




Regularised Semi-parametric Composite Likelihood Intensity Modelling of a Swedish Spatial Ambulance Call Point Pattern

Fekadu L. BAYISA, Markus ÅDAHL, Patrik RYDÉN, and Ottmar CRONIE 

Motivated by the development of optimal dispatching strategies for prehospital resources, we model the spatial distribution of ambulance call events in the Swedish municipality Skellefteå during 2014–2018 in order to identify important spatial covariates and discern hotspot regions. Our large-scale multivariate data point pattern of call events consists of spatial locations and marks containing the associated priority levels and sex labels. The covariates used are related to road network coverage, population density, and socio-economic status. For each marginal point pattern, we model the associated intensity function by means of a log-linear function of the covariates and their interaction terms, in combination with lasso-like elastic-net regularized composite/Poisson process likelihood estimation. This enables variable selection and collinearity adjustment as well as reduction of variance inflation from overfitting and bias from underfitting. To incorporate mobility adjustment, reflecting people's movement patterns, we also include a nonparametric (kernel) intensity estimate as an additional covariate. The kernel intensity estimation performed here exploits a new heuristic bandwidth selection algorithm. We discover that hotspot regions occur along dense parts of the road network. A mean absolute error evaluation of the fitted model indicates that it is suitable for designing prehospital resource dispatching strategies.

Supplementary materials accompanying this paper appear online.

Key Words: Bandwidth selection; Cyclic coordinate descent algorithm; Emergency alarm; Inhomogeneous Poisson process; Lasso-like elastic-net; Multivariate point process.

1. INTRODUCTION

In many (if not most) countries/regions, the prehospital resources available, e.g., ambulances, are scarce, and this clearly affects the availability and general quality of the pre-

F. L. Bayisa · M. Ådahl · P. Rydén, Department of Mathematics and Mathematical Statistics, Umeå University, 901 87 Umeå, Sweden. F. L. Bayisa, Department of Mathematics and Statistics, Auburn University, Auburn, AL 36849, USA. O. Cronie (✉) Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 412 96 Gothenburg, Sweden. School of Public Health and Community Medicine, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden (E-mail: ottmar@alumni.chalmers.se).

© 2023 The Author(s)

Journal of Agricultural, Biological, and Environmental Statistics, Volume 28, Number 4, Pages 664–683
<https://doi.org/10.1007/s13253-023-00534-5>

hospital service in question. For example, a small decrease in the median response time, which could be the effect of the addition of an ambulance, could save a significant number of lives annually. However, simply adding additional ambulances to a given region might not necessarily yield the expected outcome, i.e., drastically decreased response times, but instead the main challenge might be one of resource optimisation. More specifically, it may very well be the case that the existing resources are not utilised optimally; note that, given an expected performance outcome in terms of the response time distribution, one might have to both add and optimise prehospital resources. Moreover, in most/all cases, the addition of a new ambulance comes at a relatively high cost, so simply adding ambulances until the expected outcome has been achieved is likely never a viable option. Such optimisation can be quite challenging and extensive; in 2018, in Sweden, approximately 660 ambulances responded to roughly 1.2 million *ambulance/emergency alarm calls*, i.e., events defined by the deployment of an ambulance. In addition, the deployment of ambulances in 2018 cost more than 4 billion Swedish krona (Byrsell et al. 2021).

Carrying out resource optimisation within a prehospital care organisation/system, which to a large degree boils down to placing ambulance stations at locations for which, e.g., the median response time is minimised annually, requires knowledge/ understanding of where (and when) calls tend to occur spatially (and temporally). In essence, minimising the response times of ambulances is crucial for obtaining the desired clinical outcomes following ambulance calls (O'keeffe et al. 2011).

The call occurrence risk can be influenced by various underlying factors, such as demographic factors. More specifically, within the spatial context, if one is able to exploit different covariates or predictors to model the expected number of calls occurring within a given region, within an arbitrary period of time, then one can obtain an understanding of which spatial population characteristics, e.g., age distribution in a given subregion, that are driving the risk of an ambulance call occurring at a given spatial location. In addition, such a model may be extended to be exploited for predictive purposes, most notably within a spatio-temporal context. One can easily imagine that the distribution of calls (in Sweden) is complex, dynamical, and heterogeneously distributed within the spatial study region. This work is aimed at describing the spatial dynamics of ambulance calls in the municipality of Skellefteå, Sweden. The main focus of our study is to generate a risk-map for the ambulance calls, which is modelled by means of different spatial covariates, and to identify hotspot regions in Skellefteå, which can play a key role in designing optimal dispatching strategies for prehospital resources. It also focuses on identifying spatial covariates that tend to influence the occurrence of call events. Hence, we aim for a model with good predictive performance.

Note that, a priori, we do not know how many calls there will be within the spatial region studied and within the timeframe considered, which here is given by the years 2014–2018. In addition, we have access to the exact (GPS) locations of the events, as well as additional information, so-called marks, attached to each event. Such datasets are commonly referred to as (*marked*) *point patterns*, and their natural modelling framework is that of (*marked*) *spatial point processes* (Baddeley et al. 2015), which may be thought of as generalised random samples with the properties that the number of points/elements in the sample is allowed to be random and the points are allowed to be dependent. Here, in our specific dataset, each

event has two mark components attached to each spatial location: the priority/risk level (1 or 2) of the call and the gender (male or female) of the patient associated with the call.

The univariate properties of a point process are most commonly characterised through its intensity function. In most applications, it is unrealistic to assume that the underlying point process is homogeneous and there are various approaches available to deal with the spatial inhomogeneity of the points/events (Baddeley et al. 2015). Quite often, one has access to a collection of spatial covariates, which may be used to model the spatially varying intensity function. In particular, it is commonly assumed that the intensity has a log-linear form. Here, common practice is to first model the intensity function using a Poisson process log-likelihood function, which is given by a closed-form expression of only the (assumed) intensity function. In the case of a general point process, although a Poisson process is a point process with independent points/events and Poisson distributed counts within any subregion, the indicated intensity estimation approach still has good large sample properties (Coeurjolly and Lavancier 2019).

Once an intensity function estimate is obtained, one would proceed with analysing and modelling possible spatial interaction (Baddeley et al. 2000, 2015). It should here be emphasised that observed clumps of points in a point pattern may be the result of either inhomogeneity, spatial interaction (clustering/aggregation/attraction) or both. Since our main interest here is generating a spatial risk map for the observed events, i.e., the ambulance calls, which can be exploited for predictive purposes and also used as input in the problem of designing an optimal dispatching strategy of ambulances, we will solely focus on the former, i.e., modelling the intensity function as a log-linear function of a collection of spatial covariates. This allows us to address one of the main objectives of this work, which is to identify hotspot regions of the events. A further main objective of this work is to select covariates governing the intensity function. The original spatial covariates and the created spatial covariates used in this study are detailed in Sections S4 and S5.

Aside from including the individual covariates in our model framework, we also include the cross-/interaction-terms given by the products of the individual covariates, since there may potentially be interactive effects of the covariates. This results in a high-dimensional data setting (a total of nine hundred eighty-nine covariates), with possibly hard to interpret cross-terms, and it is clearly a challenge to identify which covariates sensibly explain the actual intensity of the events. As a solution, one may apply regularisation when fitting the model, which, aside from carrying out variable selection, also reduces variance inflation from overfitting and bias from underfitting. In addition, we want to adjust for the fact that the demographic covariates we have access to only reflect where different demographic groups live, and not how they move around. Since we do not have access to any human mobility covariates, we have to be pragmatic and use some proxy for such covariates. Our solution is to additionally include a nonparametric intensity estimate of the events as an additional covariate, yielding a semi-parametric approach. People may require an ambulance not only at home or at work. The nonparametric intensity estimate then provides the likelihood of incidents along a continuum, and we believe that we can catch some of this mobility by incorporating the nonparametric intensity estimate.

These added covariates would thus (to some degree) represent the portion of the intensity function that the original spatial covariates could not explain. Here we have another benefit

of the regularisation: having accounted for the other covariates, if this added covariate has little/no influence on the intensity function, then the regularisation should indicate this. Many of the spatial covariates which we deal with are most likely strongly correlated, which calls for some collinearity adjustment. Moreover, variable selection in the modelling of the spatially varying intensity function is most likely necessary in the modelling of ambulance call events. Tibshirani (1996) introduced a penalised likelihood procedure, which has been a cornerstone in the development of variable selection methods via regularisation (or penalisation). The idea of Tibshirani (1996) is to add a least absolute shrinkage and selection operator (lasso) penalty to the loss function, often the likelihood function, to shrink small coefficients of covariates to zero while retaining covariates with large coefficients in the model, thus leaving us with a sparse model with highly influential covariates. Hence, the approach simultaneously performs variable selection and parameter estimation. A plethora of regularisation methods, such as elastic-net (Zou and Hastie 2005) and adaptive lasso (Zou 2006), have been developed subsequent to the work of Tibshirani (1996). A predecessor of the lasso penalty, the ridge penalty, which is an ℓ_2 -penalty, may be combined linearly with the lasso ℓ_1 -penalty to obtain elastic-net regularisation, which is used to select variables and shrinks the coefficients of correlated variables to each other. It should be emphasised that these shrinkage/regularisation methods have the effect of balancing estimation bias and variance, which is an additional motivation for their employment. Turning to the point process context, also here regularization/variable selection has been developed to reduce variance inflation from overfitting and bias from underfitting. Yue and Loh (2015), Renner and Warton (2013) and Thurman and Zhu (2014) have used regularisation methods such as lasso, adaptive lasso, and elastic-net to model spatial point processes. Choiruddin et al. (2018) further extended the above works to include a larger range of models and penalties. We here essentially follow the track of Yue and Loh (2015).

With regard to the estimation of the regularised models, Efron et al. (2004) have developed the least angle regression to estimate the entire lasso regularisation paths. According to Friedman et al. (2007) and Friedman et al. (2010a), in comparison with the least angle regression algorithm, the cyclical coordinate descent algorithm computes the regularisation paths of different regularisation methods with lower computational costs. In this study, the cyclical coordinate descent method has been used to estimate the entire regularisation paths since it is computationally fast on large datasets (Fercoq and Richtárik 2016). The general idea of the coordinate descent algorithm is that the objective function is optimised with respect to a single parameter at a time, and the optimisation of the objective function is iteratively carried out for all parameters until a convergence criterion is fulfilled. In this work, the objective function represents a regularised quadratic approximation to the log-likelihood function of an inhomogeneous spatial Poisson process.

Two approaches have been used to evaluate the performance of the proposed model. The first approach involves training the proposed model on the whole dataset. We treat the trained (or estimated) intensity function on the whole dataset as the true intensity function of the call events. Based on the estimated intensity function, p -thinning/undersampling, using 70% of the whole dataset, has been used to generate a hundred datasets, which are then used to generate a hundred intensity images. Using quantiles and mean absolute errors between pixel-by-pixel differences of the estimated image (i.e., image based on the whole

dataset) and the hundred estimated intensity images (i.e., the intensity images based on the hundred undersampled datasets) can be used to evaluate the stability of the proposed model in estimating the intensity function of the emergency alarm call events. The second approach is to visually evaluate the performance of the proposed model. We train the proposed model on 70% and validate it on the remaining 30% of the whole dataset. Then, we compare the hotspot regions in the estimated/trained intensity images to the held out data. If the patterns of hotspot regions and the spatial locations in the respective datasets seemingly coincide with each other, then the proposed model is more likely applicable for the modelling of the call events.

In summary, the aim of this article is to explore the space-dynamics of ambulance calls and to identify spatial covariates associated with the call events. The result of this work will be used as input in designing optimal dispatching strategies for prehospital resources such as ambulances.

Throughout, entries with the prefix 'S' refer to the supplementary materials. The structure of the article is as follows. Section 2 and Section S1 provide an overview of the data and statistical methods used in this work. Section S5 and Section 4 present the created spatial features for modelling the call events and the results of the study. We evaluate the fitted model in Section 5, and discuss the implication of the results and provide a summary of the work in Section 6.

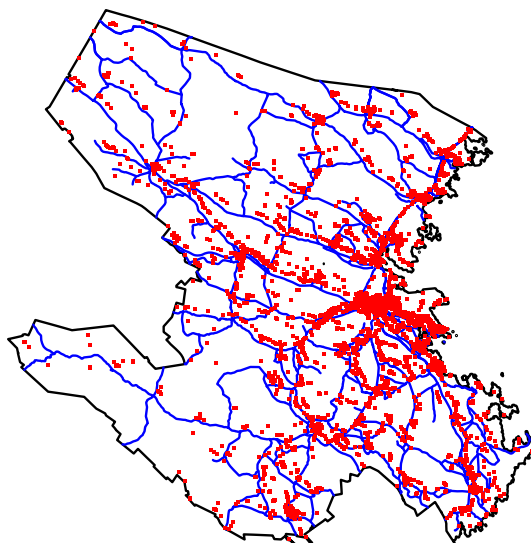
2. DATA

Given a spatial/geographical region W , which we assume to be a subset of \mathbb{R}^2 , by an *event* we will mean a location (GPS position) in W to which an ambulance is dispatched during a given time period and the total collection of events will be referred to as an *ambulance/emergency call dataset*. Here, the available information associated to an ambulance call dataset is the collection of locations $\{x_1, \dots, x_n\} \subseteq W$, $n \geq 0$, as well a *mark* m_i which is attached to each location x_i , $i = 1, \dots, n$. As we a priori do not know n , i.e., how many calls there will be within W during the period in question, such a dataset $\mathbf{y} = \{(x_1, m_1), \dots, (x_n, m_n)\}$ is most naturally classified as a *marked point pattern* (Baddeley et al. 2015). Note that when the marks are discrete, we usually refer to the point pattern as *multivariate/multitype* rather than as *marked*.

2.1. SPATIAL AMBULANCE CALL DATASET

Turning to the dataset at hand, W will represent the Swedish municipality of Skellefteå, and the time period under consideration is given by the years 2014–2018. Our dataset consists of 14,919 events where the mark structure is such that for each event, and thereby the person associated with the ambulance call in question, there is a recording of the event's priority label (1 indicates the highest severity, implying turned on sirens, and 2 is a lower severity level) and a recording of the sex/gender of the person (female or male). Among the 14,919 events, 7204 and 7715 of them have priority labels 1 and 2, respectively. Unfortunately, a missing data issue is present here: 5236 and 5238 of the events are recorded as male and

Figure 1. Ambulance emergency alarm call locations and the main road network of Skellefteå .



female, respectively, and the remaining events do not have any sex recording for the person related to the event, so they are labelled as “missing”. As our interest in the sex label lies mostly in highlighting possible structural differences related to the different covariates at hand, we have decided to proceed by studying the data as two separate marked point patterns: one where the marks are given by the priority labels and one where the marks are given by the sex labels. It should be said that there are other ways of handling this.

Figure 1 demonstrates the locations of all calls as well as the main road network of Skellefteå, Sweden. The figure highlights that the call locations are unevenly distributed over the study region and they tend to lie along the road network, which will make the statistical modelling quite challenging. We moved each location of the events randomly by at least one kilometer for data privacy reasons, i.e., to avoid the possibility of tracking plotted data to real cases/exact positions. Moving each event location at random by at least one kilometer works not only for this plot but also for the remaining data plots.

2.2. SPATIAL COVARIATES

In order to properly model the ambulance call risk, we also need a range of spatial covariates. To begin with, a closer look at Figure 1 and the relation between the road network and the call locations justifies the inclusion of road network related covariates. In addition, demographic spatial covariates should also play a role here, given that different demographic zones have different behavioural patterns. The demographic spatial covariates considered have been supplied by Statistics Sweden (SCB), and the road network related covariates considered have been provided by the Swedish Transport Administration (Trafikverket). Here, we will distinguish between two categories of covariates. The first is the collection of ‘original’ covariates, which are raw covariates retrieved from SOS alarm, Statistics Sweden (SCB), and Trafikverket. These covariates will, in turn, be used to generate a collection of

new covariates, which we will refer to as 'created' covariates. In Section S5, we outline the construction of the created covariates in detail.

All covariates considered have been based on the year 2018 and in their final forms they are given as functions of the form $z_i(s)$, $s \in W$. As the majority of these are demographic in nature, and since the demography of Sweden did not change at a high pace, we believe that covariates based on the year 2018 can be reasonably used to model the spatial point pattern data over the period 2014–2018. Moreover, in addition to including each individual covariate z_i in the analysis, we also include each interaction term $z_{ij}(s) = z_i(s)z_j(s)$, $s \in W$, which makes sense because we suspect that many of the covariates interact in one way or another. We have a total of 43 individual spatial covariates and, all in all, we consider a total of 989 ($= \binom{43}{2} + 2 \times 43$) covariates as candidates to be included in the modelling of the ambulance call risk.

2.2.1. Demographic Spatial Covariates

Each demographic spatial covariate $z_i(s)$, $s \in W$, which has been sampled on the 31st of December 2018, is piecewise constant and its value changes depending on which DeSO zone the location s belongs to. The DeSO zones, which we use to define different demographic zones, partition Sweden into 5984 smaller spatial sub-regions, which do not overlap with the borders of any of the country's 290 municipalities and are encoded based on, e.g., how rural a DeSO zone is. The spatial covariates at hand include population density, population by age, population by sex, population by Swedish/foreign background, population 20–64 years of age by occupation, population 25–64 years of age by education level, population 20+ years of age by accumulated income, population 20+ years of age by accumulated income, and population 20+ years of age by economical standard. It is worth emphasising that the covariates relating to sex (proportions of women and men in a DeSO) represent the underlying population structure, while the marks male/female relate to individual observed cases. We present a short description of these covariates in Section S4 (some are graphically illustrated in Figure S1).

2.2.2. Road Network Related Covariates

The road network data at hand have been broken down into sub-networks, which reflect the complete road network, the main road network, and the densely populated areas; a graphical illustration of these can be found in Figure S3. Each network is included in the analysis with the aim of indicating a different level of call activity: The complete road network indicates the overall spatial region where calls tend to occur, the main road network indicates the parts of the complete network where there is a reasonable amount of activity, and the densely populated area indicates on which parts of the complete road network most people live.

We would further like to adjust for the fact that the covariates above mainly reflect where different people (of different demographic groups) live, and not how they move about. Unfortunately, we do not have access to explicit mobility covariates such as aggregated movement patterns in the population, and, as one may guess, people do not only need access to ambulances when they are at home. A partial solution to this, as we see it, is to

consider an additional covariate, namely a spatial point pattern of bus stop locations. The idea is that bus stops reflect where there is a large amount of human mobility/activity. It may be noted that the above original covariates are not functions of the form $z_i(s)$, $s \in W$, but rather line segment patterns and point patterns (Baddeley et al. 2015). Our approach to including them in the analysis is to let them give rise to a set of created covariates, which are described in detail in Section S6.

2.2.3. The Spatial Domain

We finally emphasise that the x - and y -coordinates of the spatial domain $W \subseteq \mathbb{R}^2$ will be included as covariates. These are intended to explain residual and explicit spatial variation, having adjusted for the full range of spatial covariates.

3. STATISTICAL METHODS

As emphasised in Section 2, in particular through Figure 1, the ambulance calls are mainly located on/close to the underlying road network. One way to deal with the analysis is to project all events to a linear network representation of the road network and proceed with a linear network point pattern analysis (Baddeley et al. 2021). However, we here avoid such projections and instead choose to treat the spatial domain, and thereby the data point pattern, as Euclidean. As a result, we will instead introduce road network generated covariates which control that the fitted model generates events close to the underlying road network; details on the construction of such covariates can be found in Section S6.

3.1. POINT PROCESS PRELIMINARIES

In application areas such as environmental science, epidemiology, ecology, etc., aside from the spatial locations of the events, additional information about the events may be available, which can be associated to the locations of the events. Such information pieces are referred to as marks, and by including them in the analysis, we can often obtain more realistic spatial point process models for the events—note that, in contrast, a covariate reflects information which is known throughout the spatial domain before realisation of the events. For instance, as we saw above, it is both common and practically appropriate in emergency medical services to document the priority levels, the gender, the incident time, etc. for a call/patient; recall that by spatial points of the events we mean the set of spatial locations to which ambulances have been dispatched or the set of spatial locations from which calls have been made to the dispatcher/emergency alarm center.

As previously indicated, marked point patterns are modelled by *marked point processes* (Baddeley et al. 2015). Given a spatial domain $W \subseteq \mathbb{R}^2$ and a (Polish) mark space \mathcal{M} , a point process $Y = \{(x_i, m_i)\}_{i=1}^N \subseteq W \times \mathcal{M}$, $0 \leq N \leq \infty$, is a random subset such that, with probability 1, $Y \cap A \times B$ has finite cardinality $Y(A \times B)$ for any $A \times B \subseteq W \times \mathcal{M}$; throughout, any set under consideration is a Borel set. If we additionally have that $Y_g = \{x_i\}_{i=1}^N \subseteq W$ is a well-defined point process on W in its own right, then we say that Y is a marked point process. Note that each (x_i, m_i) is a random variable and note in particular that if W is

bounded then Y is automatically a marked point process. When the mark space is discrete, e.g., $\mathcal{M} = \{1, 2, \dots, k\}$, $k > 1$, we say that Y is multitype and we note that we may split Y into the marginal (purely spatial) point processes $Y_j = \{x_i : (x_i, m_i) \in Y, m_i = j\}$, $j = 1, 2, \dots, k$. This collection may formally be represented by the vector (Y_1, Y_2, \dots, Y_k) , which is referred to as a *multivariate* point process, and one commonly uses the two notions interchangeably.

3.2. SPATIAL INTENSITY FUNCTIONS

Our main interest here is to create a set of “heat maps” which describe the risk of a call occurring at a given spatial location $x \in W$, given an associated mark. This is accomplished by assuming that our data are generated by a multivariate point process $Y \subseteq W \times \{1, 2, \dots, k\}$, $k > 1$ and then modelling the *spatial intensity function* of each of the component Y_j , $j = 1, 2, \dots, k$. Formally, we define the spatial intensity function ρ_Y of Y as the Radon–Nikodym derivative of the intensity measure w.r.t. the product of a Lebesgue measure and a counting measure, i.e.,

$$A \times B \mapsto E[Y(A \times B)] = \sum_{j=1}^k \mathbf{1}\{j \in B\} \int_A \rho_Y(x, j) dx, \quad A \times B \subseteq W \times \{1, 2, \dots, k\},$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function. Heuristically, $\rho_Y(x, j)dx$ may be interpreted as the probability that Y has an event with mark j in an infinitesimal neighbourhood of x with size dx . Since $E[Y(A \times \{j\})] = E[Y_j(A)]$, it follows that the spatial intensity function of Y_j satisfies $\rho_j(x) = \rho_Y(x, j)$, $x \in W$, $j = 1, 2, \dots, k$. By modelling the marginal intensities ρ_j separately, we obtain a model for ρ_Y . Letting X denote an arbitrary Y_j , $j = 1, 2, \dots, k$, we see that its spatial intensity function ρ satisfies $E[X(A)] = \int_A \rho(x)dx$, $A \subseteq W$, or equivalently $\rho(x) = \lim_{\|dx\| \rightarrow 0} E[X(dx)]/\|dx\|$. If a point process has constant intensity function then we say that it is homogeneous, otherwise we refer to it as inhomogeneous. In a similar fashion, we may define higher-order intensity functions $\rho_n(x_1, \dots, x_n)$, $x_1, \dots, x_n \in W$, $n \geq 2$ as $\rho_n(x_1, \dots, x_n) dx_1, \dots, dx_n = E[X(dx_1) \cdots X(dx_n)] = P(X(dx_1) = 1, \dots, X(dx_n) = 1)$ for disjoint infinitesimal neighbourhoods $dx_1, \dots, dx_n \subseteq W$ of x_1, \dots, x_n . If $x_i = x_j$ for any $i \neq j$, then $\rho_n(x_1, \dots, x_n) = 0$.

3.3. POISSON PROCESSES

Poisson processes, which satisfy $\rho_n(x_1, x_2, \dots, x_n) = \rho(x_1)\rho(x_2) \cdots \rho(x_n)$, $n \geq 1$, and that $X(A)$ is Poisson distributed with mean $\int_A \rho(x)dx$ for any $A \subseteq W$, are baseline models for the case of complete spatial randomness, i.e., when there is no spatial interaction/dependence present. Poisson processes are completely governed by their spatially varying intensity functions and may be viewed as generalisations of random samples to the case where the size of the sample is random.

3.4. PARAMETRIC SPATIAL INTENSITY FUNCTION MODELLING

We have observed a high degree of inhomogeneity in the ambulance call locations and we believe that a large portion of this inhomogeneity can be attributed to (some of) the spatial covariates which we have access to. Hence, as a starting point, for a given mark, we will consider parametric modelling of the call locations.

Consider a family of spatial intensity functions $\rho_{\theta}(x)$, $x \in W$, which depends on spatial covariates through a parameter vector $\theta \in \mathbb{R}^{K+1}$. A common and convenient approach when modelling the intensity function of an arbitrary point process is to proceed as if we are considering a Poisson process, which is commonly referred to as composite likelihood estimation and is often motivated by good large sample properties (Coeurjolly and Lavancier 2019). Accordingly, consider a point pattern $\mathbf{x} = \{x_1, x_2, \dots, x_n\} \subseteq W$ which represents a realisation of a spatial point process X restricted to a bounded study region W . The associated Poisson process log-likelihood function $\log \mathcal{L}(\theta | \mathbf{x})$ can be approximated by

$$\log \mathcal{L}(\theta | \mathbf{x}) \approx \sum_{j=1}^m (y_j (\beta_0 + \mathbf{z}(s_j)\boldsymbol{\beta}') - \exp\{\beta_0 + \mathbf{z}(s_j)\boldsymbol{\beta}'\}) w_j, \quad (1)$$

where $\mathbf{z}(x) = (z_1(x), z_2(x), \dots, z_K(x))$ is a vector of spatial covariates at location $x \in W$, $s_i \in W$ represents the observed spatial location of events or a dummy point, $y_j = \frac{a_j}{w_j}$ and $a_j = \mathbf{1}\{s_j \in \mathbf{x}\}$; for details, see Section S1.1 of the supplementary material. It follows that the right-hand side of expression (1) is a weighted log-likelihood function of independent Poisson random variables Y_j , $j = 1, 2, \dots, m$. That is, for $j = 1, 2, \dots, m$, y_j are the observations, $\rho_{\theta}(s_j) = \exp\{\beta_0 + \mathbf{z}(s_j)\boldsymbol{\beta}'\}$ are the parameters of the Poisson distributions, and w_j are the weights. Thus, the weighted log-likelihood function in Eq. (1) can be maximised using standard software for fitting generalised linear models (McCullagh and Nelder 1989).

3.5. VARIABLE SELECTION: ELASTIC-NET REGULARISATION

Incorporating regularisation into the log-likelihood function in Eq. (1) can help to simultaneously select variables and estimate the parameters of the model. A penalised log-likelihood function based on Eq. (1) takes the form

$$\mathcal{L}_p(\theta) \approx \frac{1}{m} \log \mathcal{L}(\theta | \mathbf{x}) + \lambda R(\boldsymbol{\beta}), \quad (2)$$

where $R(\boldsymbol{\beta})$ is a regularisation method or penalty function, and $\lambda \geq 0$ is a tuning or smoothing parameter determining the strength of the penalty, or the amount of shrinkage. Several penalties such as garrote (Breiman 1995), least absolute shrinkage and selection operator (lasso), elastic-net and fused lasso (Tibshirani et al. 2005), group lasso (Yuan and Lin 2006), Berhu (Owen 2007), adaptive lasso (Zou 2006), and LAD-lasso (Wang et al. 2007) have been developed for penalised regression modelling. The most common regularisation methods are lasso, ridge regression, elastic-net, and adaptive lasso. In short, lasso has a tendency to shrink several coefficients to zero, leaving only the most influential ones in the model, while ridge regression shrinks the coefficients of correlated covariates towards each

other to borrow strength from each other and adjust for collinearity (Friedman et al. 2010a). The elastic-net penalty provides a mix between the ridge and the lasso penalties, and it is useful in cases where there are many correlated covariates or when the number of covariates exceeds the size of observations. It has the form

$$R(\boldsymbol{\beta}) = \sum_{k=1}^K \left(\frac{1}{2} (1 - \alpha) \beta_k^2 + \alpha \|\beta_k\| \right), \quad (3)$$

where the elastic-net parameter $\alpha \in [0, 1]$ turns (3) into a ridge penalty if $\alpha = 0$ and a lasso penalty if $\alpha = 1$. If $\alpha = 1 - \epsilon$ for small $\epsilon > 0$, then elastic-net performs like lasso but it avoids unstable behaviour due to extreme correlation (Yue and Loh 2015); empirical studies have indicated that the elastic-net technique tends to outperform lasso on data with highly correlated features (Comber and Harris 2018). Zou (2006) proposed an adaptive lasso to address the shortcomings of lasso, such as biased estimates of large coefficients and conflict between optimal prediction and consistent variable selection. According to Krämer et al. (2009), however, the performance of adaptive lasso is poor in the presence of highly correlated variables. We here follow Yue and Loh (2015) by considering elastic-net penalisation and substituting Eq. (3) into Eq. (2), whereby the optimization problem of the elastic-net penalization of the log-likelihood function can be summarized as

$$\operatorname{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{K+1}} \mathcal{L}_p(\boldsymbol{\theta}) \approx \operatorname{argmin}_{\boldsymbol{\theta} \in \mathbb{R}^{K+1}} \left\{ -\frac{1}{m} \log \mathcal{L}(\boldsymbol{\theta}) + \lambda \sum_{k=1}^K \left\{ \frac{1}{2} (1 - \alpha) \beta_k^2 + \alpha \|\beta_k\| \right\} \right\}. \quad (4)$$

To deal with the optimisation problem in (4), we refer to Section S2 in supplementary material.

3.6. SEMI-PARAMETRIC MODELLING AND A NEW BANDWIDTH SELECTION APPROACH

In the spatial intensity function modeling of our ambulance data we do not have access to explicit mobility covariates, such as aggregated movement patterns in the population. Our solution to this is quite pragmatic and simple: we consider the semi-parametric approach where we simply add a further spatial covariate to the existing collection of covariates, which is given by a nonparametric (kernel) intensity estimate $\tilde{\rho}(x)$, $x \in W$, and referred to as the *benchmark spatial intensity*; see Section S3 of the supplementary material for details. To this end, we propose a new heuristic algorithm for optimal bandwidth selection in kernel intensity estimation; see Section S3.1 of the supplementary material for more information. The idea here is that the inclusion of $\tilde{\rho}$, which is also included in the regularisation, should pick up on regions where there is an increased intensity due to human mobility, which is not reflected by the existing list of covariates.

4. DATA ANALYSIS

Having presented the semi-parametric regularised fitting procedure to be employed, we next turn to the modelling of the ambulance call data; recall the structure of the data (including the missing sex/gender label issue) and that we may model the intensity of each marginal Y_j separately. Our motivation for elastic-net regularisation is mainly that we have a large number of spatial covariates, which may be highly correlated. The choice of α is subjective, and we usually pick α between 0 and 1, and then proceed to the estimation of the proposed model (Friedman et al. 2010a). Recall further that elastic-net performs like a lasso but avoids unstable behaviour due to extreme correlation if the elastic-net parameter α is large enough but less than one (Yue and Loh 2015); $\alpha = 1$ yields lasso and $\alpha = 0$ yields ridge. To be as close as possible to lasso in terms of shrinkage/variable selection, while still reaping the benefits of the ridge component of the elastic-net penalty, which can yield more stable modelling, we here follow a recommendation of Friedman et al. (2010a) and Yue and Loh (2015) where we set $\alpha = 0.95$. Such lasso-like elastic-net regularisation results in variable selection (coefficients of less determining covariates are set to zero), but the penalty also forces highly correlated features to have similar coefficients.

4.1. MODELLING THE AMBULANCE CALL INTENSITY FUNCTION

The trace plots of the estimated coefficients for lasso and lasso-like elastic-net, as well as covariates associated with the fitted intensities are presented in Section S6.1 of the supplementary material. Also the tuning parameter selection method is shown in Section S6.1 of the supplementary material. To get an overall view of the fitted intensities, in Figure 2, for each marginal point pattern, we present a spatial interpolation of the fitted intensity over the study region, based on the intensity estimates for the corresponding data points. Note that we have scaled the intensity estimates to range between 0 and 1 so that we may compare them more easily.

The aforementioned analyses have utilised the spatial covariates $z_i(s)$, $s \in W$, and the interaction terms of the form $z_i(s)z_j(s)$, both for $i = j$ and $i \neq j$. However, with this approach it is not easy to interpret the results and present the estimated models. Therefore, we are interested in further exploring the estimation of the intensity of the emergency alarm call events using only the original spatial covariates $z_i(s)$, $s \in W$, i.e., no interactions of spatial covariates in the modelling of the spatial intensity. Given the vast number of covariates we consider when we include the interaction terms, modelling the spatial intensity of the emergency alarm call events using only the original spatial covariates can help to identify the spatial covariates that play a key role in determining the spatial distribution of the emergency alarm call events, and it also eases the interpretation and presentation of the results. Moreover, we can also compare the estimated intensities that are obtained by the two approaches (i.e., modelling with only the original spatial covariates and modelling with the original spatial covariates plus the first-order interaction terms) to discern the approach that can be practicable. Using only the original spatial covariates, a lasso-like elastic-net ($\alpha = 0.95$) with an optimal tuning parameter estimate does not provide sparse solutions, i.e., most of the coefficients of the spatial covariates are non-zero. That is to say, the lasso-

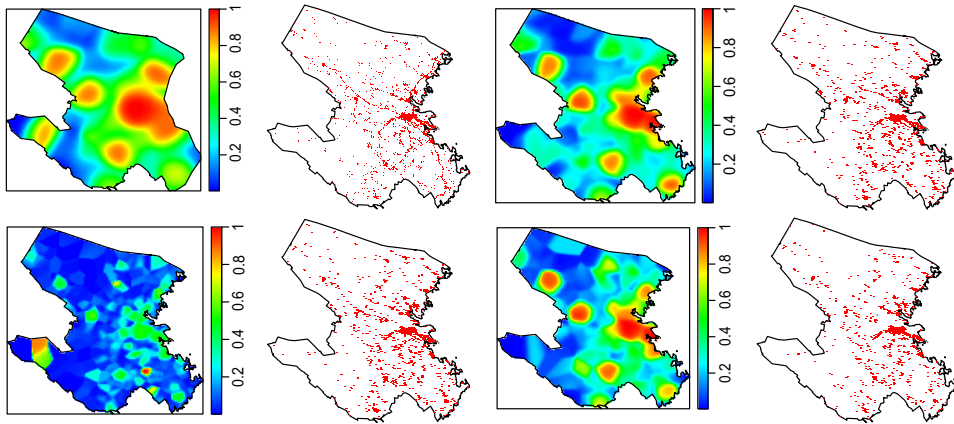


Figure 2. The estimated spatial intensities of the emergency alarm calls and their corresponding observed marginal spatial point patterns. The first column presents estimated intensities for priority levels 1 and 2, respectively, while the second column demonstrates their corresponding observed marginal spatial point patterns. In a similar manner, the third column presents the estimated intensities for genders male and female, while the fourth column shows their corresponding observed marginal spatial point patterns. Note that we have scaled the intensity estimates to range between 0 and 1 so that we may compare them more easily.

like elastic-net provides dense solutions, i.e., it has a low rate of variable exclusion; see Table S1 in Section S6.1 of the supplementary material. On the other hand, exploiting the one-standard-error rule in Friedman et al. (2010b), where we pick the most parsimonious estimated model within one standard error of the minimum, in the context of regularisation, we present the result of the estimation for each of the intensity functions of the marginal point patterns in Table S2 and also highlight the important covariates associated with the events in Section S6.1 of the supplementary material. The three approaches, which are (i) the estimated spatial intensity based on the original spatial covariates plus the first-order interaction terms, and the estimated spatial intensities that are obtained using (ii) the estimated dense model and (iii) the parsimonious/sparse model based on only the original spatial covariates, can all be compared for how well they capture the spatial variation of the events. Here, we need to remark that when we say dense model and parsimonious/sparse model, we are referring to the estimated models obtained based on only the original spatial covariates, i.e., no interaction terms in the model setting. Figure 2 indicates that the spatial distributions of the events are well captured by their corresponding estimated spatial intensities, with the exception that the distribution of the events with priority level 2 is not well captured by its corresponding estimated spatial intensity. This can be viewed as a shortcoming of incorporating the original covariates plus the first-order interaction terms when modelling the ambulance call events. Hence, we are interested in comparing the performance of the estimated model based on the original spatial covariates and the first-order interaction terms with that of the estimated dense and sparse models based on only the original covariates on the marginal point pattern corresponding to priority level 2; see Section S7 of the supplementary material. Section S7 also presents the result for the modelling of the spatial intensity function for the unmarked ambulance call data, i.e., when we ignore the marks of the ambulance call data, using only the original spatial covariates.

In Figure 2, we further see that the intensity of priority level 2 has a smooth, almost constant, shape, indicating that the risk does not change much spatially, and, in contrast, for priority level 1 we see clear peaks in the intensity, in various locations. We further see little difference in the intensities for males and females, indicating that these two groups have a similar risk. With regard to priority level 1, it seems that the spatial variation of the events is well captured by the model that involves the original and first-order interaction terms. Among others, some of the interaction terms that are associated with the spatial intensity function for priority level 1 are population density and shortest distance to bus stops, benchmark intensity and shortest distance to bus stops, benchmark intensity and shortest distance to a densely populated area, bus stop density and shortest distance to main road networks, as well as age (6–9) and population density.

5. EVALUATION OF THE FITTED MODEL

The evaluation of the fitted spatial models is generally challenging, and here we use two approaches. Firstly, we evaluate the stability of the estimated models. Arguing from the point of view of [Moradi et al. \(2019\)](#), we consider p -thinning-based subsampling/undersampling by retaining 70% of the observed points for the unmarked and marginal point patterns, to compare the stability of the fitted models. The estimates of the spatial intensities obtained for the unmarked and marginal point patterns using the fitted dense and sparse models are treated as the true spatial intensities of the corresponding point patterns. In this context, hereafter, we refer to the estimated spatial intensities based on the marginal and unmarked point patterns as the true spatial intensities. We generated one hundred undersampled random realisations/thinnings, i.e., we retained 70% for each of the unmarked and marginal spatial point patterns, in order to produce one hundred estimated spatial intensities for each of the unmarked and marginal spatial point patterns. Pixel-wise mean absolute errors, the 5% and 95% quantiles of pixel-wise absolute errors of the true intensities, and each of the one hundred estimated spatial intensities are used to assess the stabilities of the fitted dense and sparse models for each of the unmarked and marginal point patterns. Figure 3 shows the evaluation of the stabilities of the estimated dense models in estimating the spatial intensities of the emergency alarm call events. The overall means and standard deviations of the pixel-wise mean absolute errors for the dense models are 0.098 and 0.107, 0.099 and 0.123, 0.057 and 0.052, 0.060 and 0.052, and 0.052 and 0.051, respectively, for the point patterns with priority level 1, priority level 2, male, and female, as well as for the unmarked point pattern. Figure 4 demonstrates the evaluation of the stabilities of the estimated sparse models in estimating the spatial intensities of the emergency alarm call events. The overall means and standard deviations of the pixel-wise mean absolute errors for the sparse models for the priority level 1, priority level 2, male, and female point patterns, as well as for the unmarked point pattern, are 0.103 and 0.128, 0.103 and 0.140, 0.056 and 0.058, 0.062 and 0.066, and 0.054 and 0.061, respectively. Based on a comparison of the plots in Figures 3 and 4, the estimated dense models seem to be more stable than the sparse models. Even though the 95% and 5% quantiles of pixel-wise absolute errors for the estimated dense models do not entirely lie below the corresponding 95% and 5% quantiles of pixel-wise

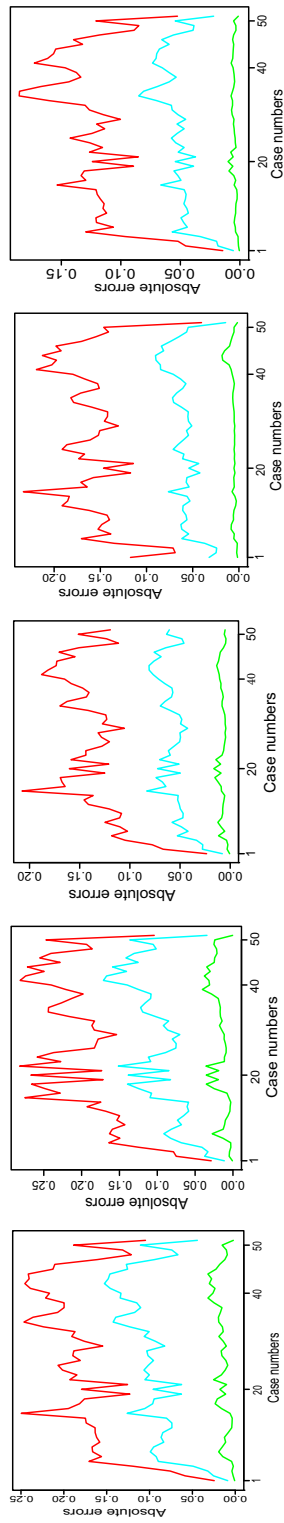


Figure 3. The smoothed plots of pixel-wise mean absolute errors, the 5% and 95% quantiles of pixel-wise absolute errors of the true intensities and the estimated spatial intensities based on the undersampled point patterns. Row-wise, the plots show the evaluation of the estimated dense models for the point patterns with priority level 1, priority level 2, male, and female, as well as for the unmarked point pattern. The top, middle, and bottom curves in each plot represent the 95% quantiles, the mean absolute errors, and the 5% quantiles .

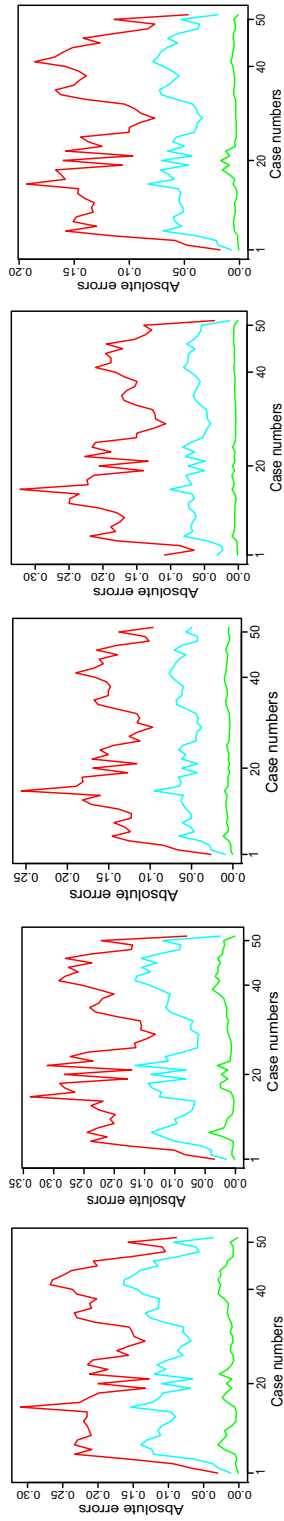


Figure 4. The smoothed plots of pixel-wise mean absolute errors, the 5% and 95% quantiles of pixel-wise absolute errors of the true intensities and the estimated spatial intensities based on the undersampled point patterns. Row-wise, the plots show the evaluation of the estimated sparse models for the point patterns with priority level 1, priority level 2, male, and female, as well as for the unmarked point pattern. The top, middle, and bottom curves in each plot represent the 95% quantiles, the mean absolute errors, and the 5% quantiles .

absolute errors for the estimated sparse models, the maximum 95% and 5% quantiles can be used to compare the stabilities of the estimated dense and sparse models. An estimated model can be unstable if it has maximum 95% and 5% quantiles and has a broader band between the two quantiles. Despite the fact that the overall means and standard deviations of the pixel-wise mean absolute errors for the estimated dense and sparse models do not differ much as the aforementioned results suggest, we still believe that the estimated dense models are more stable than the estimated sparse models. Here, we want to emphasize that the spatial covariates that continue to exist in the sparse model and have non-zero estimated coefficients have a strong association with the emergency alarm call events, and as a result, we may draw a strong interpretation based on the estimated sparse model.

We can also visually assess how well the estimated models capture the spatial distribution of the emergency alarm call events. In order to do this, following the p -thinning-based cross-validation (Monte-Carlo cross-validation) approach in Cronie et al. (2021), each of the datasets corresponding to the unmarked and marginal point patterns can be randomly divided into two parts, a training point pattern/set (retention probability 0.7) and a validation point pattern/set. That is, for each of the unmarked and marginal point patterns, we estimate the proposed model using 70% of the data and use the remaining 30% to assess/validate the performance of the estimated models; we consider only one split. The predicted intensities for the validation point patterns can be used together with the corresponding validation point patterns to visually evaluate the performance of the estimated models. As can be seen in Figure 5, the hotspot regions in the validation point patterns are well captured by their respective predicted/estimated spatial intensities. The plots in the figure also suggest that the estimated dense models seemingly perform better than their corresponding fitted sparse models, which are taken to be submodels of the estimated dense models.

6. DISCUSSION

The purpose of this study is to model the spatial distribution of ambulance/medical emergency alarm call events in order to establish a framework for developing optimal ambulance dispatching strategies/algorithms, which are related to the response times and operational costs of prehospital resources such as ambulances. Such optimal strategies can assist the concerned body in providing efficient emergency medical assistance to life-threatening emergencies as quickly as possible. Further, the locations of hotspot areas within the study region are crucial components of such designs, i.e., identifying study area subregions with a high risk of events plays a crucial role in developing the optimal dispatching strategies. Understanding what drives the associated risk is thus paramount. This work focuses on discerning such hotspot areas as well as selecting/exploring spatial covariates that are associated with the spatial distribution of the call events.

Our modeling strategy is to treat the data as a realization of a spatial point process and model its intensity function to quantify the spatially varying call risk. We assume that the intensity function is a log-linear function of the spatial covariates under consideration and their interaction terms, and we fit the intensity function by means of a Poisson process log-likelihood function, i.e., we carry out composite likelihood estimation. To carry out

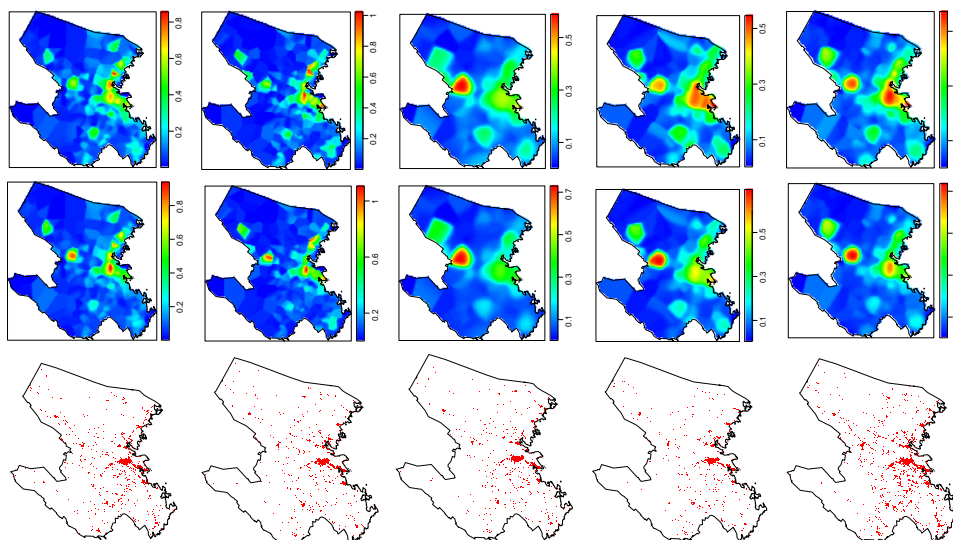


Figure 5. A cross-validation-based evaluation of the performance of the estimated dense and sparse models in estimating the spatial intensities of the emergency alarm call events. The estimated spatial intensities are shown in the first and second rows using the estimated dense and sparse models based on the corresponding test datasets in the third row. From left to right, the plots in the figure present the estimated spatial intensities for the test point patterns with priority level 1, priority level 2, male, and female, as well as for the unmarked test point pattern. Note that we have scaled the intensity estimates to range between 0 and 1 so that we may compare them more easily. We also multiplied the normalized versions by 1000 for ease of visualisation.

variable selection and adjust for over-/under-fitting, a regularisation term is added to the (approximated) log-likelihood function. Following [Yue and Loh \(2015\)](#), we have chosen to employ an elastic-net penalty, which is useful when the number of covariates considered in the modeling is large, in particular exceeds the total point count, or when the model contains several correlated spatial covariates. The elastic-net penalty is governed by a parameter, α , which controls how much weight we put on either the lasso or the ridge penalty, and by setting $\alpha = 1 - \epsilon$ for some small $\epsilon > 0$, it performs much like the lasso, but avoids any unstable behaviour caused by extreme correlations. Following a recommendation by [Friedman et al. \(2010a\)](#) and [Yue and Loh \(2015\)](#), we used $\alpha = 0.95$, i.e., we obtain a lasso-like elastic-net, and tenfold cross-validation was used to select an optimal estimate of the tuning parameter λ .

The primary challenge in this work is the nature of the spatial data; that is, the locations of the events tend to be close to/on road networks. Traditional kernel smoothing methods are incapable of displaying the spatial variation of emergency alarm call events. To address this issue, semi-parametric modelling of the spatial intensity function of the events has been considered, and it appears that the proposed approach is feasible for obtaining a good estimated spatial intensity for the events, which will be used to design optimal ambulance dispatching strategies to support emergency medical services as efficiently as possible, in particular in life-threatening situations. Even though the semi-parametric modelling of the spatial intensity function captures the spatial distribution of the emergency alarm call events well, the effect of the nature of the spatial data persists. To address this issue, we will continue

to work on the spatial intensity function modelling but focusing on the linear network setting. The proposed statistical method is tested with data gathered from Skellefteå, a municipality in northern Sweden. Long term, our goal is to model the intensity for the region covering all of the four northern regions/counties of Sweden (Norrbotten, Västerbotten, Västernorrland, and Jämtland) using the proposed statistical model.

Finally, the study developed a new heuristic algorithm for bandwidth selection and discovered that spatial covariates such as population age categories, spatial location of events, and spatial covariates related to bus stops, main road networks, complete road networks, and densely populated areas play an important role in determining the spatial distribution of emergency alarm call events. The study also demonstrated that semi-parametric modelling of the spatial intensity function can handle the spatial variation of emergency alarm call events, and the estimated spatial intensity function of the events can be used as an input in designing optimal ambulance dispatching strategies to provide better emergency medical services for life-threatening health conditions.

ACKNOWLEDGEMENTS

The authors are grateful to our project partners SOS Alarm and the regions: Västerbotten Norrbotten, Västernorrland and Jämtland-Härjedalen for fruitful discussions on the Swedish prehospital care. The authors thank the editorial board and two anonymous referees for constructive feedback.

Funding Open access funding provided by University of Gothenburg. This work was supported by Vinnova [Grant No. 2018-00422] and the regions Västerbotten, Norrbotten, Västernorrland and Jämtland-Härjedalen. The authors thank the editorial board and two anonymous referees for constructive feedback.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

[Received August 2022. Revised February 2023. Accepted February 2023. Published Online April 2023.]

REFERENCES

- Baddeley A, Nair G, Rakshit S, McSwiggan G, Davies TM (2021) Analysing point patterns on networks-a review. *Spatial Stat* 42:100435
- Baddeley A, Rubak E, Turner R (2015) *Spatial point patterns: methodology and applications* with R. Chapman and Hall/CRC
- Baddeley AJ, Møller J, Waagepetersen R (2000) Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. *Stat Neerl* 54(3):329–350

- Breiman L (1995) Better subset regression using the nonnegative garrote. *Technometrics* 37(4):373–384
- Byrsell F, Claesson A, Ringh M, Svensson L, Jonsson M, Nordberg P, Forsberg S, Hollenberg J, Nord A (2021) Machine learning can support dispatchers to better and faster recognize out-of-hospital cardiac arrest during emergency calls: a retrospective study. *Resuscitation* 162:218–226
- Choiruddin A, Coeurjolly J-F, Letué F et al (2018) Convex and non-convex regularization methods for spatial point processes intensity estimation. *Electron J Stat* 12(1):1210–1255
- Coeurjolly J-F, Lavancier F (2019) Understanding spatial point patterns through intensity and conditional intensities. In: *Stochastic geometry*. Springer, pp 45–85
- Comber A, Harris P (2018) Geographically weighted elastic net logistic regression. *J Geogr Syst* 20(4):317–341
- Cronie O, Moradi M, Biscio CA (2021) Statistical learning and cross-validation for point processes. [arXiv:2103.01356](https://arxiv.org/abs/2103.01356)
- Efron B, Hastie T, Johnstone I, Tibshirani R et al (2004) Least angle regression. *Ann Stat* 32(2):407–499
- Fercoq O, Richtárik P (2016) Optimization in high dimensions via accelerated, parallel, and proximal coordinate descent. *SIAM Rev* 58(4):739–771
- Friedman J, Hastie T, Höfling H, Tibshirani R et al (2007) Pathwise coordinate optimization. *Ann Appl Stat* 1(2):302–332
- Friedman J, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 33(1):1
- Friedman J, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordinate descent. *J Stat Softw Artic* 33(1):1–22
- Krämer N, Schäfer J, Boulesteix A-L (2009) Regularized estimation of large-scale gene association networks using graphical Gaussian models. *BMC Bioinform* 10(1):384
- McCullagh P, Nelder J (1989) *Generalised linear models*
- Moradi MM, Cronie O, Rubak E, Lachize-Rey R, Mateu J, Baddeley A (2019) Resample-smoothing of Voronoi intensity estimators. *Stat Comput* 1–16
- O’keeffe C, Nicholl J, Turner J, Goodacre S (2011) Role of ambulance response times in the survival of patients with out-of-hospital cardiac arrest. *Emerg Med J* 28(8):703–706
- Owen AB (2007) A robust hybrid of lasso and ridge regression. *Contemp Math* 443(7):59–72
- Renner IW, Warton DI (2013) Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology. *Biometrics* 69(1):274–281
- Thurman AL, Zhu J (2014) Variable selection for spatial Poisson point processes via a regularization method. *Stat Methodol* 17:113–125
- Tibshirani R (1996) Regression shrinkage and selection via the lasso. *J R Stat Soc: Ser B (Methodol)* 58(1):267–288
- Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K (2005) Sparsity and smoothness via the fused lasso. *J R Stat Soc: Ser B (Stat Methodol)* 67(1):91–108
- Wang H, Li G, Jiang G (2007) Robust regression shrinkage and consistent variable selection through the lad-lasso. *J Bus Econ Stat* 25(3):347–355
- Yuan M, Lin Y (2006) Model selection and estimation in regression with grouped variables. *J R Stat Soc: Ser B (Stat Methodol)* 68(1):49–67
- Yue Y, Loh JM (2015) Variable selection for inhomogeneous spatial point process models. *Can J Stat* 43(2):288–305
- Zou H (2006) The adaptive lasso and its oracle properties. *J Am Stat Assoc* 101(476):1418–1429
- Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *J R Stat Soc: Ser B (Stat Methodol)* 67(2):301–320