) (2015) http://dx.doi.org/10.18637/jss.v063.i18.
[36] P. Moran, Biometrika 37 (1/2) (1950) 17, http://dx.doi.org/10.2307/2332142.
[37] R. Geary, Incorporated Stat. 5 (3) (1954) 115, http://dx.doi.org/10.2307/2986645.
[38] L. Anselin, Geograph. Anal. 27 (2) (2010) 93–115, http://dx.doi.org/10.1111/j.1538-4632.1995.tb00338.x.
[39] A. Getis, J.K. Ord, Geograph. Anal. 24 (3) (2010) 189–206, http://dx.doi.org/10.1111/j.1538-4632.1992.tb00261.x.
[40] J. LeSage, R.K. Pace, Introduction to Spatial Econometrics, Chapman and Hall/CRC, 2009, http://dx.doi.org/10.1201/9781420064254.
[41] L. Anselin, Spatial Econometrics: Methods and Models, Springer Science & Business Media, 1988.
[42] J.P. Elhorst, Spatial Econ. Anal. 5 (1) (2010) 9–28, http://dx.doi.org/10.1080/17421770903541772.
[43] A.B. Golgher, P.R. Voss, Spatial Demography 4 (3) (2015) 175–205, http://dx.doi.org/10.1007/s40980-015-0016-y.
[44] A. Furková, CEJOR Cent. Eur. J. Oper. Res. 27 (3) (2018) 815–834, http://dx.doi.org/10.1007/s10100-018-0581-4.
[45] R.J. Florax, H. Folmer, S.J. Rey, Reg. Sci. Urban Econ. 33 (5) (2003) 557–579, http://dx.doi.org/10.1016/s0166-0462(03)00002-4.
[46] N. Nagelkerke, Biometrika 78 (3) (1991) 691–692, http://dx.doi.org/10.1093/biomet/78.3.691.
[47] R. Core Team, R: A language and environment for statistical computing, 2021, URL https://www.R-project.org/.
[48] E. Pebesma, the R Journal 10 (1) (2018) 439–446, http://dx.doi.org/10.32614/RJ-2018-009.
[49] R.S. Bivand, E. Pebesma, V. Gomez-Rubio, Applied Spatial Data Analysis with R, Second Edition, Springer, NY, 2013, URL https://asdar-book.org/.
[50] R. Bivand, G. Millo, G. Piras, Mathematics 9 (11) (2021) 1276, http://dx.doi.org/10.3390/math9111276.
[51] Despacho n 8903/2020 aprovação da nota técnica n° 8 - grau de prontidão dos meios de socorro, 2020, DiÁrio Da República N.° 182/2020, Série II, 2020-09-17 URL https://dre.pt/dre/detalhe/despacho/8903-2020-142961553.