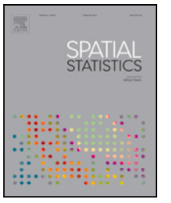
Spatial Statistics 69 (2025) 100912

Contents lists available at ScienceDirect

Spatial Statistics

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



Spatio-temporal intensity estimation for inhomogeneous Poisson point processes on linear networks: A roughness penalty method

Simone Panzeri [a], Aldo Clemente [a], Eleonora Arnone [b], Jorge Mateu [c], Laura M. Sangalli [a,*]

[a] MOX, Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci, 32, Milan, 20133, Italy
[b] Department of Management, Università degli Studi di Torino, Corso Unione Sovietica, 218 bis, Turin, 10134, Italy
[c] Department of Mathematics, Universitat Jaume I, Avinguda de Vicent Sos Baynat, Castellón de la Plana, 12071, Spain

ARTICLE INFO

Keywords:
Nonparametric likelihood estimation
Roughness penalties
Finite element method

ABSTRACT

Nowadays, a vast amount of georeferenced data pertains to human and natural activities occurring in complex network-constrained regions, such as road or river networks. In this article, our research focuses on spatio-temporal point patterns evolving over time on linear networks, which we model as inhomogeneous Poisson point processes. Within this framework, we propose an innovative nonparametric method for intensity estimation that leverages penalized maximum likelihood with roughness penalties based on differential operators applied across space and time. We provide an efficient implementation of the proposed method, relying on advanced computational and numerical techniques that involve finite element discretizations on linear networks. We validate the method through simulation studies conducted across various scenarios, evaluating its performance compared to state-of-the-art competitors. Finally, we illustrate the method through an application to road accident data recorded in the municipality of Bergamo, Italy, during the years 2017–2019.

1. Introduction

In this work, we propose an innovative method for non-separable spatio-temporal intensity estimation in inhomogeneous Poisson point processes in network-constrained regions. Spatio-temporal point pattern data on linear networks are becoming increasingly prevalent. The integration of georeferencing precision technologies, such as the Global Positioning System (GPS), into smartphones, vehicles, and wearable devices has resulted in a massive increase in the availability of georeferenced event data pertaining to human activities on road networks. A typical example concerns traffic accidents. For instance, Fig. 1 shows accident data recorded in the municipality of Bergamo, Italy, during the years 2017–2019. The point pattern under consideration is strongly influenced by the shape of the road network, exhibiting spatial variation in accident risks. Despite the higher concentration of restricted traffic zones and pedestrian areas in the city center, accidents occur more frequently there than in residential areas. As expected, a similar trend is observed along major roadways connecting peripheral areas to the city center or other cities.

Other common event data on road networks include street crimes, such as vehicle theft, vandalism, drug dealing, and robbery. Georeferenced data on linear networks are also widely studied in environmental sciences and ecology, for example, to monitor the presence of rare plant species or endangered fish populations in contaminated river networks. Important examples of this type of data also arise in industrial and life sciences, particularly in applications involving fiber-optic cables or nerve fibers.

* Corresponding author.
E-mail address: laura.sangalli@polimi.it (L.M. Sangalli).
URL: https://sangalli.faculty.polimi.it (L.M. Sangalli).

https://doi.org/10.1016/j.spasta.2025.100912
Received 17 January 2025; Received in revised form 13 May 2025; Accepted 10 June 2025
Available online 15 July 2025

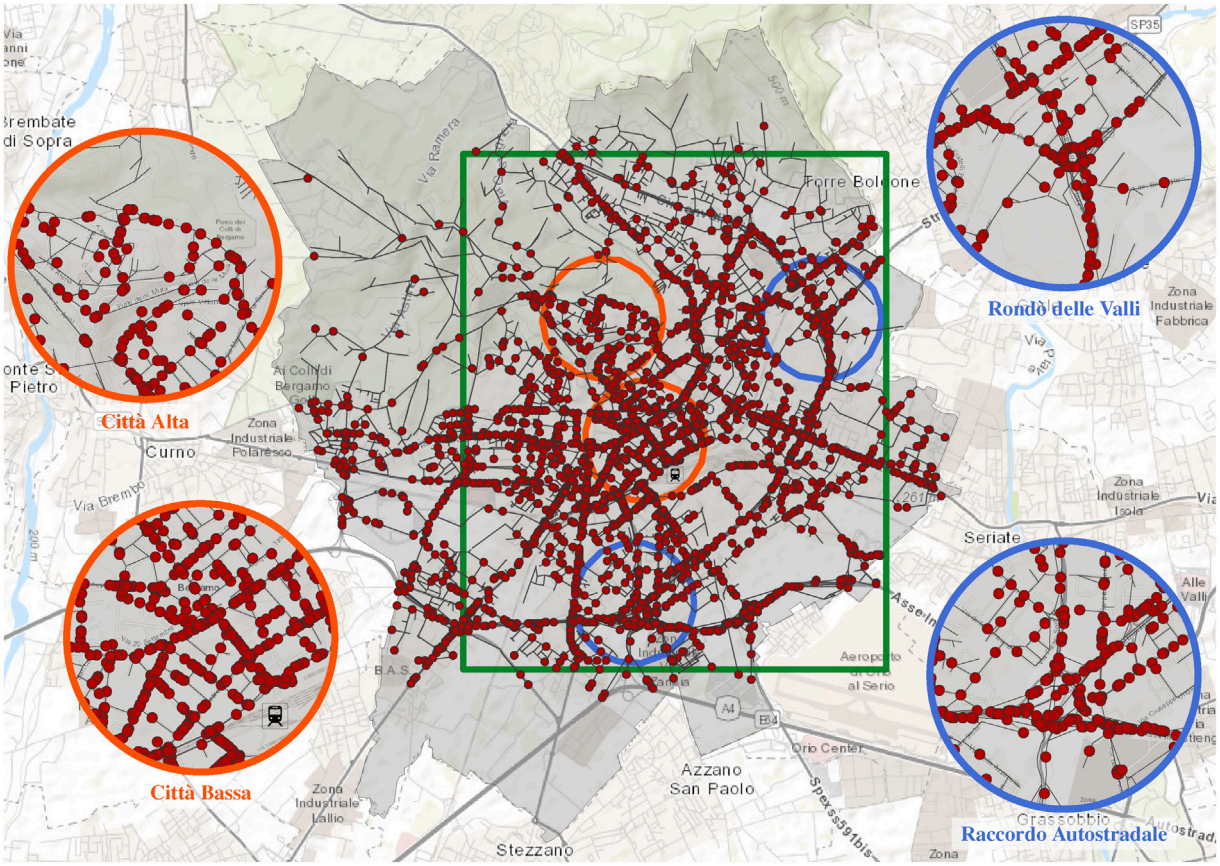


Fig. 1. Locations of 5176 road accidents that occurred in Bergamo, Italy, during the years 2017–2019. The observed point pattern exhibits different behaviors across the road network, showing spatially varying accident risks when comparing roads in urban and suburban districts. The green-bordered box contains the subnetwork relevant to the application study. The other boxes mark four zones of interest, each exhibiting distinct behaviors in terms of accident risk: the orange-bordered boxes contain roads in the upper and lower zones of the city center (*Città Alta* and *Città Bassa*), whereas the blue-bordered boxes contain ring roads and highway junctions (*Rondò delle Valli* and *Raccordo Autostradale*).

Spatial point processes over regular two-dimensional regions have been the subject of an extensive literature; see, e.g., the textbooks by Møller and Waagepetersen (2003), Daley and Vere-Jones (2003, 2008), Illian et al. (2008), Baddeley et al. (2015) and the references therein. However, the complex geometry of linear networks hinders the use of classical statistical methodologies when the point pattern is constrained to a network. One immediate reason for this is the inadequacy of the classical Euclidean metric in $\mathbb{R}^2$, which fails to provide an accurate measure of the distances between points on a network. Additionally, the stationary assumption is often unreasonable for point processes defined on linear networks (see, e.g., Baddeley et al., 2017). These factors have directed research towards the idea that the spatial domain in which events occur can significantly influence the phenomenon under study, offering valuable insights into it (Okabe and Sugihara, 2012). This idea has inspired the flourishing of dedicated literature on linear networks, primarily within the purely spatial framework. In fact, some classical methods have already been adapted to handle points constrained to network-constrained regions by computing alternative distance metrics, such as the shortest-path distance along the network. With the main goal of solving intensity estimation problems, various kernel estimators have been proposed in both the Geographical Information Systems (GIS) research field (Borruso, 2003, 2005, 2008; Xie and Yan, 2008; Okabe et al., 2009; Anderson, 2009; Okabe and Sugihara, 2012) and the spatial statistics community (Baddeley et al., 2014; McSwiggan et al., 2017; Moradi et al., 2018; Rakshit et al., 2019; Briz-Redón et al., 2019; McSwiggan et al., 2020; Moradi et al., 2019), with different proposals for selecting both the smoothing kernel and the bandwidth. Interestingly, most of these kernel-based techniques are illustrated through applications to event data observed on road networks; for a review of these techniques, the reader may refer to Baddeley et al. (2021). More recently, penalized smoothing approaches have also been developed (Schneble and Kauermann, 2022; Clemente et al., 2023a). These methods have the advantage of not requiring the computation of distances on linear networks, which can quickly become computationally prohibitive in practical applications. In such cases, even networks of moderate size often consist of thousands of vertices and edges. For example, the road network of the municipality of Bergamo, shown in Fig. 1, spans a total length of 349.4 km and comprises 3724 vertices and 4559 edges.

In addition to exhibiting complex spatial dependencies on linear networks, point patterns may also evolve over time, as in the case of the application to road accident data mentioned above. Fig. 2 shows the evolution of the point pattern on a subnetwork of

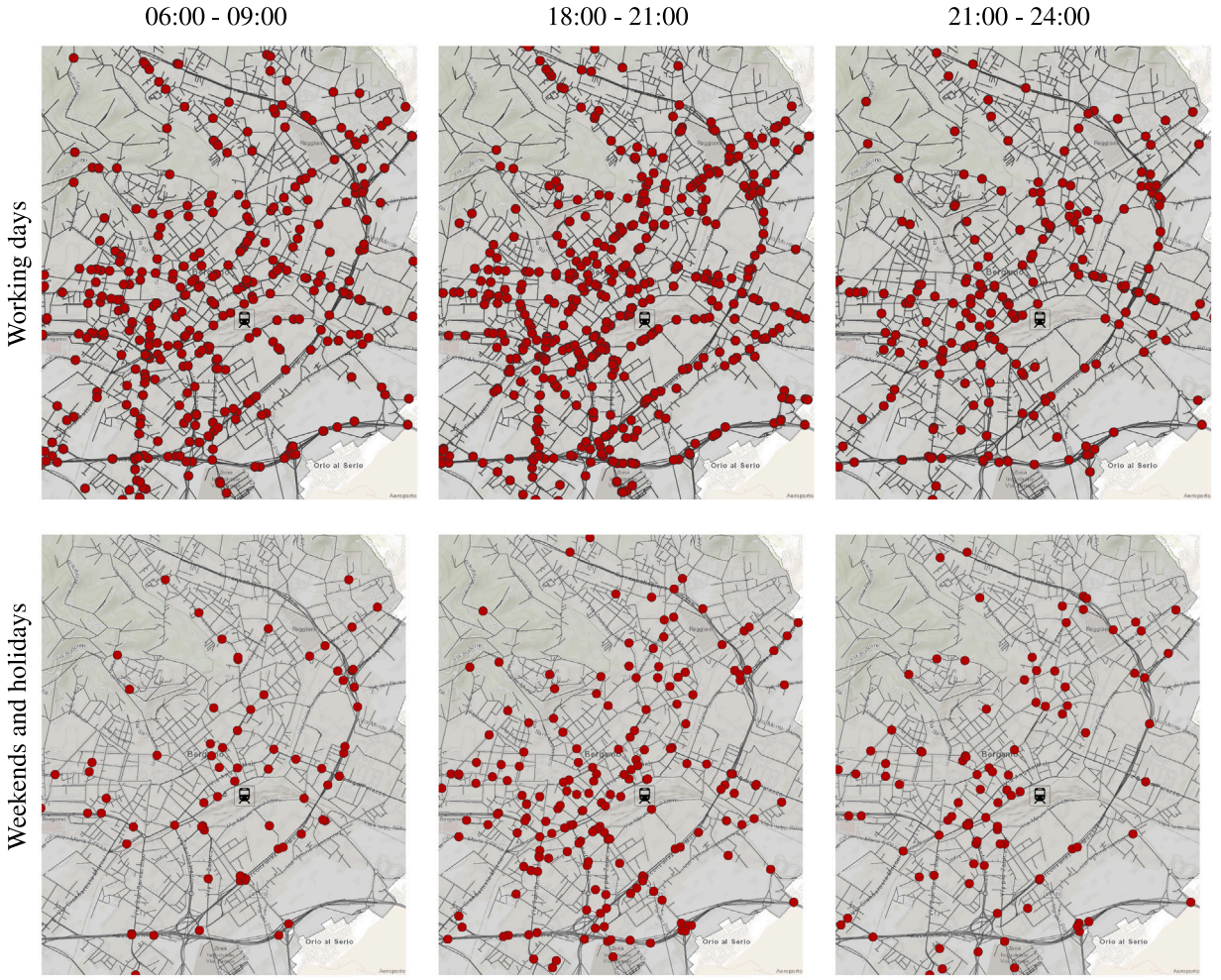


Fig. 2. Locations of 5176 road accidents that occurred in Bergamo, Italy, during the years 2017–2019, at different time slots: 3947 accidents during working days (first row) and 1229 during weekends and holidays (second row). The observed point pattern evolves differently over time across the road network, with distinct accident risks when comparing peak hours and off-peak hours. Although the plots refer to time intervals of 3 h, the data are continuously recorded over time. For clarity of visualization, plots show the subnetwork contained within the green-bordered box in Fig. 1.

interest during workdays (first row), and during weekends and holidays (second row), with all available data aggregated by day type. The phenomenon under consideration depends on time at various levels, including year, month, day, time slot, and potential interactions among these levels. As expected, the risk of road accidents varies with traffic conditions, which change over time due to intrinsic factors (e.g., workdays vs. weekends and holidays, peak vs. off-peak hours, and summer vs. winter seasons) or external influences affecting mobility trends (e.g., the COVID-19 pandemic, road closures, and construction sites).

Incorporating the time dimension into the modeling of spatio-temporal point patterns on linear networks introduces additional challenges compared to the purely spatial case, as it happens in classical studies on regular spatial regions (see, e.g., Diggle, 2013; Cressie and Wikle, 2015; González et al., 2016). These modeling complexities are further compounded by network-specific challenges, especially those related to computational costs. As a result, statistical methodologies for estimating the spatio-temporal intensity of point patterns on linear networks have lagged behind those developed for purely spatial data. To reduce the complexity of the problem being investigated, methods often adopt simplifying assumptions. The most common assumption is the first-order spatio-temporal separability of the considered point process, which requires the underlying intensity function to be expressed almost everywhere as a (non-unique) factorization into non-negative spatial and temporal components. Under the separability assumption, Moradi and Mateu (2020) presents a nonparametric kernel-based intensity estimator that leverages two-dimensional convolutions of the kernel for the spatial component of the intensity, as proposed by Rakshit et al. (2019). This estimator exploits fast Fourier transforms for efficient computation of these convolutions (Silverman, 1982a). D'Angelo et al. (2022) develops parametric estimators for the separable intensity of an inhomogeneous Poisson point process. Mateu et al. (2020) extends the purely spatial resample-smoothing technique introduced in Moradi et al. (2019) to the spatio-temporal framework, proposing a pseudo-separable Voronoi intensity estimator. The relaxation of the full separability assumption allows for a more flexible estimation method that

accounts for potential interactions between spatial and temporal components. Other works combine the network kernel density estimator with GIS-based statistics for spatial and temporal analysis of road accident and crime hotspot mapping on road networks, using a separable approach; see, e.g., Khalid et al. (2018), Vemulapalli et al. (2017). However, we point out that separability rarely holds in practice and should be rigorously assessed using statistical tests, such as those proposed by Díaz-Avalos et al. (2013), Fuentes-Santos et al. (2018), Ghorbani et al. (2021). Existing non-separable techniques for spatio-temporal point patterns typically rely on the Euclidean distance in $\mathbb{R}^2$, which clearly does not appropriately account for the network-constrained nature of the data, as already commented. To the best of our knowledge, there is currently a lack of methods for non-separable spatio-temporal intensity estimation that can successfully handle event data observed over network-constrained regions. In addition to model limitations, most existing techniques for spatio-temporal point patterns on linear networks rely on the computation of shortest-path distances along the network, which can become computationally infeasible in practical applications involving large networks.

In this paper, we overcome the main theoretical and computational limitations of existing methods for spatio-temporal point processes on linear networks. In particular, the proposed approach does not assume separability in space and time, hence offering high modeling flexibility and allowing for the estimation of strongly localized signals. Furthermore, the method does not rely on shortest-path distances and instead leverages numerical discretization and implementation strategies, making it computationally feasible even for large networks. Specifically, we propose a nonparametric penalized likelihood approach for intensity estimation. This approach is inspired by nonparametric density estimation methods developed within the purely spatial framework by Silverman (1982b), Goodd and Gaskins (1971), Tapia and Thompson (1978), Gu and Qiu (1993), Gu et al. (2013) for point patterns observed over one-dimensional or regular two-dimensional domains, and by Ferraccioli et al. (2021), Arnone et al. (2022), Das et al. (2024) for point patterns observed over irregular planar and curved regions. We begin this study by considering the spatio-temporal density estimation method recently presented by Begu et al. (2024), which employs roughness penalties involving differential operators in both space and time. To uplift this approach to the problem under investigation, we rely on a classical result concerning the equivalence between intensity estimation in inhomogeneous Poisson point processes and density estimation based on independent and identically distributed observations (see, e.g., Diggle and Marron, 1988). Moreover, following the preliminary study by Clemente et al. (2023a) in the context of purely spatial data, the proposed method employs a roughness penalty based on a differential operator suitably defined on the considered linear network, utilizing notions of metric graphs and quantum graphs (see, e.g., Berkolaiko and Kuchment, 2013). This roughness penalty is designed to yield smooth estimates that inherently respect the network geometry. To handle the differential operator defined upon the linear network and solve the estimation problem, we resort to a numerical discretization based on finite elements specifically tailored to the network (Arioli and Benzi, 2018). This discretization technique provides the proposed method with high flexibility and efficiency, ensuring feasible computational costs, also when dealing with large networks. The developed method is available in the `fdaPDE` (Arnone et al., 2025) C++/R library, which implements a comprehensive class of spatial data analysis methods with Partial Differential Equation (PDE) regularization.

The remainder of the article is organized as follows. Section 2 introduces the mathematical framework for spatio-temporal inhomogeneous Poisson point processes on linear networks. Section 3 outlines the statistical model for addressing intensity estimation problems and describes the numerical analysis and optimization techniques used in the implementation. Section 4 presents simulation studies conducted in various settings to evaluate the performance of the proposed method in comparison with state-of-the-art techniques. Section 5 demonstrates the use of the proposed method for modeling real-world data, specifically road accidents, as shown in Fig. 2. Finally, Section 6 offers a concluding discussion and delineates potential avenues for future research. Additional material is reported in Appendix A.

2. Spatio-temporal Poisson point processes on linear networks

This section presents the mathematical framework for spatio-temporal inhomogeneous Poisson point processes on linear networks. Section 2.1 introduces the necessary concepts for treating networks as spatial regions, leveraging notions of metric and quantum graphs (see, e.g., Berkolaiko and Kuchment, 2013). Section 2.2 provides the formal definition of the intensity function on the spatio-temporal region under study.

2.1. Linear networks

Let $G := (V, E)$ be a linear network embedded in $\mathbb{R}^2$, which is an undirected graph with a set of vertices $V := \{\mathbf{v}_k\}_{k=1}^K$, where $\mathbf{v}_k \in \mathbb{R}^2$, and a set of edges $E := \{e_s\}_{s=1}^S$, where $e_s \in V \times V$. For an introduction to graph theory, the reader may refer to the classical textbooks by Bondy and Murty (1976), West (2001). For potential practical applications, Marshall et al. (2018) gives an overview of key street network models, introducing detailed terminology and graph representations. The left panel of Fig. 3 shows the benchmark road network of Eastbourne, UK, which has been used in several studies on point patterns in network-constrained regions (see, e.g., Mateu et al., 2020; Moradi and Mateu, 2020), and is employed in this work for simulation studies in Section 4. This simpler geometry, with respect to that of the road network of Bergamo shown in previous figures, enables comparisons with all available competing techniques, some of which are computationally infeasible in the application to road accidents in Bergamo, and facilitates the visualization of the estimates without the need for zoomed-in views of the spatial support.

Any edge $e \in E$, connecting a node $v_1 \in V$ to another node $v_2 \in V$ in the linear network G, can be linearly reparameterized by a map $x_e : [0, \ell_e] \to \mathbb{R}^2$, which takes value v_1 at 0 and v_2 at the edge length ℓ_e. This map allows for the definition of a measure on

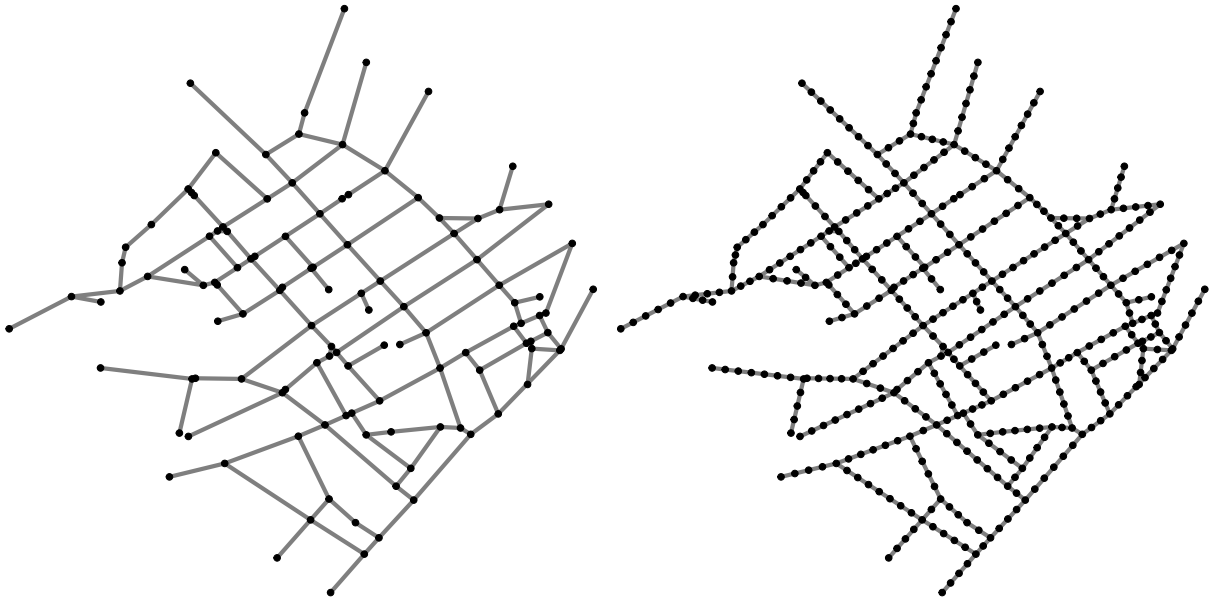


Fig. 3. Left: Road network of Eastbourne, UK, imported from the R package `stlnpp` (Moradi et al., 2022), used in the simulation studies presented in Section 4. Right: discretization of the road network of Eastbourne into a mesh with 461 nodes and 496 edges.

the edge, and, consequently, the integral of a function over the linear network. Specifically, for a function $f : G \to \mathbb{R}$, its integral over a Borel subnetwork $A = (V_A, E_A) \subseteq G$ is defined as

$$\int_A f(\mathbf{p})\,\mathrm{d}\mathbf{p} = \sum_{e \in E_A} \int_e f(x_e(\mathbf{p}))\,\mathrm{d}x_e(\mathbf{p})\,. \tag{1}$$

It is also possible to define differential operators over a linear network. In particular, for any point $\mathbf{p}$ on the edge $e \in E$, the first and second derivatives of the function f are defined as the operators

$$f \to \frac{\partial f(\mathbf{p})}{\partial x_e(\mathbf{p})}\,, \qquad f \to \frac{\partial^2 f(\mathbf{p})}{\partial x_e^2(\mathbf{p})}\,. \tag{2}$$

2.2. Inhomogeneous Poisson point processes

To model spatio-temporal point patterns on linear networks we refer to the theory of Poisson point processes; see, e.g., the textbooks by Daley and Vere-Jones (2008), Diggle (2013), Cressie and Wikle (2015) and the references therein. Let X be a spatio-temporal inhomogeneous Poisson point process on the product space $G \times [0, T]$, with $T > 0$, and $N_X(A \times B)$ be the number of points in the set $(A \times B) \cap X$, where $A \subseteq G$ is a Borel subnetwork and $B \subseteq [0, T]$. We assume that X is finite, meaning that $N_X(G \times [0, T]) < \infty$ with probability one (see, e.g., Daley and Vere-Jones, 2003). Starting from one realization of X, that is $\{(\mathbf{p}_i, t_i)\}_{i=1}^n \subseteq G \times [0, T]$, our goal is to characterize such a point process X. This corresponds to estimating the unknown inhomogeneous intensity function that governs the phenomenon under study. We recall that the intensity function of X is defined as the function $\gamma : G \times [0, T] \to \mathbb{R}^+$ satisfying

$$\mathbb{E}\left[N_X(A \times B)\right] = \int_A \int_B \gamma(\mathbf{p}, t)\,\mathrm{d}t\,\mathrm{d}\mathbf{p} =: \Gamma(A \times B)$$

for any $A \times B \subseteq G \times [0, T]$, where the integration over the subnetwork $A \subseteq G$ is defined as in Eq. (1). We refer to $\Gamma(G \times [0, T])$ as the intensity measure.

3. Proposed method: nonparametric spatio-temporal intensity estimation

This section outlines the proposed approach for intensity estimation in spatio-temporal inhomogeneous Poisson point processes on linear networks. Section 3.1 presents the model, highlighting the key contributions of the article. Section 3.2 details the estimation procedure developed to solve the estimation problem.

3.1. Model

To estimate the unknown intensity, we develop a nonparametric method that does not rely on additional assumptions, such as the first-order spatio-temporal separability of the underlying point process. In particular, we build on the work of Begu et al. (2024), which proposes a penalized likelihood approach for estimating the unknown probability density function from a sample of independent and identically distributed spatio-temporal realizations over simpler two-dimensional domains. The classical result by Diggle and Marron (1988) allows us to equivalently reformulate this density estimation problem in terms of intensity, leveraging the theory of Poisson point processes, as detailed in Appendix A. To naturally comply with the non-negativity constraint on γ, we parameterize the estimation problem in terms of the log-intensity function $u := \log \gamma$ rather than the unknown intensity γ itself. Thus, we propose to find the log-intensity u by minimizing the following penalized (negative) log-likelihood of the underlying inhomogeneous Poisson point process:

$$L(u) := -\sum_{i=1}^{n} u(\mathbf{p}_i, t_i) + \int_G \int_0^T e^{u(\mathbf{p},t)} \, \mathrm{d}t \, \mathrm{d}\mathbf{p} + R(u;\lambda)\,. \tag{3}$$

We denote by $R(u;\lambda)$ the penalization term, which ensures the regularity of the estimate in space and time, thereby preventing degenerate solutions associated with unbounded likelihoods. We define $R(u;\lambda)$ as the sum of two roughness penalties, one in space and one in time, as follows:

$$R(u;\lambda) := \lambda_G \int_G \int_0^T \left(\frac{\partial^2 u(\mathbf{p},t)}{\partial x(\mathbf{p})^2} \right)^2 \mathrm{d}t \, \mathrm{d}\mathbf{p} + \lambda_T \int_G \int_0^T \left(\frac{\partial^2 u(\mathbf{p},t)}{\partial t^2} \right)^2 \mathrm{d}t \, \mathrm{d}\mathbf{p}\,, \tag{4}$$

where the second-order differential operator in space, for functions defined on the linear network G, as well as the integration over G, are defined as in Section 2.1. The first roughness penalty is specifically designed to produce smooth estimates in space, inherently respecting the geometry of the network, while the second penalty controls the smoothness of the estimate in time. In the estimation functional in Eq. (3), the regularization $R(u;\lambda)$ is traded off against the log-likelihood. The vector of positive smoothing parameters $\lambda := \left[\lambda_G, \lambda_T\right]^\top$ regulates this trade-off, balancing the data adaptation and the smoothness of the estimate. These smoothing parameters can be calibrated using k-fold cross-validation, as detailed, e.g., by Begu et al. (2024) in the context of spatio-temporal density estimation over two-dimensional domains.

3.2. Estimation procedure

The solution to the infinite-dimensional intensity estimation problem in Eq. (3) cannot be obtained analytically and must, therefore, be approximated numerically. To achieve this, we introduce a functional basis that enables the representation of functions defined over the network. Following the works of Clemente et al. (2023a,b) on regression methods and density estimation for purely spatial data, we apply a discretization based on the finite element method, which has been adapted for solving partial differential problems over graphs by Arioli and Benzi (2018).

Let $G_h = (V_h, E_h)$ be a spatial discretization of the original linear network $G = (V, E)$, where $V_h := \{\boldsymbol{v}_k\}_{k=1}^{K_h}$ and $E_h := \{e_s\}_{s=1}^{S_h}$ denote the sets of nodes and edges of G_h. V_h is a superset of the set of nodes V of the original network G. For example, the right panel of Fig. 3 displays the regular mesh of the road network of Eastbourne, UK, used in the simulation studies, consisting of 461 nodes and 496 edges. A practical criterion for constructing regular meshes is to set a maximum allowed edge length that results in approximately equally spaced nodes along each edge of the original network. The C++/R library `fdaPDE` (Arnone et al., 2025), which implements the proposed methodology, also offers additional options for mesh refinement. Intuitively, the mesh should be fine enough to capture the spatial features of the signal. However, ultrafine meshes (i.e., meshes with a resolution significantly exceeding that of the signal) are unnecessary, as they yield comparable accuracy, at a substantially higher computational cost; see Section 4.2 for an illustration of this aspect. When dealing with large networks, adaptive meshes, whose resolution varies spatially based on the observed data, can also be considered to optimize computational cost.

Starting from the discretized network G_h, we define a functional basis over G as follows. For each node v_k, with $k = 1, \ldots, K_h$, we consider the linear finite element basis function $\psi_k(\mathbf{p})$ as the continuous piecewise linear function on G that takes value 1 at v_k and value 0 at each other node v_l, i.e., $\psi_k(\boldsymbol{v}_l) = \delta_{kl}$ for $k, l = 1, \ldots, K_h$. As an illustration, Fig. 4 shows two linear finite element basis functions on a simple linear network. The figure displays a linear finite element basis function centered at the node located at a junction of the network (left panel), and a linear finite element basis function centered at an internal node (right panel). With the vector of finite element basis functions $\boldsymbol{\psi} := \left[\psi_1, \quad \ldots, \quad \psi_{K_h}\right]^\top$, it is then possible to represent any continuous piecewise linear function on G_h. It should be noted that higher-order finite elements can also be considered within this procedure. In fact, the implementation of the proposed method in the `fdaPDE` library already supports second-order finite elements (i.e., quadratic basis) and also enables the definition of higher-order basis functions. However, considering higher-order finite elements leads to higher overall computational complexity for the same mesh. In our experience, it is computationally convenient to use linear finite element basis functions and increase the number of mesh nodes, if necessary, rather than switching to higher order basis functions while keeping the mesh fixed.

To discretize over time, we consider a vector of M_h cubic B-spline functions defined over the time interval $[0, T]$. This vector is denoted by $\boldsymbol{\varphi} := \left[\varphi_1, \quad \ldots, \quad \varphi_{M_h}\right]^\top$; see, e.g., De Boor (1978). We then seek the minimizer of the objective functional in Eq. (3)

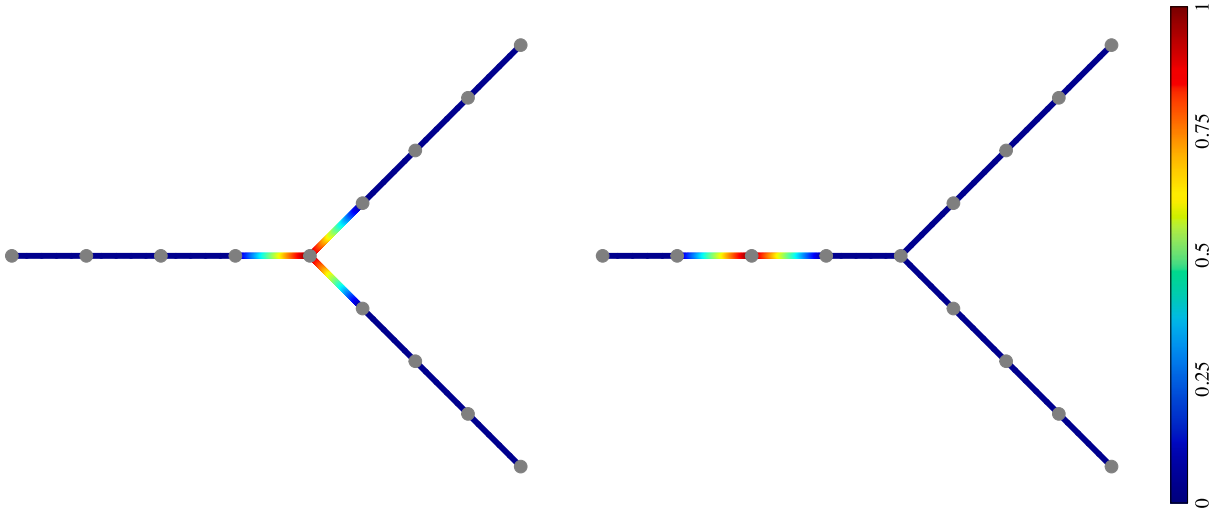


Fig. 4. Two examples of linear finite element basis functions, one at a node located at a junction (left) and the other at an internal node (right) of a simple linear network.

within the space of log-intensity functions that can be expressed as linear combinations of the basis functions in space and time, that is,

$$u(\mathbf{p}, t) = \sum_{k=1}^{K_h} \sum_{m=1}^{M_h} c_{km} \psi_k(\mathbf{p}) \varphi_m(t), \tag{5}$$

where c_{km} are the coefficients in the basis expansion, for $k = 1, \ldots, K_h$ and $m = 1, \ldots, M_h$. Let $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ denote the evaluations of the K_h spatial basis functions at the n spatial data locations $\{\mathbf{p}_i\}_{i=1}^n$ and of the M_h temporal basis functions at the n temporal data instants $\{t_i\}_{i=1}^n$, respectively, i.e., $\boldsymbol{\Psi} := (\boldsymbol{\Psi})_{ik} = \psi_k(\mathbf{p}_i) \in \mathbb{R}^{n \times K_h}$ and $\boldsymbol{\Phi} := (\boldsymbol{\Phi})_{im} = \varphi_m(t_i) \in \mathbb{R}^{n \times M_h}$. Let $\mathbf{c}$ denote the vector of basis expansion coefficients, i.e., $\mathbf{c} := \begin{bmatrix} c_{11}, & \ldots, & c_{1m}, & \ldots, & c_{k1}, & \ldots, & c_{K_h M_h} \end{bmatrix}^\top$. Moreover, we define the spatial mass matrix $R_0 \in \mathbb{R}^{K_h \times K_h}$, the temporal mass matrix $K_0 \in \mathbb{R}^{M_h \times M_h}$, and the spatial stiffness matrix $R_1 \in \mathbb{R}^{K_h \times K_h}$, defined as follows:

$$(R_0)_{ij} := \sum_{e_s \in E_h} \int_{e_s} \psi_j \psi_i \, \mathrm{d}\mathbf{p}, \qquad (K_0)_{ij} := \int_0^T \varphi_j \varphi_i \, \mathrm{d}t, \qquad (R_1)_{ij} := \sum_{e_s \in E_h} \int_{e_s} \frac{\mathrm{d}\psi_j}{\mathrm{d}x_e} \frac{\mathrm{d}\psi_i}{\mathrm{d}x_e} \, \mathrm{d}\mathbf{p}.$$

Additionally, we set $\boldsymbol{\varphi}_{tt} := \begin{bmatrix} \partial^2 \varphi_1 / \partial t^2, & \ldots, & \partial^2 \varphi_{M_h} / \partial t^2 \end{bmatrix}^\top$ as the vector of second-order derivatives of the temporal bases with respect to time. These quantities enable the discretization of the regularization term through the following penalty matrices:

$$P_G := R_1^\top R_0^{-1} R_1 \in \mathbb{R}^{K_h \times K_h}, \qquad P_T := \int_0^T \boldsymbol{\varphi}_{tt} \boldsymbol{\varphi}_{tt}^\top \in \mathbb{R}^{M_h \times M_h} \, \mathrm{d}t.$$

We also introduce the matrix $\Upsilon \in \mathbb{R}^{n \times K_h M_h}$, whose ith row is given by $(\Upsilon)_i := (\boldsymbol{\Phi})_i \otimes (\boldsymbol{\Psi})_i$, for $i = 1, \ldots, n$, where $\otimes$ denotes the Kronecker product between matrices. In order to discretize the second addendum of $L(u)$, we employ standard Gaussian quadrature rules in space and time, with q and r nodes respectively, and associated weights $\mathbf{w}_G \in \mathbb{R}^q$ and $\mathbf{w}_T \in \mathbb{R}^r$. We consistently adopt the same quadrature settings throughout all simulation and application studies; specifically, we set $q = 3$ and $r = 5$. Provided that the mesh is sufficiently fine, the choice of quadrature rule has a negligible impact on the estimates. The evaluations of the basis functions at the quadrature nodes are stored in the matrices $\boldsymbol{\Psi}_e \in \mathbb{R}^{q \times K_h}$ in space and $\boldsymbol{\Phi}_\iota \in \mathbb{R}^{r \times M_h}$ in time.

Thanks to these quantities, we can discretize the objective functional $L(u)$ in Eq. (3) as follows:

$$\tilde{L}(\mathbf{c}) := -\mathbf{1}^\top \Upsilon \mathbf{c} + n \sum_{\iota \subset [0,T]} \sum_{e \in G_h} \kappa (\mathbf{w}_T \otimes \mathbf{w}_G)^\top \exp\{(\boldsymbol{\Phi}_\iota \otimes \boldsymbol{\Psi}_e) \mathbf{c}\} + \lambda_G \mathbf{c}^\top (K_0 \otimes P_G) \mathbf{c} + \lambda_T \mathbf{c}^\top (P_T \otimes R_0) \mathbf{c}. \tag{6}$$

Hence, the problem of determining the log-intensity $u \in V$ as the minimizer of $L(u)$ is now faced by seeking the coefficients $\mathbf{c} \in \mathbb{R}^{K_h M_h}$ in the basis expansion in Eq. (5) that minimize the discretized functional $\tilde{L}(\mathbf{c})$ in Eq. (6). To address this minimization problem, we resort to classical iterative optimization algorithms; see, e.g., Nocedal and Wright (1999), Lange (2013). In particular, the default option in the `fdaPDE` library is the limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) method (Liu and Nocedal, 1989; Nocedal, 1980), with 5 correction vectors and two-loop recursion.

4. Numerical experiments

This section demonstrates the performance of the proposed method, across different scenarios, comparing it with available alternative techniques. For most simulations, we use the road network of Eastbourne as the spatial support, as it serves as a

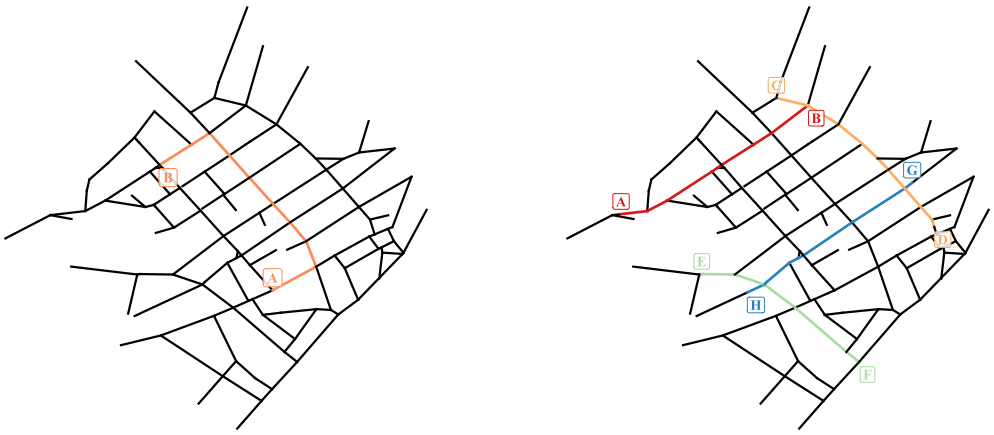


Fig. 5. Left: the intensity used for data generation in Simulation 1 is a Gaussian distribution based on distances along the network, with a large constant variance and a mean that moves over time from point A to point B. Right: the intensity used for data generation in Simulation 2 is a mixture of four Gaussian distributions based on distances along the network, each with equal and relatively small variances, and with means that move over time: from point A to point B, from point C to point D, from point E to point F, and from point G to point H, respectively.

benchmark in the literature. Section 4.1 presents a simulation study with a spatio-temporal Poisson point process with a widely distributed inhomogeneous intensity; Section 4.2 considers an inhomogeneous intensity with more strongly localized features in space and time. Finally, Section 4.3 provides additional simulation studies with separable spatio-temporal intensities. For all simulations, data generation and the computation of the corresponding estimates are repeated 50 times.

The proposed method, named Spatio-Temporal Density and intensity Estimation with PDE regularization (hereafter STDE-PDE), is implemented in the C++ / R library `fdaPDE` (Arnone et al., 2025). For the simulation studies based on the Eastbourne road network, the method uses the mesh displayed in the right panel of Fig. 3, consisting of 461 nodes and 496 edges. The use of a different mesh and its impact on the estimates are discussed in Section 4.2. For the temporal discretization, we set 4 equally spaced internal knots within the considered interval $[0, 1]$. In all simulations, we select the smoothing parameters using 10-fold cross-validation.

Unlike the proposed STDE-PDE, the competing techniques assume the first-order separability of the underlying point process. This means that they seek intensities of the form $\gamma(\mathbf{p}, t) = \gamma_G(\mathbf{p})\gamma_T(t)$, where γ_G and γ_T are two (non-unique) non-negative functions defined on G and $[0, T]$, respectively (González et al., 2016). The considered competing techniques primarily differ in the strategy adopted for estimating the spatial component γ_G. To date, most of them employ kernel functions, based on the shortest-path distance, as defined in Baddeley et al. (2021). Among the Spatio-Temporal Kernel Density Estimation methods, hereafter referred to as STKDE, we consider the following:

`STKDE-EPAN`: the method described by Okabe and Sugihara (2012) in Section 9.2.3, and implemented within the `spatstat.linnet` (Baddeley et al., 2023) R package. `STKDE-EPAN` computes the equal-split continuous rule with Epanechnikov 1D smoothing kernel (Okabe et al., 2009).

`STKDE-HEAT`: the method described by Okabe and Sugihara (2012) in Section 9.2.3, and implemented within the `spatstat.linnet` (Baddeley et al., 2023) R package. The actual computation is performed efficiently by solving the classical time-dependent heat equation on the network, following the approach outlined in McSwiggan et al. (2017); see also (Baddeley and Turner, 2005; Baddeley et al., 2015). `STKDE-HEAT` differs from `STKDE-EPAN` in the choice of the kernel function, which in this case is a Gaussian smoothing kernel.

`STKDE-QUICK`: the method presented in Moradi and Mateu (2020), and implemented within the `stlnpp` (Moradi et al., 2022) R package. `STKDE-QUICK` estimates γ_G through the fast kernel smoothing technique of Rakshit et al. (2019), based on the convolution of the point locations and the network itself with an Euclidean two-dimensional kernel. As for γ_T, `STKDE-QUICK` resorts to the Gaussian kernel density estimator. The network and time smoothing bandwidths are automatically selected using Scott's (Scott, 2015) and Silverman's (Silverman, 2018) rules of thumb, respectively.

In addition to the kernel-based methods above, we also compared STDE-PDE estimates with other approaches. These include the first-order pseudo-separable intensity estimation methods presented in Mateu et al. (2020), and implemented in the `stlnpp` (Moradi et al., 2022) R package. These methods are based on the Voronoi–Dirichlet tessellation (Barr and Schoenberg, 2010), resulting in weighted sums of intensity estimates calculated according to Moradi et al. (2019). However, these techniques proved to be less accurate than the competing methods listed above and have therefore been omitted for the sake of space.

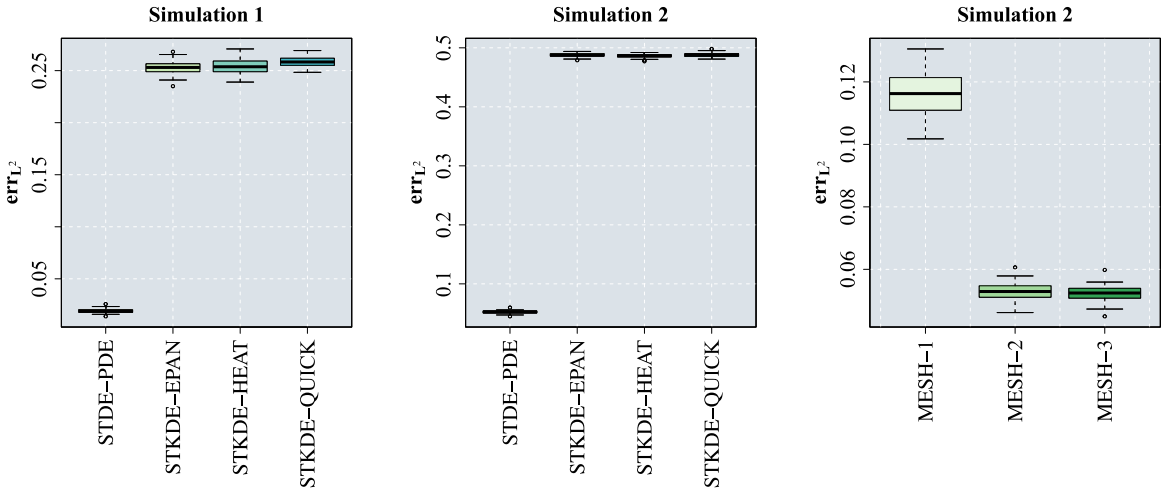


Fig. 6. Boxplots, based on 50 simulation replicates, of the L^2-norm percentage error of the estimates provided by the proposed STDE−PDE and the competing methods introduced in Section 4, namely STKDE−EPAN, STKDE−HEAT, and STKDE−QUICK, in Simulation 1 (left panel) and Simulation 2 (middle panel). The right panel shows the boxplots of the L^2-norm percentage error of the estimates provided by the proposed STDE−PDE, using three meshes with progressively finer levels of refinement, as detailed in Fig. 9.

To assess the accuracy of the competing methods, we compute the L^2-norm percentage error over the spatio-temporal domain, namely

$$\mathrm{err}_{L^2}(\hat{\gamma},\gamma) := \frac{\|\hat{\gamma}-\gamma\|^2_{L^2(0,T;L^2(G))}}{\|\gamma\|^2_{L^2(0,T;L^2(G))}} = \frac{\int_G \int_0^T \left(\hat{\gamma}(\mathbf{p},t)-\gamma(\mathbf{p},t)\right)^2 \mathrm{d}t\,\mathrm{d}\mathbf{p}}{\int_G \int_0^T \left(\gamma(\mathbf{p},t)\right)^2 \mathrm{d}t\,\mathrm{d}\mathbf{p}},$$

where integrals over $G \times [0,T]$ are numerically approximated.

4.1. Simulation 1: A widely distributed inhomogeneous intensity

In Simulation 1, we define the target intensity as a Gaussian distribution with a relatively large variance and a mean that moves along the network over time. We emphasize that this signal is based on distances computed along the network, rather than Euclidean distances in $\mathbb{R}^2$. With reference to the left panel of Fig. 5, the mean follows over time the trajectory from A to B. The standard deviation of the Gaussian distribution is set to 0.3, resulting in a widely distributed signal. We also introduce an appropriate scaling factor to ensure that the desired expected number of points, namely 5 000, is obtained. To sample data points, we use the `rpoistlpp` function (Moradi et al., 2022), with the true intensity provided as an input parameter.

The first row of Fig. 7 displays the data sampled in the first simulation replicate, over short time windows centered around the five considered time instants. The point pattern exhibits a smooth variation over space and time, consistent with the true intensity shown in the second row of the same figure. The remaining rows of the figure report the mean intensity estimates computed by the proposed STDE−PDE and the competing methods listed above, based on the 50 simulation replicates, at the five considered time instants. Since the underlying point process is not separable, STDE−PDE turns out to outperform alternative approaches. In particular, kernel-based methods are less suited for describing this signal, as they capture average spatial trends rather than the instantaneous features of the signal. These qualitative insights find quantitative validation in the left panel of Fig. 6, in terms of the L^2-norm percentage errors, that highlights the strong comparative advantage of the proposed STDE−PDE method. Concerning computational times, in this setting, STDE−PDE and STKDE−EPAN take an average of approximately 80 and 60 s per process, respectively, while the others techniques take about half as long on a standard laptop (Intel Core i7-1065G7, 3.9 GHz, quad core, 16 GB RAM) under the same working conditions.

4.2. Simulation 2: A highly localized inhomogeneous intensity

In Simulation 2, we design the true intensity to have strongly localized features in space and time, in order to mimic the behavior of road accidents. This intensity is defined as a mixture of four Gaussian distributions, each with equal and relatively small variances, and with means that move along the network over time. The right panel of Fig. 5 displays the four trajectories followed by the means of the Gaussian components of the mixture over time: from A to B, C to D, E to F, and G to H. We set the standard deviation of each Gaussian component to 0.075, which yields the desired clustered behavior. Similarly to Simulation 1, we introduce a scaling factor to ensure that, on average, 5 000 data points are sampled in each of the 50 simulation replicates.

The first row of Fig. 8 displays the data sampled in the first simulation replicate, over short time windows centered around the five considered time instants. The point pattern appears clustered in space and time, in accordance with the true intensity shown in

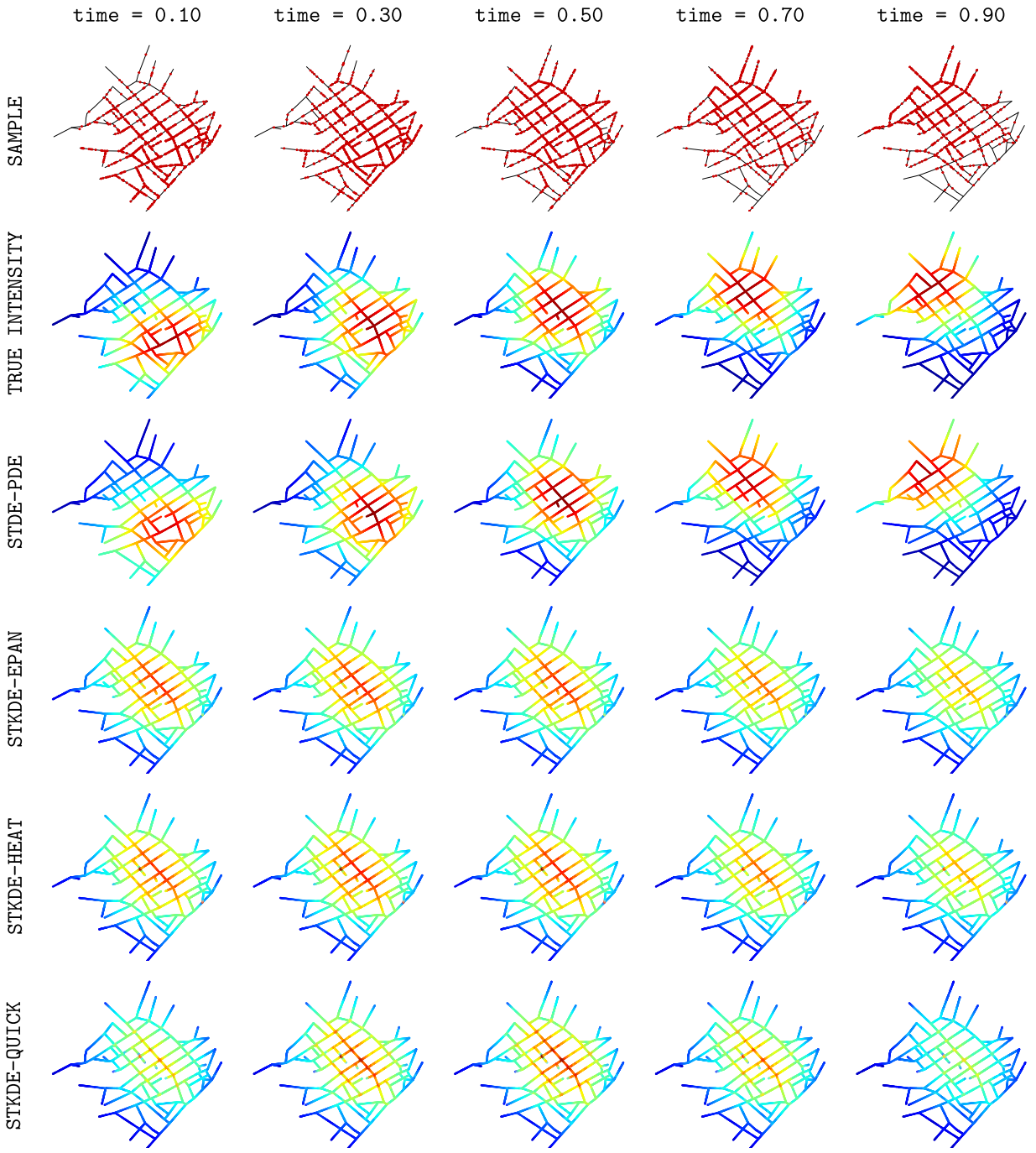


Fig. 7. Simulation 1. Views of the sample from the first simulation replicate (first row), the true intensity (second row), and the mean estimates over 50 simulation replicates obtained using STDE-PDE (third row), STKDE-EPAN (fourth row), STKDE-HEAT (fifth row), and STKDE-QUICK (sixth row), all evaluated at different time instants in the interval $[0, 1]$. All plots share the same color scale, with bluish colors corresponding to low intensity values. The competing methods are detailed in Section 4.

the second row of the same figure. The remaining rows of the figure report the mean intensity estimates computed by the proposed STDE-PDE and the competing methods listed above, based on the 50 simulation replicates, at the five considered time instants. The advantage of the proposed method over the others is even more pronounced in this non-separable setting compared to the previous simulation. Unlike kernel-based techniques, which fail to capture highly localized features in space and time, STDE-PDE demonstrates its ability to accurately estimate this inhomogeneous intensity. This is further confirmed by the L^2-norm percentage errors reported in the middle panel of Fig. 6. Computational times are alike those of Simulation 1.

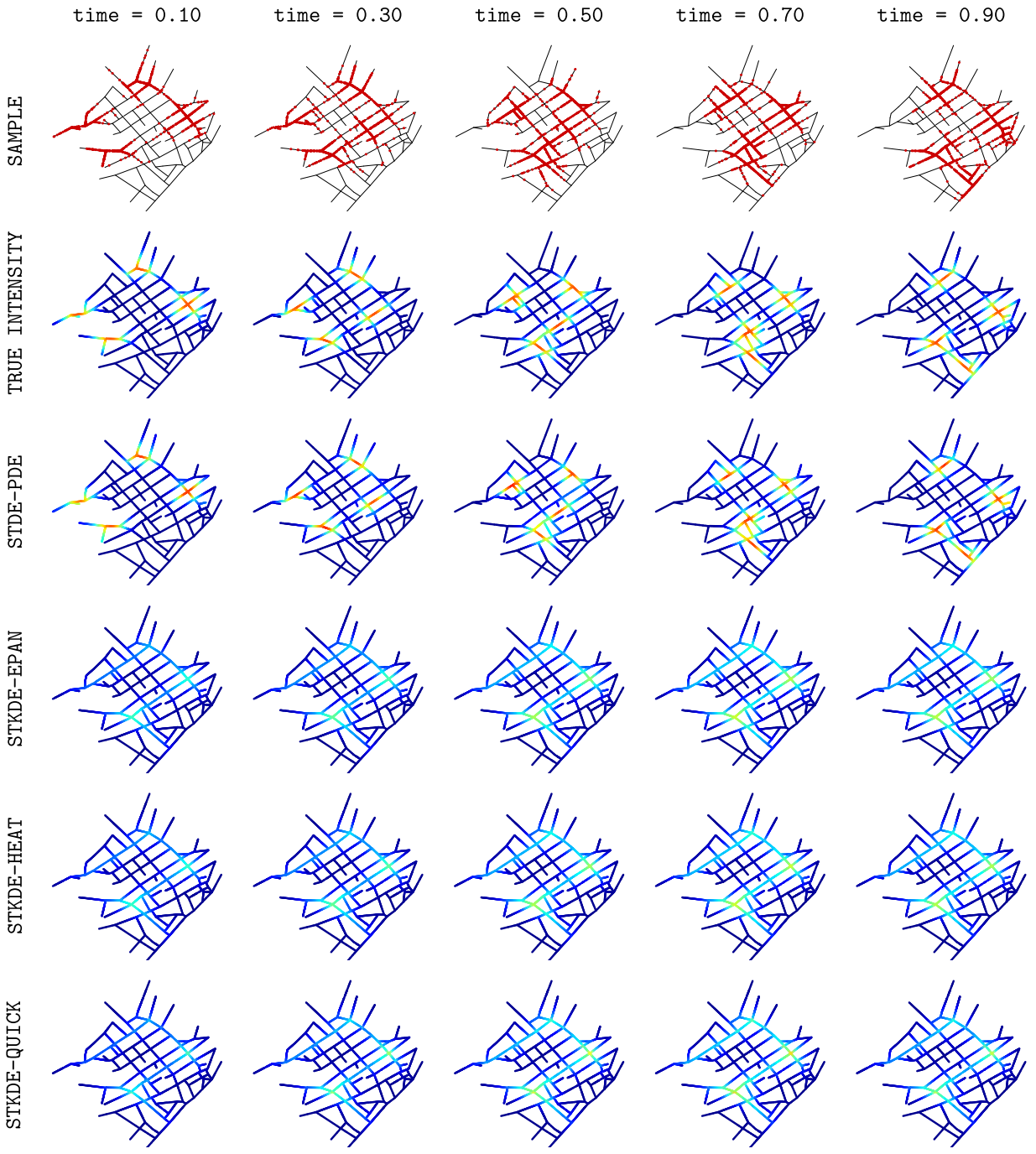


Fig. 8. Simulation 2. Views of the sample from the first simulation replicate (first row), the true intensity (second row), and the mean estimates over 50 simulation replicates obtained using STDE-PDE (third row), STKDE-EPAN (fourth row), STKDE-HEAT (fifth row), and STKDE-QUICK (sixth row), all evaluated at different time instants in the interval $[0, 1]$. All plots share the same color scale, with bluish colors corresponding to low intensity values. The competing methods are detailed in Section 4.

To illustrate the aspects concerning the definition of an appropriate discretization of the network, as discussed in Section 3.2, we here compare the estimates provided by the proposed method STDE-PDE, using three meshes with increasing levels of refinement, as shown in the top panels of Fig. 9: the first mesh, MESH-1 (117 nodes), coincides with the original network and therefore includes only the nodes necessary to define the geometry of the network, as also represented in the left panel of Fig. 3; the second mesh, MESH-2 (461 nodes), is the one used in the simulation studies and also shown in the right panel of Fig. 3; the third mesh, MESH-3 (922 nodes), contains twice as many nodes as the second. The top left panel of Fig. 9 shows the data sampled in the first replicate of

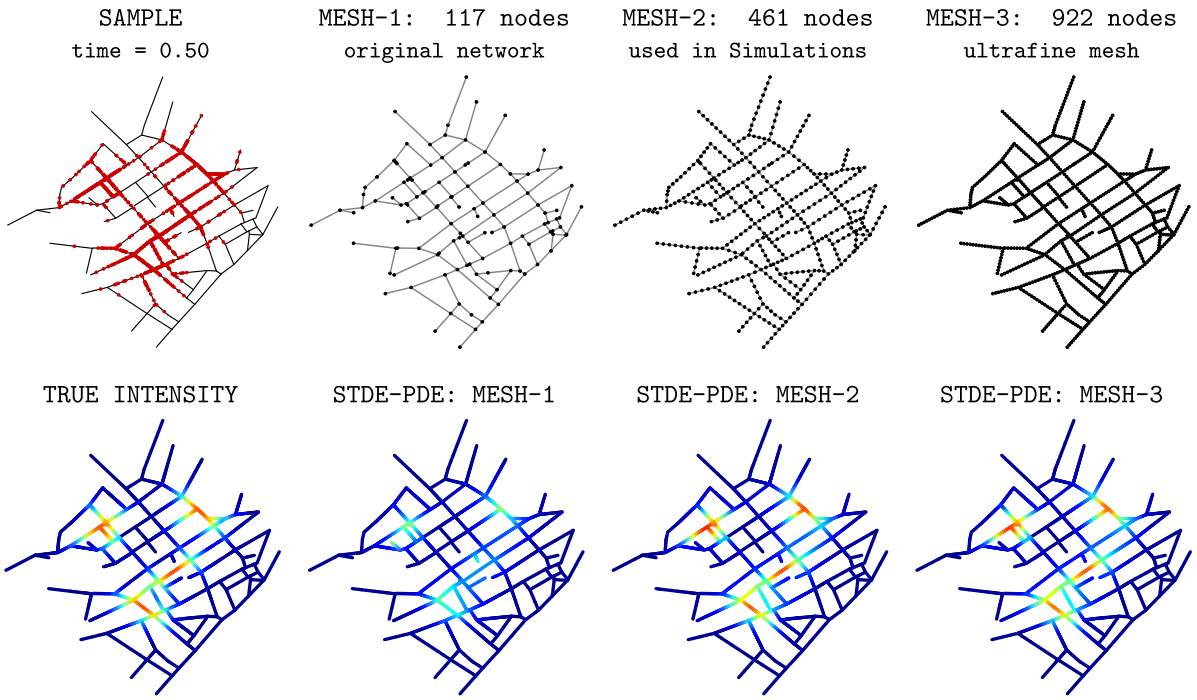


Fig. 9. Top panels: view of the sample from the first simulation replicate at time 0.50, and three meshes with progressively finer levels of refinement: MESH-1 coincides with the original network, including only the nodes necessary to define the geometry of the network; MESH-2 is the mesh used in the various simulation studies on the Eastbourne road network, as also shown in the right panel of Fig. 3; MESH-3 contains twice as many nodes as MESH-2. Bottom panels: view of the true intensity and STDE-PDE mean estimates over 50 simulation replicates, obtained using the three meshes displayed in the corresponding top panels, all evaluated at time 0.50.

Simulation 2, over a short time window centered around 0.50. The second row of that figure represents the true intensity at time 0.50, along with the STDE-PDE mean estimates over 50 simulation replicates at time 0.50, obtained using the three meshes displayed in the corresponding top panels. Qualitatively, the estimate obtained with MESH-1, which coincides with the original network, does not appear to allow the method to capture the localized features of the signal, resulting in an overly smooth intensity. The estimates obtained with MESH-2 and MESH-3 appear indistinguishable and provide a very accurate reconstruction of the true intensity. The right panel of Fig. 6 represents the L^2-norm percentage errors of STDE-PDE using the three different meshes, highlighting a strong reduction in error when moving from MESH-1 to MESH-2, while MESH-3 does not appear to offer a significant competitive advantage over MESH-2. As already discussed in Section 3.2, the results highlight that the mesh should be sufficiently fine to capture the localized features of the signal, whereas ultrafine meshes (i.e., meshes with a resolution exceeding that of the signal) are unnecessary, as they provide comparable accuracy at a substantially higher computational cost. In this regard, we report that the three simulation settings require, on average, approximately 4, 80, and 580 s per process, respectively.

4.3. Additional simulation studies

The proposed STDE-PDE method is specifically designed to estimate non-separable intensities. In non-separable frameworks, Simulations 1 and 2 demonstrate its superiority over the alternative techniques listed in Section 4. We now assess the performance of STDE-PDE when the underlying point processes satisfy instead the assumption of first-order spatio-temporal separability. To this end, we conduct two additional simulation studies on the road network of Eastbourne, applying all the considered methods with the same implementation settings used in Simulations 1 and 2.

In Simulation 3, we define the true intensity as the product of a smooth Gaussian spatial component with a fixed mean and a standard deviation of 0.3, and a temporal component with a negative exponential decay, $\exp(-t)$, on the time interval $[0, 1]$. This separable intensity function generates, on average, 5 000 data points. The left panel of Fig. 10 summarizes the L^2-norm percentage errors, based on 50 simulation replicates, while Fig. B.1 in Appendix B reports the sample generated in the first simulation replicate and the mean intensity estimates provided by the different methods. The results highlight the superiority of the proposed STDE-PDE method also in the separable setting considered here, despite the fact that STDE-PDE does not assume separability, unlike competing techniques. Computational times for the STDE-PDE, STKDE-EPAN, STKDE-HEAT, and STKDE-QUICK methods are approximately 40, 55, 30, and 35 s per process, respectively, on average.

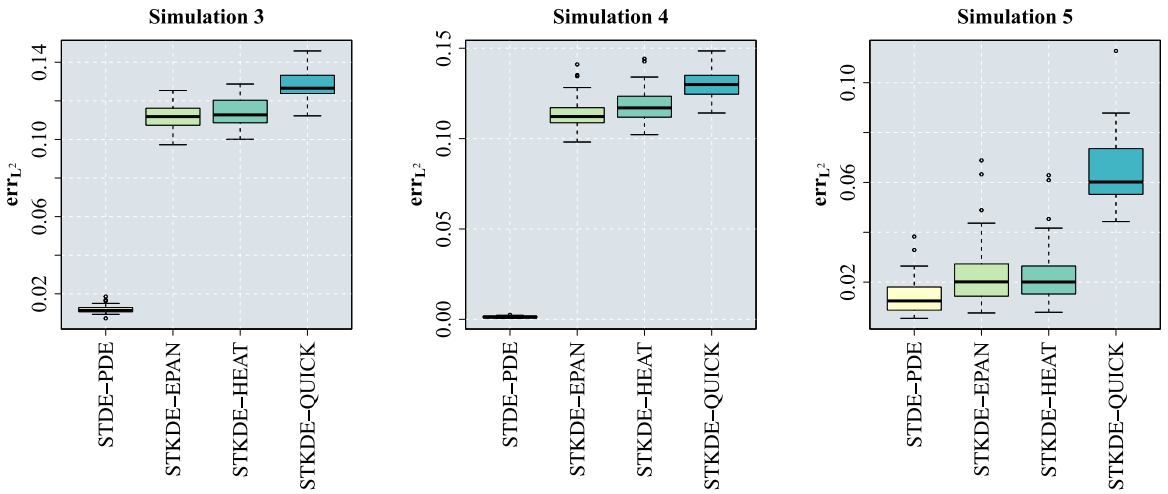


Fig. 10. Boxplots, based on 50 simulation replicates, of the L^2-norm percentage error of the estimates provided by the proposed STDE−PDE and the competing methods introduced in Section 4, namely STKDE−EPAN, STKDE−HEAT, and STKDE−QUICK, in Simulation 3 (left panel), Simulation 4 (middle panel), and Simulation 5 (right panel).

In Simulation 4, we further confirm the insight just stated for separable Poisson point processes. We consider a spatio-temporal homogeneous Poisson point process, which is inherently first-order separable, as the target intensity function is defined as a constant positive scalar. Specifically, the signal corresponds to the ratio between the expected number of sampled data points, which we set to 5 000, and the measure of the spatio-temporal domain of interest, which, in the adopted setting, is given by the total length of the road network of Eastbourne. The middle panel of Fig. 10 summarizes the L^2-norm percentage errors, based on 50 simulation replicates, while Fig. B.2 in Appendix B reports the sample generated in the first simulation replicate and the mean intensity estimates provided by the different methods. The proposed method can also accurately estimate this homogeneous intensity, whilst all competing techniques exhibit some artifacts. From a computational point of view, in this setting, the flat signal, on the linear network and over time, causes the iterative algorithm of the proposed method to reach the maximum number of iterations, thereby increasing the average computational time to approximately 3 min per process. In contrast, as expected, the computational times for the other methods remain consistent with those of the previous simulation.

Finally, in Simulation 5, we assess the performance of STDE−PDE on a network with a different geometry. In particular, we here replicate the test setting presented in Section 3.1.1 of Mateu et al. (2020). The target intensity function is given by $\gamma(\mathbf{p},t) = 1.5\exp\{\sin(t(x+y))\}$, with $(\mathbf{p},t) = ((x,y),t) \in G \times [0,1]$, where G is here the `easynet` linear network, available from the R package `stlnpp` (Moradi et al., 2022). Unlike in previous simulations, this signal is not based on a distance along the linear network; instead, it is defined on $\mathbb{R}^2$, as it depends on the x and y coordinates in $\mathbb{R}^2$. For STDE−PDE, we consider a spatial mesh with 318 nodes and 325 edges, along with the same temporal discretization previously employed. For the competing techniques, we consider the implementations outlined in Section 4. The right panel of Fig. 10 represents the L^2-norm percentage errors, based on 50 simulation replicates, while Fig. B.3 in Appendix B displays the sample in the first simulation replicate and the mean intensity estimates provided by the different methods. STDE−PDE performs better than the competitors even in this simulation setting. The comparatively small advantage of the proposed method over its competitors, compared to the results from the other simulation studies, is primarily due to the limited number of sampled data points here. The average sample size, determined by the definition of the target intensity, is approximately 500. In this setting, the high regularity of the signal on the spatial support increases the computational times for STDE−PDE to approximately 1.5 min per process on average. However, the increase in computational times for STKDE−EPAN and STKDE−HEAT is even more pronounced, as they take around 14 min and 3 min per process, on average, respectively. As stated by the developers, the computational times for these methods scale exponentially with the spatial bandwidth (Baddeley et al., 2023), which is automatically set to a relatively high value in this simulation setting, compared to previous simulations, due to the smoothness of the signal and the fact that it is defined on the plane, rather than on the network, implying longer distances between areas of the network with similar behavior of the point pattern.

5. Application to road accidents in Bergamo, Italy

This section provides an illustrative application of the proposed method to the road accident data introduced in Section 1. According to the *Global status report on road safety 2023*, published by the World Health Organization at https://www.who.int/publications/, approximately 1.19 million people died worldwide in 2021 as a result of road traffic crashes, corresponding to a rate of 15 road traffic deaths per 100 000 people. Moreover, road traffic injury is the leading cause of death for children and young

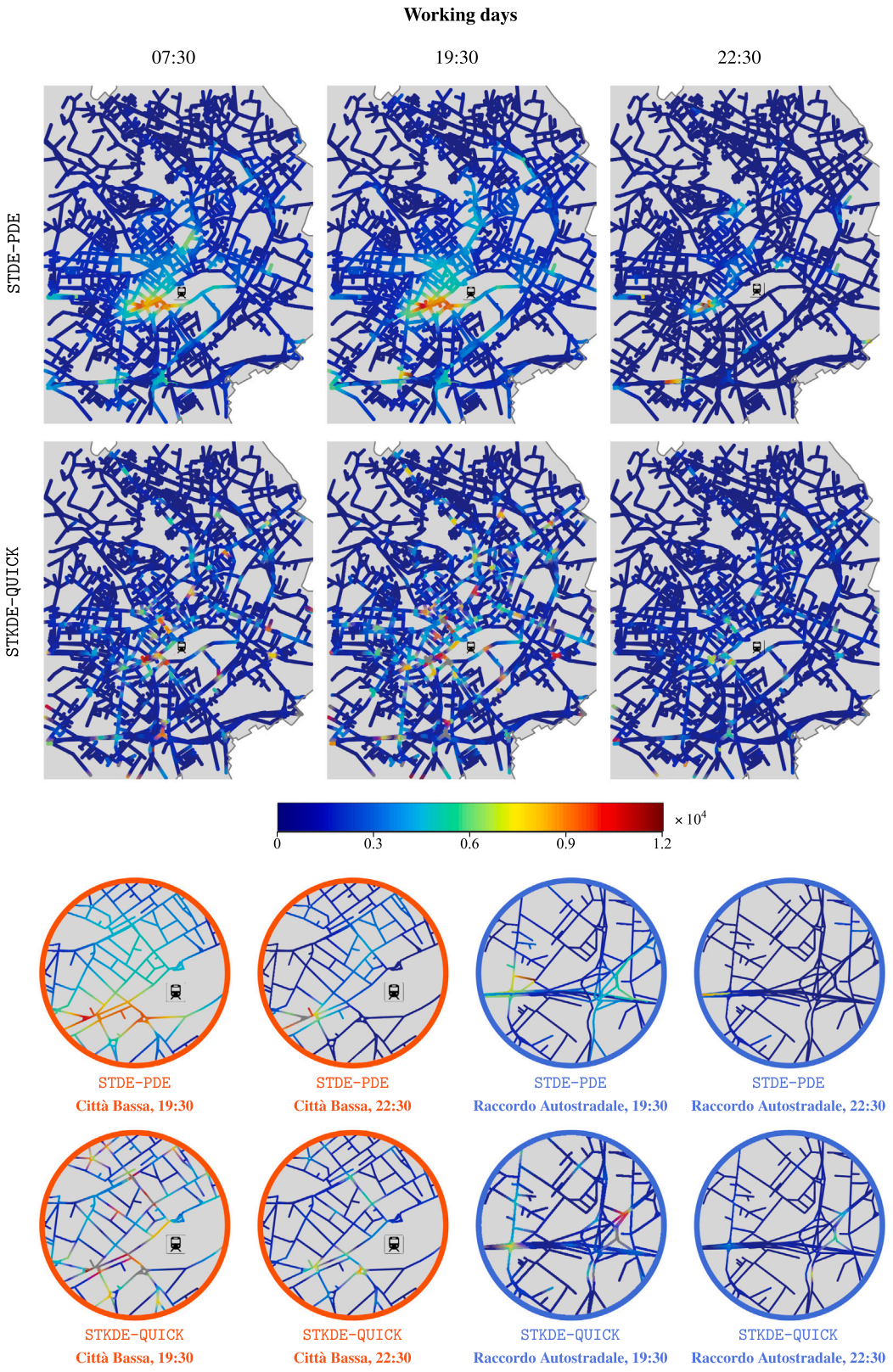


Fig. 11. Intensity estimates of road accidents in the municipality of Bergamo during working days. Top panels contrast the estimates provided by STDE-PDE and STKDE-QUICK, evaluated at different time instants. The bottom panels display zoomed-in views of these estimates in the areas of *Città Bassa* and *Raccordo Autostradale*.

adults aged 5–29 years. Developing advanced methodologies for hotspot detection on critical roads is essential for implementing prevention policies and modifying traffic codes, with the aim of enhancing safety for citizens during periods of heightened risk. For this reason, spatio-temporal modeling of road accidents has become a rapidly growing field of applied research. Various models have been developed within geostatistical frameworks. Classical studies have primarily focused on road crash incidence at the areal level. However, the geometry of road networks has prompted further research into event aggregation at the street level (Deublein et al., 2013; Wang et al., 2013; Fan et al., 2018; Kalair et al., 2021; Ramírez and Valencia, 2021; Gilardi et al., 2022; Chaudhuri et al., 2023; Göttlich and Schillinger, 2023; Gilardi et al., 2024a). Returning to point-pattern modeling, in addition to the various techniques already discussed and compared in this paper, other recent works propose deep learning approaches for predictive modeling of traffic accidents or congestion (see, e.g., Okawa et al., 2019; Zhu et al., 2021; Liu et al., 2022).

The dataset considered in this work is publicly available at https://www.dati.lombardia.it/, and it comprises 5 176 road accidents that occurred in the municipality of Bergamo, Italy, during the years 2017–2019. The original road network is not entirely linear, as some of its edges are curved. For this reason, we use the R packages `osmextract` (Gilardi et al., 2024b) and `sfnetworks` (van der Meer et al., 2024), which offer tools to extract OpenStreetMap data from various providers and convert it into accurate linear network approximations of real-world road networks. Fig. 1 shows the locations of all events recorded on the resulting network, within the administrative boundary of the municipality, while Fig. 2 displays the events on a subnetwork of interest, during different hour slots, aggregating and categorizing them according to the day type, either working day or weekend/holiday. We consider the 24 h of a day as the time domain. Concerning the implementation of the proposed method, for spatial discretization, we refine the network provided by `osmextract` into a mesh with 3 724 nodes and 4 559 edges. For temporal discretization, we use four equally spaced internal knots over the interval from 00:00 to 24:00.

Figs. 11 and 12 present the estimates of road accident risk on the subnetwork contained within the green-bordered box in Fig. 1, for working days and weekends or holidays, respectively. The estimates are shown at the same three time instants around which the three temporal windows, used to display the point pattern in Fig. 2, are centered. For clarity of visualization, the bottom panels of these figures show zoomed-in views of the results for a confined urban area in the lower part of the city center, the *Città Bassa*, and for a suburban area with a highway junction, the *Raccordo Autostradale*. We compare the `STDE-PDE` intensity estimates, in the first rows of Figs. 11 and 12, with those computed by the `STKDE-QUICK` method, in the second rows of these two figures. We do not include the results obtained from the other competing methods listed in Section 4, as their computational costs are prohibitive when dealing with the large network under consideration, and code execution for these competing techniques was halted after one day of running.

`STKDE-QUICK` estimates turns out to be heavily influenced by the number of observed accidents, which is naturally uneven when comparing the phenomenon on working days and on weekends and holidays, leading to overfitting and underfitting behaviors in the bandwidth selection, respectively. The non-separable behavior of the signal under study compromises the accuracy of this technique. In contrast, the proposed `STDE-PDE` method effectively captures strongly localized peaks of heightened criticality on certain roads and during specific hours. The estimates exhibit different spatial patterns across the network and over time. As expected, roads in the lower part of the city center (*Città Bassa*) are characterized by a high risk due to heavy urban traffic, particularly during rush hours in the morning and evening. This is further emphasized in the zoomed-in views in the bottom panel of Fig. 11. The presence of points of interest such as workplaces, schools, and services, makes this area highly frequented by cars during peak hours. The phenomenon is heightened by the railway station, which is the major cause of significant flows of commuters either arriving in the city or traveling to nearby cities, such as Milan. Conversely, the upper part of the city center (*Città Alta*) is primarily residential and touristic, with many restricted traffic zones and pedestrian-only areas that reduce the number of accidents. A moderate risk is also observed near ring road and highway junctions during rush hours (*Raccordo Autostradale* and *Rondò delle Valli*). On weekends and holidays, the trends are different. Accident risk becomes more localized to roads in the city center, with relatively high intensity estimates observed at night, likely due to weekend nightlife, as also emphasized in the zoomed-in views of *Raccordo Autostradale* in the bottom panel of Fig. 12.

6. Discussion and future research directions

As extensively validated through simulation studies and an application to road accident data, the proposed `STDE-PDE` method represents a valid and effective alternative to state-of-the-art techniques for intensity estimation in spatio-temporal Poisson processes in network-constrained regions. Indeed, `STDE-PDE` demonstrates robust performance across the various settings considered, being able to accurately estimate non-separable intensities, even in challenging scenarios characterized by sharply localized features.

The method can be further extended in several directions. An interesting direction involves incorporating additional information about the observed point pattern. For example, in the application case study, the data could be enriched with characteristics of road accidents, such as the number of vehicles involved, the number of injured individuals, or the extent of damage. For this purpose, the presented nonparametric approach for Poisson point processes could be expanded to model marked Poisson point processes, where each point is assigned additional attributes, within a spatio-temporal framework. Moreover, instead of computing two separate estimates for the aggregated data according to day type, a single intensity estimate could be computed by using all available data and encoding the day type with an appropriate dummy variable. As a side note, it is worth mentioning that time instants can be treated as positive real-valued marks of a purely spatial point pattern (see, e.g., Daley and Vere-Jones, 2008; Vere-Jones, 2009). In this sense, the proposed `STDE-PDE` could also be employed to model marked Poisson point processes within a purely spatial framework over linear networks (see, e.g., Eckardt and Moradi, 2024).

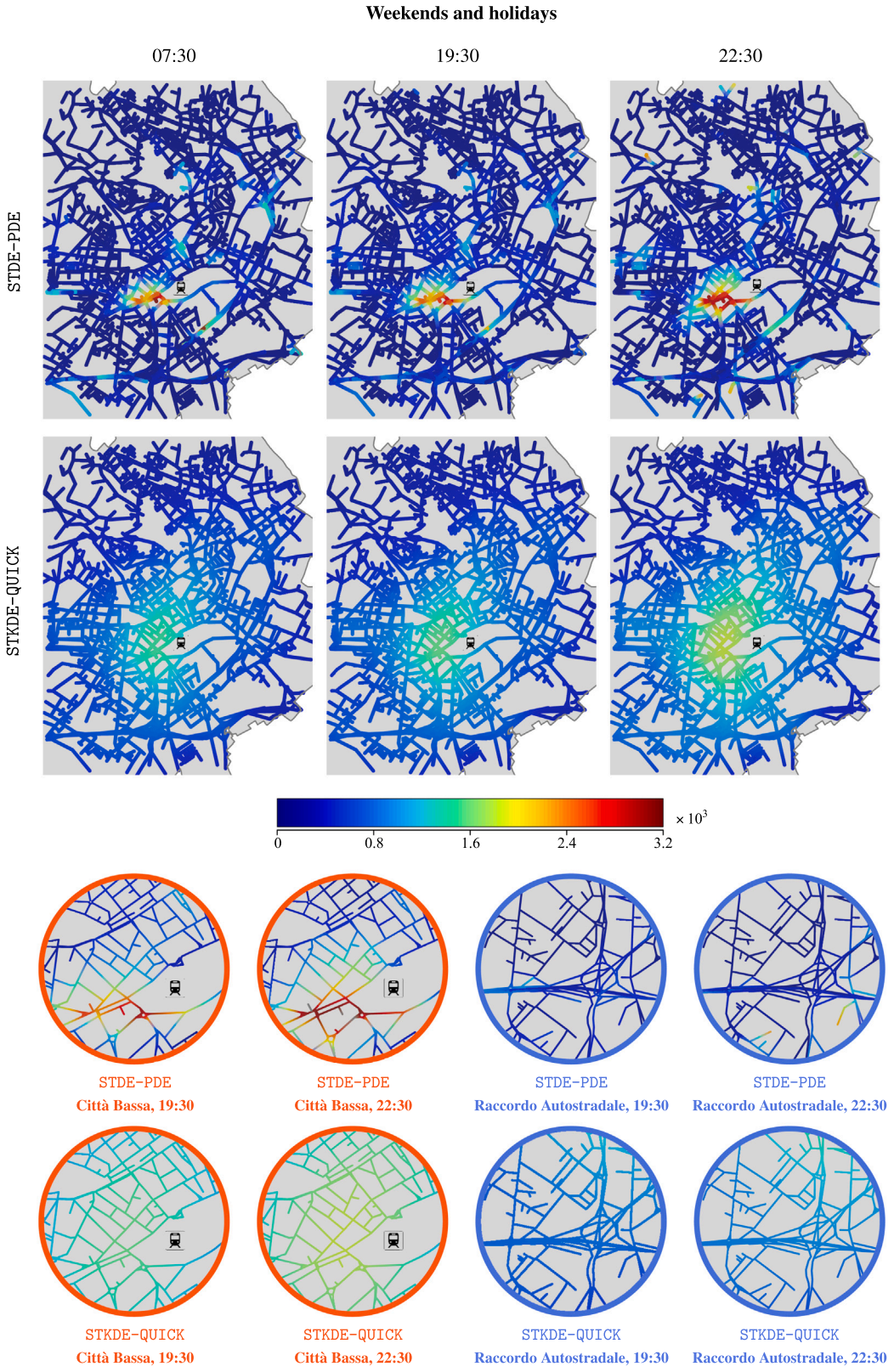


Fig. 12. Intensity estimates of road accidents in the municipality of Bergamo during weekends and holidays. The top panels contrast the estimates provided by STDE-PDE and STKDE-QUICK, evaluated at different time instants. The bottom panels display zoomed-in views of these estimates in the areas of *Città Bassa* and *Raccordo Autostradale*.

Alternatively, we could include covariates through semiparametric statistical models for point processes, extending the presented approach to accommodate different types of likelihood. For example, extensions to Gibbs point processes (Chiu et al., 2013; Illian et al., 2008) could account for both attraction and repulsion behaviors, yielding estimates of pairwise interactions based on covariates at point locations. To this end, a viable strategy would be to employ penalized composite-likelihood approaches, as discussed in Illian and Hendrichsen (2010), Guan et al. (2015), Daniel et al. (2018). A valid alternative would involve semiparametric spatio-temporal Hawkes point processes, combining parametric and nonparametric estimations of the background and triggering effects, as proposed by Alaimo Di Loro et al. (2024), D'Angelo et al. (2024), for example.

Finally, in this work, we have focused on undirected linear networks. The analysis of data observed over directed networks has been more extensively investigated in geostatistical settings, rather than in the point process literature, with contributions by Ganio et al. (2005), Skøien et al. (2006), Ver Hoef et al. (2006), Ver Hoef and Peterson (2010), Laaha et al. (2013), Som et al. (2014), Zhong et al. (2016), Ver Hoef et al. (2019), Barbi et al. (2023). However, the directionality of the network may constitute an important feature to take into account even in point pattern analysis, as it could strongly affect the phenomenon under investigation. Examples in this regard concern the spatial analysis of traffic safety in road networks (Briz-Redón et al., 2019) and environmental measurements in rivers and streams (Cressie et al., 2006). Spatio-temporal point patterns evolving on directed linear networks over time could be accommodated within the proposed methodology by using physics-informed penalties, such as those discussed in Arnone et al. (2019) in regression settings, rather than explicitly accounting for edge orientations during the discretization stage. This would call for the inclusion of transport terms defined on the graph in the partial differential equation that regularizes the negative log-likelihood provided in Section 3.1. Including such transport terms would enable the proposed method to capture the complex behaviors of spatio-temporal Poisson processes on directed networks, further broadening its utility to encompass even more challenging scenarios.

Acknowledgments

L.M. Sangalli acknowledges the research project CoEnv - Complex Environmental Data and Modeling (PRIN 2022 – CUP 2022E3RY23), funded by the European Union - NextGenerationEU programme, and by Ministero dell'Università e della Ricerca. L.M. Sangalli also acknowledges the project GRINS - Growing Resilient, INclusive and Sustainable (GRINS PE00000018 – CUP D43C22003110001), funded by the European Union - NextGenerationEU programme. The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them. A. Clemente acknowledges funding by the European Union - NextGenerationEU programme. A. Clemente, S. Panzeri and L.M. Sangalli acknowledge the research project Dipartimento di Eccellenza 2023–2027, Dipartimento di Matematica, Politecnico di Milano, Italy, funded by the Ministero dell'Università e della Ricerca. J. Mateu acknowledges research projects CIAICO/2022/191 from Generalitat Valenciana, Spain, and PID2022-141555OB-I00 from the Spanish Ministry of Science and Innovation.

Appendix A. Equivalence to density estimation

The intensity estimation problem in spatio-temporal inhomogeneous Poisson point processes, as introduced in Section 3.1, can be equivalently reformulated as a nonparametric density estimation problem. A point pattern can be modeled as a set of n mutually independent spatio-temporal pairs $(\mathbf{p}_i, t_i)_{i=1}^n$ drawn from a common distribution, with the probability density function given by:

$$f := \frac{\gamma(\mathbf{p}, t)}{\Gamma(G \times [0, T])},$$

where γ is the inhomogeneous intensity defined in Section 3.1. Thus, the log-density function $g := \log f$ can be derived by minimizing the following objective functional:

$$L_{\text{dens}}(g) := -\frac{1}{n}\sum_{i=1}^{n} g(\mathbf{p}_i, t_i) + \int_G \int_0^T e^{g(\mathbf{p},t)}\, \mathrm{d}t\, \mathrm{d}\mathbf{p} + \frac{1}{n} R(g; \lambda), \tag{A.1}$$

where the last term can also be expressed as $R(g; \lambda_{\text{dens}})$, with the positive smoothing parameter $\lambda_{\text{dens}} := \lambda \,/\, n$.

Minimizing $L_{\text{dens}}(g)$ in Eq. (A.1) is equivalent to minimizing $L(u)$ in Eq. (3), up to a constant factor of n. Specifically, $\hat{g} = \hat{u} - \log n$, or equivalently, $\hat{f} = \hat{\gamma} \,/\, n$, where n is a realization of N, and therefore an estimate of $\mathbb{E}[N]$. This result can be formally derived by exploiting the arguments in Appendix B of Begu et al. (2024), using the notion of integration over networks and the differential operators over networks provided in Section 2.1 of this work.

Appendix B. Additional simulation studies

In this section, we include Figs. B.1, B.2, and B.3, which show the results from additional simulation studies on first-order separable intensities (Simulations 3 and 4) and on a network with a different geometry (Simulation 5), as described in Section 4.3. Each figure displays the data sampled in the first simulation replicate (first row), the corresponding true intensities (second row), and the mean intensity estimates computed across 50 simulation repetitions by the proposed STDE−PDE method (third row) and its competitors (remaining rows), all evaluated at the five considered time instants.

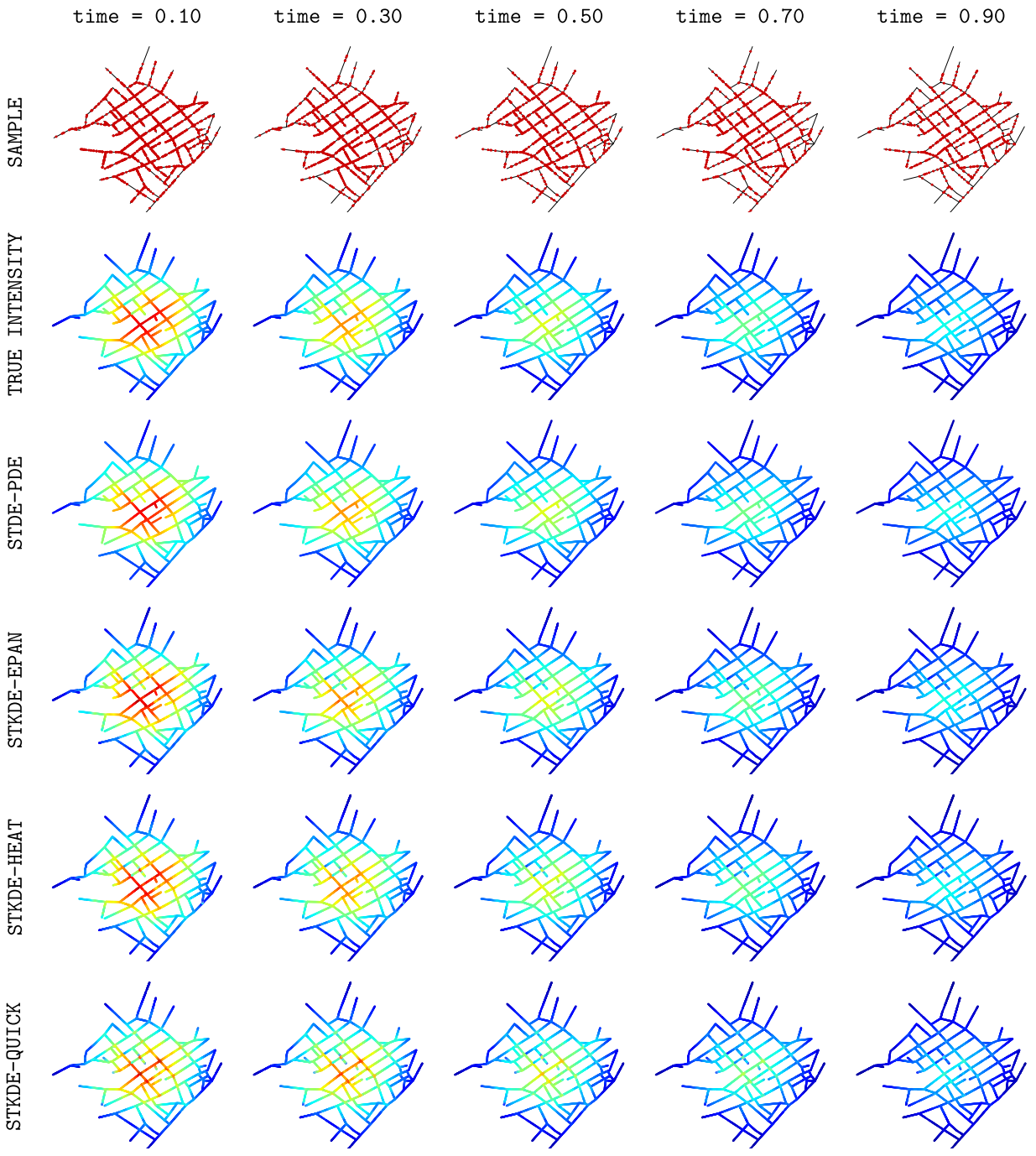


Fig. B.1. Simulation 3. Views of the sample from the first simulation replicate (first row), the true intensity (second row), and the mean estimates over 50 simulation replicates obtained using `STDE-PDE` (third row), `STKDE-EPAN` (fourth row), `STKDE-HEAT` (fifth row), and `STKDE-QUICK` (sixth row), all evaluated at different time instants in the interval $[0, 1]$. All plots share the same color scale, with bluish colors corresponding to low intensity values. The competing methods are detailed in Section 4.

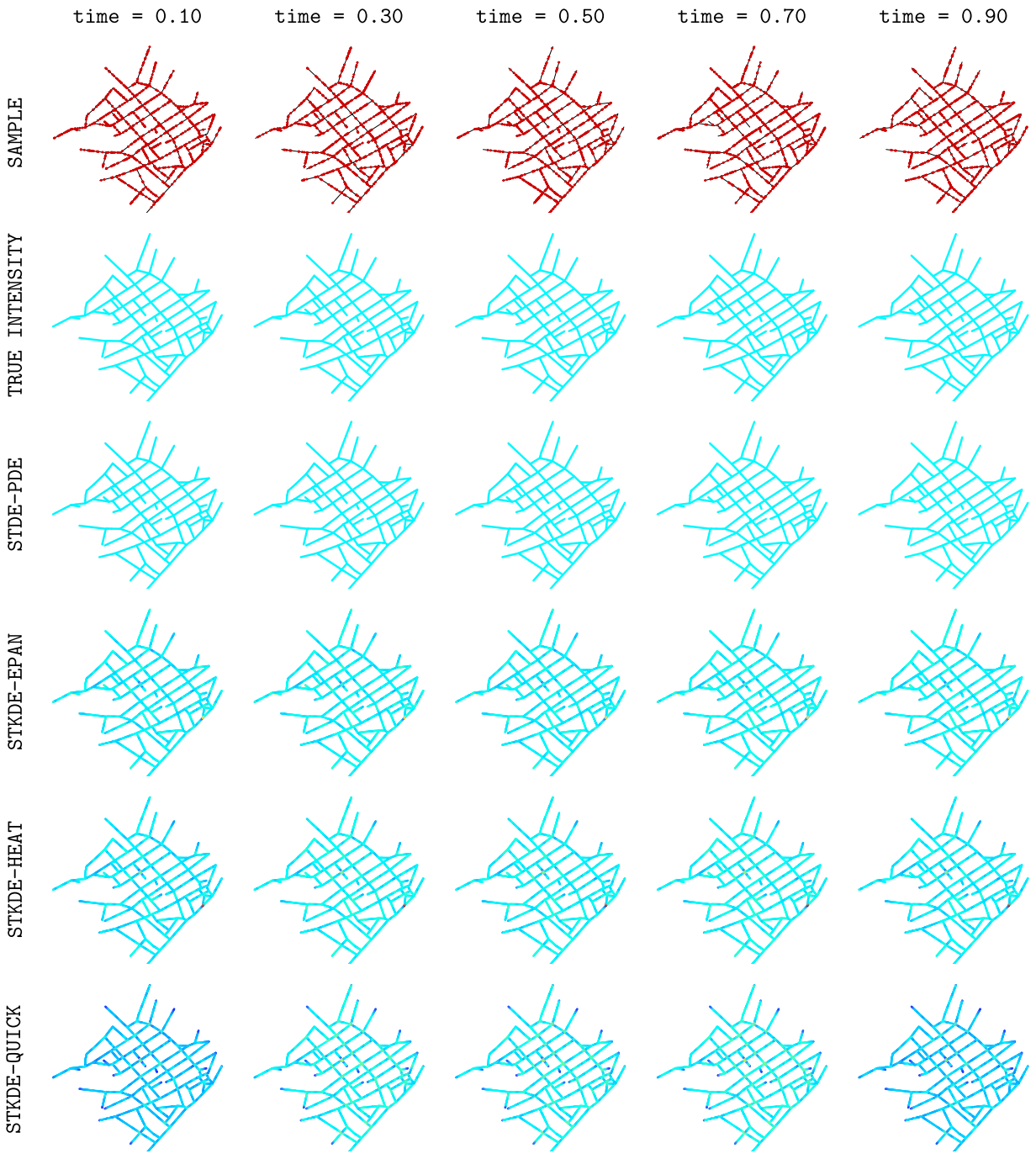


Fig. B.2. Simulation 4. Views of the sample from the first simulation replicate (first row), the true intensity (second row), and the mean estimates over 50 simulation replicates obtained using `STDE-PDE` (third row), `STKDE-EPAN` (fourth row), `STKDE-HEAT` (fifth row), and `STKDE-QUICK` (sixth row), all evaluated at different time instants in the interval $[0, 1]$. All plots share the same color scale, with bluish colors corresponding to low intensity values. The competing methods are detailed in Section 4.

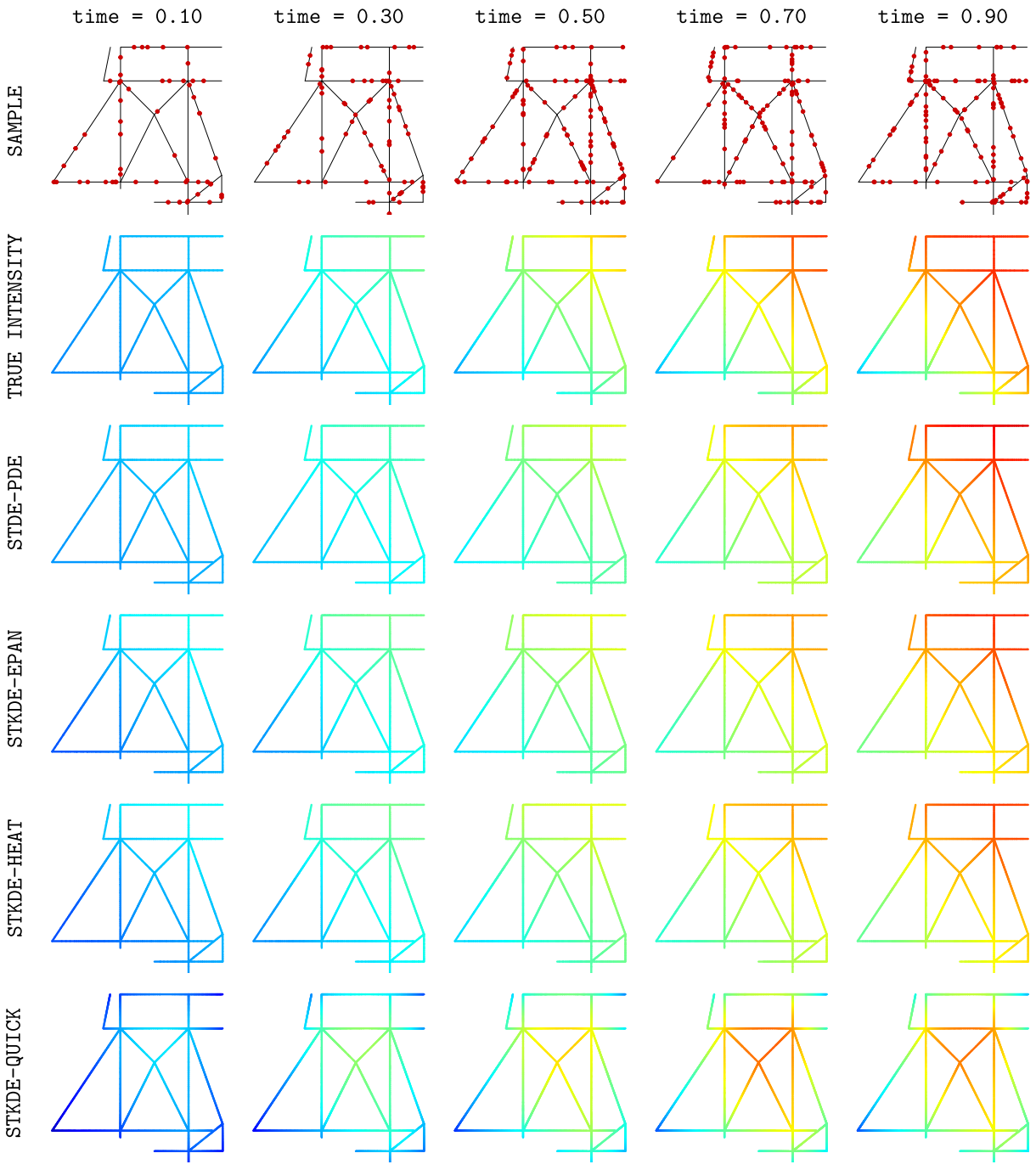


Fig. B.3. Simulation 5. Views of the sample from the first simulation replicate (first row), the true intensity (second row), and the mean estimates over 50 simulation replicates obtained using STDE-PDE (third row), STKDE-EPAN (fourth row), STKDE-HEAT (fifth row), and STKDE-QUICK (sixth row), all evaluated at different time instants in the interval $[0,1]$. All plots share the same color scale, with bluish colors corresponding to low intensity values. The competing methods are detailed in Section 4.

References

Alaimo Di Loro, P., Mingione, M., Fantozzi, P., 2024. Semi-parametric spatio-temporal hawkes process for modelling road accidents in Rome. J. Agric. Biol. Environ. Stat. 30.

Anderson, T.K., 2009. Kernel density estimation and K-means clustering to profile road accident hotspots. Accid. Anal. Prev. 41 (3), 359–364.

Arioli, M., Benzi, M., 2018. A finite element method for quantum graphs. IMA J. Numer. Anal. 38 (3), 1119–1163.

Arnone, E., Azzimonti, L., Nobile, F., Sangalli, L.M., 2019. Modeling spatially dependent functional data via regression with differential regularization. J. Multivariate Anal. 170, 275–295. http://dx.doi.org/10.1016/j.jmva.2018.09.006.

Arnone, E., Clemente, A., Sangalli, L.M., Lila, E., Ramsay, J., Formaggia, L., 2025. fdaPDE: Physics-informed spatial and functional data analysis. R package version 1.1-21, url: https://CRAN.R-project.org/package=fdaPDE.

Arnone, E., Ferraccioli, F., Pigolotti, C., Sangalli, L.M., 2022. A roughness penalty approach to estimate densities over two-dimensional manifolds. Comput. Statist. Data Anal. 174, http://dx.doi.org/10.1016/j.csda.2022.107527.

Baddeley, A., Jammalamadaka, A., Nair, G., 2014. Multitype point process analysis of spines on the dendrite network of a neuron. J. R. Stat. Soc. Ser. C. Appl. Stat. 63 (5), 673–694.

Baddeley, A., Nair, G., Rakshit, S., McSwiggan, G., 2017. "Stationary" point processes are uncommon on linear networks. Stat 6 (1), 68–78.

Baddeley, A., Nair, G., Rakshit, S., McSwiggan, G., Davies, T.M., 2021. Analysing point patterns on networks – A review. Spat. Stat. 42.

Baddeley, A., Rubak, E., Turner, R., 2015. Spatial Point Patterns: Methodology and Applications with R. CRC Press.

Baddeley, A., Turner, R., 2005. Spatstat: an R package for analyzing spatial point patterns. J. Stat. Softw. 12, 1–42.

Baddeley, A., Turner, R., Rubak, E., McSwiggan, G., Davies, T.M., Moradi, M.M., Rakshit, S., Cronie, O., 2023. Spatstat.linnet: Linear networks functionality of the "spatstat" family. R package version 3.1-5, url: https://CRAN.R-project.org/package=spatstat.linnet.

Barbi, C., Menafoglio, A., Secchi, P., 2023. An object-oriented approach to the analysis of spatial complex data over stream-network domains. Spat. Stat. 58.

Barr, C.D., Schoenberg, F.P., 2010. On the Voronoi estimator for the intensity of an inhomogeneous planar Poisson process. Biometrika 97 (4), 977–984.

Begu, B., Panzeri, S., Arnone, E., Carey, M., Sangalli, L.M., 2024. A nonparametric penalized likelihood approach to density estimation of space-time point patterns. Spat. Stat. 61, http://dx.doi.org/10.1016/j.spasta.2024.100824.

Berkolaiko, G., Kuchment, P., 2013. Introduction to Quantum Graphs. American Mathematical Society.

Bondy, J.A., Murty, U.S.R., 1976. Graph Theory with Applications. Macmillan.

Borruso, G., 2003. Network density and the delimitation of urban areas. Trans. GIS 7 (2), 177–191.

Borruso, G., 2005. Network density estimation: analysis of point patterns over a network. In: International Conference on Computational Science and Its Applications. Springer, pp. 126–132.

Borruso, G., 2008. Network density estimation: a GIS approach for analysing point patterns in a network space. Trans. GIS 12 (3), 377–402.

Briz-Redón, Á., Martínez-Ruiz, F., Montes, F., 2019. Spatial analysis of traffic accidents near and between road intersections in a directed linear network. Accid. Anal. Prev. 132.

Chaudhuri, S., Juan, P., Mateu, J., 2023. Spatio-temporal modeling of traffic accidents incidence on urban road networks based on an explicit network triangulation. J. Appl. Stat. 50 (16), 3229–3250.

Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J., 2013. Stochastic Geometry and its Applications. John Wiley & Sons.

Clemente, A., Arnone, E., Mateu, J., Sangalli, L.M., 2023a. Non-parametric density estimation over linear networks. In: Proceedings of the GRASPA 2023 Conference. pp. 141–145.

Clemente, A., Arnone, E., Mateu, J., Sangalli, L.M., 2023b. Spatial regression with differential regularization over linear networks. In: Proceedings of the Statistics and Data Science Conference. pp. 196–200.

Cressie, N., Frey, J., Harch, B., Smith, M., 2006. Spatial prediction on a river network. J. Agric. Biol. Environ. Stat. 11, 127–150.

Cressie, N., Wikle, C.K., 2015. Statistics for Spatio-Temporal Data. John Wiley & Sons.

Daley, D.J., Vere-Jones, D., 2003. An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods. Springer.

Daley, D.J., Vere-Jones, D., 2008. An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure. Springer.

D'Angelo, N., Adelfio, G., Abbruzzo, A., Mateu, J., 2022. Inhomogeneous spatio-temporal point processes on linear networks for visitors' stops data. Ann. Appl. Stat. 16 (2), 791–815.

D'Angelo, N., Payares, D., Adelfio, G., Mateu, J., 2024. Self-exciting point process modelling of crimes on linear networks. Stat. Model. 24 (2), 139–168.

Daniel, J., Horrocks, J., Umphrey, G.J., 2018. Penalized composite likelihoods for inhomogeneous Gibbs point process models. Comput. Statist. Data Anal. 124, 104–116.

Das, K., Yu, S., Wang, G., Wang, L., 2024. Nonparametric density estimation for data scattered on irregular spatial domains: A likelihood-based approach using bivariate penalized spline smoothing. arXiv preprint arXiv:2408.16963.

De Boor, C., 1978. A Practical Guide to Splines. Springer.

Deublein, M., Schubert, M., Adey, B.T., Köhler, J., Faber, M.H., 2013. Prediction of road accidents: A Bayesian hierarchical approach. Accid. Anal. Prev. 51, 274–291.

Díaz-Avalos, C., Juan, P., Mateu, J., 2013. Similarity measures of conditional intensity functions to test separability in multidimensional point processes. Stoch. Environ. Res. Risk Assess. 27, 1193–1205.

Diggle, P.J., 2013. Statistical Analysis of Spatial and Spatio-Temporal Point Patterns. CRC Press.

Diggle, P.J., Marron, J.S., 1988. Equivalence of smoothing parameter selectors in density and intensity estimation. J. Amer. Statist. Assoc. 83 (403), 793–800.

Eckardt, M., Moradi, M., 2024. Marked spatial point processes: current state and extensions to point processes on linear networks. J. Agric. Biol. Environ. Stat. 29, 346–378.

Fan, Y., Zhu, X., She, B., Guo, W., Guo, T., 2018. Network-constrained spatio-temporal clustering analysis of traffic collisions in Jianghan District of Wuhan, China. PLoS One 13 (4).

Ferraccioli, F., Arnone, E., Finos, L., Ramsay, J.O., Sangalli, L.M., 2021. Nonparametric density estimation over complicated domains. J. R. Stat. Soc. Ser. B Stat. Methodol. 83 (2), 346–368. http://dx.doi.org/10.1111/rssb.12415.

Fuentes-Santos, I., González-Manteiga, W., Mateu, J., 2018. A first-order, ratio-based nonparametric separability test for spatiotemporal point processes. Environmetrics 29 (1).

Ganio, L.M., Torgersen, C.E., Gresswell, R.E., 2005. A geostatistical approach for describing spatial pattern in stream networks. Front. Ecol. Environ. 3, 138–144.

Ghorbani, M., Vafaei, N., Dvořák, J., Myllymäki, M., 2021. Testing the first-order separability hypothesis for spatio-temporal point patterns. Comput. Statist. Data Anal. 161.

Gilardi, A., Borgoni, R., Mateu, J., 2024a. A nonseparable first-order spatiotemporal intensity for events on linear networks: An application to ambulance interventions. Ann. Appl. Stat. 18 (1), 529–554.

Gilardi, A., Lovelace, R., Rowlingson, B.S., Fernández, S., 2024b. osmextract: Download and import open street map data extracts. R package version 0.5.2, url: https://cran.r-project.org/package=osmextract.

Gilardi, A., Mateu, J., Borgoni, R., Lovelace, R., 2022. Multivariate hierarchical analysis of car crashes data considering a spatial network lattice. J. R. Stat. Soc. Ser. A: Stat. Soc. 185 (3), 1150–1177.

González, J.A., Rodríguez-Cortés, F.J., Cronie, O., Mateu, J., 2016. Spatio-temporal point process statistics: a review. Spat. Stat. 18, 505–544.
Goodd, I.J., Gaskins, R.A., 1971. Nonparametric roughness penalties for probability densities. Biometrika 58 (2), 255–277.
Göttlich, S., Schillinger, T., 2023. Data-inspired modeling of accidents in traffic flow networks. arXiv preprint arxiv:2305.03469.
Gu, C., Jeon, Y., Lin, Y., 2013. Nonparametric density estimation in high-dimensions. Statist. Sinica 1131–1153.
Gu, C., Qiu, C., 1993. Smoothing spline density estimation: Theory. Ann. Statist. 21 (1), 217–234.
Guan, Y., Jalilian, A., Waagepetersen, R., 2015. Quasi-likelihood for spatial point processes. J. R. Stat. Soc. Ser. B Stat. Methodol. 77 (3), 677–697.
Illian, J.B., Hendrichsen, D.K., 2010. Gibbs point process models with mixed effects. Environmetrics 21 (3–4), 341–353.
Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. John Wiley & Sons.
Kalair, K., Connaughton, C., Alaimo Di Loro, P., 2021. A non-parametric hawkes process model of primary and secondary accidents on a UK smart motorway. J. R. Stat. Soc. Ser. C. Appl. Stat. 70 (1), 80–97.
Khalid, S., Shoaib, F., Qian, T., Rui, Y., Bari, A.I., Sajjad, M., Shakeel, M., Wang, J., 2018. Network constrained spatio-temporal hotspot mapping of crimes in Faisalabad. Appl. Spat. Anal. Policy 11, 599–622.
Laaha, G., Skøien, J.O., Nobilis, F., Blöschl, G., 2013. Spatial prediction of stream temperatures using top-kriging with an external drift. Environ. Model. Assess. 18, 671–683.
Lange, K., 2013. Optimization. Springer Science & Business Media.
Liu, G., Jin, H., Li, J., Hu, X., Li, J., 2022. A Bayesian deep learning method for freeway incident detection with uncertainty quantification. Accid. Anal. Prev. 176.
Liu, D.C., Nocedal, J., 1989. On the limited memory BFGS method for large scale optimization. Math. Program. 45, 503–528.
Marshall, S., Gil, J., Kropf, K., Tomko, M., Figueiredo, L., 2018. Street network studies: from networks to models and their representations. Netw. Spat. Econ. 18, 735–749.
Mateu, J., Moradi, M.M., Cronie, O., 2020. Spatio-temporal point patterns on linear networks: Pseudo-separable intensity estimation. Spat. Stat. 37.
McSwiggan, G., Baddeley, A., Nair, G., 2017. Kernel density estimation on a linear network. Scand. J. Stat. 44 (2), 324–345.
McSwiggan, G., Baddeley, A., Nair, G., 2020. Estimation of relative risk for events on a linear network. Stat. Comput. 30, 469–484.
Møller, J., Waagepetersen, R.P., 2003. Statistical Inference and Simulation for Spatial Point Processes. CRC Press.
Moradi, M.M., Cronie, O., Mateu, J., 2022. stlnpp: Spatio-temporal analysis of point patterns on linear networks. R package version 0.3.9, url: https://CRAN.R-project.org/package=stlnpp.
Moradi, M.M., Cronie, O., Rubak, E., Lachieze-Rey, R., Mateu, J., Baddeley, A., 2019. Resample-smoothing of Voronoi intensity estimators. Stat. Comput. 29 (5), 995–1010.
Moradi, M.M., Mateu, J., 2020. First-and second-order characteristics of spatio-temporal point processes on linear networks. J. Comput. Graph. Statist. 29 (3), 432–443.
Moradi, M.M., Rodríguez-Cortés, F.J., Mateu, J., 2018. On kernel-based intensity estimation of spatial point patterns on linear networks. J. Comput. Graph. Statist. 27 (2), 302–311.
Nocedal, J., 1980. Updating quasi-Newton matrices with limited storage. Math. Comp. 35 (151), 773–782.
Nocedal, J., Wright, S.J., 1999. Numerical Optimization. Springer.
Okabe, A., Satoh, T., Sugihara, K., 2009. A kernel density estimation method for networks, its computational method and a GIS-based tool. Int. J. Geogr. Inf. Sci. 23 (1), 7–32.
Okabe, A., Sugihara, K., 2012. Spatial Analysis Along Networks: Statistical and Computational Methods. John Wiley & Sons.
Okawa, M., Iwata, T., Kurashima, T., Tanaka, Y., Toda, H., Ueda, N., 2019. Deep mixture point processes: Spatio-temporal event prediction with rich contextual information. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. pp. 373–383.
Rakshit, S., Davies, T., Moradi, M.M., McSwiggan, G., Nair, G., Mateu, J., Baddeley, A., 2019. Fast kernel smoothing of point patterns on a large network using two-dimensional convolution. Int. Stat. Rev. 87 (3), 531–556.
Ramírez, A.F., Valencia, C., 2021. Spatiotemporal correlation study of traffic accidents with fatalities and injuries in Bogota (Colombia). Accid. Anal. Prev. 149.
Schneble, M., Kauermann, G., 2022. Intensity estimation on geometric networks with penalized splines. Ann. Appl. Stat. 16 (2), 843–865.
Scott, D.W., 2015. Multivariate Density Estimation: Theory, Practice, and Visualization. John Wiley & Sons.
Silverman, B.W., 1982a. Algorithm AS 176: Kernel density estimation using the fast Fourier transform. J. R. Stat. Soc. Ser. C. Appl. Stat. 31 (1), 93–99.
Silverman, B.W., 1982b. On the estimation of a probability density function by the maximum penalized likelihood method. Ann. Statist. 10 (3), 795–810.
Silverman, B.W., 2018. Density Estimation for Statistics and Data Analysis. Routledge.
Skøien, J.O., Merz, R., Blöschl, G., 2006. Top-kriging-geostatistics on stream networks. Hydrol. Earth Syst. Sci. 10 (2), 277–287.
Som, N.A., Monestiez, P., Ver Hoef, J.M., Zimmerman, D.L., Peterson, E.E., 2014. Spatial sampling on streams: principles for inference on aquatic networks. Environmetrics 25 (5), 306–323.
Tapia, R.A., Thompson, J.R., 1978. Nonparametric Probability Density Estimation. Johns Hopkins University Press.
van der Meer, L., Abad, L., Gilardi, A., Lovelace, R., 2024. Sfnetworks: Tidy geospatial networks. R package version 0.6.5, url: https://cran.r-project.org/package=sfnetworks.
Vemulapalli, S.S., Ulak, M.B., Ozguven, E.E., Sando, T., Horner, M.W., Abdelrazig, Y., Moses, R., 2017. GIS-based spatial and temporal analysis of aging-involved accidents: a case study of three counties in Florida. Appl. Spat. Anal. Policy 10, 537–563.
Ver Hoef, J.M., Peterson, E.E., 2010. A moving average approach for spatial statistical models of stream networks. J. Amer. Statist. Assoc. 105 (489), 6–18.
Ver Hoef, J.M., Peterson, E.E., Isaak, D.J., 2019. Spatial statistical models for stream networks. In: Handbook of Environmental and Ecological Statistics. Chapman and Hall/CRC, pp. 421–444.
Ver Hoef, J.M., Peterson, E., Theobald, D., 2006. Spatial statistical models that use flow and stream distance. Environ. Ecol. Stat. 13, 449–464.
Vere-Jones, D., 2009. Some models and procedures for space-time point processes. Environ. Ecol. Stat. 16, 173–195.
Wang, C., Quddus, M., Ison, S., 2013. A spatio-temporal analysis of the impact of congestion on traffic safety on major roads in the UK. Transp. A: Transp. Sci. 9 (2), 124–148.
West, D.B., 2001. Introduction to Graph Theory. Prentice Hall.
Xie, Z., Yan, J., 2008. Kernel density estimation of traffic accidents in a network space. Comput. Environ. Urban Syst. 32 (5), 396–406.
Zhong, X., Kealy, A., Duckham, M., 2016. Stream Kriging: Incremental and recursive ordinary Kriging over spatiotemporal data streams. Comput. Geosci. 90, 134–143.
Zhu, S., Ding, R., Zhang, M., Van Hentenryck, P., Xie, Y., 2021. Spatio-temporal point processes with attention for traffic congestion event modeling. IEEE Trans. Intell. Transp. Syst. 23 (7), 7298–7309.