The Annals of Applied Statistics
2024, Vol. 18, No. 1, 529–554
https://doi.org/10.1214/23-AOAS1800

A NONSEPARABLE FIRST-ORDER SPATIOTEMPORAL INTENSITY FOR EVENTS ON LINEAR NETWORKS: AN APPLICATION TO AMBULANCE INTERVENTIONS

BY ANDREA GILARDI[1,a], RICCARDO BORGONI[1,b] AND JORGE MATEU[2,c]

[1] *Department of Economics, Management and Statistics, University of Milan—Bicocca,* [a]*andrea.gilardi@unimib.it,* [b]*riccardo.borgoni@unimib.it*

[2] *Department of Mathematics, Universitat Jaume I,* [c]*mateu@uji.es*

The algorithms used for the optimal management of an ambulance fleet require an accurate description of the spatiotemporal evolution of the emergency events. In the last years, several authors have proposed sophisticated statistical approaches to forecast ambulance dispatches, typically modelling the data as a point pattern occurring on a planar region. Nevertheless, ambulance interventions can be more appropriately modelled as a realisation of a point process occurring on a linear network. The constrained spatial domain raises specific challenges and unique methodological problems that cannot be ignored when developing a proper statistical approach. Hence, this paper proposes a spatiotemporal model to analyse ambulance dispatches focusing on the interventions that occurred in the road network of Milan (Italy) from 2015 to 2017. We adopt a nonseparable first-order intensity function with spatial and temporal terms. The temporal dimension is estimated semiparametrically using a Poisson regression model, while the spatial dimension is estimated nonparametrically using a network kernel function. A set of weights is included in the spatial term to capture space-time interactions, inducing nonseparability in the intensity function. A series of tests show that our approach successfully models the ambulance interventions and captures the space-time patterns more accurately than planar or separable point process models.

1. Introduction. The proper management of an ambulance fleet is of vital importance for the timely assistance of medical emergencies, particularly when, as the latest COVID-19 pandemic has demonstrated, healthcare operations are stressed by long-standing critical events, such as epidemics or natural and man-made disasters. Relevant efforts are devoted by local agencies to allocate limited human and instrumental resources, while managing an increasing demand for services, guaranteeing high levels of geographical coverage and a constant improvement of key performance metrics such as rapid responses to potentially life-threatening emergencies (Vile et al. (2012)).

Policymakers require qualitative and quantitative approaches and evidence-based studies to tackle these challenging issues. In fact, the management of an emergency medical system (EMS) is an extremely difficult task, considering the complex spatial and temporal dynamics that govern ambulance interventions, especially for large and highly populated metropolitan areas. Advanced operational research algorithms have been developed in the past years to manage the fleet size and locate the dispatch centres (Blackwell and Kaufman (2002), Henderson (2011)). However, these algorithms depend upon ad hoc inputs regarding the distribution of emergency events, and the adoption of inaccurate predictions can lead to poor deployment decisions, high response times, and, in general, low performances. Therefore, in the past years, several authors (see, e.g., Zhou and Matteson (2015), Zhou et al. (2015),

Received April 2022; revised June 2023.

Key words and phrases. Emergency interventions, linear networks, nonparametric methods, point patterns on linear networks, spatial networks, spatiotemporal data.

Bayisa et al. (2020)) proposed complex spatiotemporal models to carefully forecast the interventions.

Typically, in the aforementioned papers, the ambulance dispatches were modelled as point processes occurring on a planar surface (e.g., a polygon delimiting a city). Nevertheless, we believe that the emergency interventions can be more appropriately considered as a realisation of a point process occurring on a linear network, that is, a graph object whose nodes and edges are embedded in a space (Barthélemy (2011), Baddeley et al. (2021)). Street networks represent a particular case of linear networks where nodes and edges correspond to road junctions and street segments, respectively.

The analysis of spatial data occurring on a linear network raises geometrical, computational, and statistical complexities (Okabe and Sugihara (2012)). First, ignoring the network constraint may lead to spurious results and false positive detections (Yamada and Thill (2004), Lu and Chen (2007)). Second, the readaptation of the classical planar techniques (such as the K-function or the kernel density estimator) presents unique methodological problems due to the nonhomogeneous nature of the spatial domain.[1] Third, the length of the spatial network and volume of the data typically create additional computational problems that require ad hoc solutions (Rakshit et al. (2019), Rakshit, Baddeley and Nair (2019)). We refer to Baddeley et al. (2021) and the references therein for more details.

The goal of this paper is to analyse all ambulance interventions that occurred in Milan (Italy) from 2015 to 2017, using a spatiotemporal point pattern model developed at the road network level. Starting from the assumption that the emergency events can be modelled by a nonhomogeneous Poisson Process, we propose a nonseparable structure for the first-order intensity function with spatial and temporal terms. The temporal component is modelled semiparametrically using a Poisson regression with deterministic covariates, while the spatial dimension is modelled using a nonparametric kernel estimator. The nonseparability of the intensity function is induced by a set of weights that are included in the spatial component to capture space-time interactions. To the best of our knowledge, this paper represents the first attempt to model EMS data on an extensive road network via a nonseparable intensity function.

The rest of the paper is organised as follows. Section 2 examines the ambulance interventions data and presents the procedures used to build the computational representation of the street network. We introduce the spatiotemporal framework and the first-order nonseparable intensity function in Section 3, providing an overview of the spatial and temporal statistical models. The main results are presented in Section 4; while in Section 5, we validate the performances of the proposed methodology. Section 6 compares the suggested approach with alternative specifications that include planar or separable approaches, discusses the scalability of our model to large linear networks, and exemplifies a real-world application. Finally, Section 7 concludes the article summarising the most important findings and the main contributions.

2. Data: Ambulance interventions.

In this section we describe the characteristics of the data, its peculiarities, and the procedures used to transform the raw data into a computational structure suitable for fitting the proposed model, which is detailed in Section 3.

The dataset was provided by the official regional EMS and included all ambulance dispatches in the city of Milan from 2015-01-01 to 2017-12-31. Milan is the second largest city in Italy, after Rome, with a total population of 1,386,235 in 2021 and an area of about 183 square kilometres. It represents one of the most important Italian metropolitan areas where

[1] A street network is not a homogeneous spatial domain since each edge (i.e., each road segment) is surrounded by different configurations of the neighboring streets.

hundreds of thousands of people pass through every day. For this reason the management of the ambulances in Milan requires ad hoc modelling and planning. The dataset includes four fields recording the day, the hour, and the GPS coordinates of the ambulance interventions (stored using an official Italian coordinate reference system named Gauss–Boaga projection). These columns provided the necessary information to estimate the spatiotemporal model introduced below.

Approximately 50,000 observations representing errors, outliers, and spurious or duplicated ambulance dispatches were excluded from the raw data. More precisely:

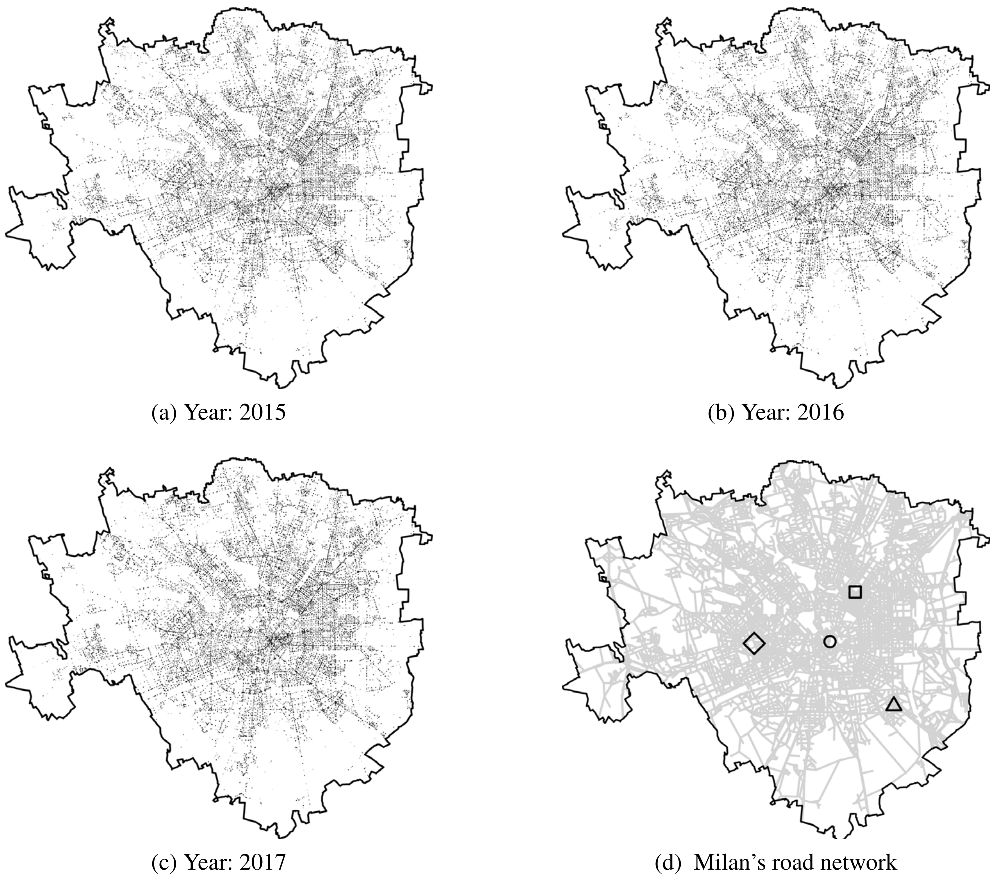
1. We did not consider 10,000 outlier interventions that occurred during EXPO 2015, the World Expo hosted in Milan from May 1 to October 31 and dedicated to food and life themes. The event attracted about 2.15 million visitors from all around the world, and the emergency interventions that occurred in the area where EXPO took place were handled in a different way than usual.
2. We ignored 11,500 erroneous calls (i.e., situations where someone requested an ambulance but an error was recorded in the EMS database).
3. We removed 30,000 records linked to multiple ambulance dispatches. These situations typically occur in particular circumstances, like serious or life-threatening emergencies (e.g., heart failures or severe car crashes). In these cases duplicated observations are recorded for the same event, but we retained only the first ambulance dispatch.
4. We filtered out 7000 observations linked to ambulance reroutings (i.e., an ambulance going to location A gets redirected to another, typically more pressing, emergency at location B), and we considered only those records that correspond to the actual interventions.

Similar preprocessing steps were also implemented in Matteson et al. (2011), Zhou and Matteson (2015). The remaining sample included 494,614 interventions, 163,075 occurred in 2015, 164,871 in 2016, and 166,668 in 2017.

The spatial distribution of the EMS events is depicted in Figures 1a–1c. The spatial patterns look stable among the three years, and the interventions clearly mirror the skeleton of a road network (see Figure 1d), highlighting the city ring road and some of the most important arterial thoroughfares. Most of the white areas represent nonurban places, mainly located in southern and western parts of the municipality. Given the spatial distribution of the emergency interventions, we believe that a network-approach is more appropriate than a planar approach since it takes into account the nature of the data and the particular constrains of their geolocations.

We also explored the temporal dimension of the data, examining the daily and weekly dynamics that govern the total number of emergency interventions. Figure 2 shows the daily number of ambulance dispatches. The data exhibit a clear trend, with a global minimum registered in August, the typical period for summer holidays in Italy. Other local peaks and minima could be linked with the most important religious holidays (such as Christmas or Easter), national celebrations (New Year's Eve), or other occasional events (such as the heatwave in July 2015 or the ice storms in January 2017). The three years are characterised by similar temporal patterns.

Figure 3 displays the temporal dynamics of the emergency interventions within a day. The panel summarises the average number of hourly ambulance interventions split by the hour of the day and the day of the week. A similar pattern is found in the three years: after rapidly increasing in the early morning, the time series peaks around 10:00, slowly falls until 15:00, and remains stable until 20:00 when it rapidly drops until the next day. The hourly seasonalities are different between weekends and weekdays. In fact, possibly due to the city's nightlife, the regional EMS registers, on average, more interventions during the first hours of the day at the weekend with respect to the rest of the week. Instead, lower frequencies are detected during the rest of the day.


(a) Year: 2015

(b) Year: 2016

(c) Year: 2017

(d) Milan’s road network

FIG. 1. *(a) to (c): Locations of ambulance interventions that occurred in Milan from* 2015-01-01 *to* 2017-12-31. *Each map represents one year. (d): Milan’s road network. The empty square symbol denotes* Stazione Centrale *(i.e., the Central Station), the empty circle denotes the Duomo of Milan, the triangle denotes an important square, and the diamond denotes a famous retirement house.*

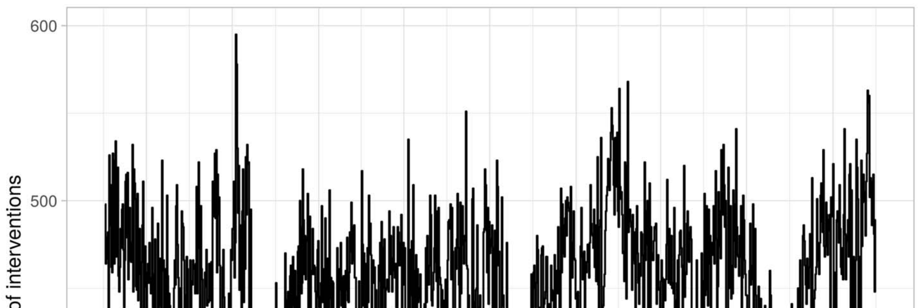


FIG. 2. *Daily number of ambulance interventions that occurred in Milan from* 2015-01-01 *to* 2017-12-31.

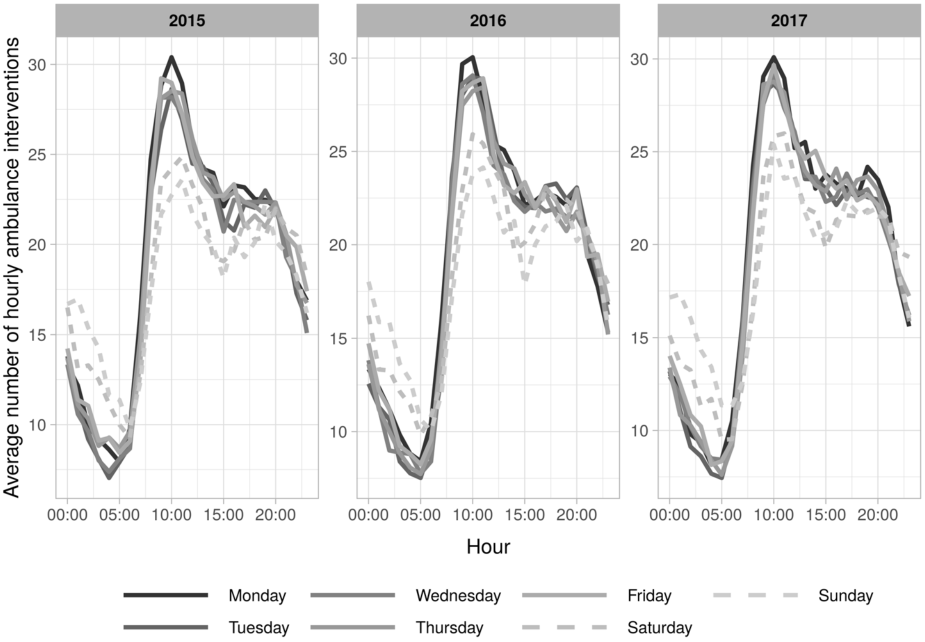


FIG. 3. *Average number of hourly ambulance interventions divided by the day of the week. There are clear seasonal patterns that characterise weekends and weekdays.*

We completed the temporal exploratory analysis examining the autocorrelation function (ACF) of the hourly number of EMS interventions. A graphical output, considering two weeks of lagged counts, is reported in Figure 4. The plot clearly highlights hourly, daily, and weekly seasonalities. We can also notice that the ACF is negative when considering lags

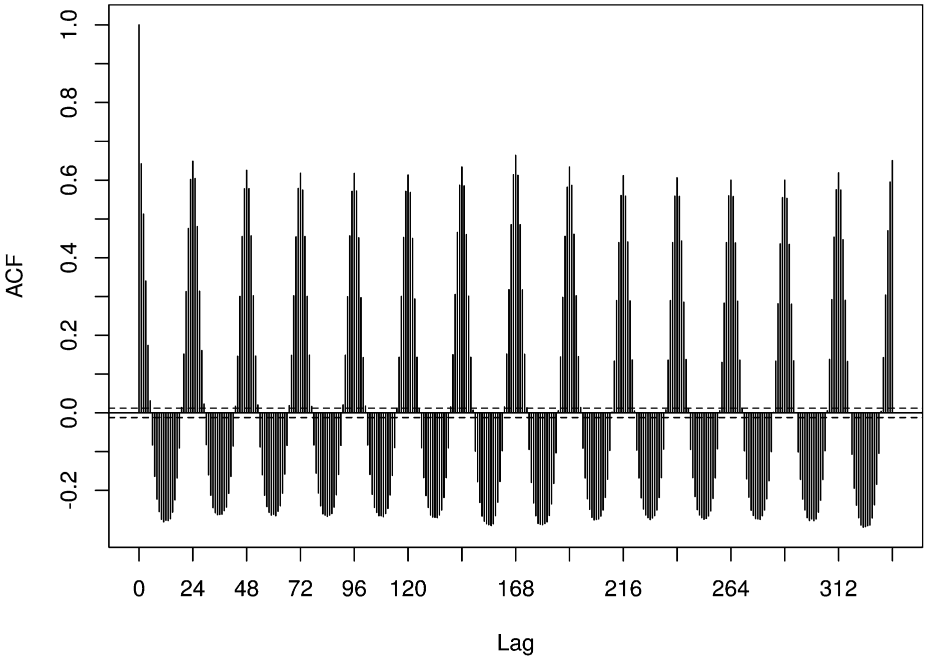


FIG. 4. *Autocorrelation function of the hourly number of EMS interventions occurred in the street network of Milan from* 2015 *to* 2017*. It clearly displays daily and weekly seasonalities.*

that are multiples of 12, pointing out an opposite behaviour during mornings, afternoons, and nights. The statistical model, proposed in Section 3, takes into account the temporal patterns detected in the data.

As already mentioned, we analysed the ambulance interventions as a spatial point process occurring on a restricted one-dimensional spatial domain, which represents Milan's road network. More generally, a *linear network*, denoted hereinafter by L, is defined as the union of a finite set of segments, say l_i, lying in a planar region S (Ang, Baddeley and Nair (2012), Baddeley et al. (2021))

$$l_i = [\mathbf{u}_i, \mathbf{v}_i] = \{\mathbf{s} : \mathbf{s} = t\mathbf{u}_i + (1-t)\mathbf{v}_i; 0 \le t \le 1\},$$

where $\mathbf{u}_i$ and $\mathbf{v}_i$ denote the endpoints of l_i stored using an appropriate coordinate reference system (CRS). In this paper we adopt a projected CRS (EPSG code: 3003)[2] that expresses the units in metres.

The computational representation of Milan's road network adopted in this paper was created using data downloaded from Open Street Map (OSM) servers, and, in particular, we used the *openstreetmap.fr*[3] provider via the `R` package `osmextract` (R Core Team (2020), Gilardi and Lovelace (2021)). OSM is a project that aims to build an open and editable map of the world (OpenStreetMap contributors (2017), Barrington-Leigh and Millard-Ball (2017)). The basic components of OSM data are called *elements*, and they are divided into *nodes*, which represent points on the earth's surface, *ways*, which are ordered lists of nodes, and *relations*, which are lists of nodes, ways, and other relations where each member has additional information that describes its relationship with the other elements.

We downloaded OSM road data for Lombardia (the region of Northern Italy where Milan is located), and using a spatial operation, we retained only the OSM elements that lay inside Milan's polygonal boundary. Then we selected only those segments that correspond to the most important streets of Milan, focusing on the following classes[4] (listed in descending order of importance): *motorways*, *trunks*, *primary*, *secondary*, *tertiary*, *unclassified*, and *residential*. We created a road network with 20,064 segments that longs approximately 194 km, including the majority of the most important streets in Milan.

The road network spreads all around the city and is depicted in Figure 1d. The white areas clearly identify suburb/nonurban places and some of the most iconic locations in Milan, like Parco Sempione, Giardini Indro Montanelli, or City Life.

After creating the road network, we excluded all emergency calls whose GPS locations were found farther than 50 meters away from the closest segment of the network,[5] since they probably occurred on other minor streets not included in the considered network. Approximately 14,000 events were discarded, and the remaining interventions were projected to their nearest point of the network. The final sample included 480,252 events.

Finally, we explored the spatiotemporal dynamics, testing the presence of space-time interactions. First, we split all EMS interventions into 12 two-hours classes according to their occurrence times. Then we calculated (independently for each class) a smoothed intensity surface using the convolution kernel estimator detailed in Rakshit et al. (2019). The result is reported in Figure 5. We notice that, from 10 a.m. to 6 p.m., the smoothed intensity peaks in the proximity of Duomo and the city centre, whereas during the night hours, the interventions

[2] See https://epsg.io/3003 for more details. Last access: 2023-03-06.

[3] See http://download.openstreetmap.fr/. Last access: 2021-12-09.

[4] We refer to https://wiki.openstreetmap.org/wiki/Highways for a comprehensive description of road network data in Open Street Map and guidelines for its classification system. We also refer to https://wiki.openstreetmap.org/wiki/IT:Key:highway for a comparison between the Italian classification system and the classes defined by OSM.

[5] The distance between a point and a segment was measured using the shortest euclidean perpendicular distance.

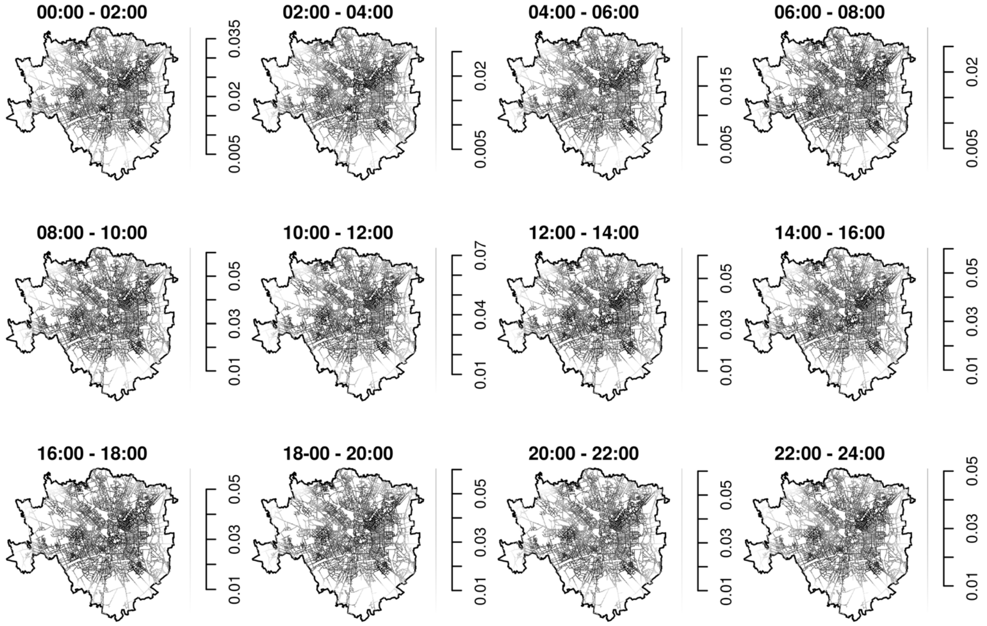


FIG. 5. *Smoothed intensity functions of EMS interventions in Milan estimated after classifying the events into two-hours classes.*

are concentrated in the proximity of night-life areas. In both scenarios there are some clear hot spots near the main train station (i.e., Milano Centrale) and several nursing homes. The spatial model detailed in Section 3 also includes these potential space-time interactions.

3. Statistical model.

Let L be a continuous one-dimensional spatial region and $\mathcal{T} = \{1, 2, \dots, T\}$ a discrete temporal dimension divided into intervals of one hour. As mentioned before, in this paper L represents the street network of Milan while T is equal to 26,304, that is, the number of hours from 2015-01-01 00:00 to 2018-01-01 00:00.

Let y_t denote the number of ambulance interventions that occurred in the network L at time $t \in \mathcal{T}$, and let $\boldsymbol{s}_{t,i}$, $i = 1, \dots, y_t$, denote the location of the ith event. We assume that, for each $t \in \mathcal{T}$, $\{\boldsymbol{s}_{t,i} : i = 1, \dots, y_t\}$ is a realisation of a *nonhomogeneous Poisson Process* (NHPP) on a linear network with *intensity function* $\lambda_t(\boldsymbol{s})$ (Okabe and Sugihara (2012), Diggle (2014), Baddeley et al. (2021)). A NHPP on a linear network satisfies the following two properties:

- The number of events occurring in $L' \subseteq L$, which is denoted by $N(L')$, follows a Poisson distribution with parameter $\int_{L'} \lambda_t(\boldsymbol{s})\, \mathrm{d}_1\boldsymbol{s}$, where L' represents a finite portion of the network L and $\mathrm{d}_1\boldsymbol{s}$ denotes integration with respect to arc-length measure;
- Let $N(L) = y_t$; then the y_t events represent a random sample from a distribution whose probability density function is proportional to $\lambda_t(\boldsymbol{s})$.

We assume that the intensity function of the process can be decomposed as

$$\lambda_t(\boldsymbol{s}) = \mu_t g_t(\boldsymbol{s}) \quad \text{for } \boldsymbol{s} \in L, \tag{1}$$

where μ_t and $g_t(\boldsymbol{s})$ represent the temporal and spatial dimension at time t, respectively. We also assume that $g_t(\boldsymbol{s})$ satisfies the following two conditions:

- $g_t(\boldsymbol{s}) > 0 \ \forall t \in \mathcal{T}$ and $\forall \boldsymbol{s} \in L$,
- $\int_L g_t(\boldsymbol{s})\, \mathrm{d}_1\boldsymbol{s} = 1 \ \forall t \in \mathcal{T}$,

which imply $\mu_t = \int_L \lambda_t(\boldsymbol{s})\, \mathrm{d}_1\boldsymbol{s}$. Therefore, considering the NHPP hypothesis, we notice that $y_t | \lambda_t \sim \text{Poisson}(\mu_t) \ \forall t \in \mathcal{T}$, where μ_t represents the *total volume* of ambulance dispatches at time t. Moreover, under the same assumptions, we have $\boldsymbol{s}_{t,i} | \lambda_t, y_t \overset{\text{i.i.d.}}{\sim} g_t(\boldsymbol{s})$ for $i = 1, \dots, y_t$, highlighting that $g_t(\boldsymbol{s})$ denotes the spatial density of ambulance interventions at time t. Finally, we remark that our modelling scheme entails a temporal correlation among the interventions, but it assumes that, for a fixed period t, the spatial locations are independent, given $\lambda_t(\boldsymbol{s})$, and conform to an inhomogeneous Poisson process.

Equation (1), despite being similar to the classical separability assumption for spatiotemporal point processes (Møller and Waagepetersen (2004), Diggle (2014)), suggests that the spatial component evolves over time. This particular functional form is motivated by the space-time interactions observed in the hourly evolution of ambulance interventions that were displayed in Figure 5. In fact, as reported by several authors (see, e.g., González et al. (2016) and references therein), the separability of the first-order intensity function is usually taken as a working assumption in order to simplify the estimation process. However, given the exploratory analysis detailed before, we believe that this is not appropriate for our case. Therefore, following the ideas in Zhou and Matteson (2015), in this paper we propose adopting a nonseparable first-order intensity function readapted to analyse ambulance interventions as a point pattern on a linear network. As explained in Section 3.2, the space-time interactions are modelled using an appropriate set of weights.

Hereinafter, we introduce two statistical models for μ_t and $g_t(\boldsymbol{s})$, respectively. The temporal component is modelled using a semiparametric Poisson regression with smoothed deterministic calendar covariates, namely, the hour of the day, the day of the week, the week of the year, and an additional term allowing yearly fluctuations in the expected EMS counts. The spatial dimension is modelled nonparametrically using a network readaptation of a weighted kernel density estimator (KDE).

3.1. *The temporal model.* As mentioned above, the term μ_t represents the expected number of interventions that occurred over the network L at time $t \in \mathcal{T}$. Following the suggestions in Diggle, Rowlingson and Su (2005), Bayisa et al. (2020), we model y_t using a Poisson regression. To incorporate smoothness into the model, generalized additive models (GAMs) are used in the estimation of μ_t (Wood (2011)). GAMs extend generalized linear models, allowing for nonlinear relationships between the response variable and the covariates.

Being y_t the observed number of emergency interventions in the linear network L at time t, under a Poisson distribution assumption one has $\mu_t = E(y_t)$, and the log-linear Poisson additive regression model is given by

$$\log(\mu_t) = \beta_0 + \beta_1 \cdot \text{year}_t + \beta_2(\text{hour}_t) + \beta_3(\text{week}_t),$$

where β_0 denotes the intercept and $\text{year}_t = 0, 1, 2$ represents the year of the event occurred at time t with respect to 2015. In addition, hour_t represents the hour of the day (taking values from 0 to 23), while week_t represents the week of the year (taking values from 1 to 53). The notation $\beta_j(x),\ j = 2, 3$ represent a spline transformation, that is, $\beta_j(x) = \sum_{r=1}^{k_j} b_{jr}\gamma_{jr}(x)$, where $\gamma_{jr}(x), r = 1, \dots, k_j$ are the basis functions and b_{jr} the unknown coefficients. In particular, a cyclic cubic regression spline is adopted since in our context it is appropriate to assume a smooth transition between the last hour of one day and the first hour of the next day as well as between the last week of one year and the first week of the next year (see Figures 2 and 3).

To account for potential different impacts among the days of the week (see Figure 3), an interaction term was also included in the linear predictor. Hence, the final model fitted to the data writes as follows:

$$\log \mu_t = \beta_0 + \beta_1 \cdot \text{year}_t + \text{dow}_t + \text{dow}_t \times \beta_2(\text{hour}_t) + \beta_3(\text{week}_t). \tag{2}$$

The term dow_t is a factor variable that represents the day of the week.

3.2. *The spatial model.* The spatial component of the intensity function, previously denoted by $g_t(\boldsymbol{s})$, is modelled nonparametrically using a network readaptation of a Jones–Diggle corrected weighted kernel density estimator (Diggle (1985), Jones (1993), Rakshit et al. (2019)). The weights are computed using a weight function that takes into account the space-time interactions described in Section 2.

More precisely, given a set of observed time periods $\mathcal{T}$, an hour u, and a location $\boldsymbol{s}$ on the network, the weighted kernel estimator can be written as

$$\hat{g}_u(\boldsymbol{s}) = \frac{\sum_{t\in\mathcal{T}} \sum_{i=1}^{y_t} w(t,u) K_N(\boldsymbol{s}, \boldsymbol{s}_{t,i}; h)}{\sum_{t\in\mathcal{T}} \sum_{i=1}^{y_t} w(t,u)}, \tag{3}$$

where $w(t,u)$ represents the weight associated to the $\boldsymbol{s}_{t,i}$ ambulance intervention and $K_N(\boldsymbol{s}, \boldsymbol{s}_{t,i}; h)$ denotes the Jones–Diggle corrected network kernel function. Following the proposal in Rakshit et al. (2019), the kernel function is defined as

$$K_N(\boldsymbol{s}, \boldsymbol{s}_{t,i}; h) = \frac{K(\boldsymbol{s} - \boldsymbol{s}_{t,i}; h)}{c_L(\boldsymbol{s}_{i,t})}, \tag{4}$$

where $K(\boldsymbol{s} - \boldsymbol{s}_{t,i}; h)$ denotes a planar Gaussian kernel with bandwidth h and $c_L(\boldsymbol{s}_{i,t}) = \int_L K(\boldsymbol{s} - \boldsymbol{s}_{i,t}, h)\,\mathrm{d}_1\boldsymbol{s}$ represents the convolution of kernel K with arc-length measure on the network.

The KDE in equation (4) is one of the most relevant examples of statistical estimators for spatial network data that is based on euclidean distances instead of shortest path distances. For this reason, it can be expressed in terms of convolutions of two-dimensional planar kernels and can be computed extremely efficiently using the fast Fourier transformation (FFT) (Silverman (1982)). More precisely, the numerator in (4) can be expressed as the convolution of a kernel K with respect to the counting measure on the data points, whereas the denominator can be expressed as a convolution with respect to arc-length measure on the network. In both cases the estimators can be computed rapidly using the FFT after discretising the point pattern and the linear network via a fine pixel grid.

We end this section observing that, although the suggested kernel approach does not consider shortest-path distances computed on the network, the structure of the spatial domain is still taken into account by the denominator in equation (4). Moreover, as discussed in Rakshit et al. (2019), the proposed technique consistently estimates the intensity function of a point process on a linear network, and its statistical efficiency is only slightly suboptimal with respect to other approaches (see, e.g., McSwiggan, Baddeley and Nair (2017)), whereas the computational advantages are enormous for large networks as the one considered in this paper.

3.2.1. *Defining the weight function.* The weight function $w(t,u)$ is used to capture the contribution of each past observation to predict the future ambulance demand by taking advantage of EMS data temporal patterns to improve the forecasting of future interventions. It incorporates space-time interactions into the weighted kernel estimator, creating a nonseparable structure in the spatiotemporal intensity $\lambda_t(\boldsymbol{s})$. In particular, we assumed that $w(t,u)$ can be modelled as a function of the time lag between u and t, say $m = u - t$. The following functional form, first proposed by Zhou and Matteson (2015), was adopted:

$$w(t,u) = w(u-t) = w(m) = \rho_1^{(m)} + \rho_2^{(m)} \rho_3^{\sin^2(\frac{\pi m}{24})} \rho_4^{\sin^2(\frac{\pi m}{168})}. \tag{5}$$

Equation (5) includes a separate coefficient for each seasonal pattern displayed in Figure 4: ρ_1 captures the short-term dependence while ρ_3 and ρ_4 measure the daily and weekly seasonalities with a periodicity equal to 24 and $24 \times 7 = 168$ hours, respectively. The coefficient

ρ_2 represents a discount factor added to fade out the product of daily and seasonal terms, whereas the term π denotes the constant value 3.1415.... The four parameters are bounded between 0 and 1 to avoid (unrealistic) exponential growths. Consequently, $w(m)$ takes values in $(0, 2)$, which also prevents negative weights that would potentially result in a negative kernel estimate of the density function.

Unfortunately, estimating $\rho_1, \dots, \rho_4$ in a full likelihood-based approach entails a nontrivial computational burden. Consequently, we implemented an algorithm first suggested in Zhou and Matteson (2015). As mentioned above, the weight function aims to grasp the time regularities of EMS interventions (also displayed in Figure 5), giving more importance to those events that occurred in the proximity of the seasonality peaks. Therefore, $w(m)$ should reflect the temporal dependency depicted by the ACF of the hourly number of EMS interventions, mirroring the behaviour displayed in Figure 4 and assigning a negligible weight to those observations that, from a temporal perspective, are unlikely to be important for future predictions.

For this reason and after calculating the empirical hourly ACF up to lag M and taking its positive part, denoted by $\text{ACF}^+ = \max(0, \text{ACF})$, the parameters $\rho_1, \dots, \rho_4$ were estimated by minimising the following loss function

$$(6) \qquad \frac{1}{M}\sum_{m=1}^{M}\left(\text{ACF}^+(m) - \rho_0 w(m)\right)^2 \quad \text{s.t. } 0 < \rho_j < 1 \ \forall j = 0, \dots, 4.$$

The coefficient ρ_0 represents a further discount factor without any practical interpretation. It is used to scale $w(\cdot)$ between 0 and 1, in order to make it consistent with the ordinate range of ACF^+. In this paper we choose $M = 672$, which represents four weeks of historical temporal data, whereas the minimisation problem was solved using the *box-constrained* method implemented in the `R` function `optim()`, initialising all parameters at a random value between 0 and 1 (Byrd et al. (1995), R Core Team (2020)).

4. Results.

We now present the results obtained when estimating the spatial and temporal models described in Sections 3.1 and 3.2. All procedures were implemented using the software `R` (R Core Team (2020)) and several contributed packages. More precisely, the smooth temporal components were estimated using the package `mgcv` (Wood (2017)), while the network-version of the Gaussian weighted kernel is implemented in the package `spatstat` (Baddeley (2015)). We fitted the temporal model, the weight function, and the spatial kernel using a laptop with an AMD Ryzen 5 3500U with Radeon Vega Mobile Gfx 2.10 GHz processor, four cores and eight GB of RAM. After downloading the OSM data, it took approximately five minutes to build the computational representation of the point pattern on the street network, two minutes to estimate the temporal model and the parameters in the weight function, and three minutes to compute the spatial kernel estimator considering two different future time periods.

4.1. *The temporal component.*

As detailed in Section 3.1, the temporal component was estimated using a GAM with deterministic predictors representing yearly fluctuations and hourly, daily, and weekly seasonal components. The cyclic cubic spline terms are included to capture the smooth intraday dynamics and the weekly temporal trends. The daily effects are taken into account by considering an interaction term and a set of dummy variables. The observed counts originally presented five missing values from 2015-04-01 at 00:00 to 2015-04-01 at 04:00. These values were imputed using a GAM as in equation (2) that was trained using the interventions until 2015-03-31 at 23:00.

The estimates of β_0, β_1, and dow_t coefficients for the complete model are summarised in Table 1. The intercept represents the (logarithm of) the mean number of interventions per

TABLE 1

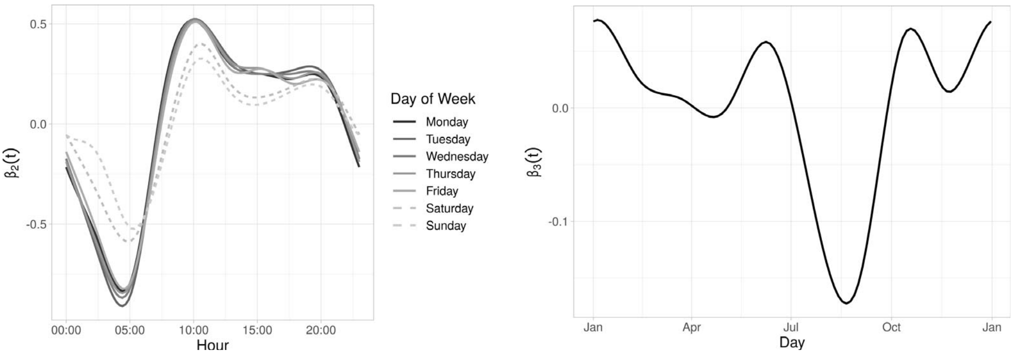
Estimates of the effects obtained after fitting the model described in equation (2). The reference category for the daily seasonal term is Sunday

	Estimate	Standard error	z-value	p-value
(Intercept)	2.8333	0.0044	648.9070	<0.001
Year	0.0116	0.0018	6.5792	<0.001
Monday	0.0193	0.0057	3.3835	<0.001
Tuesday	−0.0244	0.0058	−4.2161	<0.001
Wednesday	−0.0177	0.0058	−3.0807	0.002
Thursday	−0.0137	0.0057	−2.3937	0.017
Friday	0.0024	0.0057	0.4292	0.668
Saturday	−0.0052	0.0057	−0.9218	0.357

hour, β_1 is the annual variation in the EMS counts with respect to 2015, and the remaining set of coefficients represents the deviation from the reference level, that is, Sunday. The estimates highlight the presence of a tiny but significant and positively increasing trend in the hourly number of interventions per year. Moreover, we can notice that the behaviour of the ambulance interventions during the weekend looks quite different from the first days of the week (e.g., Monday to Thursday), with Monday being the day when the majority of interventions take place, whereas Friday and Saturday are found to be not significantly different from Sunday at the usual significance levels.

The smooth seasonal terms are depicted in Figure 6. In particular, Figure 6a reports the (smoothed) daily effects for each day of the week. The seven curves mirror the shapes displayed in Figure 3, and a clear distinction exists between weekends and weekdays. In all cases we found a peak around 10 a.m. Figure 6b shows the smoothed weekly effects, which look similar to the patterns displayed in Figure 2. We observe a drop in the expected number of ambulance dispatches around August. In both cases the cyclic cubic regression splines were fitted using $K = 10$ knots evenly placed throughout the values of the covariates.

Finally, we explored the predictive accuracy of the GAM using the following strategy. First, we trained the model considering all ambulance interventions up to 2017-10-01 at 00:00, and then we forecasted the EMS counts until the end of the year. We compared the observed counts with the out-of-sample fitted values, and the result is displayed in Figure 7 that suggests a good agreement between the two time series.



(a) Intra-day effects

(b) Weekly seasonality

FIG. 6. *Estimates of the smooth seasonal terms obtained after fitting the model described in Equation (2). Figure (a) represents the intraday effects divided by the day of the week. Figure (b) displays the weekly temporal trends.*

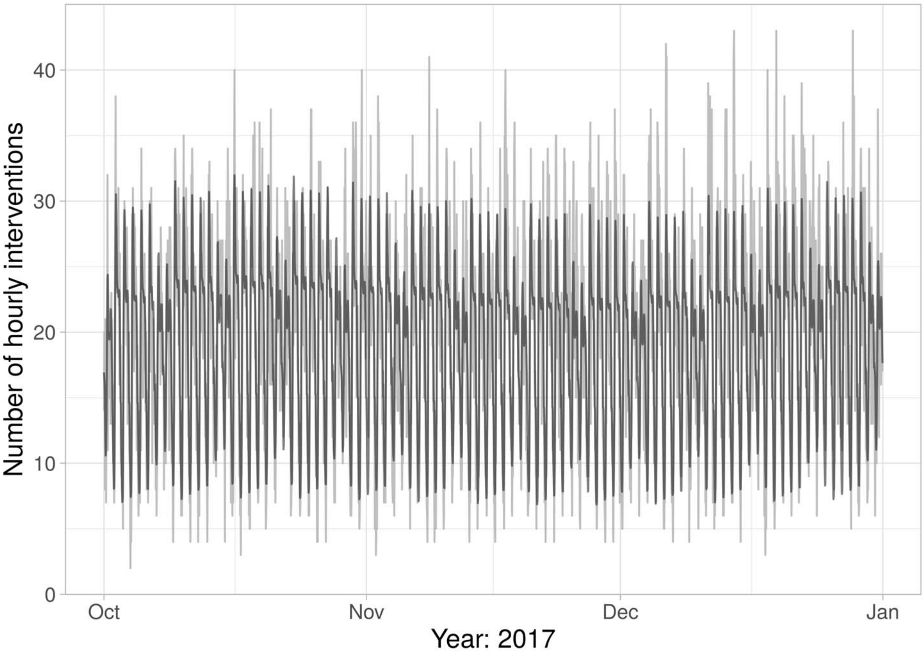


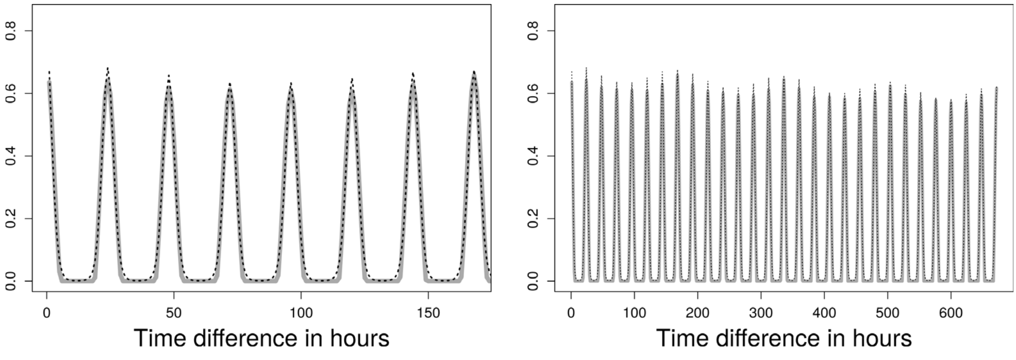
FIG. 7. *Graphical comparison between observed counts* (*in grey*) *and out-of-sample fitted values* (*in black*) *considering EMS data from* 2017-10-01 *to* 2017-12-31.

4.2. *The spatial component.* As introduced in Section 3.2, the spatial component was estimated combining a network readaptation of a Gaussian smoothing kernel with a weight function that measures the predictive importance of each past EMS intervention. The weights are used to mimic the interactions between two ambulance dispatches separated by l temporal lags, replicating the hourly, daily, and weekly seasonalities in the ACF displayed in Figure 4.

4.2.1. *Estimating the weight function.* As reported in equation (5), the weight function depends on four parameters that represent the three seasonal components plus a discount factor. They were estimated solving the minimisation problem detailed in equation (6). We found ρ_1 as big as 0.213, pointing out a mildly strong short-term correlation in the EMS counts. The second seasonal parameter, that is, ρ_3, was found equal to 0.002, which means that the effects related to the daily component range between 0.002 and 1. Given the periodic behaviour of the sinusoid function, the maximum value of $\hat{\rho}_3^{\sin^2(\frac{\pi l}{24})}$ is obtained when the lag l is approximately a multiple of 24, while the minimum is reached when the time difference is close to 12 or its odd multiples. The value of $\hat{\rho}_4$ was found equal to 0.927, pointing out that the weekly effects are smoother and oscillate between 0.927 and 1. Finally, the fitted values of ρ_0 and ρ_2, that is, the two discount factors, were found equal to 0.695 and 0.999, respectively, meaning that daily and weekly seasonality vanish slowly.

We display in Figure 8 a graphical comparison between the observed positive part of the hourly ACF and the estimates of the weight function. Figure 8a shows one week of lagged counts, while Figure 8b shows the complete set of lags up to four weeks. In both cases the weight function successfully fits the ACF.

4.3. *Spatial and spatiotemporal component.* After estimating the weight function, we applied equation (3) to obtain the predicted spatial density $\hat{g}_u(\boldsymbol{s})$ for a time period u. In particular, considering that the data at hand included the EMS interventions from 2015-01-01 at 00:00 to 2017-12-31 at 23:59, we decided to forecast $\hat{g}_u(\boldsymbol{s})$ considering two randomly



(a) One week of lagged counts

(b) Four weeks of lagged counts

FIG. 8. *The observed positive part of ACF* (*grey solid line*) *and the estimated weight function* (*black dashed line*) *considering lagged counts for* (*a*) *one week and* (*b*) *four weeks.*

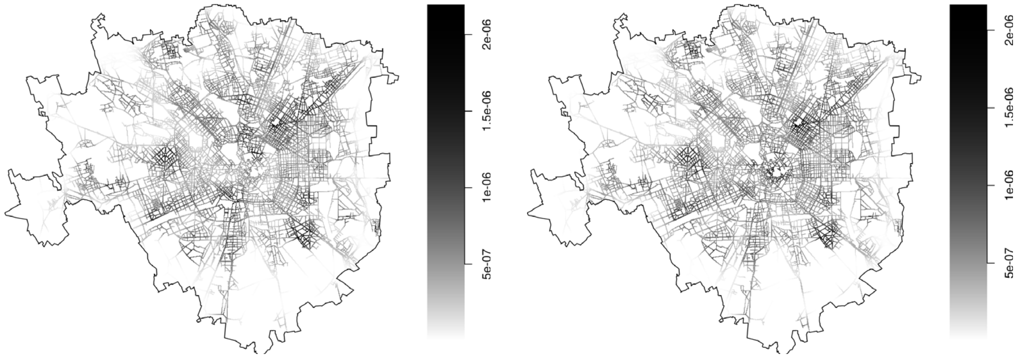
selected future time periods: 2018-01-03 at 03:00 and 2018-01-03 at 15:00. The first one falls at night, while the other one falls in the early afternoon. In both cases the value of the bandwidth h was chosen using the rule of thumb suggested by Rakshit et al. ((2019), Section 8). The bandwidth is given by

$$h = (3n)^{-1/5}\bar{s}, \tag{7}$$

where $\bar{s} = \sqrt{s_1^2 + s_2^2}$ and s_j, $j = 1, 2$, denote the sample standard deviation of the jth Cartesian coordinate values for the locations of the ambulance interventions. Equation (7) adapts the rule of thumb proposed by Scott ((1992), page 152) in kernel density estimation to the analysis of spatial data lying on a one-dimensional domain.

The results are reported in Figure 9. Figure 9a shows that during the night the EMS interventions are spread in several parts of the municipality and highlights some roads of the network nearby night-life areas. Figure 9b underlines that ambulance dispatches are concentrated in the areas close to Duomo and other relevant working places during daytime, whereas nightlife areas are no longer highlighted by the model. In both cases the central station, a popular square (Piazzale Corvetto), and several retirement houses (such as Pio Albergo Trivulzio) are pointed out.

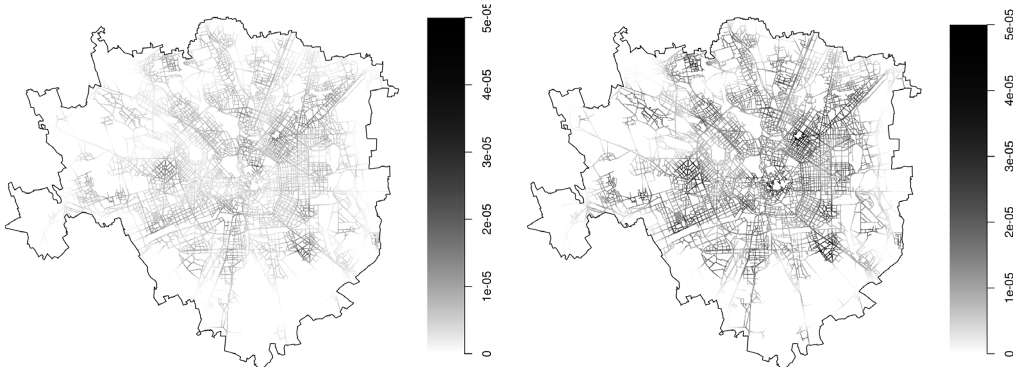
The values of $\hat{g}_u(\boldsymbol{s})$, displayed in Figure 9, represent only the spatial dimension of the data. Hence, to compare and visualise the spatiotemporal evolution of $\lambda_u(\boldsymbol{s})$, we estimated



(a) 2018-01-03 03:00

(b) 2018-01-03 15:00

FIG. 9. *Estimates of spatial density function* $\hat{g}_u(\boldsymbol{s})$ *considering two future time periods*: 2018-01-03 *at* 03:00 (*a*) *and* 2018-01-03 *at* 15:00 (*b*). *The unit for the color scale is* 1/*m*.

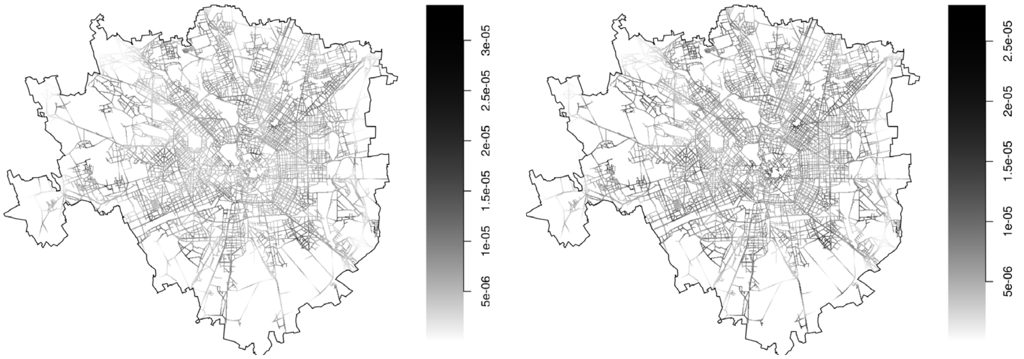


(a) 2018-01-03 03:00 (b) 2018-01-03 15:00

FIG. 10. *Estimates of the intensity function* $\hat{\lambda}_u(\boldsymbol{s})$ *considering two future time periods*: 2018-01-03 *at* 03:00 (*a*) *and* 2018-01-03 *at* 15:00 (*b*). *The two maps highlight the temporal patterns in spatial locations of emergency interventions. The unit for the color scale is* 1/*m. The values represent the expected number of ambulance dispatches occurring in a small linear neighbourhood around a point of the network.*

the expected number of interventions at time u and calculated $\hat{\lambda}_u(\boldsymbol{s})$ by multiplying the spatial and the temporal components. The results are depicted in Figure 10. Although the two maps highlight areas similar to those displayed in Figure 9, they now account for the temporal patterns of EMS interventions. In particular, considering that the majority of ambulance dispatches occur between 8 a.m. and 6 p.m., the intensity function at 15:00 was found higher than in the other scenario.

Finally, using the approach described in Rakshit et al. ((2019), Section 6.2), we estimated the pointwise standard errors of the spatial density function $\hat{g}_u(\boldsymbol{s})$ considering the same two future time periods. The results are reported in Figure 11. In both cases the two maps highlight certain areas of the municipality in the proximity of nightlife neighbourhoods or train and metro stations. We can also clearly recognise the shape of some arterial thoroughfares (e.g., the A51 motorway located on the rightmost part of the maps and far from any residential neighbourhood) passing through the city. Unsurprisingly, the standard error estimates generally increase in the proximity of the city boundary since they are based on fewer data points.



(a) 2018-01-03 03:00 (b) 2018-01-03 15:00

FIG. 11. *Estimates of the standard errors of the spatial density function* $\hat{g}_u(\boldsymbol{s})$ *considering two future time periods*: 2018-01-03 *at* 03:00 (*a*) *and* 2018-01-03 *at* 15:00 (*b*). *The values were obtained using the formulas described in Rakshit et al.* ((2019), *Section* 6.2).

5. Spatial validation. Section 4 presented the results obtained when fitting the spatial and temporal components of $\lambda_t(\boldsymbol{s})$ and omitted any consideration on the spatial predictive accuracy. Nevertheless, as mentioned in the Introduction, the algorithms used to minimise the ambulance response times require a model that can produce reliable forecasts of the spatial distribution of the EMS events at the road network level. In this section we discuss the procedure adopted to validate our proposal.

The predictive accuracy of $\hat{g}_t(\boldsymbol{s})$ was first inspected graphically by comparing observed and predicted EMS interventions at different levels of temporal aggregation via a network-readaptation of the *relative-risk* function (McSwiggan, Baddeley and Nair (2020)). More precisely, given two point patterns $\boldsymbol{A}$ and $\boldsymbol{B}$ that occur on the same network L and a point $\boldsymbol{s} \in L$, we define the *(normalised) relative-risk function* (also named *probability distribution of one type* in Baddeley ((2015), Chapter 14)) as

$$\rho(\boldsymbol{s}) = \frac{g_{\boldsymbol{A}}(\boldsymbol{s})}{g_{\boldsymbol{A}}(\boldsymbol{s}) + g_{\boldsymbol{B}}(\boldsymbol{s})},$$

where $g_{\boldsymbol{A}}(\boldsymbol{s})$ and $g_{\boldsymbol{B}}(\boldsymbol{s})$ denote the spatial densities of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. The plug-in estimator of $\rho(\boldsymbol{s})$ is given by

$$\hat{\rho}(\boldsymbol{s}) = \frac{\hat{g}_{\boldsymbol{A}}(\boldsymbol{s})}{\hat{g}_{\boldsymbol{A}}(\boldsymbol{s}) + \hat{g}_{\boldsymbol{B}}(\boldsymbol{s})}, \tag{8}$$

where $\hat{g}_{\boldsymbol{A}}(\boldsymbol{s})$ and $\hat{g}_{\boldsymbol{B}}(\boldsymbol{s})$ are kernel estimates of $g_{\boldsymbol{A}}(\boldsymbol{s})$ and $g_{\boldsymbol{B}}(\boldsymbol{s})$, respectively. The value of $\hat{\rho}(\boldsymbol{s})$ represents the probability that a point $\boldsymbol{s} \in L$ belongs to $\boldsymbol{A}$ instead of $\boldsymbol{B}$. Values of $\hat{\rho}(\boldsymbol{s})$ around 0.5 highlight that the relative risk function cannot discern the two processes. We refer to McSwiggan, Baddeley and Nair (2020) for more details and extensive theoretical and computational considerations regarding the estimation of the relative risk for point patterns on linear networks.

As mentioned before, the forecasting accuracy of $\hat{g}_t(\boldsymbol{s})$ was tested by comparing observed points and (out-of-sample) predictions. More precisely, we first selected the EMS interventions that occurred before the end of September 2017 and trained the weight function to estimate $\rho_0, \dots, \rho_4$. Then, considering the temporal evolution of the weights, we derived the spatial KDE $\hat{g}_u(\boldsymbol{s})$ for each hour u of a given (out-of-sample) time period $\mathcal{U}$, and finally, we obtained an out-of-sample prediction of the emergency events by sampling y_u points from each density $\hat{g}_u(\boldsymbol{s}),\ u \in \mathcal{U}$. Finally, we aggregated observed occurrences and predicted points over $\mathcal{U}$ and compared the two types of events by means of the relative risk function. The pseudo-code that summarises this procedure is reported in Algorithm 1.

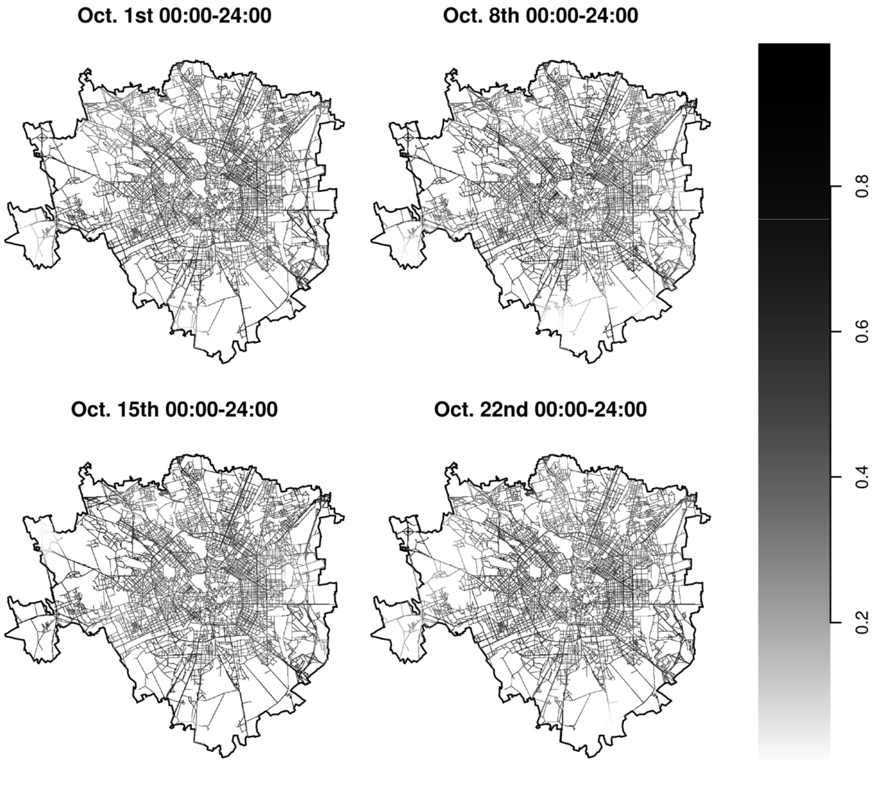
In the analysis reported below, we tested the spatial accuracy considering four days placed farther and farther in time from the end of the training period, namely, 2017-10-01, 2017-10-08, 2017-10-15, and 2017-10-22. We decided to focus on several days spread over a month close to the end of the training set since that represents a realistic scenario to organise the ambulance shifts.

After simulating the ambulance interventions for each time period and extracting the corresponding observed EMS events, the relative risk function was computed as in McSwiggan, Baddeley and Nair (2020). The same bandwidth h was used when applying equation (3) to the two types of points, and its value was estimated using a readaptation of Scott's rule of thumb for one-dimensional coordinates data, as suggested in McSwiggan, Baddeley and Nair ((2020), page 5). We did not explore the other techniques for bandwidth selection due to the prohibitive computational costs of applying leave-one-out cross-validation to large road networks with hundreds of thousands of points and time-consuming out-of-sample simulations.

The relative risk functions $\hat{\rho}(\boldsymbol{s})$ for the four days under analysis are depicted in Figure 12. Following the notation adopted in equation (8), the object $\boldsymbol{A}$ denotes the observed EMS

Algorithm 1: Pseudocode describing the procedure used to simulate future events for a nonseparable model. The algorithm can be adapted to sample from a separable model (see Section 6) skipping steps 2 and 3. In both cases, after the for loop, we aggregate all predicted events collapsing the temporal dimension

```
Input: data: ambulance interventions data; $\mathcal{U}$: set of future time periods.
Output: An estimate of the normalised relative risk function.
/* 1. subset EMS data that occurred before 2017-10-01 at
   00:00.                                                    */
train ← subset(data, date-occurrence < 2017-10-01 00:00);
/* 2. estimate $\rho_0, \ldots, \rho_4$ using the methods described in
   Section 3.2.1                                             */
$\{\hat{\rho}_0, \ldots, \hat{\rho}_4\}$ ← estimate_coefs(train);
for each $u \in \mathcal{U}$ do
    /* 3. Estimate the weights $w(t, u)$ given $\hat{\rho}_0, \ldots, \hat{\rho}_4$. See
       equation (5).                                         */
    $w$ ← estimate_weights($u$, $\hat{\rho}_0, \ldots, \hat{\rho}_4$);
    /* 4. Estimate $\hat{g}_u(\boldsymbol{s})$ using equation (3).             */
    $\hat{g}_u$ ← estimate_KDE(ems_train, $w$);
    /* 5. Simulate $y_u$ events sampling from a probability
       density function on a linear network equal to $\boldsymbol{g}_u$    */
    pred_events(u) ← simulate_points($\hat{g}_u$, $y_u$)
end
```


Oct. 1st 00:00-24:00

Oct. 8th 00:00-24:00

Oct. 15th 00:00-24:00

Oct. 22nd 00:00-24:00

FIG. 12. *Spatial representations of the (normalised) relative risk function considering four different days.*

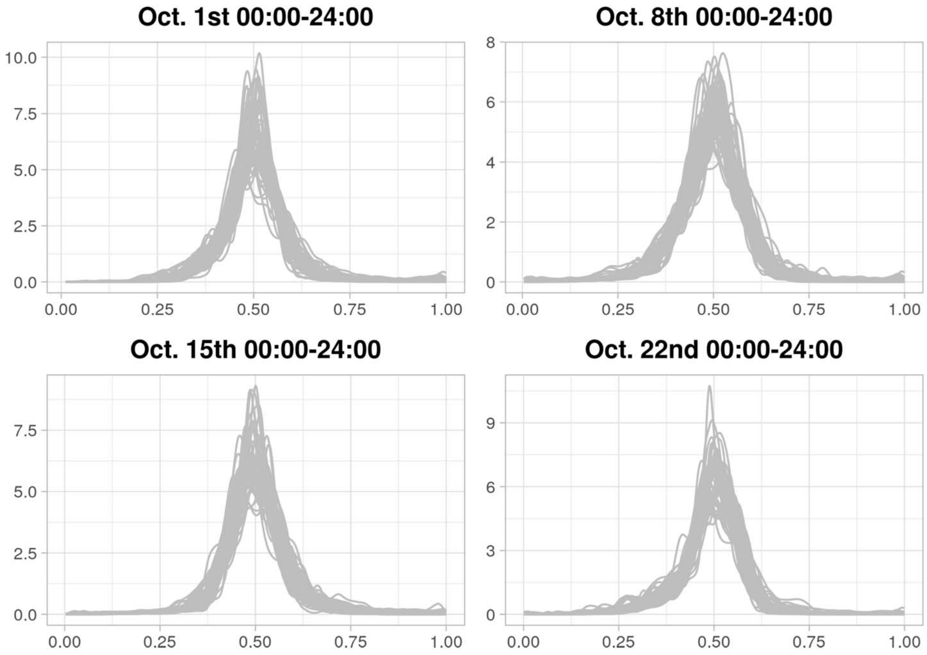


FIG. 13. *Density curves representing* 50 *simulations of the* (*normalised*) *relative risk function considering four days*.

events, while $\boldsymbol{B}$ represents the predicted points. A spatiotemporal EMS model successfully predicts future emergency events, when $\hat{\rho}(\boldsymbol{s})$ is close to 0.5, since that implies the relative risk function cannot distinguish between observed and predicted cases. Figure 12 shows that, in the four cases, the relative risk functions are always concentrated around 0.5 but for a few parts in the suburban areas, suggesting that our approach can be employed for EMS events forecasting.

We repeated the procedure listed in Algorithm 1 50 times obtaining, for each pixel composing the road network, several estimates of the relative risk function. The smoothed curves displayed in Figure 13 represent the distribution of $\hat{\rho}(\boldsymbol{s})$ for each time period and for each simulation. In all cases these curves are concentrated around 0.5 and the value of $\hat{\rho}(\boldsymbol{s})$ lies between 0.4 and 0.6 for approximately 80% of all road pixels. Moreover, the Supplementary Material (Gilardi, Borgoni and Mateu (2024)) reports the results obtained when testing the spatial accuracy for different time periods and training sets. In particular, the procedure detailed above was replicated by considering two alternative time intervals of six and twelve hours, respectively. We also tested the sensitivity of our results by shifting the training and test sets ahead of one and two months, respectively. We found that the suggested approach successfully predicts future events in all scenarios under consideration.

To conclude, we readapted the ideas in Kelsall and Diggle (1998, 1995) to our context by constructing a procedure that investigates the spatial variation of $\rho(\boldsymbol{s})$ and tests whether two processes, defined on a common network L (e.g., observed and simulated future emergency interventions), have the same intensity function. More formally, given two point patterns $\boldsymbol{A}$ and $\boldsymbol{B}$ with, respectively, $n_{\boldsymbol{A}}$ and $n_{\boldsymbol{B}}$ observations, the aforementioned papers proposed a Monte Carlo test for departure from a null hypothesis of random labelling, that is, H_0 : $\tilde{\rho}(\boldsymbol{s}) = \log\{g_{\boldsymbol{A}}(\boldsymbol{s})/g_{\boldsymbol{B}}(\boldsymbol{s})\} = 0$ or, equivalently, $H_0 : \rho(\boldsymbol{s}) = \frac{g_{\boldsymbol{A}}(\boldsymbol{s})}{g_{\boldsymbol{A}}(\boldsymbol{s})+g_{\boldsymbol{B}}(\boldsymbol{s})} = 0.5$. The test was implemented by generating m new datasets which are consistent with H_0 but, otherwise, have similar characteristics with respect to the original processes. The authors used the following test statistics:

$$t_j = \int \log\left\{\frac{\hat{g}_{j,\boldsymbol{A}}(\boldsymbol{s})}{\hat{g}_{j,\boldsymbol{B}}(\boldsymbol{s})}\right\}^2 \mathrm{d}(\boldsymbol{s}), \quad j = 1, \dots, m, \tag{9}$$

where $\hat{g}_{j,\boldsymbol{A}}(\boldsymbol{s})$ and $\hat{g}_{j,\boldsymbol{B}}(\boldsymbol{s})$, respectively, represent kernel estimates of the density function of the two processes $\boldsymbol{A}$ and $\boldsymbol{B}$ for the jth simulated dataset and t_O is the observed value of the test statistics on the original point patterns. The p-value can be computed as $p = \frac{k+1}{m+1}$, where k is the number of times that $t_j > t_O$.

Conditional on the location of the points, the null hypothesis states that the probability that a given event belongs to $\boldsymbol{A}$ instead of $\boldsymbol{B}$ does not depend on the spatial location and is constant over the region. Therefore, the generation of datasets under H_0 can be performed by combining the two original point patterns into a unique object and randomly labelling $n_{\boldsymbol{A}}$ of them as coming from process $\boldsymbol{A}$ and the remaining ones as type $\boldsymbol{B}$.

The algorithm described in the previous paragraphs was adjusted for the comparison of observed and simulated EMS data on a linear network, as performed in this paper, by adopting the following two modifications:

1. The integral in equation (9) is computed over the network L with arc-length measure, and the density estimates are derived applying the weighted kernel approach described in Section 3.2;
2. Considering the intrinsic variation of simulated future ambulance interventions, the Monte Carlo test was repeated $\tilde{n}$ times, and for each simulation, the p-value was derived generating $\tilde{m}$ datasets under the null hypothesis of random labelling using the technique described before.

The results are reported in Figure 14. The four boxplots summarise the p-values obtained by comparing observed and predicted EMS data considering the same four time periods as before. For each time period, we generated $\tilde{n} = 64$ possible future scenarios, using the procedure described in Algorithm 1, and for each scenario, we simulated $\tilde{m} = 96$ datasets under the random labelling hypothesis. We can clearly notice that, in all four cases, the relative risk function cannot distinguish between observed and predicted future interventions, and we cannot reject the null hypothesis of random labelling. Therefore, this test highlights that the proposed model can successfully predict the future distribution of emergency data.

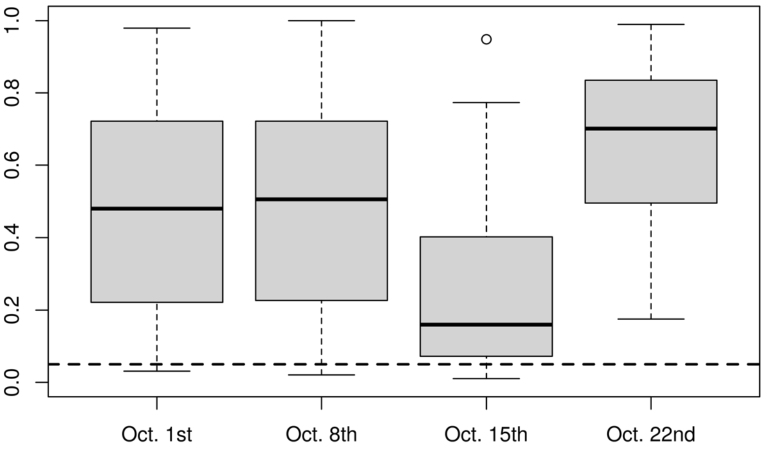


FIG. 14. *Boxplots displaying the p-values of Monte Carlo tests exploring the spatial variation of the relative risk function that compares observed and depicted EMS events. The dashed horizontal black line denotes the $\alpha = 0.05$ level.*

6. Additional comparisons with separable and planar models and further considerations.

As mentioned in Section 1, this paper represents the first attempt to model ambulance interventions on a linear network considering an NHPP with a nonseparable first-order intensity function. Although the proposed approach was found to perform reasonably well, a few aspects deserve further consideration. First, some comparisons to the previously proposed methods should be considered in order to appreciate how the conceptual improvement provided by our methodology translates into a practical improvement in a real-world application. Second, although our model efficiently deals with the considered network (which includes the most important roads of Milan), we should also test whether this methodology effectively scales to larger networks, such as the complete road network of cities like Milan which can be composed by hundreds of thousands of road segments. Third, some evaluations are in order regarding how much of a difference the methodological improvements of the proposed approach make for the ultimate application and how they translate into practical interventions on ambulance dispatch policy. These three points are discussed in the rest of this section.

6.1. *Comparison with separable and planar models.*

Sections 6.1.1 and 6.1.2 highlight the importance of the extensions considered in this paper, namely, modelling the spatiotemporal intensity in a nonseparable manner while explicitly accounting for the network structure of the spatial domain. Hereinafter, the suggested methodology is compared to two different approaches adopted in previous papers dealing with EMS data: (a) assuming separability of the first-order intensity function and (b) ignoring the network structure of the road system and using a planar spatial support instead.

6.1.1. *Separable first-order network-based intensity function.*

The impact of the nonseparability assumption was tested by comparing our proposal to a simpler model that assumes a separable first-order intensity function and estimates the spatial dimension of the process via equation (3), assigning a unitary weight to each past observation. To sample from the simpler model, we implemented a strategy analogous to the one described in the previous section, and we compared the discrepancies between observed and predicted events at different levels of temporal aggregation. More precisely, the two models were trained using the events that occurred before 2017-10-01 at 00:00, and we predicted y_u observations for each hour u of a set $\mathcal{U}$ of future time periods, as described in Algorithm 1. Then the predicted and observed point patterns were aggregated in the considered temporal period, and the predictive performances of the two strategies were evaluated by computing the integrated squared error (ISE), defined by

$$\text{ISE} = \int_L \left(\hat{\eta}_{\text{pred}}(\boldsymbol{s}) - \hat{\eta}_{\text{obs}}(\boldsymbol{s})\right)^2 \mathrm{d}_1(\boldsymbol{s}), \tag{10}$$

where $\hat{\eta}_{\text{pred}}(\boldsymbol{s})$ and $\hat{\eta}_{\text{obs}}(\boldsymbol{s})$ denote the smoothed spatial density at location $\boldsymbol{s} \in L$ obtained by the network convolution kernel (Rakshit et al. (2019)).

The procedure was repeated 150 times, obtaining several estimates of the ISE for the separable and nonseparable strategy. Figure 15 reports the empirical cumulative distribution function (ECDF) of ISE criterion considering three time windows of seven days located farther and farther from the end of the training set. The black curve denotes the nonseparable model, whereas the grey curve represents the separable one. We notice that, in all cases, the nonseparable approach has superior predictive performances exhibiting lower ISE values and pointing out that the nonseparability assumption plays a key role in the prediction of EMS events.

In the Supplementary Material (Gilardi, Borgoni and Mateu (2024)), we have reported the results obtained when applying the same procedure to temporal windows of five and 14 days.

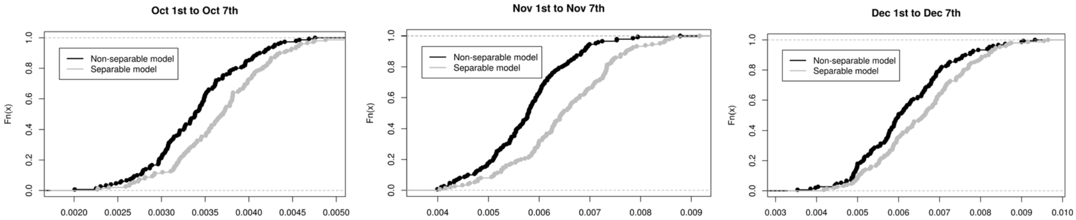


FIG. 15. *Comparison of separable and nonseparable approaches for three different time periods. The three figures represent the ECDF of ISE obtained using* 150 *simulations.*

In all cases the ECDFs show the same behaviour as in Figure 15, highlighting the stability of our findings. Furthermore, we compared each pair of curves using a Kolmogorov–Smirnov test, and the nonseparable model always outperformed the separable counterpart in all tested scenarios but one where the Kolmogorov–Smirnov was not significant.

6.1.2. *Planar intensity function.* Hereinafter, we evaluate the importance of taking the street network into account. To this end, we compared the approach suggested in this paper to another model developed using the same statistical structure but on a planar spatial domain (i.e., the polygon delimiting the city of Milan). Following the procedure detailed in Algorithm 1 and adopting the same time windows considered in Section 6.1.1, we trained the two models and sampled y_u points for each hour u of the future time period $\mathcal{U}$. Finally, after aggregating the points, we compared predicted and observed interventions using the relative integrated squared error (rISE) criterion that reads

$$\text{rISE}_{\text{net}} = \int_L \left(\frac{\hat{\eta}_{\text{pred}}(\boldsymbol{s}) - \hat{\eta}_{\text{obs}}(\boldsymbol{s})}{\hat{\eta}_{\text{obs}}(\boldsymbol{s})} \right)^2 \mathrm{d}_1(\boldsymbol{s}), \tag{11}$$

in case of point patterns defined on network support and

$$\text{rISE}_{\text{planar}} = \int_W \left(\frac{\hat{\eta}_{\text{pred}}(\boldsymbol{s}) - \hat{\eta}_{\text{obs}}(\boldsymbol{s})}{\hat{\eta}_{\text{obs}}(\boldsymbol{s})} \right)^2 \mathrm{d}(\boldsymbol{s}) \tag{12}$$

in case of a planar domain.

The quantities in equation (11) have been introduced before, whereas the terms $\hat{\eta}_{\text{pred}}(\boldsymbol{s})$ and $\hat{\eta}_{\text{obs}}(\boldsymbol{s})$ in equation (12), respectively, denote the planar smoothed spatial density at location $\boldsymbol{s} \in W$ (where W denotes the two-dimensional domain) for predicted and observed interventions obtained via a classical planar kernel with Jones–Diggle's edge correction (Jones (1993)). We adopted the relative ISE criterion to compare the network and planar estimates since the two processes are defined on incompatible spatial domains, implying that the corresponding density functions have different orders of magnitude and making a direct comparison unfeasible. More precisely, the intensity function on a linear network has dimensions $1/m$ while its planar counterpart has dimensions $1/m^2$. Therefore, the ISE criterion, as defined in equation (10) would compare two quantities having dimensions $1/m^2$ and $1/m^4$, respectively. On the other hand, as we can see from equations (11) and (12), the rISE criterion includes an additional term at the denominator of the two equations that creates a unitless ratio of intensities and removes the effects due to the different natures of the corresponding spatial domains.

We repeated the procedure described above 150 times, obtaining several values of the rISE index for each time window. The results are summarised in Table 2. We can clearly notice that the average rISE for the network approach is several times smaller than its planar counterpart, highlighting that the analysis of EMS interventions always requires appropriate considerations regarding the spatial support of the events. The same procedure was repeated for different time intervals of five and 14 days, obtaining similar conclusions. The results are summarised in the Supplementary Material (Gilardi, Borgoni and Mateu (2024)).

TABLE 2

Numerical summary of the comparisons between network and planar approaches using the rISE criterion defined in equations (11) and (12) considering three time-windows of seven days

Time window	Type	Mean	Std. Dev.	0.25 Quantile	0.75 Quantile
Oct. 1st to Oct. 7th	Network	1.0×10^5	4.0×10^4	7.3×10^4	1.2×10^5
	Planar	1.7×10^7	1.2×10^7	1.1×10^7	1.9×10^7
Nov. 1st to Nov. 7th	Network	2.1×10^5	6.3×10^4	1.7×10^5	2.4×10^5
	Planar	3.0×10^7	1.4×10^7	2.1×10^7	3.5×10^7
Dec. 1st to Dec. 7th	Network	1.9×10^5	1.2×10^5	1.2×10^5	2.2×10^5
	Planar	2.4×10^7	1.1×10^7	1.7×10^7	2.7×10^7

6.2. *Scalability.* As already mentioned in Section 2, the analyses reported in this paper are based on a subset of Milan's street network that includes the most important road types, since the majority of ambulance interventions were georeferenced on their proximity. This may rise some concerns about the scalability of the proposed methodology, that is, the ability of our procedure to maintain effectiveness when applied to a larger network if this would be necessary under different circumstances. In the Supplementary Material (Gilardi, Borgoni and Mateu (2024)), we summarise the results obtained when applying the statistical model detailed in Section 3 to a larger spatial network composed by 118,720 segments covering 4636 km. It was created considering all road segments located in Milan that are available from OSM servers. These additional analyses allow us to assess the robustness to different spatial networks of the approach proposed in this paper for EMS data modelling and prove the excellent scalability of the suggested methodology with large spatial domains. More precisely, after downloading the network data from OSM servers, it took approximately 17 minutes to estimate the temporal model, the weight function, and the KDE in equation (3) on the extended network. Although the extended network is more than two times longer than the original one, the computational time required to perform the statistical analysis was definitely reasonable, as compared to the time requested, 10 minutes in total, to analyse the data on the restricted network considered in the previous sections of this paper. In particular, thanks to the fast Fourier transform algorithm adopted in the kernel estimator, fitting the statistical model on the two networks requires roughly the same computational effort. Further details on these comparisons are reported in the Supplementary Material (Gilardi, Borgoni and Mateu (2024)).

6.3. *Use for policy interventions.* Ambulance dispatch planning requires careful consideration of a number of different factors, and the model proposed in this paper can support this activity by providing an estimate of the potential demand for interventions in different parts of the city network.

Hereinafter, we have considered how the analysis described in the previous sections may help local EMS agencies to manage the ambulance rescue points. In fact, in the city of Milan there are 42 locations (such as dedicated squares or parking spots) where the ambulances can park during the day or the night while waiting for a request of intervention. As shown in Figure 16, these rescue points are placed in strategic areas of the city network identified to optimise access time to patients and provide good coverage of the territory of the city.

Using the model described in the previous sections, we are able to anticipate the pressure that each of these stations would suffer in terms of requests for intervention on a given day. We exemplified this point using the data collected between 2015-01-01 and 2017-12-31. We fitted our model to the data and used it to simulate the locations of the ambulance interventions a few days ahead. We then calculated the minimum distance on the road network

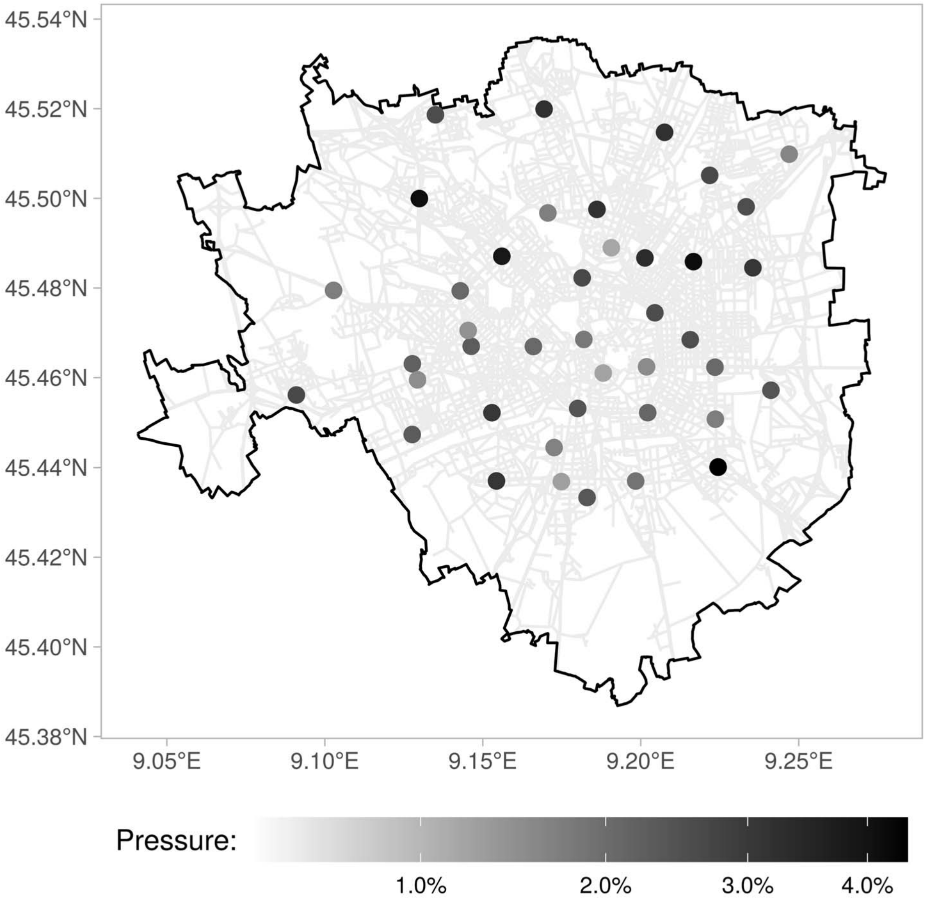


FIG. 16. *Predicted pressure on the* 42 *ambulance stations of Milan for January* 3, 2018.

between each simulated event and the closest ambulance station, assuming that this station would be the one to be activated in order to provide the most efficient reaction. In this way we were in the position to estimate the potential pressure on each station that is expected on the considered day, where the potential pressure is calculated as the percentage of interventions that are closer to this station than to any other station in the city. In order to compensate for the simulation variability, we simulated 150 point patterns, using the fitted model, and averaged the percentage over the performed simulations. The results are reported in Figure 16.

This information may allow the authorities to allocate some extra ambulance crews, if available, to those stations that are expected to be exposed to high pressure or move ambulances stationed on low-stress points there. Given the low computational time (the previous simulations take only a few minutes), this analysis can be conducted on a daily basis to maintain an up-to-date intervention units system or at least to have a benchmark to compare the spatial allocation of ambulance crews.

In addition, the model developed in this paper can be potentially adopted to evaluate how efficient ambulance dispatches have been on a given day by comparing, retrospectively, the actual pressure observed in the considered day to the one estimated using the approach described above. We performed this analysis on October 1, 2017, finding only a mild agreement between the two quantities: only eight of the top 15 stressed stations were in the top 15 predicted pressures. Although several other factors may impact on the observed pressure of a given station in a certain day (e.g., rerouting of ambulances from/to other parts of the city or temporary traffic jams occurring nearby the ambulance point or the location of intervention), discrepancies between potential and actual pressure may suggest a rethinking of the criteria adopted for ambulance dispatches in order to speed up the service and improve its efficiency.

7. Conclusions.

In this paper, we investigated the spatiotemporal distribution of approximately 480,000 ambulance interventions that occurred in the City of Milan from 2015-01-01 to 2017-12-31. Unlike several previous approaches, we assumed that the emergency events represent a realisation of a spatiotemporal point process occurring on a road network, that is, a geolocated graph structure representing road segments and street junctions.

A preliminary exploratory analysis, summarised in Section 2, revealed that the temporal evolution of the events presents several types of seasonalities due to hourly, daily, and weekly patterns. We also observed the presence of space-time interactions in the hourly distribution of the events, which motivated the adoption of a nonseparable statistical model. More precisely, after dividing the interventions into one-hour intervals, we assumed that, for each time period, the ambulance dispatches represented a realisation of an NHPP on a linear network with a nonseparable first-order intensity function. The temporal component was modelled via a semiparametric Poisson regression with deterministic temporal covariates. Considering the results of the exploratory analysis, the annual patterns were included with a linear term, while the hourly and weekly trends were smoothed using cyclic cubic regression splines, whereas the daily effects are included using dummy variables. The spatiotemporal component was modelled by a weighted kernel estimator. The weights were used to capture the space-time interactions of EMS data, trying to grasp the temporal regularities in the emergency interventions and induce nonseparability into the spatiotemporal intensity.

We found that the temporal Poisson model fits the EMS counts well and the deterministic temporal components successfully approximate the hourly, daily, and weekly patterns. The weight function also adequately mirrors the temporal seasonalities displayed by the ACF of EMS counts. The spatial and spatiotemporal dynamics were exemplified considering two future time periods: 2018-01-03 at 03:00 and 2018-01-03 at 15:00. Our results highlight that ambulance interventions are more spread in the municipality during the night, whereas they tend to cluster in the city centre during working hours. In both cases the main train station, a few popular squares, and a retirement house stand out.

The predictive accuracy of our proposal was tested using the relative risk function by comparing observed and predicted ambulance interventions for four different days. In all cases the relative risk is concentrated around 0.5, implying that the model successfully predicts future events. A series of Monte Carlo tests confirmed that conclusion.

Finally, we demonstrated that the approach proposed in this paper improves over the methodologies previously adopted for modelling EMS data, taking into account both the network structure of the spatial domain and the nonseparability of the spatiotemporal intensity function. We also found that this approach scales well to very large networks, hence it proves to be particularly suitable to manage real-world applications.

To conclude, we remark that the main challenges in this paper stem from the spatiotemporal dynamics and the specific spatial support of our data. First, the exploratory analysis suggested an interaction in the spatial and temporal components, requiring a nonseparable structure when modelling the ambulance intervention process. Second, the spatial nature of the data also suggested that linear networks are the most appropriate spatial domain for modelling EMS data. Third, the geographic region represents a large metropolitan area, and the huge number of interventions required the adoption of fast statistical techniques. This latter point may also imply that the spatial and temporal variability can be impacted by secondary variables possibly measured both at the areal level (e.g., population density) and at the network level (e.g., road types, traffic flows, commuting patterns, or other variables representing specific anthropic activities at a given point of the network). However, the main purpose of this paper is the spatial and short-term temporal prediction of ambulance interventions, and according to our experience, regionalised time-varying covariates are difficult to obtain at the desired spatial and temporal levels, and their inclusion is scarcely impactful. Nevertheless, in future works it might be desirable to develop parametric or semiparametric models that allow the introduction of such explanatory spatial covariates in the intensity function. Furthermore, it should be pointed out that some of the aforementioned variables are typically recorded only at the areal level (e.g., census wards or traffic zones), and their inclusion in a network model presents several layers of complexity. In fact, the projection of areal data into a linear network may induce abrupt changes in the covariate (every time a segment intersects different areas) or imprecise measurements, hence requiring further modelling care.

We point out that, considering the complexities detailed before, machine learning (ML) methods (such as classification trees or neural networks) may represent a promising approach to analyse spatial and spatiotemporal point patterns. However, to the best of our knowledge, the literature is extremely scarce in this field, and only a few recent papers exist addressing this aspect. For example, Yang et al. (2019) merge the theory of classical kernel density estimation with variational autoencoders to develop a model for the analysis of spatial inhomogeneous Poisson processes. Mateu and Jalilian (2022) provide a mathematical framework for coupling neural network models with the statistical analysis of planar point patterns focusing on point processes with multiple groups observed for $T \geq 2$ times, whereas Jalilian and Mateu (2023) develop a Siamese neural network discriminant model to evaluate the similarities between spatial point patterns obtaining superior performances than the classical statistical tools (i.e., the K function). However, it should be noted that the deep learning methods introduced in the aforementioned papers typically require that the point pattern is reduced to a two-dimensional grid of cell counts that is treated as an image, that is, a set of pixel values; hence, planar spatial support is assumed for the data. The adaptation of those techniques to the analysis of linear network data requires substantial methodological improvements, which are beyond the scope of this paper.

A further extension is to move towards statistical models that account for data clustering following different routes, for instance, a double stochastic process, such as the Cox process where a stochastic component is included in the intensity function to deal with the unexplained space-time variation. A natural solution, in this case, would be to adopt an inhomogeneous log-Gaussian Cox process (Møller, Syversveen and Waagepetersen (1998)), already proposed for modelling ambulance interventions by Bayisa et al. (2020). However, moving to the latter approach requires a substantial amount of methodological development since defining a proper covariance function for the stochastic component of the intensity function on a linear network spatial support is not straightforward.

Acknowledgments. Map data copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org. We greatly acknowledge Doct. Piero Brambilla, Doct. Andrea Pagliosa and Doct. Rodolfo Bonora, the regional experts of ambulance interventions working for the local EMS (AREU: Azienda Regionale Emergenza Urgenza). They provided us the ambulance intervention data used in this paper and, more importantly, priceless and continuous assistance to properly understand the complexity of an EMS system and the peculiarities of this problem. We also greatly acknowledge the DEMS Data Science Lab for supporting this work by providing computational resources. This study was carried out within the MOST—Sustainable Mobility National Research Center and received funding from the European Union Next-Generation EU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR)—MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4—D.D. 1033 17/06/2022, CN00000023). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them. Andrea Gilardi's present address is Politecnico di Milano, Departimento di Matematica.

Funding. J. Mateu is partially supported by grant PID2019-107392RB-I00 from the Spanish Ministry of Science.

SUPPLEMENTARY MATERIAL

Supplementary material (DOI: 10.1214/23-AOAS1800SUPP; .pdf). The supplementary material summarises the results obtained when testing the spatial predictive accuracy considering different time periods and alternative training sets. Moreover, it includes additional details on the comparison with planar and separable modelling approaches. Finally, we also report more precise details regarding the computing times on the extended road network.

REFERENCES

ANG, Q. W., BADDELEY, A. and NAIR, G. (2012). Geometrically corrected second order analysis of events on a linear network, with applications to ecology and criminology. *Scand. J. Stat.* **39** 591–617. MR3000837 https://doi.org/10.1111/j.1467-9469.2011.00752.x

BADDELEY, A., RUBAK, E. and TURNER, R. (2015). *Spatial Point Patterns*: *Methodology and Applications with R*. CRC Press, Boca Raton, FL.

BADDELEY, A., NAIR, G., RAKSHIT, S., MCSWIGGAN, G. and DAVIES, T. M. (2021). Analysing point patterns on networks—A review. *Spat. Stat.* **42** Paper No. 100435, 35 pp. MR4233256 https://doi.org/10.1016/j.spasta.2020.100435

BARRINGTON-LEIGH, C. and MILLARD-BALL, A. (2017). The world's user-generated road map is more than 80% complete. *PLoS ONE* **12** e0180698. https://doi.org/10.1371/journal.pone.0180698

BARTHÉLEMY, M. (2011). Spatial networks. *Phys. Rep.* **499** 1–101. MR2770962 https://doi.org/10.1016/j.physrep.2010.11.002

BAYISA, F. L., ÅDAHL, M., RYDÉN, P. and CRONIE, O. (2020). Large-scale modelling and forecasting of ambulance calls in northern Sweden using spatio-temporal log-Gaussian Cox processes. *Spat. Stat.* **39** Paper No. 100471, 22 pp. MR4158686 https://doi.org/10.1016/j.spasta.2020.100471

BLACKWELL, T. H. and KAUFMAN, J. S. (2002). Response time effectiveness: Comparison of response time and survival in an urban emergency medical services system. *Acad. Emerg. Med.* **9** 288–295.

BYRD, R. H., LU, P., NOCEDAL, J. and ZHU, C. Y. (1995). A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* **16** 1190–1208. MR1346301 https://doi.org/10.1137/0916069

OPENSTREETMAP CONTRIBUTORS (2017). Planet dump. Available at https://planet.osm.org. https://www.openstreetmap.org.

DIGGLE, P., ROWLINGSON, B. and SU, T. (2005). Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics* **16** 423–434. MR2147534 https://doi.org/10.1002/env.712

DIGGLE, P. J. (2014). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*, 3rd ed. *Monographs on Statistics and Applied Probability* **128**. CRC Press, Boca Raton, FL. MR3113855

DIGGLE, P. J. (1985). A kernel method for smoothing point process data. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **34** 138–147.

KELSALL, J. E. and DIGGLE, P. J. (1998). Spatial variation in risk of disease: A nonparametric binary regression approach. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **47** 559–573.

GILARDI, A., BORGONI, R. and MATEU, J. (2024). Supplement to "A nonseparable first-order spatiotemporal intensity for events on linear networks: An application to ambulance interventions." https://doi.org/10.1214/23-AOAS1800SUPP

GILARDI, A. and LOVELACE, R. (2021). osmextract: Download and read OpenStreetMap data extracts.

GONZÁLEZ, J. A., RODRÍGUEZ-CORTÉS, F. J., CRONIE, O. and MATEU, J. (2016). Spatio-temporal point process statistics: A review. *Spat. Stat.* **18** 505–544. MR3575505 https://doi.org/10.1016/j.spasta.2016.10.002

HENDERSON, S. G. (2011). Operations research tools for addressing current challenges in emergency medical services. In *Wiley Encyclopedia of Operations Research and Management Science* Wiley, Inc, Hoboken, NJ. https://doi.org/10.1002/9780470400531.eorms0605

JALILIAN, A. and MATEU, J. (2023). Assessing similarities between spatial point patterns with a Siamese neural network discriminant model. *Adv. Data Anal. Classif.* **17** 21–42. MR4552824 https://doi.org/10.1007/s11634-021-00485-0

JONES, M. C. (1993). Simple boundary correction for kernel density estimation. *Stat. Comput.* **3** 135–146.

KELSALL, J. E. and DIGGLE, P. J. (1995). Kernel estimation of relative risk. *Bernoulli* **1** 3–16. MR1354453 https://doi.org/10.2307/3318678

LU, Y. and CHEN, X. (2007). On the false alarm of planar K-function when analyzing urban crime distributed along streets. *Soc. Sci. Res.* **36** 611–632.

MATEU, J. and JALILIAN, A. (2022). Spatial point processes and neural networks: A convenient couple. *Spat. Stat.* **50** Paper No. 100644, 19 pp. MR4439344 https://doi.org/10.1016/j.spasta.2022.100644

MATTESON, D. S., MCLEAN, M. W., WOODARD, D. B. and HENDERSON, S. G. (2011). Forecasting emergency medical service call arrival rates. *Ann. Appl. Stat.* **5** 1379–1406. MR2849778 https://doi.org/10.1214/10-AOAS442

MCSWIGGAN, G., BADDELEY, A. and NAIR, G. (2017). Kernel density estimation on a linear network. *Scand. J. Stat.* **44** 324–345. MR3658517 https://doi.org/10.1111/sjos.12255

MCSWIGGAN, G., BADDELEY, A. and NAIR, G. (2020). Estimation of relative risk for events on a linear network. *Stat. Comput.* **30** 469–484. MR4064631 https://doi.org/10.1007/s11222-019-09889-7

MØLLER, J., SYVERSVEEN, A. R. and WAAGEPETERSEN, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Stat.* **25** 451–482. MR1650019 https://doi.org/10.1111/1467-9469.00115

MØLLER, J. and WAAGEPETERSEN, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. *Monographs on Statistics and Applied Probability* **100**. CRC Press/CRC, Boca Raton, FL. MR2004226

OKABE, A. and SUGIHARA, K. (2012). *Spatial Analysis Along Networks*: *Statistical and Computational Methods*. Wiley, New York.

RAKSHIT, S., BADDELEY, A. and NAIR, G. (2019). Efficient code for second order analysis of events on a linear network. *J. Stat. Softw.* **90** 1–37.

RAKSHIT, S., DAVIES, T., MORADI, M. M., MCSWIGGAN, G., NAIR, G., MATEU, J. and BADDELEY, A. (2019). Fast kernel smoothing of point patterns on a large network using two-dimensional convolution. *Int. Stat. Rev.* **87** 531–556. MR4043351 https://doi.org/10.1111/insr.12327

SCOTT, D. W. (1992). *Multivariate Density Estimation*: *Theory*, *Practice*, *and Visualization*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. Wiley, New York. MR1191168 https://doi.org/10.1002/9780470316849

SILVERMAN, B. W. (1982). Algorithm AS 176: Kernel density estimation using the fast Fourier transform. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **31** 93–99.

R CORE TEAM (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

VILE, J. L., GILLARD, J. W., HARPER, P. R. and KNIGHT, V. A. (2012). Predicting ambulance demand using singular spectrum analysis. *J. Oper. Res. Soc.* **63** 1556–1565.

WOOD, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 3–36. MR2797734 https://doi.org/10.1111/j.1467-9868.2010.00749.x

WOOD, S. N. (2017). *Generalized Additive Models*: *An Introduction with R*, 2 ed. Chapman and Hall/CRC, Boca Raton, FL.

YAMADA, I. and THILL, J.-C. (2004). Comparison of planar and network K-functions in traffic accident analysis. *J. Transp. Geogr.* **12** 149–158.

YUAN, B., WANG, X., MA, J., ZHOU, C., BERTOZZI, A. L. and YANG, H. (2019). Variational autoencoders for highly multivariate spatial point processes intensities. In *International Conference on Learning Representations*.

ZHOU, Z. and MATTESON, D. S. (2015). Predicting ambulance demand: A spatio-temporal kernel approach. In *Proceedings of the* 21*th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. *KDD* '15 2297–2303. Association for Computing Machinery, New York, NY, USA.

ZHOU, Z., MATTESON, D. S., WOODARD, D. B., HENDERSON, S. G. and MICHEAS, A. C. (2015). A spatio-temporal point process model for ambulance demand. *J. Amer. Statist. Assoc.* **110** 6–15. MR3338482 https://doi.org/10.1080/01621459.2014.941466