

Analysis of massive marked point patterns with Stochastic Partial Differential Equations

Virgilio Gómez-Rubio^a, Michela Cameletti^{b,*}, Francesco Finazzi^c

^a*Dept. of Mathematics, School of Industrial Engineering, University of Castilla-La Mancha, 02071 Albacete, Spain.*

^b*Dept. of Management, Economics and Quantitative Methods, University of Bergamo, 24127, Bergamo, Italy.*

^c*Dept. of Management, Economics and Quantitative Methods, University of Bergamo, 24127, Bergamo, Italy.*

Abstract

In this paper we describe a novel approach to modelling marked point patterns based on recent computational developments for Bayesian inference. We use the flexible class of log-Gaussian Cox Processes to model the intensity of the different observed point patterns. We propose several types of models to account for spatial variability and provide a modelling framework that allows for a common spatial component to all point processes (regardless of the mark) and also for a mark-specific spatial components. In this way, we provide a method of assessing whether all processes share a common spatial distribution or there are specific features.

In order to fit these models, we have resorted to the Integrated Nested Laplace Approximation (INLA) method and the Stochastic Partial Differential Equation (SPDE) approach. This defines a connection between point process and geostatistics, as we model a point pattern by means of a continuous spatial process.

Our new approach to spatial modelling is applied to a massive dataset on the occurrence of tornados in the United States. We have divided the tornados in the 1950-2013 period according to their magnitude and fitted

*Corresponding author

Email addresses: `Virgilio.Gomez@uclm.es` (Virgilio Gómez-Rubio),
`michela.cameletti@unibg.it` (Michela Cameletti), `francesco.finazzi@unibg.it`
(Francesco Finazzi)

our proposed models.

Keywords: INLA, Matèrn spatial covariance function, Tornados

1. Introduction

Marked point patterns (Gelfand et al., 2010, Chap. 21) play an important role in the analysis of different natural phenomena. They have been used, for example, in veterinary epidemiology, in ecology, spatial disease mapping and in different applications where spatial locations are supplied with extra information - i.e. marks or point types - which could be related to the spatial point pattern. In doing so, a number of different approaches have been followed. For example, Diggle et al. (2005) analyse the spatial distribution of different strains of bovine tuberculosis (BTB) in the United Kingdom using a multivariate point process. Here, the marks are associated to different BTB genotypes that may have different spatial distributions. A different spatial Poisson process is fit to each pattern using kernel estimation, and a Monte Carlo test (based on random labelling) is proposed to assess equal spatial distribution of the different BTB genotypes.

Liang et al. (2009) analyze colon and rectum cancer data by taking into account spatially-referenced and non-spatial covariates. One of the latter is represented by cancer type which is considered as the mark of the point pattern. In the class of log-Gaussian Cox processes, they consider a model with mark-varying covariate coefficients and spatial continuous random processes. The estimation is performed by Markov Chain Monte Carlo (MCMC) combined with a knot-based predictive process approximation.

In the ecological context, Illian et al. (2013) model the locations of a plant species together with health status, which is represented by a categorical mark and is assumed to depend on the spatial pattern through a shared spatial effect. In particular, they jointly fit the spatial pattern and the marks with a bivariate model with Poisson-Binomial distributions using the Integrated Nested Laplace Approximation (INLA) approach. This work is extended to the case of multiple marks in Illian et al. (2012a) with an ecological application for studying the koala-eucalyptus relationship.

Following these ideas, we have developed a novel approach for the analysis of marked point patterns using recent statistical developments that provide a powerful computational advantage for the analysis of large point patterns.

In particular, we model each point type using a log-Gaussian Cox process (LGCP, see, for example, Diggle et al., 2013) using a smooth spatial term of the Matérn family which is fitted using the Stochastic Partial Differential Equation approach (SPDE, Lindgren et al., 2011) and fitted by means of the Integrated Nested Laplace Approximation method (INLA, Rue et al., 2009). This is inspired by the disease mapping models with shared components for two different diseases (Knorr-Held and Best, 2001), but we have extended these models to several patterns.

A critical issue connected with the analysis of point patterns is the computational burden that arises when the observed number of points is large (see, for example, Illian et al., 2012b). If the point process is marked, then this computational burden increases as models become more complex. In a Bayesian context, these prohibitive computational costs are mainly due to MCMC simulation methods which can be very time-consuming when the number of unknown parameters and latent process is high and/or the amount of available data is huge. In this work we avoid MCMC by employing the INLA approach through the R-INLA package (<http://www.r-inla.org/>). INLA is an algorithm for Bayesian inference which is especially designed for latent Gaussian hierarchical models (a class of models which includes also LGCPs) and, compared to MCMC, it provides accurate results in shorter computing time. Moreover, INLA can be combined with the SPDE approach to overcome the so-called “big n problem” (Jona Lasinio et al., 2012) which occurs in geostatistics. The strength of SPDE stems from representing a continuous spatial Gaussian field with Matérn covariance function as a discrete spatial random process (i.e. a Gaussian Markov random field) which enjoys nice computational properties and leads to fast inference.

In the past few years the number of papers reporting usage and applications of the INLA-SPDE method has increased considerably in several scientific fields, e.g. Blangiardo et al. (2013) for spatial and spatio-temporal models for disease mapping and geostatistics, Bivand et al. (2014) for spatial econometrics models, Grilli et al. (2014) for binary logit mixed models, Muff et al. (2014) for measurement error models, Ingebrigtsen et al. (2014) for non-stationary geostatistical models, Cameletti et al. (2013) for air pollution modeling, Cosandey-Godin et al. (2014) for fisheries bycatch studies, Papoila et al. (2014) for modeling stomach cancer incidence, Musenge et al. (2013) for zero-inflated spatio-temporal models with child mortality data, Saez et al. (2012) for space-time interpolation of daily air temperatures.

We illustrate our new modelling approach using tornado data from the

United States. This is a quite huge dataset including 56908 tornado trajectories registered from 1950 to 2013. In our application we assume that the point pattern mark is given by the tornado magnitude as measured in the Fujita Tornado Damage Scale with categories from 0 to 5 (with 0 being the weakest tornado type and 5 the strongest). By assuming one of the marks as the reference pattern, we set a baseline for the spatial distribution of tornadoes. This simple model is then extended to include mark-specific spatial patterns and covariates. Hence, we can use model comparison methods to assess differences in the spatial distribution of tornadoes based on their scale and the influence of environmental covariates.

This paper is structured as follows. Section 2 introduces marked point patterns and some methods for their analysis. Section 3 deals with our methodological approach based on SPDE. Next, Section 4 develops an example on data from tornados in the United States. Finally, we provide a short discussion and some important remarks in Section 5.

2. Analysis of marked point patterns

2.1. Summary of relevant models

Given a point pattern that occurs in a bounded region $\Omega \subset \mathbb{R}^2$, we will assume that the observed pattern is a realisation of a log-Gaussian Cox Process (LGCP, Møller et al., 1998; Simpson et al., 2013; Diggle et al., 2013), so that the intensity at location $s \in \Omega$ is given by

$$\log(\lambda(s)) = \xi(s), \tag{1}$$

where $\xi(s)$ is a stationary Gaussian random field. Conditioning on a realisation of $\xi(s)$, the LGCP is an inhomogeneous Poisson process with intensity $\lambda(s)$. Hence, the observed number of points in any region $\mathcal{D} \subseteq \Omega$ is distributed as a Poisson with mean $\int_{\mathcal{D}} \lambda(s) ds$.

For the case of a marked point pattern, we will assume that we have a categorical mark, so that the point pattern can be split into $K + 1$ groups. We will represent the intensities of each group by $\lambda_j(s)$, $j = 0, \dots, K$. The number of observations attached to each mark will be denoted by n_j , $j = 0, \dots, K$. In addition, other statistics could be computed for these groups, such as, the cross K -function to assess clustering of one pattern as compared to another.

Diggle et al. (2005) describe the use of marked point patterns for the analysis of the spatial distribution of different genotypes of bovine tuberculosis

in the United Kingdom. They assume a different intensity for each pattern, which is estimated using non-parametric methods, and propose a test for equal spatial distribution. Furthermore, they suggest the use of relative risk surfaces to compare the spatial distribution of the different types using

$$\rho_{jk}(s) = \frac{\lambda_j(s)}{\lambda_k(s)}, \quad \forall j \neq k. \quad (2)$$

Finally, conditional probability surfaces can be defined for each of the point types to estimate the probability that a point at a given location is of a particular type. These *type-specific probabilities* can be computed as

$$p_j(s) = \frac{\lambda_j(s)}{\sum_{k=1}^K \lambda_k(s)}, \quad j = 0, \dots, K. \quad (3)$$

Diggle et al. (2005) use non-parametric methods for the estimation of the intensities and type-specific probabilities.

Similarly, Diggle et al. (2007) focus on a two-type point pattern (cases and controls) and propose the estimation of the intensity of the cases ($\lambda_1(s)$) by modulating the intensity of the controls ($\lambda_0(s)$) using a set of covariates, so that

$$\lambda_1(s) = \frac{n_1}{n_0} \lambda_0(s) \exp(\alpha + \boldsymbol{\beta} \boldsymbol{x}) \quad (4)$$

where \boldsymbol{x} is a vector of covariates and $\boldsymbol{\beta}$ their associated coefficients. Diggle et al. (2007) use this modulated intensity to propose a test for clustering of the cases as compared to the controls and accounting for the covariates. In this particular problem with two point types only, the coefficients can be estimated using a typical logistic regression on the marks and the covariates.

Diggle et al. (2005, 2007), Møller and Waagepetersen (2007) and Bivand et al. (2013), among others, suggest that other options are available to estimate the intensity of a point pattern. For example, the intensity of the cases can be modelled by means of a smooth spatial term $\xi(s)$, so that it becomes

$$\lambda_1(s) = \frac{n_1}{n_0} \lambda_0(s) \exp(\xi(s)). \quad (5)$$

Note that $\xi(s)$ can take different forms, such as a smoothing spline, polynomial on the coordinates, etc. (for instance, see the examples in Bivand et al., 2013).

2.2. Our new proposal

For the analysis of marked point patterns we combine these two ideas to develop a new approach to estimating and testing for the intensity of marked point patterns. We will consider one of the marks as the reference pattern and define its intensity $\lambda_0(s)$ using a smooth term, such as the one provided by the SPDE approach that we describe in Section 3.2. Hence, this intensity, in the log-scale, will become

$$\log(\lambda_0(s)) = \alpha_0 + \xi_0(s) \quad (6)$$

with α_0 an intercept and $\xi_0(s)$ a smooth term based on SPDE. It is worth noting that other additive terms (e.g., covariates) can be included in the right hand side of Equation (6) to account for other linear or non linear effects. The intensities $\lambda_j(s)$ for the other marks ($j = 1, \dots, K$) can be defined similarly.

In order to provide a way of estimating the intensity and testing for any particular differences in the patterns, we will define the intensity of the other marks using $\lambda_0(s)$ as a baseline. If all marks have the same spatial pattern it is known that the intensities can be written down as

$$\lambda_j(s) = \frac{n_j}{n_0} \lambda_0(s), \quad j = 1, \dots, K. \quad (7)$$

This can be extended to propose a shared component model (Knorr-Held and Best, 2001) to account for an overall spatial trend (common to all patterns) plus a specific spatial process (of a single pattern), as follows:

$$\log(\lambda_j(s)) = \log\left(\frac{n_j}{n_0}\right) + \alpha_0 + \xi_0(s) + \xi_j(s), \quad j = 1, \dots, K. \quad (8)$$

Note that $\log\left(\frac{n_j}{n_0}\right)$ is an offset and that α_0 and $\xi_0(s)$ are estimated solely on the baseline point pattern. The term $\xi_j(s)$ can be any smooth term but in our setting it is based on a SPDE and it is in fact tested for any spatial departure from the baseline. If this term is equal to zero then the reference and the j -th pattern will have the same spatial distribution. Otherwise, they will not. This new term $\xi_j(s)$ will also help us to determine the regions where there is a departure from the baseline. Testing for the significance of $\xi_j(s)$ is explained in Section 2.3.

Finally, the model in Equation (8) can be extended to account for relevant covariates. This will lead to the following expression for the log-intensity:

$$\log(\lambda_j(s)) = \log\left(\frac{n_j}{n_0}\right) + \alpha_0 + \boldsymbol{\beta}\mathbf{x}(s) + \xi_0(s) + \xi_j(s), \quad j = 1, \dots, K. \quad (9)$$

Other types of effects can be considered as well, such as, temporal trends but this will probably make model fitting more complex.

Also, the previous model can be simplified and change the term $\log\left(\frac{n_j}{n_0}\right) + \alpha_0$ by a pattern-specific intercept α_j as

$$\log(\lambda_j(s)) = \alpha_j + \boldsymbol{\beta}\mathbf{x}(s) + \xi_0(s) + \xi_j(s), \quad j = 1, \dots, K. \quad (10)$$

All models that we will discuss in this paper are based on the previous one. The models can be simpler versions that do not include covariates and/or pattern-specific spatial effects. Hence, we will fit the following joint model for a given $j = 1, \dots, K$:

$$\begin{aligned} \log(\lambda_0(s)) &= \alpha_0 + \boldsymbol{\beta}\mathbf{x}(s) + \xi_0(s) \\ \log(\lambda_j(s)) &= \alpha_j + \boldsymbol{\beta}\mathbf{x}(s) + \xi_0(s) + \xi_j(s). \end{aligned} \quad (11)$$

Note that fitting this model requires a two-stage approach. First, $\xi_0(s)$ needs to be fitted from the baseline point pattern alone and then it must be plugged-in to the estimation of $\log(\lambda_j(s))$. Because R-INLA is not capable of this type of modeling, we will use an estimate of the terms in $\log(\lambda_0(s))$ so that we will be fitting

$$\begin{aligned} \log(\lambda_j(s)) &= \log\left(\frac{n_j}{n_0}\right) + \hat{\alpha}_0 + \hat{\boldsymbol{\beta}}\mathbf{x}(s) + \hat{\xi}_0(s) + \xi_j(s) = \\ &= \log\left(\frac{n_j}{n_0}\right) + \log(\hat{\lambda}_0(s)) + \xi_j(s) = \\ &= \log\left(\frac{n_j}{n_0}\hat{\lambda}_0(s)\right) + \xi_j(s). \end{aligned} \quad (12)$$

for the other point patterns. Also note that the first term in the last line of the previous equation is simply an offset in the model. Note that $\hat{\lambda}_0(s)$ can be easily obtained at any grid location and at the other tornado locations by making a prediction once we have fitted the model.

Although this new model is just an approximation to the actual model, it will provide a convenient way of separating the estimation of the large scale variation - measured by $\log\left(\frac{n_j}{n_0}\hat{\lambda}_0(s)\right)$ - from local effects (measured by $\xi_j(s)$). Diggle et al. (2007) discuss the issue of disentangling first and second-order variation and how it is not possible to do so unless we impose further assumptions. We have done so when including $\log\left(\frac{n_j}{n_0}\hat{\lambda}_0(s)\right)$ in the model.

2.3. Testing for significance of the spatial effects

In the models that we have described so far, the different spatial effects that appear are estimated using SPDEs. In order to assess whether a spatial effect is significantly different from zero we could consider the credible intervals provided at locations on a grid, and whether zero is contained in their interval. From this, we could produce a map highlighting the points in the grid where the spatial effect is different from zero.

A similar approach can be followed if the interest is in comparing the spatial distribution of two types of patterns. Now, we can take the difference between two SPDEs and make inference on the resulting effect, i.e., inference is based now on the difference $\xi_i(s) - \xi_j(s)$, $i \neq j, i, j = 1, \dots, K$. Significant departures from zero will indicate regions where the spatial variation of the patterns is different.

3. INLA and SPDE for spatial point pattern

For making inference with the LGCP introduced in Section 2, we assume that the Gaussian random field of Equation (1) is defined by a Matérn covariance function given by

$$\text{Cov}(\xi(s_i), \xi(s_j)) = \text{Cov}(\xi_i, \xi_j) = \frac{\sigma^2}{\Gamma(\lambda)2^{\lambda-1}} (\kappa\|s_i - s_j\|)^\lambda K_\lambda(\kappa\|s_i - s_j\|), \quad (13)$$

where $\|s_i - s_j\|$ is the Euclidean distance between two generic locations $s_i, s_j \in \Omega$ and σ^2 is the marginal variance. The term $K_\lambda(\cdot)$ denotes the modified Bessel function of second kind and order $\lambda > 0$, which measures the degree of smoothness of the process and is usually kept fixed due to poor identifiability. Conversely, $\kappa > 0$ is a scaling parameter related to the range r , i.e., the distance at which the spatial correlation becomes almost null.

Typically, the empirically derived definition for the range is $r = \frac{\sqrt{8\lambda}}{\kappa}$ (see Section 2 in Lindgren et al., 2011), with r corresponding to the distance at which the spatial correlation is close to 0.1, for each $\lambda \geq 1/2$.

Given the point pattern data $\mathbf{y} = (y(s_1), \dots, y(s_N))$ the log-likelihood is given by

$$\log(p(\mathbf{y} | \boldsymbol{\xi})) = |\Omega| - \int_{\Omega} \exp(\xi(s)) ds + \sum_{i=1}^N \xi(s_i) \quad (14)$$

which is an intractable function due to the presence of a stochastic integral. The usual approach for inference (Diggle et al., 2013; Møller and Waagepetersen, 2007) is based on the construction of a fine regular lattice defined over the spatial domain Ω and on the count of number of points y_{ij} for each grid cell s_{ij} : this gives rise to independent Poisson random variables with mean Λ_{ij} depending on the Gaussian field $\xi(s)$ through the relationship $\Lambda_{ij} = \int_{s_{ij}} \lambda(s) ds$. This integral can be approximated by $\Lambda_{ij} \approx |s_{ij}| \exp(\xi_{ij})$, where $|s_{ij}|$ is the cell area and ξ_{ij} is the value of the zero-mean multivariate Gaussian random variable $\boldsymbol{\xi}$, whose dimension is given by the grid resolution, with (dense) covariance matrix defined by the Matérn function of Equation (13).

3.1. Integrated Nested Laplace Approximation

In the Bayesian framework, inference for this kind of modeling is usually performed through MCMC methods but giving rise to computationally challenging tasks, as discussed in Diggle et al. (2013). Recently, a computationally effective alternative to MCMC has been introduced by Rue et al. (2009) with the name of Integrated Nested Laplace Approximation (INLA). INLA, which is not a simulation-based algorithm, is designed for the class of latent Gaussian models, where the response variable y_i observed for the i -th unit is assumed to belong to a distribution family (not necessarily part of the exponential family) characterized by a parameter ϕ_i (usually the mean) which is linked to a structured additive predictor η_i through a link function $g(\cdot)$, such that $g(\phi_i) = \eta_i$. A general way for specifying the linear predictor η_i is the following

$$\eta_i = \beta_0 + \sum_{m=1}^M \beta_m x_{mi} + \sum_{l=1}^L f_l(z_{li}), \quad (15)$$

where β_0 is a scalar representing the intercept, the coefficients $\boldsymbol{\beta} = (\beta_1, \dots, \beta_M)$ quantify the (linear) effect of some covariates $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_M)$ on the response, and $\boldsymbol{f} = \{f_1(\cdot), \dots, f_L(\cdot)\}$ is a collection of functions defined in terms of a set of covariates $\boldsymbol{z} = (\boldsymbol{z}_1, \dots, \boldsymbol{z}_L)$. The terms $f_l(\cdot)$ can assume different forms such as smooth and non-linear effects of covariates, time trends and seasonal effects, random intercept and slopes as well as temporal or spatial random effects. For this reason, the class of latent Gaussian models can accommodate a wide range of models including the LGCP introduced in Section 2. By default, a Gaussian prior is assigned to α , $\boldsymbol{\beta}$ and \boldsymbol{f} , and then all the latent Gaussian components are collected in the vector of parameters (or latent field) $\boldsymbol{\theta} = \{\alpha, \boldsymbol{\beta}, \boldsymbol{f}\}$, which is a function of some hyper-parameters $\boldsymbol{\psi} = (\psi_1, \dots, \psi_K)$.

The objectives of Bayesian inference are the marginal posterior distributions for each element of the parameter and hyper-parameter vector denoted by $p(\theta_i | \boldsymbol{y})$ and $p(\psi_k | \boldsymbol{y})$, respectively. INLA provides accurate approximations to these distributions in a short computing time. More details regarding INLA can be found in Martins et al. (2013), Simpson et al. (2012a), Simpson et al. (2012b) and Gómez-Rubio et al. (2014).

3.2. Fitting spatial point patterns with SPDE

Illian et al. (2012a) and Illian et al. (2013) use INLA for fitting complex spatial point process and describe some applications regarding ecological data. In particular, they fit a log-Gaussian Cox model with the grid-based approach described previously, where the observed number of points for each cell s_{ij} is distributed as

$$y_{ij} | \eta_{ij} \sim \text{Poisson}(|s_{ij}| \exp(\eta_{ij})), \quad (16)$$

where η_{ij} is the linear predictor of Equation (15) which includes a spatially structured effect for modeling the discrete spatial process at the grid level (i.e. a conditional autoregressive model) which approximates the continuous spatial process $\xi(s)$. In this paper instead, we follow the alternative inference strategy adopted in Simpson et al. (2013), where the Gaussian field $\xi(s)$ is approximated through the SPDE method introduced by Lindgren et al. (2011). A similar approach has been used also in Paula et al. (2013) for modeling annual wildfire risk.

Briefly, the SPDE approach represents the continuous Matérn Gaussian field as a discrete indexed Gaussian Markov random field (GMRF, Rue and

Held, 2005), which is characterized by a sparse precision matrix and enjoys computational benefits in terms of fast inference. This representation is based on the following finite combination of piecewise linear functions defined over a triangulation of the domain Ω with G vertexes

$$\xi(s) = \sum_{g=1}^G \varphi_g(s) \tilde{\xi}_g, \quad (17)$$

where $\{\varphi_g\}$ is the set of deterministic basis functions and $\{\tilde{\xi}_g\}$ are zero-mean Gaussian distributed weights. Lindgren et al. (2011) show that the vector of basis weights $\tilde{\boldsymbol{\xi}} = (\tilde{\xi}_1, \dots, \tilde{\xi}_G)$ is a GMRF with sparse precision matrix depending on the Matérn parameters κ and σ^2 , for $\alpha = 1, 2, 3, \dots$ where $\alpha = \lambda + 1$ (recall that λ is the smoothness parameter).

Given the SPDE representation the log-likelihood in Equation (14) can be rewritten as

$$\log \left(p \left(\mathbf{y} \mid \tilde{\boldsymbol{\xi}} \right) \right) \approx |\Omega| - \sum_{j=1}^p \tilde{\alpha}_j \exp \left(\sum_{g=1}^G \varphi_g(\tilde{s}_j) \tilde{\xi}_g \right) + \sum_{i=1}^N \sum_{g=1}^G \varphi_g(s_i) \tilde{\xi}_g, \quad (18)$$

where the second term on the right hand side derives from a numerical approximation of the stochastic integral by using p integration points $\tilde{s}_1, \dots, \tilde{s}_p$ with weights $\tilde{\alpha}_1, \dots, \tilde{\alpha}_p$. As shown in Simpson et al. (2013), this log-likelihood can be referred to the standard Poisson case. In particular, by introducing some fake observations (equal to 0) for the p integration points and exponentiating Equation (18), we obtain the following likelihood function

$$p \left(\mathbf{y} \mid \tilde{\boldsymbol{\xi}} \right) \approx c \prod_{i=1}^{p+N} \eta_i^{y_i} \exp(-\alpha_i \eta_i) \quad (19)$$

which is the product of $p+N$ conditionally independent Poisson distributions with mean equal to $\alpha_i \eta_i$. In Equation (19), c is a constant, y_i is the i -th element of the observation vector \mathbf{y} defined by assigning 0 for each integration point and 1 to the actual N locations, as follows:

$$\mathbf{y} = \underbrace{(0, \dots, 0)}_{p \text{ times}}, \underbrace{(1, \dots, 1)}_{N \text{ times}}.$$

In the same way, α_i is the generic element of the weight vector $\boldsymbol{\alpha}$ defined as

$$\boldsymbol{\alpha} = (\tilde{\alpha}_1, \dots, \tilde{\alpha}_p, \underbrace{0, \dots, 0}_{N \text{ times}}),$$

and finally η_i comes from the following vector

$$\boldsymbol{\eta} = \exp \left(\left(\mathbf{A}_1 \tilde{\boldsymbol{\xi}} \right)', \left(\mathbf{A}_2 \tilde{\boldsymbol{\xi}} \right)' \right)',$$

where \mathbf{A}_1 and \mathbf{A}_2 are the $(p \times G)$ - and $(N \times G)$ -dimensional matrices with elements $\varphi_g(\tilde{s}_i)$ and $\varphi_g(s_i)$, respectively. In Section 4.3 we will describe how to implement this using INLA and its associated R package R-INLA.

4. Example: Tornado in the United States

4.1. Introduction

The Storm Prediction Center provides accurate information on the occurrence of tornados in the United States of America. We will use this dataset¹ to illustrate the use and application of the new approach described before. The dataset is available as a shapefile that contains the tornado trajectories from 1950 to 2013. We will focus on the strength of the tornado as measured in the Fujita Tornado Damage Scale². This scale divides the tornados in 6 groups from 0 to 5, with 0 being the weakest tornado type and 5 the strongest.

This dataset has been previously analysed by several authors. For example, Karpman et al. (2013) propose a number of spatio-temporal models, implemented using a nonparametric kernel-based approach, to study tornado point pattern by accounting for topographic variation and performing model selection using the Akaike information criterion (AIC). Elsner et al. (2013) propose a Poisson point process on the non-violent tornadoes to predict violent tornados during springtime across the US central Great Plains. In addition to predict violent tornadoes using the nearest distance to non-violent tornado, they correct for population bias (i.e. there are fewer tornado reports in lower populated areas) by including the distance to the nearest city. In Elsner et al. (2014) the tornado intensity is assumed to be distributed like a Weibull with log-mean depending linearly on the path length and width which are strongly correlated to the Fujita categories. The relationship of the length and width of tornadoes to the intensity is also analyzed in Akers et al. (2014) by means of a multinomial logistic model (without spatial random

¹Available from <http://www.spc.noaa.gov/gis/svrgis/zipped/tornado.zip>

²See <http://www.spc.noaa.gov/faq/tornado/f-scale.html>

effects) to calculate the odds that a particular tornado with a given Fujita category occurs. Wikle and Anderson (2003) consider a Bayesian hierarchical spatio-temporal model on the tornado dataset based on count data. In particular, they employ a zero-inflated Poisson likelihood in order to model the excess of zeros. The mean of the Poisson process is then modelled using different spatial and non-spatial random effects.

We have used the R software (R Core Team, 2014) to conduct all the analysis described in this section. We will only consider tornados whose starting point has occurred inside the Contiguous United States. Boundaries have been obtained from the U.S. Census Bureau (<http://www.census.gov/>) at the state level and then internal boundaries dissolved by means of R package `mapproj` (Bivand and Lewin-Koh, 2014). This leaves a total of 56908 tornados in 6 categories, according to the magnitude of the tornado. The number of tornadoes ranges from 26088 (scale 0) to 59 (scale 5), and the number decreases monotonically as the scale number increases. The locations of the starting points of the tornados are plotted in Figure 1 according to their intensity. It can be seen that most of the tornados occur in the central and eastern part of US. Moreover, while some non-violent events (with magnitude lower than 4) take place in western areas, destructive tornados (with magnitude 4 or 5) are located just in the midwest and in the south.

4.2. Modeling details

We have fitted several of the models described in Section 2 to the tornado data. First of all, we have considered a separate model for each tornado type with 6 different intensity models (named MODEL 1A) given by

$$\log(\lambda_j(s)) = \alpha_j + \xi_j(s), \quad j = 0, \dots, 5 \quad (20)$$

where α_j is the intercept and $\xi_j(s)$ is a spatial term that we have modelled using a SPDE latent effect. We extend this case by including also a linear effect on the altitude, giving rise to the following model (named MODEL 1B):

$$\log(\lambda_j(s)) = \alpha_j + \beta_j \mathbf{x}(s) + \xi_j(s), \quad j = 0, \dots, 5. \quad (21)$$

Note that since these models are fitted independently for each type of point pattern, we allow the covariates' coefficients to be different for different values of j .

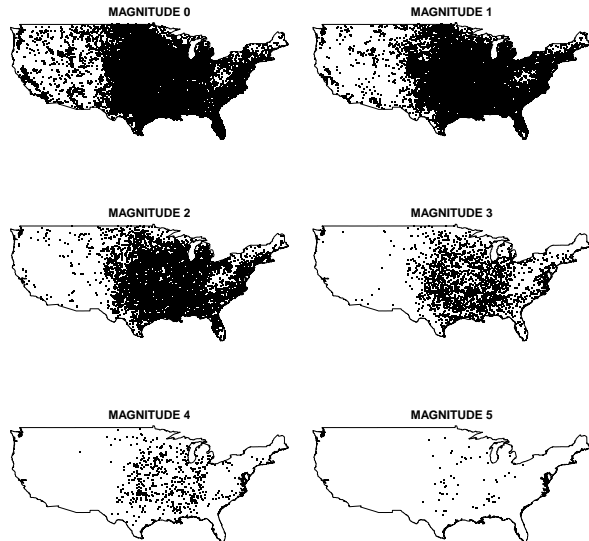


Figure 1: Map of the tornado locations according to the magnitude.

Then we consider the case, denoted by MODEL 2A, given by Equation (12) but with $\hat{\beta} = \mathbf{0}$ and $\xi_j(s) = 0$. This means that the tornados with magnitude 0 represents the baseline point pattern which is shared by all the other marks. The same model is then implemented by including the covariate term, as stated in Equation (12), and it will be denoted as MODEL 2B. So both models can be expressed as

$$\log(\lambda_j(s)) = \log\left(\frac{n_j}{n_0}\right) + \log(\hat{\lambda}_0(s)), \quad j = 1, \dots, K \quad (22)$$

with the difference that MODEL 2A does not include the term on the covariates in $\hat{\lambda}_0(s)$.

The last models we fit (MODEL 3A and 3B) are given by Equation (12), that may include covariates and specific spatial variation through $\xi_j(s)$, to assess different spatial distributions for each magnitude. As in the previous cases, MODEL 3B is MODEL 3A plus a linear term on the covariates (altitude, in our case).

All models are compared by using the Deviance Information Criterion (DIC) given by the sum of two components, one for quantifying the model fit and the other for evaluating the complexity of the model (the latter measured

through the effective number of parameters). Note that since models 2A and 2B are based on models 1A and 1B, respectively, fitted for the baseline point pattern we have not included the DIC.

4.3. Implementation details

To estimate models of type 1A/1B, 2A/2B and 3A/3B introduced in the previous section, we have used the SPDE approach discussed in Section 3.2 and implemented through R-INLA (Rue et al., 2014; Lindgren and Rue, 2015). In particular, the mesh used to compute the SPDE is available in Figure 2. It covers the study area at a reasonable resolution to estimate the SPDE effects. We have considered the same mesh when estimating the different SPDE effects as they all occurred within the same study region. As we will see in Section 4.4, we believe that this mesh is good because we obtain very similar estimates of the intensity between kernel smoothing and model 1A.

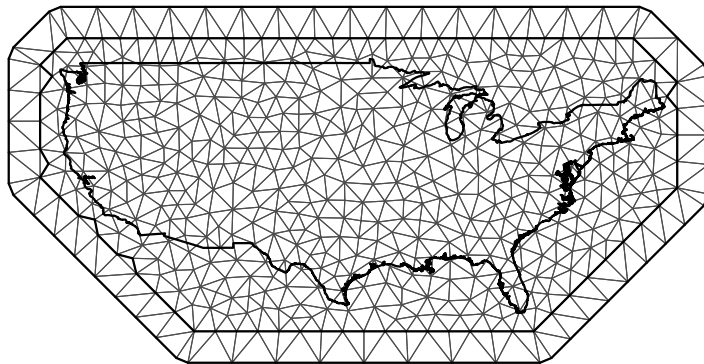


Figure 2: Mesh over the Contiguous United States that has been used to compute SPDE spatial effects.

As regards the integration points discussed in Section 3.2, we use the approach described in Krainski and Lindgren (2014): the integration points coincide with the mesh vertices with weights given by the relative area of the polygons obtained through a Voronoi tessellation.

To implement models 2A/2B we have simply taken models 1A/1B fitted on the baseline point pattern and then used a prediction at the locations of the other point patterns. Then, we have re-scaled the predictions to match the scale of the intensity, as shown in Equation (22).

We have included these predictions as an offset to implement models 3A/3B, that also include a specific spatial term using SPDEs. Hence, the offset will account for the large scale variation whilst we hope that the SPDE will account for the small scale variation.

When estimating the SPDE we have further imposed the condition to integrate-to-zero. This will help to make the model identifiable and will provide a more stable model fitting. We have also reduced the value of the step-length for the gradient calculations for the hyperparameters and set it to 0.001 to avoid numerical problems.

4.4. Results

An initial point pattern analysis has been conducted using the R package `spatstat` (Baddeley and Turner, 2005). First of all, we have estimated the intensity of each point pattern using a kernel smoothing based on a bivariate Normal distribution with bandwidth equal to 1.6. This bandwidth has been obtained by using the `bw.scott` function, that implements the bandwidth method proposed in Scott (1992). This criterion provides a bandwidth for each dimension and we have taken the largest of these two values. Other methods for the selection of the bandwidth were not feasible given the large number of points for low magnitudes. The estimated log-intensity, separately for each point pattern, is shown in Figure 3. We have also included the estimate of the intensity using independent SPDE models (MODEL 1A) for each point pattern. It can be seen how both approaches provide very similar estimates of the intensity. The kernel smoothing provides very small estimates in the regions with no points. These results also confirm that our mesh is good enough to estimate the intensity of the different point patterns.

Tables 1 and 2 contain a summary of the parameter estimates for models 1A and 1B as well as the DIC and the effective number of parameters.

Regarding the spatial variation account for the smooth spatial term, Figure 4 shows the point estimates and 95% credible intervals for the spatial

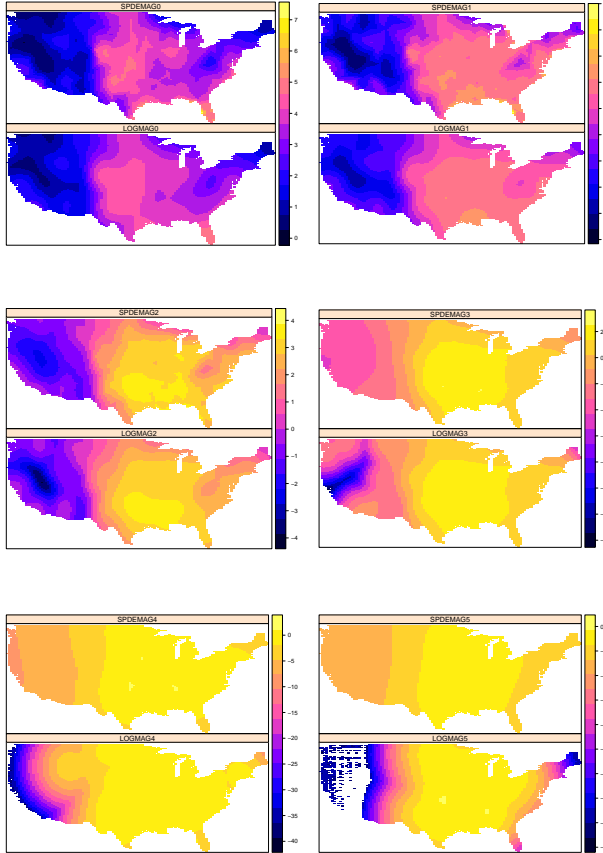


Figure 3: Intensities (log-scale) of the marked patterns using kernel smoothing and SPDE with MODEL 1A.

effects at the mesh vertices that lay inside the Contiguous United States. In order to visualise the spatial trend better, we have sorted the value in increasing order according to the point estimates. This will place the points with negative values on the left and values with positive values on the right.

According to the credible intervals, it is clear that the SPDEs pick some of the spatial variation present in the data. Credible intervals are also wider for the patterns with a smaller number of points. This is not surprising as estimating the spatial distribution is harder when only a few points are present. For example, we only have 59 tornados of Magnitude 5 and, in this case, credible intervals are very wide with only a few points of the mesh

	α (mean)	α (sd)	σ^2 (mean)	σ^2 (sd)	r (mean)	r (sd)	DIC	P.EFF
MAG0	3.24	0.12	41.04	56.17	88.86	53.85	-155216.20	216.14
MAG1	2.94	0.20	4.30	1.48	14.25	2.66	-102509.93	248.76
MAG2	1.22	0.14	4.44	2.92	33.43	11.29	-33597.80	154.22
MAG3	-0.77	0.17	114.12	390.32	200.67	242.02	-3413.16	71.42
MAG4	-3.03	0.44	569.05	2865.89	290.77	478.75	704.88	45.47
MAG5	-5.58	1.00	88.61	257.46	124.30	159.95	337.43	21.11

Table 1: Summary statistics: posterior mean and standard deviation (sd) for the parameters of MODEL 1A, DIC and effective number of parameters (P.EFF).

	α (mean)	α (sd)	β (mean)	β (sd)	σ^2 (mean)	σ^2 (sd)	r (mean)	r (sd)	DIC	P.EFF
MAG0	27.27	0.39	-0.06	0.0005	6352.72	7.07	415.62	0.20	-209355.65	240.81
MAG1	9.33	0.25	-0.02	0.0003	566143.08	3637854.07	2523.98	5053.89	-111157.95	223.86
MAG2	14.30	0.65	-0.03	0.0007	2384343.80	14860179.93	1981.09	4098.49	-38862.33	228.53
MAG3	0.05	0.20	-0.002	0.0002	21.89	44.82	105.20	90.03	-3466.91	72.87
MAG4	-1.90	0.43	-0.002	0.0005	77.12	246.94	139.34	163.16	693.36	45.10
MAG5	-4.36	1.11	-0.002	0.001	15.84	26.62	62.35	58.20	338.37	21.05

Table 2: Summary statistics: posterior mean and standard deviation (sd) for the parameters of MODEL 1B, DIC and effective number of parameters (P.EFF).

showing a significant (positive) spatial effect. The effect of including covariates reduces the width of the credible intervals when the number of points in the pattern is large, but it seems to have no impact on the patterns with a smaller number of points.

With regards to the results shown in Table 1 and Table 2 the values of the spatial range are usually smaller for MODEL 1A (at least for the magnitudes 0-2). This means that the spatial term is picking the small scale spatial variation in MODEL 1A, but that this spatial variation is accounted for by the covariate, and so the ranges in MODEL 1B are larger, i.e., the scale of the spatial effect is larger. Furthermore, the coefficient of the altitude is negative with a 95% credible interval that leaves the null value out in all cases but for tornadoes of magnitude 5. We believe this happens because of the small number of points which is not enough to obtain accurate estimates of the coefficient.

Table 3 and Table 4 show the estimates of the parameters of the SPDE spatial term for MODEL 3A and MODEL 3B, respectively. These models only include the baseline point pattern, through an offset with the predicted intensity using MODEL 1A and MODEL 1B respectively, plus a specific SPDE spatial effect. For tornadoes of Magnitude 1 the baseline pattern does not seem to account for small scale variations as the estimates of the range of the spatial smooth term indicate that it is picking small scale spatial variation. On the other hand, including the prediction in the model seems to be important because now the estimates of the parameters indicate that

	σ^2 (mean)	σ^2 (sd)	r (mean)	r (sd)	DIC	P.EFF
MAG1	1.20	0.73	32.24	10.96	-102790.41	150.13
MAG2	335.86	753.70	296.22	390.76	-33786.87	162.96
MAG3	275.23	117.62	102.22	37.61	-3327.26	151.99
MAG4	1492.22	1894.93	280.58	309.96	836.15	129.56
MAG5	763.01	570.17	66.75	24.37	458.32	88.74

Table 3: Summary statistics: posterior mean and standard deviation (sd) for the parameters of MODEL 3A, DIC and effective number of parameters (P.EFF).

	σ^2 (mean)	σ^2 (sd)	r (mean)	r (sd)	DIC	P.EFF
MAG1	8.64	0.13	4.65	0.02	-111572.25	288.19
MAG2	4.76	0.93	6.04	0.88	-35875.66	247.78
MAG3	259.80	160.21	100.41	70.14	-3706.02	159.05
MAG4	1505.04	2064.36	278.77	316.72	750.68	128.61
MAG5	5277.91	11100.82	738.68	956.84	465.92	92.68

Table 4: Summary statistics: posterior mean and standard deviation (sd) for the parameters of MODEL 3B, DIC and effective number of parameters (P.EFF).

the spatial term is accounting for large scale spatial variation. This is seen by noticing how the range estimates are higher than in MODEL 1A for magnitudes 1 and 2. Opposite results are observed for magnitudes 3-5 and this could be an effect of the decreasing number of tornados.

Figure 5 shows the estimates of the SPDE effects for MODEL 3A and MODEL 3B for all types of point patterns but the baseline one. For tornados of Magnitude 1 it seems reasonable to use the ones with Magnitude 0 as a baseline, as the residual spatial variation is very small. For all the other types of tornado there is clear evidence that the baseline pattern does not account for all the spatial variation, as many credible intervals do not include the null value. In fact, point estimates are larger than for the case of MODEL 1A and MODEL 1B, which means that there is important spatial difference between these tornado types and tornados of Magnitude 0.

4.5. Testing for equal spatial distribution

In Section 2.3 we have discussed some ideas to assess whether the spatial distribution of two point patterns is the same. In particular, when using MODEL 1A it could be possible to compute the difference between the spatial smooth terms used to estimate the models and check whether credible intervals contain the null value. Hence, we can plot the estimates of $\xi_i(s) - \xi_j(s), i \neq j$ at the mesh points, that provide a summary of the estimates of the spatial distribution, and see whether the two spatial terms differ.

Note that this implies computing 15 differences. In order to provide a summary of the difference we have computed point estimates and 95%

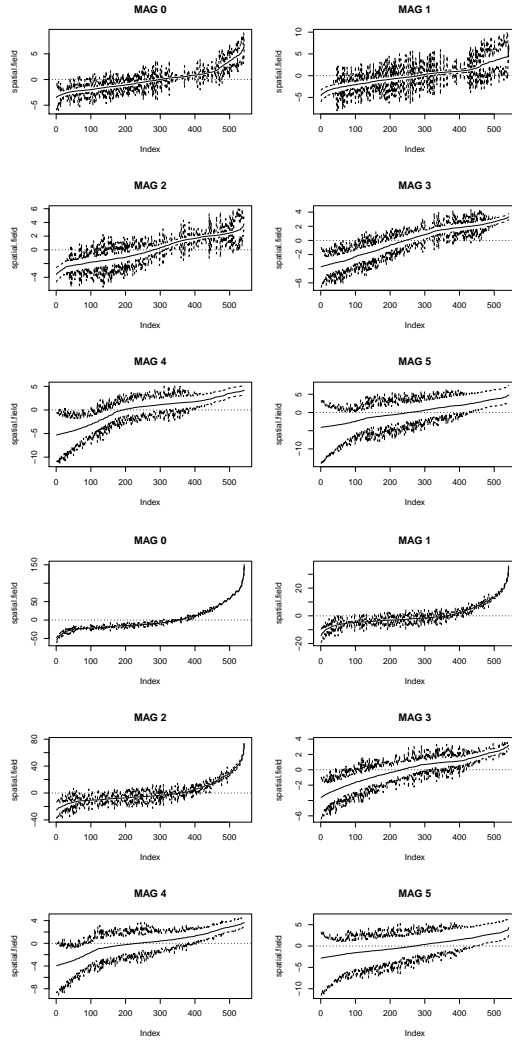


Figure 4: Point estimates and 95% credible intervals for the spatial effects using SPDEs for MODEL 1A (top) and MODEL 1B (bottom).

credible intervals at the mesh points. In Table 5 we have summarised the results by including the proportions of points at which the 95% credible intervals contain the null value. Also, Figure 6 shows the pairs of marks with the highest and least agreement.

The two marks that have the least agreement are Magnitudes 0 and 4. As it can be seen in Figure 6 (left plot), there are clearly a few points where the

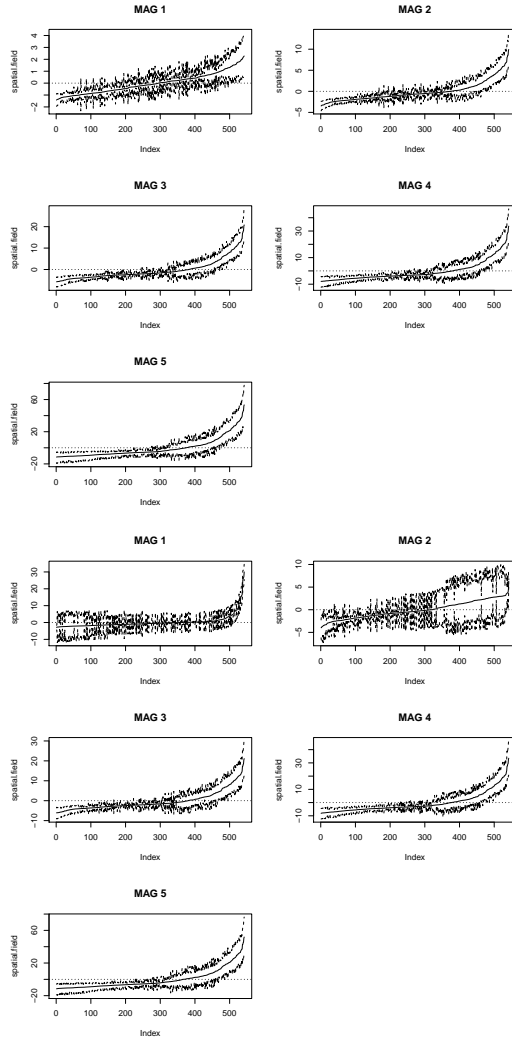


Figure 5: Point estimates and 95% credible intervals for the spatial effects using SPDEs for MODEL 3A (top) and MODEL 3B (bottom).

difference is either too low or too high. On the other hand, the two marks that have the highest agreement are Magnitudes 3 and 4. As it can be seen in the same figure, now it looks like most of the credible intervals contain the zero value. These results can be used to argue about the convenience of using a baseline point pattern to model all the other types of tornadoes.

Finally, we have produced some maps with type-specific probabilities as

	MAG0	MAG1	MAG2	MAG3	MAG4	MAG5
MAG0	1.00	0.75	0.72	0.59	0.58	0.74
MAG1	0.75	1.00	0.72	0.63	0.68	0.78
MAG2	0.72	0.72	1.00	0.84	0.78	0.92
MAG3	0.59	0.63	0.84	1.00	0.95	1.00
MAG4	0.58	0.68	0.78	0.95	1.00	1.00
MAG5	0.74	0.78	0.92	1.00	1.00	1.00

Table 5: Proportion of points at the mesh for which 95% credible intervals of the difference of the spatial effects contains the null value.

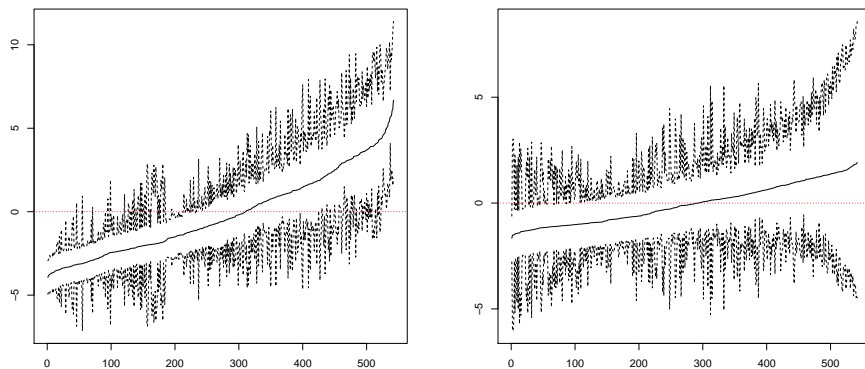


Figure 6: Difference of the spatial effects, for the two marks with the highest (Magnitudes 3 and 4, right) and least (Magnitudes 0 and 4, left) agreement.

in Equation (3). Instead of computing the posterior distribution of each probability surface $p_j(s)$, $j = 0, \dots, 5$, we have used posterior means of the intensities (from MODEL 1A in this case) at the points of a grid. The resulting plots can be seen in Figure 7. As expected, the tornadoes with the lowest magnitudes are the ones with the highest probabilities. Nevertheless, the spatial distributions of the tornadoes with the highest magnitudes can be seen in the plots. For example, tornadoes of magnitude 5 are more likely to be observed in the central-east part of the United States.

5. Discussion

In this paper we have considered the analysis of marked point patterns using recent methodological development, namely, the Integrated Nested

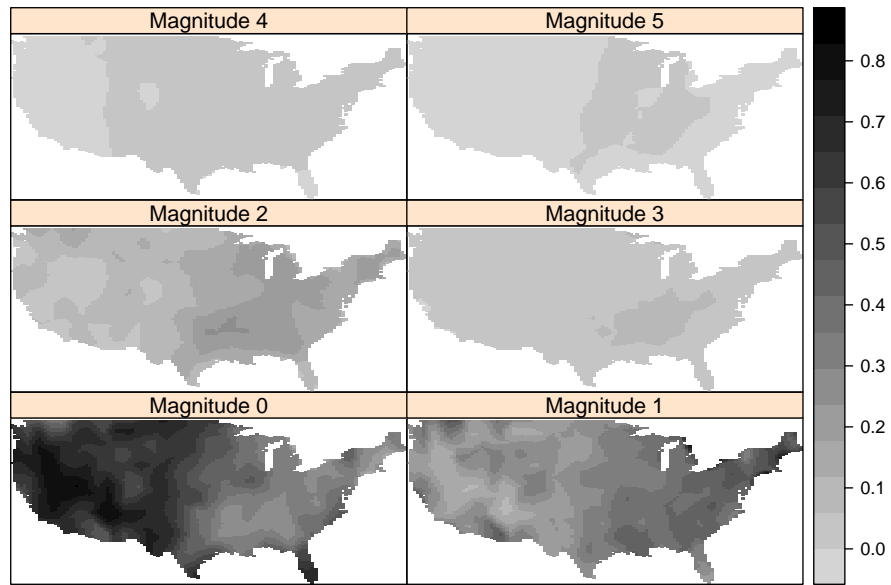


Figure 7: Conditional probabilities of finding each type of tornado obtained from MODEL 1A.

Laplace Approximation and spatial smooth effects estimated with Stochastic Partial Differential Equations. We have described how these new developments can be used to estimate the intensity of point patterns using some recent works and we have tackled the problem of comparing the spatial distributions of point with different marks.

Our approach considers a point pattern (with a particular mark) as a baseline for the remainder of the types of point patterns. We have proposed different types of models to account for covariates and to use an estimate of the baseline point pattern to build the intensity of the other point patterns. This in addition provides a way of assessing whether this is a reasonable

modelling or not by looking at the credible intervals of the spatial smooth term point estimates and credible intervals.

Furthermore, we have shown an application of these methods in the analysis of the spatial distribution of tornados. Being able to predict the intensity of a tornado would allow for more effective forecasting and increased weather safety in extreme storms.

Although we have only focused on the spatial nature of the data in our example, it is very easy to extend the methodology presented in this paper to include a (separable) temporal trend. However, including a separable temporal effect will only influence the number of tornados in time and it is not of interest in this case. An interesting new direction is that of having a spatio-temporal interaction term but it is not evident how this can be combined with our SPDE approach.

Acknowledgements

Virgilio Gómez-Rubio has been supported by the University of Bergamo through the Italy[®] 2013 project (Grants for Visiting Professors). Michela Cameletti and Francesco Finazzi have been supported by the FIRB Project StEPHl (project no. RBFR12URQJ, <http://stephiproject.it/>) provided by the Italian Ministry for Education, University and Research. We would like to thank Håvard Rue and Finn Lindgren for helpful comments on this work.

Bibliography

- Akers, C. M., Smith, N. J., Shifa, N., 2014. Multinomial logistic regression model for predicting tornado intensity based on path length and width. *World Environment* 4 (2), 61–66.
- Baddeley, A., Turner, R., 2005. *spatstat*: An R Package for Analyzing Spatial Point Patterns. *Journal of Statistical Software* 12 (6), 1–42.
- Bivand, R., Gómez-Rubio, V., Rue, H., 2014. Approximate bayesian inference for spatial econometrics models. *Spatial Statistics* 9, 146–165.
- Bivand, R., Lewin-Koh, N., 2014. *maptools*: Tools for reading and handling spatial objects. R package version 0.8-30.

- Bivand, R. S., Pebesma, E. J., Gómez-Rubio, V., 2013. *Applied Spatial Data Analysis with R*, 2nd Edition. Springer.
- Blangiardo, M., Cameletti, M., Baio, G., Rue, H., 2013. Spatial and spatio-temporal models with R-INLA. *Spatial and Spatio-temporal Epidemiology* 4, 33 – 49.
- Cameletti, M., Lindgren, F., Simpson, D., Rue, H., 2013. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Advances in Statistical Analysis* 97 (2), 109–131.
- Cosandey-Godin, A., Teixeira Krainski, E., Worm, B., Mills Flemming, J., 2014. Applying Bayesian spatio-temporal models to fisheries bycatch in the Canadian Arctic. *Canadian Journal of Fisheries and Aquatic Sciences* 0, null.
- Diggle, P., Zheng, P., Durr, P., 2005. Nonparametric estimation of spatial segregation in a multivariate point process: bovine tuberculosis in cornwall, uk. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 54 (3), 645–658.
- Diggle, P. J., Gómez-Rubio, V., Brown, P. E., Chetwynd, A. G., Gooding, S., 2007. Second-order analysis of inhomogeneous spatial point processes using case-control data. *Biometrics* 63 (2), pp. 550–557.
- Diggle, P. J., Moraga, P., Rowlingson, B., Taylor, B. M., 11 2013. Spatial and spatio-temporal log-gaussian cox processes: Extending the geostatistical paradigm. *Statistical Science* 28 (4), 542–563.
- Elsner, J., Jagger, T., Elsner, I., 2014. Tornado intensity estimated from damage path dimensions. *PLoS ONE* 9 (9), n/a–n/a.
- Elsner, J., Murnane, R., Jagger, T., Widen, H., 2013. A Spatial Point Process Model for Violent Tornado Occurrence in the US Great Plains. *Mathematical Geosciences* 45 (6), 667–679.
- Gelfand, A., Diggle, P., Fuentes, M., Guttorp, P. (Eds.), 2010. *Handbook of Spatial Statistics*. Chapman & Hall.
- Gómez-Rubio, V., Bivand, R. S., Rue, H., 2014. Spatial models using laplace approximation methods. In: Fischer, M. M., Nijkamp, P. (Eds.), *Handbook of Regional Science*. Springer-Verlag, New York, pp. 1401–1417.

- Grilli, L., Metelli, S., Rampichini, C., 2014. Bayesian estimation with integrated nested laplace approximation for binary logit mixed models. *Journal of Statistical Computation and Simulation*.
URL <http://dx.doi.org/10.1080/00949655.2014.935377>
- Illian, J., Sørbye, S., Rue, H., 2012a. A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). *Journal of the Royal Statistical Society: Series B* 6 (4), 1499–1530.
- Illian, J., Sørbye, S. H., Rue, H., Hendrichsen, D. K., 2012b. Using inla to fit a complex point process model with temporally varying effects - a case study. *Journal of Environmental Statistics* 3 (7), 1–25.
- Illian, J. B., Martino, S., Sørbye, S. H., Gallego-Fernández, J. B., Zunzunegui, M., Esquivias, M. P., Travis, J. M. J., 2013. Fitting complex ecological point process models with integrated nested laplace approximation. *Methods in Ecology and Evolution* 4 (4), 305–315.
- Ingebrigtsen, R., Lindgren, F., Steinsland, I., 2014. Spatial models with explanatory variables in the dependence structure. *Spatial Statistics* 8 (0), 20 – 38.
- Jona Lasinio, G., Mastrantonio, G., Pollice, A., 2012. Discussing the “big n problem”. *Statistical Methods and Applications*, 1–16.
- Karpman, D., Ferreira, M. A., Wikle, C. K., 2013. A point process model for tornado report climatology. *Stat* 2 (1), 1–8.
- Knorr-Held, L., Best, N. G., 2001. A shared component model for detecting joint and selective clustering of two diseases. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 164 (1), pp. 73–85.
- Krainski, E. T., Lindgren, F., 2014. The R-INLA tutorial: SPDE models. R-INLA.
URL <http://www.math.ntnu.no/inla/r-inla.org/tutorials/spde/spde-tutorial.pdf>
- Liang, S., Carlin, B. P., Gelfand, A. E., 2009. Analysis of minnesota colon and rectum cancer point patterns with spatial and nonspatial covariate information. *The Annals of Applied Statistics* 3 (3), 943–962.

- Lindgren, F., Rue, H., 2015. Bayesian spatial modelling with r-inla. *Journal of Statistical Software* 63 (19), ??–??
- Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach (with discussion). *Journal of the Royal Statistical Society: Series B* 73 (4), 423–498.
- Martins, T. G., Simpson, D., Lindgren, F., Rue, H., 2013. Bayesian computing with INLA: New features. *Computational Statistics & Data Analysis* 67, 68–83.
- Møller, J., Syversveen, A. R., Waagepetersen, R. P., 1998. Log Gaussian Cox Processes. *Scandinavian Journal of Statistics* 25 (3), 451–482.
- Møller, J., Waagepetersen, R. P., 2007. Modern statistics for spatial point processes. *Scandinavian Journal of Statistics* 34 (4), 643–684.
- Muff, S., Riebler, A., Held, L., Rue, H., Saner, P., 2014. Bayesian analysis of measurement error models using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, n/a–n/a.
URL <http://dx.doi.org/10.1111/rssc.12069>
- Musenge, E., Chirwa, T. F., Kahn, K., Vounatsou, P., 2013. Bayesian analysis of zero inflated spatiotemporal HIV/TB child mortality data through the INLA and SPDE approaches: Applied to data observed between 1992 and 2010 in rural North East South Africa. *International Journal of Applied Earth Observation and Geoinformation* 22 (0), 86 – 98.
- Papoila, A. L., Riebler, A., Amaral-Turkman, A., São João, R., Ribeiro, C. a., Geraldés, C., Miranda, A., 2014. Stomach cancer incidence in Southern Portugal 1998-2006: A spatio-temporal analysis. *Biometrical Journal* 56 (3), 403–415.
- Paula, P., Kamil Feridun, T., Maria Antónia, A. T., Ana, S., José M.C., P., 2013. Quantification of annual wildfire risk; a spatio-temporal point process approach. *Statistica* 73 (1), 55–68.
URL <http://rivista-statistica.unibo.it/article/view/3985>

- R Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
URL <http://www.R-project.org/>
- Rue, H., Held, L., 2005. Gaussian Markov Random Fields. Theory and Applications. Chapman & Hall.
- Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society Series B* 2 (71), 1–35.
- Rue, H., Martino, S., Lindgren, F., Simpson, D., Riebler, A., Krainski, E. T., 2014. INLA: Functions which allow to perform full Bayesian analysis of latent Gaussian models using Integrated Nested Laplace Approximation. R package version 0.0-1406288160.
- Saez, M., Barcelò, M. A., Garcia, A. T., Varga, D., Ocaña Riola, R., Verdoy, P. J., Mateu, J., 2012. Space-time interpolation of daily air temperatures. *Journal of Environmental Statistics* 3 (5), 1–15.
- Scott, D., 1992. Multivariate Density Estimation. Theory, Practice and Visualization. Wiley, New York.
- Simpson, D., Illian, J., Lindgren, F., Sørbye, S., Rue, H., 2013. Going off grid: Computationally efficient inference for log-Gaussian Cox processes. ArXiv e-prints.
- Simpson, D., Lindgren, F., Rue, H., 2012a. In order to make spatial statistics computationally feasible, we need to forget about the covariance function. *Environmetrics* 23 (1), 65–74.
- Simpson, D., Lindgren, F., Rue, H., 2012b. Think continuous: Markovian Gaussian models in spatial statistics. *Spatial Statistics* 1, 16–29.
- Wikle, C. K., Anderson, C. J., 2003. Climatological analysis of tornado report counts using a hierarchical Bayesian spatiotemporal model. *Journal of Geophysical Research: Atmospheres* 108, n/a–n/a.