



Weighted Poisson intensity modelling for marked point processes with real-valued marks

Downloaded from: <https://research.chalmers.se>, 2026-10-05 04:02 UTC

Citation for the original published paper (version of record):

Tarantino, M., D'Angelo, N., Cronie, O. et al (2026). Weighted Poisson intensity modelling for marked point processes with real-valued marks. Statistics, In Press.
<http://dx.doi.org/10.1080/02331888.2026.2726452>

N.B. When citing this work, cite the original published paper.



Weighted Poisson intensity modelling for marked point processes with real-valued marks

Marco Tarantino , Nicoletta D'Angelo , Ottmar Cronie & Giada Adelfio

To cite this article: Marco Tarantino , Nicoletta D'Angelo , Ottmar Cronie & Giada Adelfio (07 Sep 2026): Weighted Poisson intensity modelling for marked point processes with real-valued marks, *Statistics*, DOI: [10.1080/02331888.2026.2726452](https://doi.org/10.1080/02331888.2026.2726452)

To link to this article: <https://doi.org/10.1080/02331888.2026.2726452>



© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 07 Sep 2026.



Submit your article to this journal [↗](#)



Article views: 291



View related articles [↗](#)



View Crossmark data [↗](#)

Weighted Poisson intensity modelling for marked point processes with real-valued marks

Marco Tarantino^a, Nicoletta D'Angelo^a, Ottmar Cronie^b and Giada Adelfio^a

^aDepartment of Economics, Business and Statistics, University of Palermo, Palermo, Italy; ^bDepartment of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden

ABSTRACT

Real-valued marks in spatial point processes are often analyzed separately from first-order intensity estimation or through assumptions specific to their generating mechanism. We propose a Weighted Poisson Intensity approach for incorporating real-valued marks into first-order intensity modelling without requiring a full probabilistic specification of their marginal distribution. The proposed framework represents the intensity as a baseline Poisson component combined with a non-negative mark-related contribution. Since marks are observed as attributes of observed events, we consider two modelling views: a Parametric Marked Poisson model, in which each observed event is represented by a location–mark pair on the joint space and the mark enters as an explicit intensity argument, and a Semi-parametric Marked Poisson model, in which the observed marks are interpreted as spatially varying values sampled at event locations, which are reconstructed to form a mark-related adjustment over the ground domain. These models allow us to investigate how real-valued marks are associated with the fitted first-order intensity under these two alternative interpretations. Their performance is evaluated through a simulation study based on Poisson and clustered point processes under different mark-generating mechanisms, and their practical usefulness is illustrated through applications to real data.

ARTICLE HISTORY

Received 10 December 2025
Accepted 20 August 2026

KEYWORDS

Real-valued marks; weighted intensity; Poisson point process

2020 MATHEMATICS

SUBJECT

CLASSIFICATIONS

60G55; 62H11

1. Introduction

Point patterns are often observed together with additional attributes attached to each event, commonly referred to as marks. Such marks arise in many application domains, including seismology, where earthquakes are associated with magnitudes [1,2], forestry and ecology, where trees may be characterized by size-related variables [3,4], astronomy, where stars may be described by physical quantities such as age or effective temperature [5], disease mapping [6], and image analysis [7]. Marks may be categorical, continuous, or even function-valued [8,9]. In this paper, we focus on marked point processes with real-valued marks, that is, settings in which each observed event carries a continuous mark.

CONTACT Ottmar Cronie  ottmar@chalmers.se  Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Göteborg, Sweden.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

The analysis of marked point processes has traditionally relied on summary characteristics tailored to the mark structure, such as mark correlation functions, mark covariance functions, and mark variograms [9–11]. These tools are useful for exploring mark dependence and mark–location association. More recent work has extended such summaries to inhomogeneous marked point processes and to settings beyond the classical planar case [9,12,13]. However, these approaches are primarily descriptive and do not directly provide a methodology for modelling the first-order intensity while explicitly incorporating real-valued marks.

A further difficulty is conceptual, since marks do not necessarily play the same role as ordinary spatial covariates in intensity-based analyses. Spatial covariates are typically defined over the whole observation window and can therefore be evaluated both at observed events and at arbitrary spatial locations. By contrast, a mark is observed only at realized event locations, as part of the corresponding location–mark pair. Consequently, marks may be assigned different modelling roles depending on the scientific interpretation of the data. Hence, in some applications, marks are naturally viewed as part of the event description, so that the model is formulated on the product space of locations and marks, as in the standard marked point process framework [14,15]. In other applications, observed marks are more naturally interpreted as values of an underlying spatially structured quantity sampled at event locations [16,17].

A central issue is therefore whether the marks carry information that is relevant for the first-order structure of the spatial pattern. In some applications, marks may be only weakly related to the event rate, whereas in others they plausibly contain information about the spatial mechanism generating the observed pattern [17,18]. For example, in seismology earthquake magnitude can affect the occurrence of future earthquakes [1], while in forest data tree size may reflect local competitive or environmental conditions and may therefore vary with local event concentration [4]. In such settings, incorporating real-valued marks into first-order intensity estimation may help capture spatial variation that is not explained by location alone.

For marked point processes with categorical marks, several well-established modelling strategies are available [15,19,20]. By contrast, for continuous real-valued marks, existing approaches are often application-specific, rely on additional assumptions on the mark-generating mechanism, or are developed in conditional-intensity settings tailored to particular contexts. In seismology, for instance, ETAS-type models incorporate earthquake magnitude through a multiplicative contribution to the conditional intensity [1,21]. Likewise, Bayesian Hawkes formulations treat marks jointly with the conditional intensity within a model-specific inferential framework [22]. These approaches are effective in their own domains, but they are not designed as a flexible methodology for first-order Poisson intensity modelling with real-valued marks.

In this work, our goal is to estimate the first-order intensity of a marked point process while explicitly accounting for the contribution of real-valued marks. To this end, we introduce a Weighted Poisson Intensity (WPI) approach, in which a baseline Poisson intensity on the ground space is combined with a non-negative mark-related component. Within this framework, the role of the mark can be specified in two different ways, leading to two distinct modelling implementations.

The first implementation is the *Parametric Marked Poisson model* (PMP), which treats the real-valued mark as an explicit argument of the intensity and formulates the model on

the product space of locations and marks. Under this view, the mark enters the model as an additional modelling dimension rather than as an ordinary spatial covariate, in close analogy with product-domain representations used in spatio-temporal point process modelling [4]. The second implementation is the *Semi-parametric Marked Poisson model* (SMP), which instead interprets the observed marks as values of a spatially structured signal associated with the point pattern and incorporates them through a reconstructed mark-related adjustment defined on the spatial ground domain [16,17].

The PMP approach provides an interpretable parametric contribution of the mark by modelling the intensity on the joint space of locations and marks, while the SMP approach constructs a spatially varying mark-related adjustment through kernel smoothing on the ground domain, offering a more flexible alternative in settings where a fully parametric marked specification is less convenient. Neither approach requires a fully specified parametric probability model for the marginal distribution of the continuous mark itself. Rather, both approaches are designed to assess how real-valued marks are associated with the fitted first-order intensity under two different modelling interpretations.

Our primary focus is on Poisson point processes, since the proposed methodology is built around first-order Poisson intensity modelling. Nevertheless, we also explore its behaviour for clustered point process settings, such as Hotspot processes and log-Gaussian Cox processes (LGCPs), where real-valued marks may still contain useful information about the spatial distribution of events even when the observed process is not Poisson [23,24]. In this sense, the proposed methods are intended as practical first-order modelling tools that may also be informative under moderate departures from the Poisson assumption.

The proposed methods were implemented and the analyses were carried out in R [25]. All code and analyses are available from the authors.

The remainder of the paper is organized as follows. Section 2 reviews basic preliminaries on marked Poisson point processes. Section 3 introduces the Weighted Poisson Intensity framework and presents its two specific implementations, namely PMP and SMP. Section 4 reports a simulation study comparing the classical Poisson model with the two proposed approaches under Poisson, Hotspot, and LGCP scenarios. Section 5 presents two real-data applications. In Section 6, conclusions are drawn and avenues for future research are discussed.

2. Marked Poisson point process preliminaries

A spatial point process is a random element in the measurable space $\mathcal{N}$ of locally finite point patterns. When restricted to a bounded observation window $W_2 \subset \mathbb{R}^2$, its realization is called a two-dimensional spatial point pattern. Thus, the point process represents the underlying stochastic mechanism, while the point pattern is the observed configuration of points within W_2 . A marked point pattern is obtained when each observed point s_i carries an associated mark taking values in a mark space $\mathcal{M}$. In the most general case, the mark associated with point s_i is a vector $\mathbf{m}_i \in \mathcal{M} \subset \mathbb{R}^k$, so that an observed marked point pattern can be written as

$$A = \{(s_i, \mathbf{m}_i) : i = 1, \dots, n\},$$

that is, as a finite set of events on the product space $W_2 \times \mathcal{M}$. In this work, however, we focus on the case of a single continuous real-valued mark. Therefore, from this point on, we simply write the mark as $m_i \in \mathbb{R}$.

We begin by briefly recalling the definition of a Poisson point process on $\mathbb{R}^2$, which serves as the unmarked reference model for the marked setting considered below. A Poisson point process on $\mathbb{R}^2$ is characterized by the property that, for any bounded Borel set $B \subset \mathbb{R}^2$, the count $N(B)$ follows a Poisson distribution with mean

$$\Lambda(B) = \int_B \lambda(s) \, ds,$$

where $\lambda(s)$ is the first-order intensity function. Moreover, counts in disjoint regions are independent. Let the parametric intensity function of a two-dimensional Poisson point process model be $\lambda(s; \theta)$, with θ a vector of unknown parameters to be estimated. The log-likelihood of the Poisson model is

$$\log L(\theta) = \sum_i \log \lambda(s_i; \theta) - \int_{W_2} \lambda(s; \theta) \, ds, \quad (1)$$

where the sum is over all points in the point pattern [14].

A marked Poisson point process may be represented as a Poisson point process on the product space $\mathbb{R}^2 \times \mathcal{M}$, with intensity function $\lambda(s, m)$. Its restriction to $W_2 \times \mathcal{M}$ yields the observed marked point pattern. For a marked point process on $\mathbb{R}^2 \times \mathcal{M}$, the first-order intensity is

$$\lambda(s, m) = \lim_{|ds| \rightarrow 0, |dm| \rightarrow 0} \frac{\mathbb{E}[N(ds \times dm)]}{|ds \times dm|},$$

where $ds \times dm$ is an infinitesimal region defined by the product of an infinitesimal two-dimensional region ds around the location s and an infinitesimal element in the mark space dm , while $N(ds \times dm)$ is the number of events in that region.

When the conditional mark density exists, the joint intensity can be factorized as

$$\lambda(s, m) = \lambda(s)f(m | s), \quad (2)$$

where $\lambda(s)$ is the ground intensity and $f(m | s)$ is the conditional density of the mark given location. Consequently, the ground process is a Poisson process on $\mathbb{R}^2$ with intensity

$$\lambda(s) = \int_{\mathcal{M}} \lambda(s, m) \, dm,$$

where

$$f(m | s) = \frac{\lambda(s, m)}{\lambda(s)},$$

whenever $\lambda(s) > 0$. This decomposition is useful as a conceptual reference for distinguishing the spatial contribution of the ground process from the contribution associated with the mark. Equivalently, a marked point process may be described either directly on the product space $\mathbb{R}^2 \times \mathcal{M}$ or through its ground process together with the conditional mark distribution $f(m | s)$; these two descriptions refer to the same collection of location–mark pairs [26].

A common distinction is between independently marked Poisson processes, in which the mark distribution does not depend on location, and Poisson processes with a location-dependent mark distribution, in which the conditional distribution of the marks may vary with s [26]. If $\lambda(s, m)$ is constant over $\mathbb{R}^2 \times \mathcal{M}$, the process is homogeneous; otherwise it is inhomogeneous. In such a case, a common modelling choice is the log-linear specification

$$\lambda(s, m) = \lambda(s, m; \theta) = \exp(\theta^\top \mathbf{Z}(s, m)), \quad s \in W_2, \quad m \in \mathcal{M}, \quad \theta \in \Theta, \quad (3)$$

where $\mathbf{Z}(s, m)$ denotes a vector of predictors constructed from spatial location and, where appropriate, from the mark.

3. Weighted Poisson intensity methodology

Let X denote a spatial marked point process and let $A = \{(s_i, m_i) : i = 1, \dots, n\}$ be the observed marked point pattern in a bounded window $W_2 \subset \mathbb{R}^2$, with mark space $\mathcal{M} \subset \mathbb{R}$. As recalled in Section 2, a marked point process is described in the product space $\mathbb{R}^2 \times \mathcal{M}$, while its spatial ground process is defined on $\mathbb{R}^2$.

The proposed Weighted Poisson Intensity (WPI) approach aims to incorporate real-valued mark information into first-order intensity modelling. The general idea is to start from a baseline Poisson intensity for the spatial ground process and to enrich it by means of an additional non-negative mark-related contribution. Let u denote the generic argument of the intensity, so that $u = (s, m)$ when the model is defined on the product space $W_2 \times \mathcal{M}$, and $u = s$ when the model is defined directly on the spatial window W_2 . Under the WPI framework, the intensity may be written as

$$\lambda(u; \theta) = \nu(s; \theta) \Phi(u), \quad (4)$$

where $\nu(s; \theta)$ represents a baseline Poisson intensity defined on W_2 , and $\Phi(u)$ is a non-negative mark-related component. The continuous mark itself is not required to be non-negative. Throughout, $m \in \mathcal{M} \subset \mathbb{R}$, so both positive and negative mark values are admissible. Hence, the WPI formulation separates a baseline spatial component from an additional term carrying the information associated with the continuous mark.

Two distinct modelling views of the real-valued mark are considered within the WPI framework. In the first, the mark is treated as an explicit coordinate of the event representation. The corresponding Parametric Marked Poisson model (PMP) is therefore formulated on the product space $W_2 \times \mathcal{M}$, with $u = (s, m)$, and models the joint first-order intensity of location–mark pairs. Under this interpretation, the mark is not treated as an ordinary spatial covariate defined at every location in W_2 , but as an additional modelling dimension of the marked event itself. The associated mark-related contribution is denoted by $\Phi_P(m)$. In the second view, adopted by the Semi-parametric Marked Poisson model (SMP), the observed marks are interpreted as sampled values of a spatially structured quantity associated with the ground process. The model is then defined directly on the spatial domain W_2 , with $u = s$, and the observed marks are used to reconstruct a spatially varying mark-related component, denoted by $\Phi_S(s)$. Thus, the two approaches differ both in the domain on which the intensity is defined and in the modelling status assigned to the mark.

3.1. Parametric marked Poisson model

The PMP model is defined on the product space $W_2 \times \mathcal{M}$, so that the argument in (4) is $u = (s, m)$. The observed marked pattern is accordingly represented as a point configuration $A = \{(s_i, m_i) : i = 1, \dots, n\}$ on the joint domain $W_2 \times \mathcal{M}$. Hence, $m \in \mathcal{M}$ is an additional coordinate of a potential event in the product space. Thus, $\lambda(s, m; \theta)$ describes the expected local density of events with spatial location close to s and mark value close to m . At the first-order level, this representation is formally analogous to that of a spatio-temporal point process, in which each event is represented by a location–time pair and the intensity is defined on the corresponding joint domain. Under this interpretation, the mark-related component is specified parametrically. We further impose a separability restriction between the spatial and mark components, so that the general WPI formulation in (4) becomes

$$\lambda(s, m; \theta) = v(s; \theta_s) \Phi_P(m; \theta_m), \quad (5)$$

where $v(s; \theta_s)$ denotes the baseline Poisson intensity on W_2 , $\Phi_P(m; \theta_m)$ is the mark-related contribution, and $\theta = (\theta_s, \theta_m)$ is the full parameter vector.

Thus, the PMP specification assumes that the joint first-order intensity on $W_2 \times \mathcal{M}$ factorizes into a spatial component and a mark component, where such separability is an additional modelling assumption.

Let $z_1(s), \dots, z_p(s)$ denote spatial predictors defined on W_2 . A general PMP specification is

$$\lambda(s, m; \theta) = \exp(\theta_0 + \theta_1 z_1(s) + \dots + \theta_p z_p(s) + g(m; \theta_m)), \quad (6)$$

where $g : \mathcal{M} \rightarrow \mathbb{R}$ is a specified parametric function describing the effect of the mark on the log-intensity scale, and the two terms in Equation (5) correspond to

$$v(s; \theta_s) = \exp(\theta_0 + \theta_1 z_1(s) + \dots + \theta_p z_p(s)), \quad \Phi_P(m; \theta_m) = \exp(g(m; \theta_m)).$$

Hence, in the PMP case, the general WPI mark-related component $\Phi(u)$ in Equation (4) is represented by $\Phi_P(m; \theta_m)$, that is, by a non-negative parametric function of the continuous mark. A simple special case is obtained by taking $g(m; \theta_m) = \theta_m m$, yielding

$$\lambda_P(s, m; \theta) = \exp(\theta_0 + \theta_1 z_1(s) + \dots + \theta_p z_p(s) + \theta_m m), \quad (7)$$

in which the mark has a globally constant parametric effect on the log-intensity scale.

Since the model is defined directly on $W_2 \times \mathcal{M}$, inference is based on the Poisson log-likelihood of the marked process:

$$\log L(\theta) = \sum_{i=1}^n \log \lambda(s_i, m_i; \theta) - \int_{W_2} \int_{\mathcal{M}} \lambda(s, m; \theta) \, dm \, ds. \quad (8)$$

To maximize (8), the integral term must be approximated numerically. In the present setting, this is done by a cubature construction on the joint domain $W_2 \times \mathcal{M}$, extending the general idea of Berman–Turner type likelihood approximation to the marked setting [19,27–29]. Thus, the observed marked points are augmented with dummy points generated over the joint domain of spatial coordinates and marks, and suitable weights are used to approximate the integral contribution in (8). This provides a computationally tractable

likelihood approximation while preserving the interpretation of the mark as an explicit modelling dimension.

When the primary interest concerns the spatial intensity over W_2 , rather than the full joint intensity on $W_2 \times \mathcal{M}$, it is natural to consider the marginal intensity obtained by integrating out the mark:

$$\lambda_p^{\text{marg}}(s; \boldsymbol{\theta}) = \int_{\mathcal{M}} \lambda(s, m; \boldsymbol{\theta}) \, dm, \quad (9)$$

whenever the integral is finite. Under the separable PMP specification, the marginal intensity can be written explicitly as

$$\lambda_p^{\text{marg}}(s; \boldsymbol{\theta}) = v(s; \boldsymbol{\theta}_s) \int_{\mathcal{M}} \Phi_P(m; \boldsymbol{\theta}_m) \, dm.$$

Provided that the integral exists, the marginal spatial intensity is therefore proportional to the baseline spatial component $v(s; \boldsymbol{\theta}_s)$, with a proportionality factor depending only on the mark parameters. In particular, within this separable PMP formulation, the inclusion of the mark affects the overall scale of the spatial intensity, but not its spatial shape.

This marginalization produces a first-order intensity defined on the spatial window only, while retaining the contribution of the continuous mark through the fitted PMP model. In particular, the existence of (9) depends on both the support of $\mathcal{M}$ and the chosen parametric form $g(m; \boldsymbol{\theta}_m)$.

A useful way to assess the role of the mark in fitted spatial terms is then to compare the marginal PMP intensity with the baseline Poisson intensity fitted on W_2 using spatial predictors only. Let

$$v(s; \boldsymbol{\gamma}) = \exp(\gamma_0 + \gamma_1 z_1(s) + \cdots + \gamma_p z_p(s)) \quad (10)$$

denote the corresponding spatial Poisson baseline model. For clarity, we denote the parameter vector of the separately fitted baseline Poisson model by $\boldsymbol{\gamma}$, to distinguish it from the PMP parameter vector $\boldsymbol{\theta}$. We then define the fitted relative intensity ratio for the PMP approach as

$$\mathcal{R}_P(s) = \frac{\lambda_p^{\text{marg}}(s; \boldsymbol{\theta})}{v(s; \boldsymbol{\gamma})}. \quad (11)$$

This ratio measures how the inclusion of the real-valued mark, through the marked model on $W_2 \times \mathcal{M}$, modifies the fitted spatial intensity relative to a model based only on the spatial ground process. Values greater than 1 indicate that the marked specification yields a higher fitted spatial intensity at location s , whereas values below 1 indicate the opposite.

Within the marked model itself, the direct contribution of the mark is instead represented by the parametric term $g(m; \boldsymbol{\theta}_m)$, or equivalently by $\Phi_P(m; \boldsymbol{\theta}_m) = \exp(g(m; \boldsymbol{\theta}_m))$. In the linear specification (7), for example, a one-unit increase in the mark changes the log-intensity by θ_m , so that the multiplicative contribution on the intensity scale is $\exp(\theta_m)$. More generally, the form of $g(m; \boldsymbol{\theta}_m)$ determines how the real-valued mark contributes to the intensity on the product space $W_2 \times \mathcal{M}$.

The PMP formulation can be extended in a straightforward way when more than one continuous mark is available, by replacing m with a mark vector and specifying the intensity on the corresponding higher-dimensional product space. In the present paper, however, we restrict attention to a single real-valued mark.

3.2. Semi-parametric marked Poisson model

The SMP model corresponds to the case where the real-valued mark is not treated as an explicit argument of the intensity on $W_2 \times \mathcal{M}$, but rather as a spatially structured quantity associated with the ground process. Accordingly, the model is specified directly on the spatial domain W_2 . Hence, the generic argument in the WPI representation (4) is $u = s$, rather than $u = (s, m)$. The mark-related component is therefore not a function of a free mark coordinate m . Instead, it is a spatial surface reconstructed from the observed marked pattern $A = \{(s_i, m_i) : i = 1, \dots, n\}$. Conditional on A , this reconstructed component is a function of spatial location s only. In this case, the WPI formulation becomes

$$\lambda(s; \theta) = v(s; \theta) \Phi_S(s; A) = v(s; \theta) \Phi_S(s), \quad (12)$$

where $\Phi_S(s)$ is a non-negative mark-related contribution defined over the ground domain. The final expression suppresses the dependence on the observed marked pattern A for notational simplicity. If $\Phi_S(s)$ were known, inference for θ could be based on the Poisson log-likelihood associated with the intensity $\lambda(s; \theta) = v(s; \theta) \Phi_S(s)$. In the proposed WPI approach, however, $\Phi_S(s)$ is unknown and replaced by a first-step estimate $\hat{\Phi}_S(s)$ obtained from the same observed point pattern. The resulting criterion is therefore not a genuine likelihood, but a two-step plug-in Poisson pseudo-likelihood. Specifically

$$\begin{aligned} \log L(\theta) &= \sum_i \log(v(s_i; \theta) \hat{\Phi}_S(s_i)) - \int_{W_2} v(s; \theta) \hat{\Phi}_S(s) \, ds \\ &= \sum_i \log v(s_i; \theta) - \int_{W_2} v(s; \theta) \, ds \\ &\quad + \left[\sum_i \log \hat{\Phi}_S(s_i) + \int_{W_2} v(s; \theta) (1 - \hat{\Phi}_S(s)) \, ds \right]. \end{aligned} \quad (13)$$

From (13), it follows that if $\hat{\Phi}_S(s) = 1$ for all $s \in W_2$, then the additional term vanishes and the criterion reduces to the standard Poisson log-likelihood.

Equation (13) should therefore be interpreted as a two-step plug-in Poisson pseudo-likelihood rather than as an exact likelihood. The mark-related component $\hat{\Phi}_S$ is first reconstructed from the observed marks and is then treated as fixed in the optimization with respect to θ . Consequently, the standard errors and p -values obtained from the Poisson fit are conditional on the estimated surface and should be regarded as naive, since they do not account for the additional variability introduced by the first-step smoothing. A fully valid unconditional inference would require a joint asymptotic treatment of $(\hat{\Phi}_S, \hat{\theta})$, or a sandwich-type correction accounting for the generated mark-related component, which is beyond the scope of this paper [30,31]. Nevertheless, the use of such a pseudo-likelihood is in line with point process practice, where Poisson likelihoods, composite likelihoods, and pseudo-likelihood criteria are often used as practical estimating functions for first-order modelling [19,32]. Under suitable regularity conditions, the proposed SMP criterion can therefore be viewed as a plug-in estimating function targeting a pseudo-true first-order intensity structure.

To maximize Equation (13), the integral term must be approximated numerically by means of a quadrature scheme [27]. This requires values of the mark-related contribution

not only at the observed event locations, but also at the dummy points generated over W_2 by the quadrature algorithm. Here, a crucial distinction arises: ordinary spatial covariates are defined over the whole observation window, whereas marks are observed only at event locations. Therefore, the observed marks cannot be directly assigned to quadrature points as if they were standard spatial covariates. To address this issue, we introduce a modelling assumption. The observed real-valued marks are treated as values of an underlying spatially structured surface, observed at the event locations and reconstructed over the whole window W_2 . This perspective is related to the geostatistical marking or random-field view of marked point processes, in which marks may be interpreted as values of an underlying spatial field sampled at the locations of the ground process [16,17]. A useful analogy is provided by quadrature methods for multitype point processes [19]: in that setting, dummy points are replicated across the finite set of mark categories in order to approximate the likelihood contribution over the mark space, whereas with continuous marks this is replaced here by the construction of a continuous mark-related surface over the spatial domain. In our setting, $\Phi_S(s)$ is represented through this reconstructed surface, which provides the mark-related information entering the intensity specification and should be interpreted as a smoothing-based proxy for the spatial signal carried by the marks, rather than as a uniquely identified latent field.

In this work, the reconstruction is obtained through a Nadaraya–Watson kernel smoother [33–35], so that for any $s \in W_2$

$$\hat{\kappa}(s; A) = \frac{\sum_{i=1}^n K_h(s - s_i) m_i}{\sum_{i=1}^n K_h(s - s_i)},$$

where $K_h(\cdot)$ is a bivariate kernel function with bandwidth h . Thus, $\hat{\kappa}(s; A)$ is a function of the spatial argument s , constructed from the observed locations and mark values in A . In the implementation, h is selected separately for each observed point pattern using the data-driven bandwidth selection criterion of Cronie and Van Lieshout [36], and is then kept fixed when evaluating the reconstructed mark-related surface at both observed and quadrature points. The resulting smooth surface provides a spatial surrogate for the observed mark values and can be evaluated at any location in W_2 , including quadrature points. To ensure identifiability with respect to the intercept parameter and to guarantee that the mark-related adjustment vanishes when the reconstructed surface is approximately constant, we use a centred version: $\kappa_c(s) = \hat{\kappa}(s) - \bar{\kappa}$, where $\bar{\kappa} = \frac{1}{n} \sum_{i=1}^n \hat{\kappa}(s_i)$, and n denotes the number of observed points. Under this centring, if the reconstructed mark-related surface is approximately flat over W_2 , then $\kappa_c(s)$ is approximately zero and the model reduces to a standard Poisson specification based only on the spatial predictors. The SMP intensity is then specified as

$$\lambda_S(s; \theta) = \exp(\theta_0 + \theta_1 z_1(s) + \cdots + \theta_p z_p(s) + \kappa_c(s)), \quad (14)$$

where θ is the vector of unknown parameters and $z_j(s)$ are the model covariates, and the two terms in Equation (12) correspond to

$$\nu(s; \theta) = \exp(\theta_0 + \theta_1 z_1(s) + \cdots + \theta_p z_p(s)), \quad \Phi_S(s) = \exp(\kappa_c(s)).$$

Hence, in the SMP case, the general WPI mark-related component $\Phi(u)$ is represented by $\Phi_S(s)$, that is, by the exponential of the reconstructed centred mark-related surface.

From this specification, $\exp(\kappa_c(s))$ represents the direct multiplicative contribution of the reconstructed mark-related signal to the first-order intensity.

In contrast to the separable PMP specification on the product space $W_2 \times \mathcal{M}$, the SMP formulation does not impose a decomposition in which the mark contribution enters as a function of m only. Instead, after reconstruction, the mark-related component is represented directly as a location-dependent surface $\Phi_S(s)$ on W_2 . As a consequence, the inclusion of the mark in the SMP model may modify not only the overall level of the fitted intensity, but also its spatial shape. This is a key distinction from the separable PMP case, where marginalization over the mark provides a spatial intensity that remains proportional to the baseline component $\nu(s; \theta_s)$.

To assess the contribution of the reconstructed mark-related component in fitted spatial terms, we recall from Equation (10) the corresponding baseline spatial Poisson model fitted without the mark-related term. We then define the fitted relative intensity ratio for the SMP approach as

$$\mathcal{R}_S(s) = \frac{\lambda_S(s; \theta)}{\nu(s; \gamma)} = \exp \left[(\theta_0 - \gamma_0) + \sum_{j=1}^p (\theta_j - \gamma_j) z_j(s) + \kappa_c(s) \right]. \quad (15)$$

This ratio measures how the inclusion of the reconstructed mark-related component modifies the fitted spatial intensity relative to the baseline model.

The SMP formulation is particularly useful when the real-valued mark is more naturally interpreted as a spatially structured signal associated with the point locations, rather than as an explicit additional dimension of the process. In that setting, the model remains defined on W_2 , while the real-valued mark information is incorporated through a flexible, location-dependent adjustment to the baseline spatial intensity.

4. Simulation study

This section presents a simulation study aimed at evaluating the performance of the proposed Weighted Poisson Intensity approach and of its two specific implementations, namely the PMP and the SMP. The main goal is to assess whether incorporating real-valued marks into first-order intensity modelling can improve first-order intensity estimation relative to a classical Poisson model fitted without mark information.

Although the proposed methodology is developed for Poisson first-order intensity modelling, the simulation study is not restricted to Poisson point processes only. In addition to inhomogeneous Poisson processes, we also consider two classes of clustered point patterns, namely Hotspot processes and log-Gaussian Cox processes (LGCPs). This broader design makes it possible to evaluate the proposed methods not only under correct Poisson specification, but also in settings where the observed pattern departs from Poisson independence while the first-order intensity remains of primary interest.

The reason for including these three process classes is the following. Poisson scenarios provide the natural benchmark, since they are fully aligned with the modelling framework introduced in Section 3. Hotspot scenarios allow us to study situations in which the intensity surface exhibits strong local heterogeneity through spatially concentrated peaks, while still remaining within a first-order Poisson construction. LGCP scenarios instead represent genuinely clustered processes driven by a latent random field. In this case, the

Poisson model is no longer correctly specified for the full data-generating mechanism, but first-order intensity estimation remains an important step for practical modelling. In particular, for LGCP scenarios, we do not compare the proposed marked Poisson models with a fully specified Cox-process fit in terms of overall model adequacy, nor do we aim to assess second-order properties such as clustering or pairwise dependence. The purpose of including LGCP patterns is instead to evaluate whether the proposed mark-informed first-order approaches can still provide competitive or improved spatial intensity estimates in settings where the observed process departs from Poisson independence. The simulation study should therefore be interpreted as an assessment of first-order modelling performance, rather than as a comparison between full generative models for spatial point processes.

A key aspect of the simulation design concerns the role played by the continuous mark. As discussed in Section 3, a real-valued mark may enter first-order intensity modelling under two different interpretations: either as an explicit argument of the intensity on the product space, as in the PMP specification, or as a spatially structured signal attached to the ground process, as in the SMP specification. The simulation study is constructed to examine the behaviour of the proposed methods under these two modelling views. More specifically, the scenarios are based on three distinct mark-generating mechanisms. First, in the random-labelling benchmark, the spatial point pattern is generated independently of the mark and the observed marks carry no information about the first-order intensity. Second, in the product-space scenarios, the marked point process is generated jointly on $W_2 \times \mathcal{M}$ according to an intensity depending on both spatial location and mark value; these scenarios are designed to reflect the PMP interpretation. Third, in the spatial-field scenarios, a mark surface $M(s)$ is generated over the whole observation window before the point pattern is simulated, and the spatial process is then generated conditionally on an intensity depending on $M(s)$; the observed mark at a point s_i is therefore $m_i = M(s_i)$. These scenarios are designed to reflect the SMP interpretation. Thus, except for the random-labelling benchmark, the marks are not treated as post hoc attributes attached to an already generated spatial pattern. For simplicity, throughout the study, we consider a single real-valued mark associated with each observed event. For each scenario, we generate $n_{\text{sim}} = 250$ independent marked point patterns on the unit square, considering configurations with different expected sample sizes, specifically $\mathbb{E}[N] \in \{100, 500, 1000\}$. The exact data-generating mechanisms are described separately for Poisson, Hotspot, and LGCP processes in the following subsections. For each simulated pattern, we fit three competing first-order models: the baseline Poisson model, the PMP model, and the SMP model. All comparisons are carried out on the spatial ground domain W_2 . In the product-space scenarios, the target is the marginal spatial intensity obtained by integrating the joint intensity over the mark space.

To assess intensity estimation performance, let $\lambda_0(s)$ denote the true first-order intensity function of the generating process on the spatial domain W_2 , and let $\hat{\lambda}(s)$ denote a fitted spatial intensity obtained from one of the competing models. The overall discrepancy between $\hat{\lambda}(s)$ and $\lambda_0(s)$ is measured through the Mean Integrated Squared Error (MISE)

$$\text{MISE}_{\hat{\lambda}} = \mathbb{E} \left[\int_{W_2} \left(\hat{\lambda}(s) - \lambda_0(s) \right)^2 ds \right] = \int_{W_2} \text{Var} \left(\hat{\lambda}(s) \right) ds + \int_{W_2} \text{Bias}^2(\hat{\lambda}(s)) ds,$$

where $\text{Bias}(\hat{\lambda}(s)) = \mathbb{E}[\hat{\lambda}(s)] - \lambda_0(s)$. In practice, this quantity is approximated over the simulated replications as

$$\text{MISE}(\hat{\lambda}) = \frac{1}{n_{\text{sim}}} \sum_{j=1}^{n_{\text{sim}}} \int_{W_2} (\hat{\lambda}(s; A_j) - \lambda_0(s))^2 ds, \quad (16)$$

where $A_1, \dots, A_{n_{\text{sim}}}$ denote the simulated marked point patterns. Hence, the MISE summarizes the average integrated discrepancy between the fitted and true spatial intensities over repeated samples.

In addition, when model parameters are directly estimated, we assess their accuracy through the Root Mean Squared Error (RMSE). Let θ denote the true parameter of interest and $\hat{\theta}$ its estimator. The RMSE is computed as

$$\text{RMSE}(\hat{\theta}) = \left[\frac{1}{n_{\text{sim}}} \sum_{j=1}^{n_{\text{sim}}} (\hat{\theta}_j - \theta)^2 \right]^{1/2}, \quad (17)$$

where $\hat{\theta}_j$ is its estimate in the j th replication. While the MISE evaluates the quality of first-order intensity reconstruction, the RMSE provides complementary information on the stability and accuracy of parameter estimates.

4.1. Marked inhomogeneous Poisson processes

We begin the simulation study with marked inhomogeneous Poisson processes, since they provide the setting most directly aligned with the framework introduced in Section 3. In this subsection, the aim is to evaluate the behaviour of the baseline Poisson model, PMP, and SMP under the three distinct mark-generating mechanisms described in Section 4. Focussing on marked Poisson point processes, we consider four different scenarios. For Scenarios 1 and 2, the marked process is specified on the product domain $W_2 \times \mathcal{M}$, with $\mathcal{M} = [0, 1]$, through the joint intensity

$$\lambda_0(x, y, m) = \exp(\theta_0 + \theta_1 x + \theta_2 y + \theta_3 m), \quad \theta_1 = 1, \quad \theta_2 = 1. \quad (18)$$

Thus, in these scenarios, the observed marked pattern consists of pairs (s_i, m_i) generated according to a joint intensity on $W_2 \times \mathcal{M}$.

Scenarios 3 and 4 follow a different construction. A spatial mark field is first specified over the whole observation window as

$$M(x, y) = \exp(b_0 + b_1 x + b_2 y), \quad b_0 = 3, \quad b_1 = 0.5, \quad b_2 = 0.5,$$

where b_0 controls the overall scale of the mark field, whereas b_1 and b_2 determine its spatial gradient. Conditional on this field, the spatial point pattern is generated from an inhomogeneous Poisson process with intensity

$$\lambda_0(x, y | M) = \exp\{\theta_0 + \theta_1 x + \theta_2 y + \theta_3 M(x, y)\}, \quad \theta_1 = 1, \quad \theta_2 = 1. \quad (19)$$

For each scenario and each target value of $\mathbb{E}[N]$, the intercept parameter θ_0 is calibrated so that the expected number of points in the observation window matches the desired sample size. More specifically, the four scenarios are defined as follows:

- (1) Random labelling: the process is generated according to (18) with $\theta_3 = 0$. Since the joint intensity does not depend on m , the mark is uniformly distributed on $\mathcal{M} = [0, 1]$, independently of the spatial component, and carries no information about the spatial intensity.
- (2) Mark as an additional modelling dimension: the marked process is generated on $W_2 \times \mathcal{M}$ according to (18) with $\theta_3 = 1$. The mark is therefore an explicit argument of the joint intensity, consistent with the PMP interpretation.
- (3) Spatially structured mark, positive contribution: the field $M(x, y)$ is specified before the point pattern is generated and enters the conditional spatial intensity in (19) with $\theta_3 = 1$.
- (4) Spatially structured mark, negative contribution: the field $M(x, y)$ is specified before the point pattern is generated and enters the conditional spatial intensity in (19) with $\theta_3 = -1$.

Since the possible simulation designs are numerous, both in terms of mark-generating mechanisms and of the functional specifications adopted within each mechanism, the set of scenarios considered here is not intended to be exhaustive. Rather, we focus on a small number of representative cases that reflect the main modelling situations of interest for the proposed methodology. Figure 1 shows one realization of the marked Poisson point processes for each of the scenarios described above. The first row of Figure 1 reports the patterns in the W_2 space, with points coloured according to the mark value, while the second row shows the corresponding three-dimensional representations in the $W_2 \times \mathcal{M}$ domain in which the third axis corresponds to the mark. The three-dimensional plots are especially useful for distinguishing Scenario 1 from Scenario 2. In both cases, the two-dimensional spatial pattern may look broadly similar, but the three-dimensional representation makes clear that in Scenario 2 the mark contributes directly to the intensity

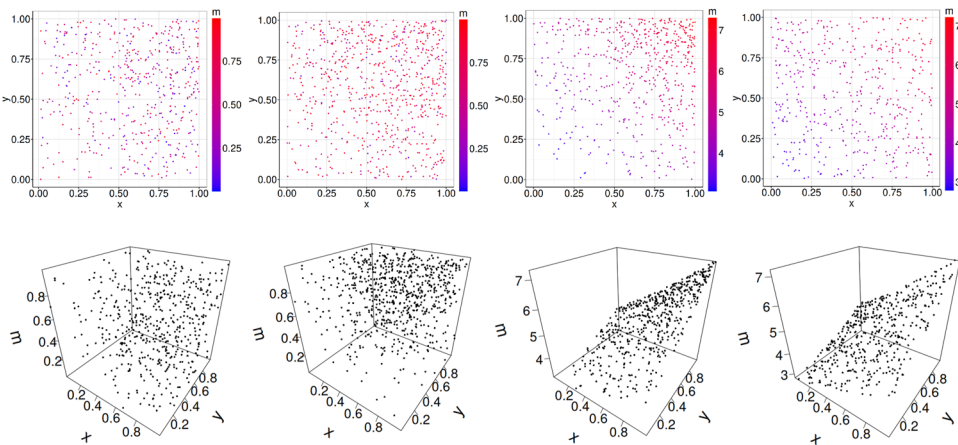


Figure 1. Simulated marked Poisson point patterns under the four considered scenarios. The first row reports the two-dimensional projections of the spatial point patterns, where points are coloured according to the mark value m . The second row reports the corresponding three-dimensional representations, where the third axis corresponds to the mark m . Each column refers to the same scenario.

Table 1. Results of the simulation study for the Poisson point processes scenarios: MISE and its decomposition over 250 simulations. All values are reported in scale $\times 10^3$.

Scenario	θ_3	$g(m)$	$\mathbb{E}[N]$	MISE $_{\hat{\lambda}}$			Var($\hat{\lambda}(s)$)			Bias $^2(\hat{\lambda}(s))$		
				Pois	PMP	SMP	Pois	PMP	SMP	Pois	PMP	SMP
1	0	U(0,1)	100	0.35	0.51	0.37	0.35	0.39	0.36	0.01	0.12	0.01
			500	1.93	2.60	2.05	1.91	2.11	2.04	0.01	0.49	0.01
			1000	3.32	5.04	3.54	3.29	2.66	3.51	0.03	2.38	0.03
2	1	U(0,1)	100	0.43	0.93	0.47	0.35	0.78	0.43	0.08	0.16	0.03
			500	3.81	3.83	4.18	1.78	3.42	2.45	2.03	0.41	1.72
			1000	11.34	8.53	12.29	3.77	3.06	5.40	7.58	5.48	6.90
3	1	exp(x, y)	100	0.53	0.58	0.54	0.51	0.46	0.50	0.02	0.12	0.04
			500	2.80	3.65	2.45	2.28	2.30	2.36	0.52	1.34	0.09
			1000	7.48	15.46	5.29	4.88	5.98	5.05	2.60	9.48	0.23
4	-1	exp(x, y)	100	0.34	0.53	0.36	0.32	0.35	0.34	0.02	0.18	0.02
			500	1.76	3.97	1.47	1.38	1.63	1.43	0.38	2.34	0.04
			1000	4.37	15.23	2.94	2.89	2.84	2.92	1.48	12.39	0.03

and effectively acts as an additional modelling dimension. By contrast, in Scenarios 3 and 4, the mark exhibits a clear spatial gradient induced by its dependence on the coordinates.

For each simulated pattern, we fit the baseline Poisson model, the PMP model, and the SMP model. In the PMP case, comparison with the other models is performed through the marginal spatial intensity obtained from the fitted marked specification. In the SMP case, the mark-related surface is reconstructed by kernel smoothing, with the bandwidth selected separately for each simulated pattern according to the criterion proposed by Cronie and Van Lieshout [36]. This data-driven choice allows the amount of smoothing to adapt to the observed pattern, rather than imposing a common smoothing parameter across all scenarios and sample sizes. The resulting comparison, based on MISE and its decomposition into variance and squared bias over 250 simulations, is summarized in Table 1 for the four marked Poisson point process scenarios.

To complement this comparison, we also report in Table 2 the parameter recovery of Scenarios 2 and 3, the two cases in which the role of the mark is most informative for understanding the behaviour of the competing estimators.

As expected, scenario 1 demonstrates that when the mark has no effect on the intensity of the process, the proposed models do not offer any improvement in terms of MISE compared to the standard Poisson process model. Scenario 2 corresponds to the setting most closely aligned with the PMP specification, because the marked process is generated on the product space and the mark enters the joint intensity as an additional modelling dimension independently of the spatial coordinates. Although the intensity comparisons are carried out on the spatial ground domain, the PMP model is the only approach among those considered that explicitly represents the joint location–mark intensity. As the expected number of points increases, the PMP model becomes increasingly competitive and, for the largest sample size, provides the lowest MISE among the three approaches. The corresponding reduction in MISE is mainly associated with a lower squared-bias component relative to the baseline Poisson model. Table 2 complements the spatial-intensity comparison by showing that the PMP model recovers the mark coefficient θ_3 reasonably well, together with the parameters of the joint intensity specification. By contrast, the baseline Poisson and SMP models are fitted on the spatial ground domain and do not represent the

Table 2. Parameter estimates in Scenarios 2 and 3 across sample sizes. For each parameter, the table reports the true value together with the mean estimate, empirical bias, and RMSE over 250 simulation replicates.

Scenario	$\mathbb{E}[N]$	Parameter	True	Pois			PMP			SMP		
				Mean	Bias	RMSE	Mean	Bias	RMSE	Mean	Bias	RMSE
2	100	θ_0	3.0	3.494	—	—	3.235	0.254	0.450	2.570	—	—
		θ_1	1.0	1.007	0.007	0.368	0.947	−0.053	0.331	0.986	−0.014	0.368
		θ_2	1.0	1.014	0.014	0.365	0.864	−0.136	0.377	0.956	−0.044	0.380
		θ_3	1.0	—	—	—	0.898	−0.102	0.343	—	—	—
	500	θ_0	4.6	5.121	—	—	4.773	0.173	0.237	4.126	—	—
		θ_1	1.0	1.011	0.011	0.167	0.912	−0.088	0.180	1.011	0.011	0.173
		θ_2	1.0	1.017	0.017	0.152	0.932	−0.068	0.155	1.011	0.011	0.159
		θ_3	1.0	—	—	—	0.909	−0.091	0.177	—	—	—
	1000	θ_0	5.3	5.826	—	—	5.494	0.194	0.223	4.824	—	—
		θ_1	1.0	1.014	0.014	0.112	0.914	−0.086	0.137	1.018	0.018	0.115
		θ_2	1.0	1.010	0.010	0.120	0.918	−0.082	0.139	1.008	0.008	0.125
		θ_3	1.0	—	—	—	0.907	−0.093	0.140	—	—	—
3	100	θ_0	1.7	2.420	0.720	0.950	2.280	0.580	0.810	1.890	0.190	0.460
		θ_1	1.0	1.910	0.910	1.120	1.770	0.770	0.980	1.080	0.080	0.390
		θ_2	1.0	1.930	0.930	1.140	1.790	0.790	1.000	1.070	0.070	0.400
		θ_3	1.0	—	—	—	0.740	−0.260	0.350	—	—	—
	500	θ_0	3.25	3.980	0.730	0.810	3.860	0.610	0.700	3.340	0.090	0.210
		θ_1	1.0	1.930	0.930	1.000	1.810	0.810	0.890	1.040	0.040	0.190
		θ_2	1.0	1.940	0.940	1.010	1.820	0.820	0.900	1.030	0.030	0.180
		θ_3	1.0	—	—	—	0.770	−0.230	0.150	—	—	—
	1000	θ_0	4.0	4.741	0.741	0.749	4.663	0.663	0.682	4.060	0.060	0.140
		θ_1	1.0	1.925	0.925	0.932	1.822	0.822	0.831	1.020	0.020	0.090
		θ_2	1.0	1.937	0.937	0.944	1.836	0.836	0.843	1.018	0.018	0.085
		θ_3	1.0	—	—	—	0.785	−0.215	0.070	—	—	—

Note. In Scenario 2, the Poisson and SMP models are fitted on the spatial ground domain. Their intercepts are therefore associated with the marginal spatial intensity and are not directly comparable with the joint-model PMP intercept. Accordingly, bias and RMSE are not reported for these intercepts.

joint location–mark intensity. Their intercepts therefore refer to the marginal spatial intensity obtained after integrating over the mark space and are not directly comparable with the intercept θ_0 of the joint PMP model. Thus, Scenario 2 illustrates that, when the data-generating mechanism is genuinely defined on the product space, explicitly modelling the mark allows recovery of the marked first-order structure, while the MISE results describe the corresponding finite-sample performance for the marginal spatial intensity.

The picture changes in Scenarios 3 and 4, where the mark is a function of the spatial coordinates and therefore behaves more like a spatially structured signal than like an independent additional modelling dimension. In these settings, the PMP approach performs poorly, especially in terms of squared bias, whereas the SMP approach is either comparable to or clearly better than the baseline Poisson model. The difference is particularly evident when the expected number of points is at least 500, and becomes strongest in Scenario 4, where the SMP model achieves the lowest MISE by a substantial margin. The parameter recovery results in Scenario 3 support this interpretation. Both the baseline Poisson model and the PMP model show substantial bias in the estimates of θ_1 and θ_2 , whereas the SMP model recovers these parameters much more accurately. This suggests that, when the mark is strongly aligned with the spatial coordinates, the main difficulty is not only estimating the overall intensity level, but also allocating the contribution correctly between the spatial covariates and the mark-related component. In such cases, the semi-parametric

structure of SMP is better suited to the data-generating mechanism than the fully parametric PMP specification. However, this similarity in aggregate performance should be interpreted carefully. When the mark is spatially structured, the Poisson model can still approximate the overall intensity surface reasonably well by absorbing the mark-related contribution into the spatial coefficients, but this comes at the cost of a distorted interpretation of the underlying components. The SMP approach, by construction, is better able to separate the baseline spatial effect from the contribution associated with the reconstructed mark-related signal.

4.2. Marked Hotspot Poisson processes

To simulate a Hotspot process, we first construct a latent spatial surface on the unit square by combining a deterministic hotspot component with a Gaussian random field. The deterministic component is centred at $(0.5, 0.5)$, so that the main peak of the surface is located at the centre of the observation window, and is specified through a Gaussian-shaped hotspot with amplitude 2 and standard deviation 0.1. Here, the amplitude controls the height of the hotspot peak, whereas the standard deviation determines how concentrated or spread out the hotspot is around its centre. On top of this deterministic structure, we superimpose a Gaussian random field defined on a regular grid over W_2 . The Gaussian field has an exponential covariance function, a marginal standard deviation of 0.8, and a range parameter of 0.15. In this setting, the marginal standard deviation controls the overall magnitude of the random fluctuations around the deterministic hotspot surface, while the range parameter governs the spatial scale of dependence, with larger values corresponding to smoother and more spatially persistent local deviations. The role of this Gaussian field is therefore to introduce random spatial heterogeneity around the main hotspot pattern. The resulting latent field is then transformed through a logistic map and rescaled to the interval $[0, 1]$, providing a bounded mark surface $M(s)$. This transformation ensures that the simulated mark remains within a fixed range while preserving the spatial structure induced by the combination of the deterministic hotspot and the Gaussian random field. In this way, the mark is generated as a spatially structured surface defined over the whole observation window before the point pattern is simulated. This simulation design is particularly relevant for the methodological distinction introduced in Section 3. Specifically, these scenarios are naturally closer to the interpretation underlying the SMP approach, where the mark is treated as a spatially structured signal attached to the ground process, rather than to the PMP view in which the mark acts as an explicit additional modelling dimension independent of space.

Conditional on this realized mark surface M , the spatial point pattern is generated as an inhomogeneous Poisson process with conditional intensity $\lambda_0(x, y \mid M)$ as in (19). More specifically, candidate points are first generated from a homogeneous Poisson process with rate equal to the maximum of $\lambda_0(x, y \mid M)$ over the simulation grid. A candidate point at location s is retained with probability proportional to $\lambda_0(s \mid M)$, and its observed mark is then defined as $m_i = M(s_i)$. Thus, the observed marks are values of the surface used to generate the conditional spatial intensity, rather than values assigned after the spatial point pattern has been generated. We consider two Hotspot scenarios:

- (5) the intensity depends on the mark only, with $\theta_1 = \theta_2 = 0$ and $\theta_3 = 2$.

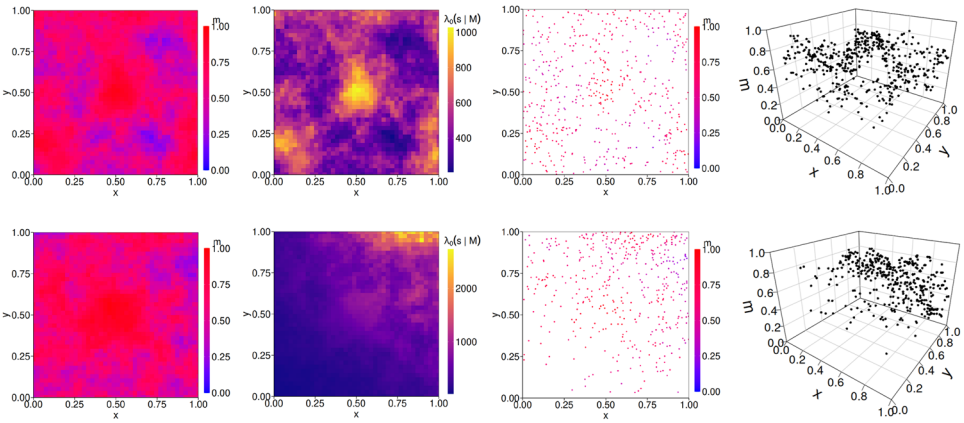


Figure 2. Illustration of the two Hotspot scenarios considered in the simulation study. For each scenario, the first panel reports the realized Gaussian-field mark surface $M(s)$, the second panel shows the corresponding conditional spatial intensity $\lambda_0(s | M)$ used to generate the point pattern, the third panel displays the resulting two-dimensional marked point pattern with points coloured according to their mark values, and the fourth panel shows the corresponding three-dimensional representation, where the third axis corresponds to the observed mark m . *First row: Scenario 5. Second row: Scenario 6.*

Table 3. Results of the simulation study for the Hotspot Poisson point processes scenarios: MISE and its decomposition over 250 simulations. All values are reported in scale $\times 10^3$.

Scenario	$\mathbb{E}[N]$	MISE $_{\hat{\lambda}}$			Var($\hat{\lambda}(s)$)			Bias $^2(\hat{\lambda}(s))$		
		Pois	PMP	SMP	Pois	PMP	SMP	Pois	PMP	SMP
5	100	1.25	5.40	0.82	0.89	1.38	0.63	0.36	4.01	0.18
	500	29.26	64.70	15.28	20.22	18.53	11.82	9.04	46.17	3.47
	1000	118.29	253.90	58.55	82.47	69.80	45.42	35.83	184.11	13.12
6	100	1.44	6.19	1.11	1.11	2.03	0.92	0.33	4.16	0.18
	500	29.74	76.74	17.47	21.34	27.00	14.10	8.40	49.74	3.37
	1000	111.78	280.01	61.44	79.37	87.93	49.32	32.41	192.08	12.12

- (6) the intensity depends both on the spatial coordinates and on the mark, with $\theta_1 = \theta_2 = 1$ and $\theta_3 = 2$.

In both scenarios, the intercept θ_0 is chosen so as to obtain the desired expected number of points, i.e., $\mathbb{E}[N] \in \{100, 500, 1000\}$. Specifically, for Scenario 5 we set $\theta_0 \in \{3.5, 5.1, 5.8\}$, while for Scenario 6 we set $\theta_0 \in \{2.4, 4.0, 4.7\}$. The corresponding realisations and conditional intensities are displayed in Figure 2, while Table 3 summarizes the MISE results and Table 4 reports parameter recovery.

In both Hotspot scenarios, the SMP model provides the lowest MISE across all sample sizes. The improvement over the baseline Poisson model is already visible when $\mathbb{E}[N] = 100$, and becomes increasingly pronounced as the expected number of points grows. This gain is driven by a substantial reduction in both variance and squared bias, indicating that the SMP approach is better able to reconstruct the true spatial intensity surface when the mark acts as a spatially structured component of the process.

By contrast, the PMP approach performs systematically worse than both the baseline Poisson model and the SMP model. This is particularly evident in the bias component of

Table 4. Parameter estimates in Hotspot Scenarios 5 and 6 across sample sizes. For each parameter, the table reports the true value together with the mean estimate, empirical bias, and RMSE over 250 simulation replicates.

Scenario	$\mathbb{E}[N]$	Parameter	True	Pois			PMP			SMP		
				Mean	Bias	RMSE	Mean	Bias	RMSE	Mean	Bias	RMSE
5	100	θ_0	3.485	4.607	1.122	1.122	4.804	1.320	1.415	3.655	0.171	0.223
		θ_1	0.0	—	—	—	—	—	—	—	—	—
		θ_2	0.0	—	—	—	—	—	—	—	—	—
		θ_3	2.0	—	—	—	0.419	−1.581	1.807	—	—	—
	500	θ_0	5.101	6.215	1.114	1.114	6.193	1.092	1.147	5.238	0.137	0.176
		θ_1	0.0	—	—	—	—	—	—	—	—	—
		θ_2	0.0	—	—	—	—	—	—	—	—	—
		θ_3	2.0	—	—	—	0.588	−1.412	1.571	—	—	—
	1000	θ_0	5.793	6.910	1.117	1.117	6.796	1.002	1.057	5.933	0.139	0.177
		θ_1	0.0	—	—	—	—	—	—	—	—	—
		θ_2	0.0	—	—	—	—	—	—	—	—	—
		θ_3	2.0	—	—	—	0.744	−1.256	1.439	—	—	—
6	100	θ_0	2.405	3.519	1.114	1.180	3.812	1.407	1.530	2.601	0.196	0.410
		θ_1	1.0	0.937	−0.063	0.462	0.879	−0.121	0.428	0.924	−0.076	0.418
		θ_2	1.0	1.011	−0.011	0.444	0.876	−0.124	0.431	0.969	−0.031	0.375
		θ_3	2.0	—	—	—	0.438	−1.562	1.819	—	—	—
	500	θ_0	4.022	5.216	1.195	1.224	5.272	1.251	1.347	4.202	0.180	0.272
		θ_1	1.0	0.911	−0.089	0.343	0.839	−0.161	0.349	0.950	−0.050	0.231
		θ_2	1.0	0.919	−0.081	0.335	0.867	−0.133	0.337	0.950	−0.050	0.233
		θ_3	2.0	—	—	—	0.558	−1.442	1.652	—	—	—
	1000	θ_0	4.712	5.891	1.179	1.208	5.886	1.175	1.261	4.876	0.164	0.246
		θ_1	1.0	0.935	−0.065	0.310	0.867	−0.133	0.313	0.972	−0.028	0.186
		θ_2	1.0	0.939	−0.061	0.313	0.858	−0.142	0.321	0.967	−0.033	0.195
		θ_3	2.0	—	—	—	0.685	−1.315	1.501	—	—	—

the MISE, which becomes very large for the PMP model as the expected number of points increases. The parameter recovery results help explain this behaviour. In both scenarios, the PMP model strongly underestimates the mark coefficient θ_3 , while at the same time overestimating the intercept. In Scenario 6, where the spatial coordinates also enter the intensity directly, the PMP estimates of θ_1 and θ_2 are also more biased than those obtained under SMP. These results indicate that, when the mark is generated as a spatial surface and is therefore strongly aligned with the spatial domain, the fully parametric marked specification is not able to allocate the different contributions properly between the mark term and the spatial terms.

The baseline Poisson model remains competitive relative to PMP, but it is consistently outperformed by SMP. This suggests that, although the Poisson model can partially absorb the effect of the mark through its spatial component, it cannot recover the true intensity surface as accurately as a model that explicitly incorporates the mark-related spatial signal. In this respect, the Hotspot scenarios provide strong support for the modelling rationale of the SMP approach. Overall, the Hotspot simulation study confirms that the relative performance of the competing methods depends on the way in which the mark enters the data-generating mechanism. Under the present construction, which is specifically designed so that the mark surface contributes directly to the process intensity, the SMP model is the most appropriate specification and provides the most accurate first-order intensity estimates.

4.3. Marked LGCP processes

If the Poisson and Hotspot scenarios were primarily aimed at assessing the behaviour of the proposed approaches under first-order Poisson intensity, the LGCP scenarios allow us to evaluate their performance in a setting where the observed process exhibits genuine extra-Poisson clustering. The purpose of this part of the simulation study is therefore not to compare the proposed methods with a fully specified Cox-process fit, but rather to assess whether mark-informed first-order intensity models can still provide competitive or improved spatial intensity estimates when the underlying process is clustered with dependencies among points.

For the LGCP scenarios, clustering is introduced through a latent Gaussian random field. Let G denote a zero-mean Gaussian field with an exponential covariance function $C(s, s') = \sigma^2 \exp(-\|s - s'\|/\alpha)$, where σ^2 controls the variability of the latent field and α determines the spatial range of dependence. Conditional on a realization of G , the point process is Poisson with stochastic intensity given by the exponential of a deterministic trend plus the latent field.

For the random-labelling benchmark and the spatial-field scenarios, the conditional ground intensity is specified as

$$\Lambda_0(x, y \mid G, M) = \exp \{ \theta_0 + \theta_1 x + \theta_2 y + \theta_3 M(x, y) + G(x, y) \}, \quad (20)$$

where $M(x, y)$ denotes the mark-related component. In the random-labelling benchmark, $M(x, y)$ is omitted by setting $\theta_3 = 0$. In the spatial-field scenarios, $M(x, y)$ is generated over the whole observation window before the LGCP realization is simulated. Conditional on the realized fields G and M , points are then generated from (20), and the observed mark at a retained point s_i is $m_i = M(s_i)$.

Scenario 8 follows instead a product-space LGCP construction. Let $\tilde{G}(x, y, m)$ denote a zero-mean Gaussian field on $W_2 \times \mathcal{M}$, with $\mathcal{M} = [0, 1]$, and exponential covariance defined with respect to the Euclidean distance on the three-dimensional product domain. Conditional on $\tilde{G}$, the marked process is Poisson on $W_2 \times \mathcal{M}$ with joint intensity

$$\Lambda_0(x, y, m \mid \tilde{G}) = \exp \{ \theta_0 + \theta_1 x + \theta_2 y + \theta_3 m + \tilde{G}(x, y, m) \}. \quad (21)$$

Thus, in this scenario, locations and marks are generated jointly on the product domain, and the spatial intensity used for comparison is obtained by marginalizing (21) over $\mathcal{M}$.

Throughout the LGCP scenarios, $\theta_1 = \theta_2 = 1$, and the intercept θ_0 is calibrated separately for each target sample size so that the expected number of points is approximately $\mathbb{E}[N] \in \{100, 500, 1000\}$. The coefficient θ_3 controls the contribution of the continuous mark, or of the spatial mark field, to the deterministic component of the conditional log-intensity. Specifically, we consider the following five scenarios:

- (7) *Random labelling*: a spatial LGCP is generated with $\theta_3 = 0$. Conditional on the generated spatial pattern, the marks are sampled independently from a Uniform distribution on $[0, 1]$.
- (8) *Mark as an additional modelling dimension*: the marked LGCP is generated jointly on $W_2 \times \mathcal{M}$ according to (21), with $\theta_3 = 2$.
- (9) *Spatially structured mark from a Gaussian-field surface*: a spatially structured Gaussian-field mark surface $M(x, y)$ is first generated over W_2 . Conditional on this

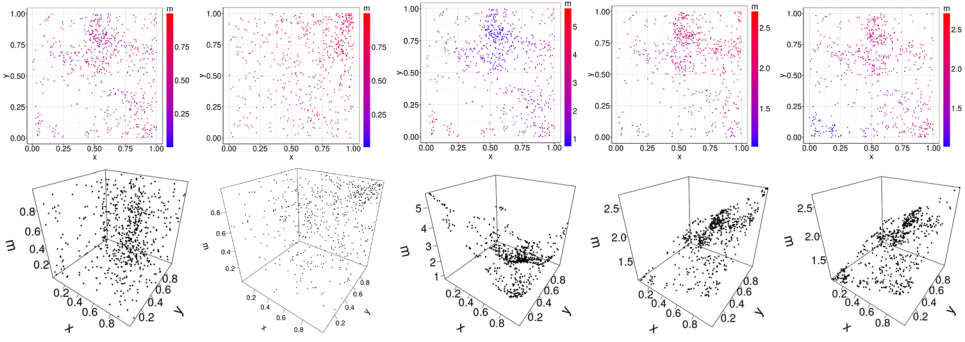


Figure 3. Illustration of the simulated marked LGCP point patterns under scenarios 7–11. The first row reports the two-dimensional spatial patterns, where points are coloured according to the mark value m . The second row reports the corresponding three-dimensional representations, where the third axis corresponds to the mark m . Each column refers to the same simulation scenario: from left to right, scenarios 7, 8, 9, 10 and 11.

Table 5. Results of the simulation study for the LGCP processes scenarios: MISE and its decomposition over 250 simulations. All values are reported in scale $\times 10^3$.

Scenario	θ_3	$g(m)$	$\mathbb{E}[N]$	MISE $_{\hat{\lambda}}$			Var($\hat{\lambda}(s)$)			Bias $^2(\hat{\lambda}(s))$		
				Pois	PMP	SMP	Pois	PMP	SMP	Pois	PMP	SMP
7	0	$U(0,1)$	100	2.85	2.99	3.24	2.83	2.89	3.21	0.02	0.10	0.03
			500	47.42	48.67	50.22	46.98	47.40	49.63	0.45	1.26	0.59
			1000	199.39	207.75	205.30	197.74	201.56	203.43	1.65	6.18	1.88
8	2	$U(0,1)$	100	1.72	1.33	1.74	0.55	0.47	0.63	1.17	0.86	1.10
			500	14.33	11.58	15.77	1.66	2.04	2.27	12.67	9.54	13.50
			1000	69.04	53.11	77.51	3.71	4.42	5.69	65.33	48.70	71.82
9	2	GRF	100	3.12	3.38	3.44	0.87	1.03	1.36	2.25	2.35	2.08
			500	106.75	98.52	95.53	15.79	16.53	16.72	90.96	81.99	78.81
			1000	619.83	695.18	584.86	409.84	427.20	419.18	209.98	267.98	165.68
10	1	$\exp(x,y)$	100	4.81	4.76	4.83	4.80	4.63	4.80	0.01	0.14	0.03
			500	81.50	87.53	81.03	81.29	84.98	80.74	0.21	2.55	0.29
			1000	358.97	397.55	357.02	357.74	381.95	355.18	1.23	15.60	1.84
11	-1	$\exp(x,y)$	100	2.38	2.97	2.52	2.33	2.76	2.48	0.05	0.21	0.03
			500	35.33	44.38	35.47	34.43	41.02	35.10	0.90	3.36	0.37
			1000	137.88	172.69	137.38	134.29	158.24	135.92	3.59	14.46	1.46

surface and on the clustering field G , the LGCP is generated according to (20) with $\theta_3 = 2$.

- (10) *Spatially structured mark as an exponential function of the coordinates, positive contribution:* the mark field is specified as $M(x,y) = \exp(b_0 + b_1x + b_2y)$, $b_0 = 3$, $b_1 = 0.5$, $b_2 = 0.5$, and enters (20) with $\theta_3 = 1$.
- (11) *Spatially structured mark as an exponential function of the coordinates, negative contribution:* the same mark field $M(x,y)$ is specified before simulating the LGCP, but it enters (20) with $\theta_3 = -1$.

Figure 3 provides an illustration of the five considered scenarios, with the same structure as Figures 1 and 2, while Table 5 reports the MISE results for this subsection.

Scenario 7 provides the natural benchmark. When the mark has no effect on the intensity and is generated independently of the spatial structure, the baseline Poisson model

remains the best-performing approach, while PMP and SMP do not provide any systematic improvement. This is the expected outcome, since in this case the mark does not contain information that is useful for first-order intensity estimation. Scenario 8 is instead the case that is most favourable to the PMP specification. Here, the mark acts as an explicit additional dimension of the process, and the data-generating mechanism is therefore fully coherent with the product-domain interpretation underlying the PMP model. The results reflect this clearly: PMP provides the lowest MISE for all three sample sizes, and the gain relative to the baseline Poisson model becomes substantial as $\mathbb{E}[N]$ increases. The improvement is mainly driven by a marked reduction in squared bias, which indicates that explicitly modelling the mark in the joint intensity helps reconstruct the correct marginal spatial intensity even in the presence of LGCP-type clustering.

Scenario 9 is conceptually different. The mark is generated from a Gaussian-field surface and is therefore spatially structured over the observation window. In this setting, the relative performance of the competing methods depends on the sample size. For $\mathbb{E}[N] = 100$, the baseline Poisson model still provides the lowest MISE, suggesting that the additional mark-related components may introduce extra variability when the number of observed points is limited. However, for $\mathbb{E}[N] = 500$ and $\mathbb{E}[N] = 1000$, SMP becomes the best-performing method. This improvement is mainly associated with a reduction in squared bias, which is consistent with the modelling rationale of SMP: when the mark behaves as a spatially structured signal over the observation window, reconstructing it through a mark-related surface can be more effective than treating it as an explicit additional modelling dimension. The PMP model also improves over the baseline Poisson model for $\mathbb{E}[N] = 500$, but this advantage does not persist for the largest sample size, where the semi-parametric reconstruction provides the most accurate first-order intensity estimate.

The same general pattern emerges in Scenarios 10 and 11, where the mark is an exponential function of the coordinates and then contributes to the intensity with positive and negative signs, respectively. In these cases, PMP is again the worst-performing method, with larger MISE values than both the baseline Poisson model and SMP. This behaviour is consistent with what was already observed in the Poisson scenarios: when the continuous mark is strongly associated with the spatial coordinates, the fully parametric marked specification struggles to separate the contribution of the mark from that of the spatial covariates. By contrast, SMP remains competitive with the baseline Poisson model. In Scenario 10, SMP slightly outperforms the Poisson model for medium and large sample sizes, while in Scenario 11 the two approaches are essentially comparable, with SMP becoming marginally better only for $\mathbb{E}[N] = 1000$. These results suggest that, under clustered processes with spatially structured marks, SMP can incorporate the additional mark information without the instability that affects PMP, although the gains over the baseline model may be modest when the Poisson fit already captures much of the first-order trend. The results on parameter recovery for the LGCP scenarios follow the same pattern observed in the Poisson and Hotspot settings. In particular, in scenarios where the mark contributes to the first-order intensity, the baseline Poisson model can still provide MISE values comparable to those of the mark-informed approaches, but this occurs at the cost of more biased coefficient estimates. By contrast, the PMP and SMP specifications, when aligned with the underlying role of the mark, tend to recover the model components more accurately, leading to lower parameter bias and RMSE. Hence, even when the gain in terms

of fitted spatial intensity is limited, incorporating the mark remains important for a more correct reconstruction of the first-order structure of the process.

5. Applications to real data

In this section, we illustrate the proposed methodology through two real-data applications based on the *nbfires* and *bronzefilter* datasets available in the *spatstat* package [37].

The goal is to compare the fitted spatial intensities obtained from the baseline Poisson model, from the PMP model after marginalization over the mark space, and from the SMP model defined directly on the ground space. In addition, we inspect the fitted relative intensity ratios introduced in Section 3, namely $\mathcal{R}_P(s)$ and $\mathcal{R}_S(s)$, in order to assess how the inclusion of the real-valued mark modifies the fitted spatial intensity relative to the baseline Poisson specification.

The two applications are chosen to reflect two different empirical situations already introduced in the simulation study. In the first case, the mark does not appear to be strongly aligned with the spatial coordinates, so that the empirical setting is closer to the interpretation underlying PMP. In the second case, the mark is strongly associated with the spatial domain, making the data more similar to the scenarios in which SMP proved more appropriate.

5.1. Fires data

In the first application, we consider wildfire data provided by the New Brunswick Department of Natural Resources. Specifically, we use the point pattern corresponding to all recorded fires within their jurisdiction for the year 1992, extracted from a larger dataset covering the period from 1987 to 2003 [38]. The year 1992 was selected for illustrative purposes, since it provides a relatively large number of recorded fires within the available time series. The dataset consists of 268 events observed in a rectangular window covering the central region of New Brunswick. Each point represents the spatial location of a fire, and the associated real-valued mark is the fire size, measured in hectares. A total of 59 fires have been recorded with a size equal to 0.0 hectares. Since the burned area is rounded to the nearest 0.1 hectare, these values likely correspond to very small fires, recorded as zero because of rounding. On the left panel of Figure 4, the marked point pattern is shown. As a preliminary descriptive step, we also compute a kernel-based estimate of the spatial intensity and a kernel-smoothed spatial surface of the marks [33–35], using bandwidth selection based on the Cronie and van Lieshout criterion [36]. These two exploratory surfaces are shown in the central and right panels of Figure 4. They suggest a negative association between local fire concentration and observed fire size: areas with higher estimated point intensity tend to be associated with lower mark values, which indicates that the mark may contain relevant information for the first-order intensity estimate.

Because the fire-size distribution is highly skewed, with many zero values and a few very large observations, we apply the transformation $m^* = \log(m + 1)$, which stabilizes the scale and also accommodates zeros. The transformed marks are then rescaled to the interval $[0, 1]$.

For this application, we fit a baseline Poisson model on W_2

$$\hat{v}(x, y) = \exp(\hat{\theta}_0 + \hat{t}_p(x, y)),$$

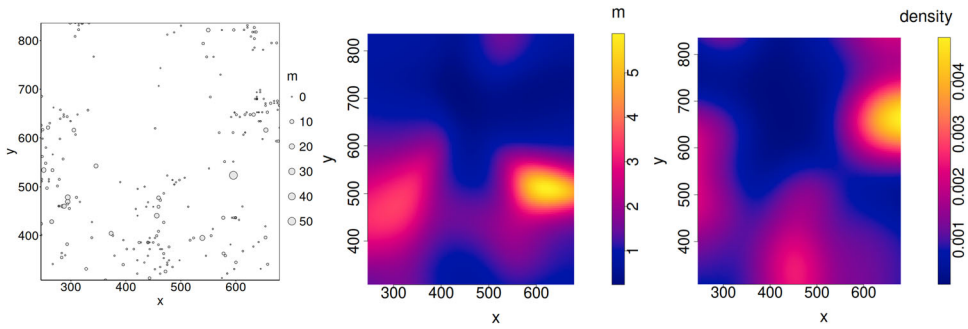


Figure 4. Left-panel: Marked fires point pattern. Central-panel: Kernel-smoothed spatial surface of the mark values. Right-panel: Kernel estimate of the spatial intensity.

Table 6. Parametric coefficient estimates for the Poisson, PMP, and SMP models fitted to the fires data.

Model	Term	Estimate	Std. Error	z value	Pr(> z)
Pois	Intercept	−7.193	0.101	−71.29	$< 2 \cdot 10^{-16}***$
PMP	Intercept	−6.651	0.108	−61.26	$< 2 \cdot 10^{-16}***$
	m^*	−1.707	0.109	−15.68	$< 2 \cdot 10^{-16}***$
SMP	Intercept	−8.228	0.101	−81.28	$< 2 \cdot 10^{-16}***$

Table 7. Significance of smooth spatial terms for the Poisson, PMP, and SMP models fitted to the fires data.

Model	Term	edf	Ref.df	χ^2	Pr(> χ^2)
Poisson	$t_p(x, y)$	26.61	28.52	322.6	$< 2 \cdot 10^{-16}***$
PMP	$t_p(x, y)$	26.25	28.48	205.0	$< 2 \cdot 10^{-16}***$
SMP	$t_p(x, y)$	26.33	28.47	240.3	$< 2 \cdot 10^{-16}***$

a PMP model on $W_2 \times \mathcal{M}$

$$\hat{\lambda}_P(x, y, m) = \exp(\hat{\theta}_0 + \hat{t}_p(x, y) + \hat{\theta}_3 m^*),$$

and an SMP model on W_2

$$\hat{\lambda}_S(x, y) = \exp(\hat{\theta}_0 + \hat{t}_p(x, y) + \hat{\kappa}_c(x, y)),$$

where $\hat{t}_p(x, y)$ is a thin plate regression spline surface [39] with basis dimension equal to 30, and $\hat{\kappa}_c(x, y)$ denotes the centred reconstructed mark-related surface used in the SMP specification. For graphical comparison with the baseline Poisson fit, the PMP model is reported through its marginal spatial intensity $\hat{\lambda}_P^{\text{marg}}(x, y)$, obtained by integrating out the mark as in (9).

The parametric coefficients are reported in Table 6, while Table 7 summarizes the fitted smooth spatial terms.

Several remarks emerge from Tables 6 and 7. First, the smooth spatial term is highly significant in all three models, confirming that the fire pattern has an important first-order spatial structure. Second, in the PMP specification, the coefficient associated with the transformed mark is strongly negative and highly significant. This indicates that, after

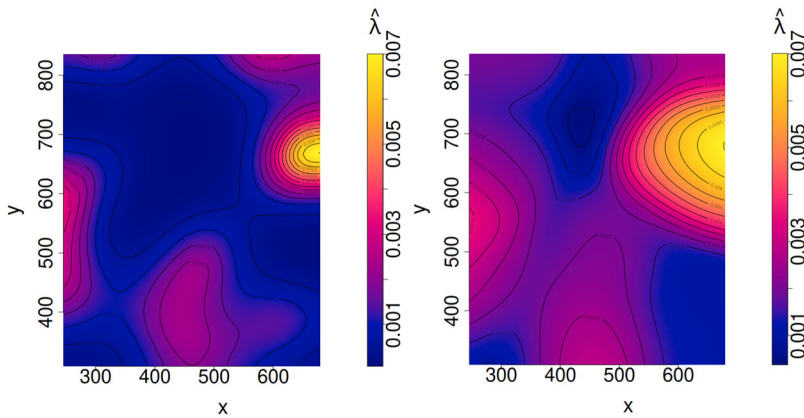


Figure 5. Fitted spatial intensity surfaces for the fires data: marginal PMP intensity (left-panel) and SMP intensity (right-panel).

controlling for the smooth spatial component, larger observed fire sizes are associated with lower fitted first-order intensity on the joint domain. Finally, the intercept changes across models, but these differences should not be over-interpreted in isolation: they mainly reflect how the baseline level is reallocated once the mark contribution is incorporated either parametrically, as in PMP, or through the reconstructed spatial adjustment, as in SMP.

Figure 5 compares the fitted spatial intensities obtained from the marginal PMP fit and from the SMP fit.

Both fitted surfaces recover the main regions of high fire concentration. The two maps are broadly similar in terms of global first-order structure, but the marginal PMP fit appears slightly more contrasted, with a sharper separation between high- and low-intensity areas. This is consistent with the idea that, in this application, the mark is not strongly collinear with the spatial coordinates and can therefore be incorporated more directly without the instability observed in the simulation scenarios with strongly spatially structured marks.

To assess more explicitly how the inclusion of the mark modifies the fitted spatial intensity relative to the baseline Poisson model, Figure 6 reports the fitted relative intensity ratios $\hat{\mathcal{R}}_P(s)$ and $\hat{\mathcal{R}}_S(s)$, following (11) and (15).

The two ratios convey rather different messages. The PMP ratio shows non-negligible departures from 1 over the observation window: relative to the baseline Poisson model, the marginal spatial intensity under the PMP model is lower in the South-Western part of the window and higher in central and North-Eastern areas. By contrast, the SMP ratio remains extremely close to 1 almost everywhere, indicating that once the SMP model is refitted jointly with the reconstructed mark-related surface, its fitted spatial intensity is globally very similar to that of the baseline Poisson model.

This, however, does not imply that the reconstructed mark-related component is spatially constant in the SMP specification. To display the reconstructed mark-related component separately, Figure 7 reports the fitted surface $\hat{\Phi}_S(s) = \exp(\hat{\kappa}_c(s))$, which is the mark-related multiplier entering the SMP intensity.

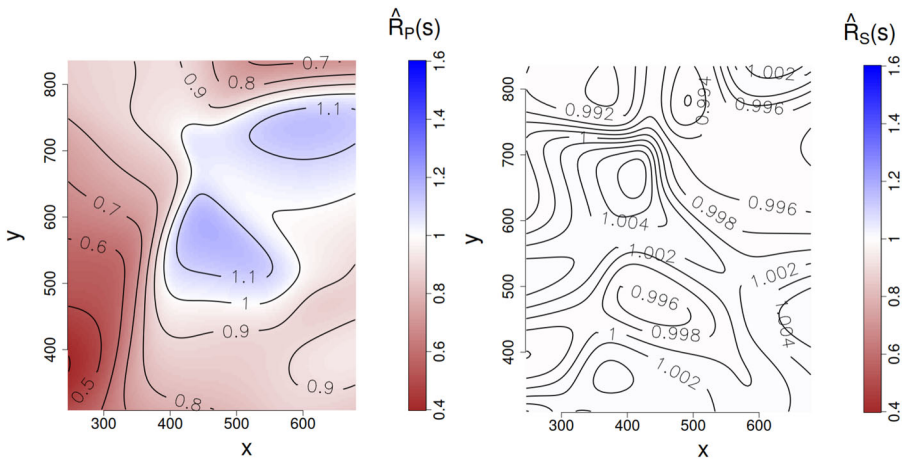


Figure 6. Fitted relative intensity ratios for the fires data: $\hat{\mathcal{R}}_P(s)$ from the PMP model (left-panel) and $\hat{\mathcal{R}}_S(s)$ from the SMP model (right-panel).

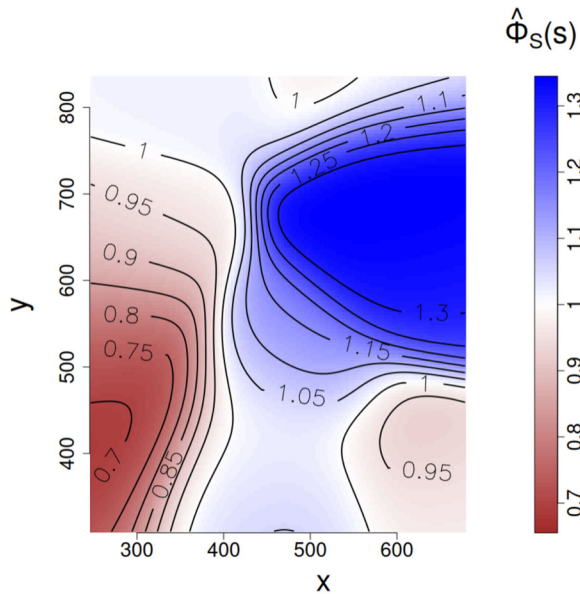


Figure 7. Estimated direct mark-related component in the SMP model for the fires data, $\hat{\Phi}_S(s) = \exp(\hat{\kappa}_c(s))$.

The isolated SMP mark-related component shows a much clearer spatial pattern: it is below 1 in the South-Western portion of the window and above 1 in the eastern and North-Eastern areas. Hence, the reconstructed mark-related multiplier displays systematic spatial variation, with values below 1 in the South-Western portion of the window and above 1 in the Eastern and North-Eastern areas. The reason why $\hat{\mathcal{R}}_S(s)$ remains close to 1 is that, when the SMP model is estimated, part of the same spatial variation is reallocated between the smooth baseline component and the reconstructed mark-related component. Thus, the

reconstructed mark-related component varies systematically over the observation window, but spatial variation is largely offset in the comparison with the separately fitted baseline Poisson model.

Overall, this application is more consistent with the empirical setting in which the real-valued mark can be treated as an additional modelling dimension without severe collinearity with the spatial coordinates. In this sense, the results are closer to the simulation scenarios favourable to PMP. The PMP fit provides a more pronounced modification of the baseline spatial intensity and a clear negative parametric association between transformed fire size and the fitted joint intensity. At the same time, the SMP fit remains informative by revealing a spatially varying mark-related signal, even though its final fitted intensity is globally very close to that of the unmarked Poisson model.

5.2. Bronzefilter data

In the second application, we analyse the *bronzefilter* data, consisting of an inhomogeneous spatial point pattern of 678 points representing circular section profiles of particles observed in a longitudinal plane section through a gradient sinter filter made from bronze powder. The data are in the *spatstat* package [37] and are supplied as a marked point pattern of circle centres, with circle radii as continuous marks. Coordinates and radii are measured in millimetres, and the observation window is an 18×7 mm rectangle. The pattern is shown in the left panel of Figure 8.

As for the fires data, we begin with a kernel-based exploratory analysis, using the Nadaraya–Watson reconstruction of the mark-related surface [33–35]. Figure 8 suggests two main features. First, the spatial intensity exhibits a strong large-scale gradient along the x -direction. Second, the mark is clearly associated with the spatial domain, especially with the horizontal coordinate. This makes the application much closer to the simulation scenarios in which the mark behaves as a spatially structured signal rather than as an independent additional modelling dimension. Accordingly, we fit a baseline Poisson model with a linear spatial trend in x

$$\hat{v}(x) = \exp(\hat{\theta}_0 + \hat{\theta}_1 x),$$

a PMP model on the product space

$$\hat{\lambda}_P(x, m) = \exp(\hat{\theta}_0 + \hat{\theta}_1 x + \hat{\theta}_2 m),$$

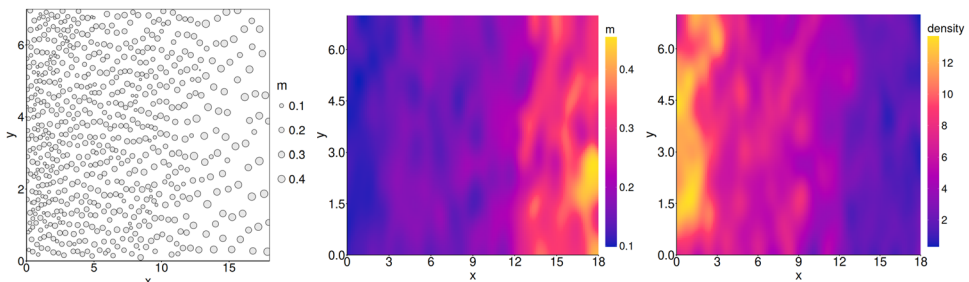
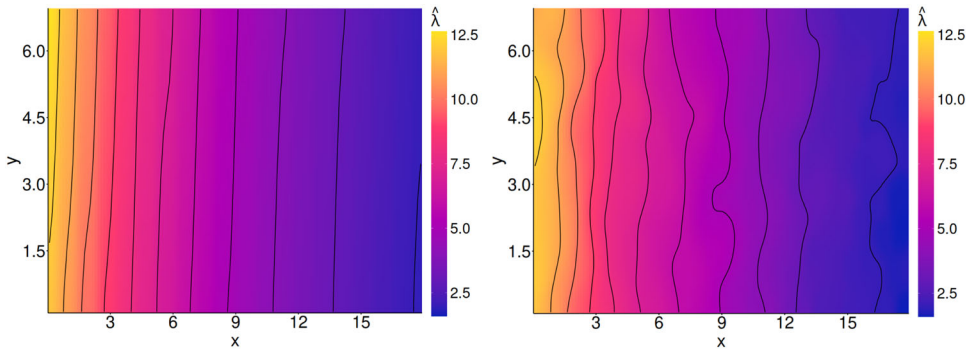


Figure 8. Left-panel: Marked bronzefilter point pattern. Central-panel: Kernel-smoothed spatial surface of the mark values. Right-panel: Kernel estimate of the spatial intensity.

Table 8. Parameter estimates of the Poisson, PMP, and SMP models on the bronze-filter data.

Model	Coefficients	Estimate	Std. Error	z value	Pr(> z)
Poisson	Intercept	2.454	0.093	26.464	< 2·10 ^{−16} ***
	<i>x</i>	−0.108	0.008	−13.353	< 2·10 ^{−16} ***
PMP	Intercept	3.299	0.104	31.656	< 2·10 ^{−16} ***
	<i>x</i>	−0.102	0.0078	−12.894	< 2·10 ^{−16} ***
	<i>m</i>	−1.957	0.146	−13.441	< 2·10 ^{−16} ***
SMP	Intercept	2.482	0.094	26.365	< 2·10 ^{−16} ***
	<i>x</i>	−0.082	0.008	−9.941	< 2·10 ^{−16} ***

**Figure 9.** Fitted spatial intensity surfaces for the bronzefilter data: marginal PMP intensity (left-panel) and SMP intensity (right-panel).

and an SMP model on the spatial window

$$\hat{\lambda}_S(x, y) = \exp(\hat{\theta}_0 + \hat{\theta}_1 x + \hat{\kappa}_c(x, y)).$$

As before, the PMP fit is compared with the other models through its marginal spatial intensity $\hat{\lambda}_p^{\text{marg}}(x, y)$. The estimated coefficients are reported in Table 8.

The PMP fit provides a strongly negative and highly significant coefficient for the mark. This would suggest that larger particle radii are associated with lower intensity on the joint domain. However, because the mark is strongly associated with the spatial coordinates, the parametric coefficients of the PMP model should not be interpreted too literally one by one. This is precisely the type of setting in which the simulation study showed that PMP may have difficulty separating the contribution of the mark from that of the spatial trend.

This is also reflected in the remaining coefficients. The estimated slope for *x* is very similar in the Poisson and PMP models, whereas it is attenuated in the SMP fit. This suggests that, under SMP, part of the large-scale horizontal gradient is absorbed by the reconstructed mark-related surface rather than being forced into the linear spatial term. At the same time, the PMP intercept is markedly larger than the corresponding Poisson and SMP intercepts, again pointing to a different reallocation of the baseline level once the mark is included parametrically in a setting with strong mark–space association.

Figure 9 compares the fitted spatial intensities from the marginal PMP model and from the SMP model.

Both models reproduce the decrease in the first-order intensity from left to right, which is the main large-scale feature of the pattern. However, the two fits differ in how the mark

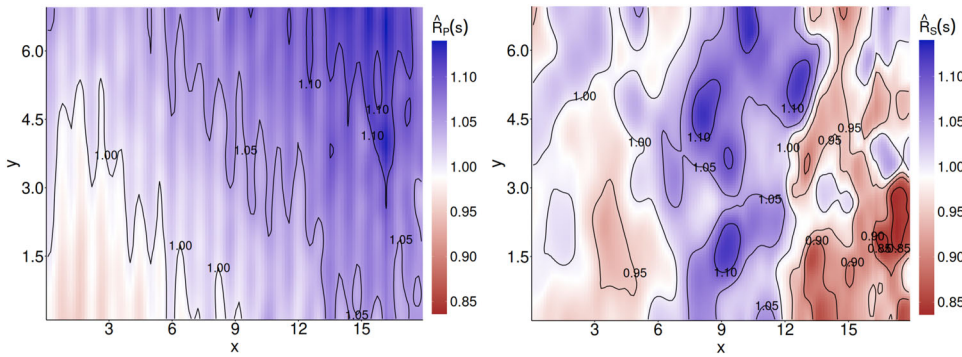


Figure 10. Fitted relative intensity ratios for the bronzefilter data: $\hat{\mathcal{R}}_P(s)$ from the PMP model (left-panel) and $\hat{\mathcal{R}}_S(s)$ from the SMP model (right-panel).

information is incorporated. The marginal PMP intensity remains largely driven by the global x -gradient and appears overall higher and more regular. By contrast, the SMP fit preserves the same global behaviour while allowing for more local spatial modulation through the reconstructed mark-related component. This result is more in line with the exploratory evidence, which suggested that the mark is spatially structured over the observation window. To make this comparison explicit, Figure 10 reports the fitted relative intensity ratios $\hat{\mathcal{R}}_P(s)$ and $\hat{\mathcal{R}}_S(s)$.

The PMP ratio is mostly above 1 and tends to increase from left to right. This is an important feature of the fitted comparison. Although the PMP coefficient associated with the mark is negative, the fitted marginal PMP spatial intensity is, relative to the baseline Poisson fit, generally higher. In other words, once the model is estimated and then marginalized over the mark space, the difference between the fitted marginal PMP intensity and the baseline Poisson intensity cannot be inferred directly from the sign of the parametric mark coefficient alone. This is consistent with the simulation results for spatially structured marks, where the PMP often struggled to separate mark-related variation from the large-scale spatial trend and could obtain a biased fitted spatial intensity. Focussing on the SMP model, it remains centred around 1, but with clear local departures: the fitted SMP ratio is above 1 in a central band of the window and below 1 in parts of the left and right regions, indicating systematic spatial variation in the fitted mark-related component relative to the separately fitted baseline model. This is exactly the type of behaviour expected when the mark acts as a spatially varying signal rather than as an independent modelling dimension.

Taken together, these results suggest that the bronzefilter data are much closer to the empirical context where SMP was favoured in the simulation study. Since the mark is strongly associated with the spatial domain, the main issue is not simply whether the mark coefficient is significant, but whether the model can separate the large-scale spatial trend from the contribution associated with the mark. In this application, the SMP specification appears more coherent with that objective and provides a more plausible first-order decomposition than the PMP alternative.

6. Conclusion

In this paper, we introduced a Weighted Poisson Intensity (WPI) framework for modelling first-order intensity of marked point processes with real-valued marks. The main idea is to combine a baseline Poisson intensity on the ground space with a non-negative mark-related component, so as to assess how the inclusion of the mark modifies the fitted first-order structure.

Within this framework, we proposed two specific modelling implementations. The Parametric Marked Poisson model (PMP) treats the mark as an explicit argument of the intensity on the product space and therefore provides a direct parametric representation of the mark contribution. The Semi-parametric Marked Poisson model (SMP) instead interprets the observed marks as a spatially structured signal attached to the ground process and incorporates them through a reconstructed mark-related surface. The two approaches therefore correspond to two different modelling views of the real-valued mark. The simulation study showed that the relative performance of the competing approaches depends on how the mark enters the data-generating mechanism. When the mark carries no relevant information for the intensity, neither PMP nor SMP offers systematic gains over the baseline Poisson fit. On the other hand, when the mark acts as an additional modelling dimension and is not strongly associated with the spatial coordinates, PMP is generally the most appropriate specification and can improve first-order intensity estimates with respect to a baseline Poisson model. When the mark is instead strongly aligned with the spatial domain and behaves more like a spatially structured signal, SMP is more stable and typically provides better first-order reconstruction, while PMP may suffer from bias. These findings remained qualitatively consistent not only under Poisson scenarios, but also in the Hotspot and LGCP settings. The two real-data applications were coherent with these results. In the fires data, where the mark does not appear to be strongly collinear with the spatial coordinates, the PMP fit produced the clearest modification of the baseline spatial intensity and a clear negative parametric association between the transformed mark and the fitted joint intensity. In the bronzefilter data, where the mark is strongly associated with the spatial domain, the SMP approach provided a more plausible decomposition between the large-scale spatial trend and the mark-related contribution, whereas PMP appeared less stable. Taken together, the applications illustrate that the choice between PMP and SMP should be guided by the empirical role of the mark in relation to the spatial structure of the pattern.

These findings also provide practical guidance for choosing between the two approaches. PMP is most appropriate when the real-valued mark is substantively regarded as part of the event description and can plausibly be treated as an additional modelling dimension on the product space. This choice is particularly natural when the mark is not strongly aligned with the spatial coordinates and when interest focuses on a parametric description of the association between the mark and the fitted joint intensity. SMP is instead more appropriate when the observed marks are plausibly interpreted as observations of an underlying spatially varying signal, especially when exploratory analyses reveal marked spatial structure or strong association between the mark and the spatial trend. In the latter setting, the reconstruction of a mark-related surface can provide a more stable first-order decomposition than a parametric product-space specification. In practice, exploratory maps of the marks and their relation to the spatial coordinates can help guide this choice.

When the modelling interpretation remains ambiguous, fitting both approaches can be informative, provided that their fitted spatial intensities, the stability of the estimated spatial components, and the plausibility of their respective interpretations are assessed on the common ground domain.

Several directions remain open for future research. A first important step concerns inference for the SMP estimator, since the current implementation relies on a two-step plug-in pseudo-likelihood and its standard errors are conditional on the first-stage reconstructed surface. Another possible direction would be to extend the proposed approach to settings with multiple or functional continuous marks, where suitable low-dimensional representations of the mark information may be needed. A further relevant extension would be to allow for more flexible specifications of the mark contribution within the WPI framework, for instance, through interaction terms between the mark and the spatial coordinates or through more general non-linear mark effects. Such developments could help accommodate settings in which the role of the mark is more heterogeneous or only partially separable from the spatial structure.

Overall, the proposed methodology provides a practical way to incorporate real-valued marks into first-order intensity modelling, while also clarifying that the modelling nature assigned to the mark is crucial for obtaining an appropriate and interpretable fitted intensity.

Author contributions

CRedit: **Marco Tarantino**: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing; **Nicoletta D'Angelo**: Conceptualization, Methodology, Resources, Supervision, Writing – review & editing; **Ottmar Cronie**: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Supervision, Writing – review & editing; **Giada Adelfio**: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The authors gratefully acknowledge financial support of the following projects: Fondi di Ateneo - Bando EUROSTART 2025 - progetto di ricerca ASPIRE - Advances in Spatio-temporal Point process models Innovations and Research P. I. Nicoletta D'Angelo; the Targeted Research Funds 2026 (FFR 2026) of the University of Palermo; and the Ministero dell'Università e della Ricerca (MUR) PRIN 2022: Spatio-temporal Functional Marked Point Processes for probabilistic forecasting of earthquakes, grant number 2022BN7CJP, P. I. Giada Adelfio, CUP 53C24006340006.

Data availability statement

All the data used in this paper are available in the R package [spatstat](#).

References

- [1] Ogata Y. Statistical models for earthquake occurrences and residual analysis for point processes. *J Am Stat Assoc.* 1988;83(401):9–27. doi: [10.1080/01621459.1988.10478560](https://doi.org/10.1080/01621459.1988.10478560)

- [2] D'Angelo N, Adelfio G, Mateu J, et al. Local inhomogeneous weighted summary statistics for marked point processes. *J Comput Graph Stat.* **2024**;33(2):588–602. doi: [10.1080/10618600.2023.2206441](https://doi.org/10.1080/10618600.2023.2206441)
- [3] Stoyan D, Penttinen A. Recent applications of point process methods in forestry statistics. *Stat Sci.* **2000**;15(1):61–78.
- [4] Møller J, Ghorbani M, Rubak E. Mechanistic spatio-temporal point process models for marked point processes, with a view to forest stand data. *Biometrics.* **2016**;72(3):687–696. doi: [10.1111/biom.12466](https://doi.org/10.1111/biom.12466)
- [5] Drinkwater MJ, Parker QA, Proust D, et al. The large scale distribution of galaxies in the Shapley supercluster. *Publ Astron Soc Aust.* **2004**;21(1):89–96. doi: [10.1071/AS03057](https://doi.org/10.1071/AS03057)
- [6] Mancini L, Paganoni AM. Marked point process models for the admissions of heart failure patients. *Stat Anal Data Min.* **2019**;12(2):125–135. doi: [10.1002/sam.2019.12.issue-2](https://doi.org/10.1002/sam.2019.12.issue-2)
- [7] Descombes X, Zerubia J. Marked point process in image analysis. *IEEE Signal Process Mag.* **2002**;19(5):77–84. doi: [10.1109/MSP.2002.1028354](https://doi.org/10.1109/MSP.2002.1028354)
- [8] Ghorbani M, Cronie O, Mateu J, et al. Functional marked point processes: a natural structure to unify spatio-temporal frameworks and to analyse dependent functional data. *Test.* **2021**;30(3):529–568. doi: [10.1007/s11749-020-00730-2](https://doi.org/10.1007/s11749-020-00730-2)
- [9] Eckardt M, Moradi M. Marked spatial point processes: current state and extensions to point processes on linear networks. *J Agric Biol Environ Stat.* **2024**;29(2):346–378. doi: [10.1007/s13253-024-00605-1](https://doi.org/10.1007/s13253-024-00605-1)
- [10] Stoyan D. On correlations of marked point processes. *Math Nachr.* **1984**;116(1):197–207. doi: [10.1002/\(ISSN\)1522-2616](https://doi.org/10.1002/(ISSN)1522-2616)
- [11] Schlather M. On the second-order characteristics of marked point processes. *Bernoulli.* **2001**;7(1):99–117. doi: [10.2307/3318604](https://doi.org/10.2307/3318604)
- [12] Cronie O, Van Lieshout MNM. Summary statistics for inhomogeneous marked point processes. *Ann Inst Stat Math.* **2016**;68(4):905–928. doi: [10.1007/s10463-015-0515-z](https://doi.org/10.1007/s10463-015-0515-z)
- [13] Moradi M, Eckardt M. Inhomogeneous mark correlation functions for general marked point processes. *Biometrics.* **2026**;82(1):ujaf177. doi: [10.1093/biometc/ujaf177](https://doi.org/10.1093/biometc/ujaf177)
- [14] Daley DJ, Vere-Jones D. An introduction to the theory of point processes. Volume II: general theory and structure. 2nd ed. New York: Springer-Verlag; **2007**.
- [15] Baddeley A. Multivariate and marked point processes. In: Gelfand AE, Diggle PJ, Fuentes M, Guttorp P, editors. *Handbook of spatial statistics*. Boca Raton, FL: Chapman & Hall/CRC; **2010**. p. 371–402. (Handbooks of Modern Statistical Methods).
- [16] Mateu J, Ribeiro PJ Jr. Geostatistical data versus point process data: analysis of second-order characteristics. In: *geoENV II—Geostatistics for Environmental Applications: Proceedings of the Second European Conference on Geostatistics for Environmental Applications held in Valencia, Spain, November 18–20, 1998*. Springer; **1999**. p. 213–224.
- [17] Schlather M, Ribeiro PJ Jr., Diggle PJ. Detecting dependence between marks and locations of marked point processes. *J R Stat Soc Ser B Stat Methodol.* **2004**;66(1):79–93. doi: [10.1046/j.1369-7412.2003.05343.x](https://doi.org/10.1046/j.1369-7412.2003.05343.x)
- [18] Renshaw E, Mateu J, Saura F. Disentangling mark/point interaction in marked-point processes. *Comput Stat Data Anal.* **2007**;51(6):3123–3144. doi: [10.1016/j.csda.2006.07.035](https://doi.org/10.1016/j.csda.2006.07.035)
- [19] Baddeley A, Turner TR. Practical maximum pseudo likelihood for spatial point patterns. *Aust N Z J Stat.* **2000**;42(3):283–322. doi: [10.1111/anzs.2000.42.issue-3](https://doi.org/10.1111/anzs.2000.42.issue-3)
- [20] Baddeley A, Rubak E, Turner R. *Spatial point patterns: methodology and applications with R*. Boca Raton, FL: Chapman & Hall/CRC; **2015**. (Chapman & Hall/CRC Interdisciplinary Statistics).
- [21] Ogata Y. Space-time point-process models for earthquake occurrences. *Ann Inst Stat Math.* **1998**;50(2):379–402. doi: [10.1023/A:1003403601725](https://doi.org/10.1023/A:1003403601725)
- [22] Rasmussen JG. Bayesian inference for Hawkes processes. *Methodol Comput Appl Probab.* **2013**;15:623–642. doi: [10.1007/s11009-011-9272-5](https://doi.org/10.1007/s11009-011-9272-5)
- [23] Cox DR. Some statistical methods connected with series of events. *J R Stat Soc Ser B.* **1955**;17(2):129–157. doi: [10.1111/j.2517-6161.1955.tb00188.x](https://doi.org/10.1111/j.2517-6161.1955.tb00188.x)

- [24] Møller J, Syversveen AR, Waagepetersen RP. Log Gaussian Cox processes. *Scand J Statist.* 1998;25(3):451–482. doi: [10.1111/sjos.1998.25.issue-3](https://doi.org/10.1111/sjos.1998.25.issue-3)
- [25] R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2024. Available from: <https://www.R-project.org/>.
- [26] Daley DJ, Vere-Jones D. An introduction to the theory of point processes: volume I: elementary theory and methods. New York: Springer; 2003.
- [27] Berman M, Turner TR. Approximating point process likelihoods with glim. *J R Stat Soc Ser C.* 1992;41(1):31–38.
- [28] Tarantino M, D'Angelo N, Adelfio G. Modelling three-dimensional point patterns. In: Scientific Meeting of the Italian Statistical Society. Springer; 2024. p 588–594.
- [29] Tarantino M, D'Angelo N, Adelfio G Recommendations on cubature parameter selection for likelihood approximation in three-dimensional poisson point processes. *J Stat Comput Simul.* 2026; 1–44. doi:[10.1080/00949655.2026.2681025](https://doi.org/10.1080/00949655.2026.2681025).
- [30] Hardin JW. The robust variance estimator for two-stage models. *Stata J.* 2002;2(3):253–266. doi: [10.1177/1536867X0200200302](https://doi.org/10.1177/1536867X0200200302)
- [31] Coeurjolly J-F, Drouilhet R. Asymptotic properties of the maximum pseudo-likelihood estimator for stationary Gibbs point processes including the Lennard-Jones model. *Electron J Stat.* 2010;4:677–706. doi: [10.1214/09-EJS494](https://doi.org/10.1214/09-EJS494)
- [32] Cronie O, Moradi M, Biscio CAN. A cross-validation-based statistical theory for point processes. *Biometrika.* 2024;111(2):625–641. doi: [10.1093/biomet/asad041](https://doi.org/10.1093/biomet/asad041)
- [33] Nadaraya EA. On estimating regression. *Theory Probab Appl.* 1964;9(1):141–142. doi: [10.1137/1109020](https://doi.org/10.1137/1109020)
- [34] Watson GS. Smooth regression analysis. *Sankhyā.* 1964;26:359–372.
- [35] Nadaraya EA. Nonparametric estimation of probability densities and regression curves. Dordrecht: Springer; 1989.
- [36] Cronie O, Van Lieshout MNM. A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions. *Biometrika.* 2018;105(2):455–462. doi: [10.1093/biomet/asy001](https://doi.org/10.1093/biomet/asy001)
- [37] Baddeley A, Turner R. spatstat: an R package for analyzing spatial point patterns. *J Stat Softw.* 2005;12(6):1–42. doi: [10.18637/jss.v012.i06](https://doi.org/10.18637/jss.v012.i06)
- [38] Turner R. Point patterns of forest fire locations. *Environ Ecol Stat.* 2009;16:197–223. doi: [10.1007/s10651-007-0085-1](https://doi.org/10.1007/s10651-007-0085-1)
- [39] Wood SN. Thin plate regression splines. *J R Stat Soc Ser B Stat Methodol.* 2003;65(1):95–114. doi: [10.1111/1467-9868.00374](https://doi.org/10.1111/1467-9868.00374)