

# Unemployment estimation: Spatial point referenced methods and models

Soraia Pereira<sup>a,\*</sup>, K F Turkman<sup>a</sup>, Luís Correia<sup>b</sup>, Håvard Rue<sup>c</sup>

<sup>a</sup>*CEAUL-Faculdade de Ciências, University of Lisbon, Portugal.*

<sup>b</sup>*Instituto Nacional de Estatística, Lisboa, Portugal.*

<sup>c</sup>*King Abdullah University of Science and Technology, Saudi Arabia*

---

## Abstract

From the fourth quarter of 2014, the Portuguese Labour Force Survey (PLFS) started geo-referencing the sampling units, namely the dwellings, in which the surveys are carried out. This opens new possibilities in analysing and estimating unemployment and its spatial distribution across any region by employing point referenced methods and models. According to a preestablished sampling criteria, the labour force survey selects a certain number of dwellings from across the nation to study, and establishes the number of unemployed people in each dwelling. Based on this survey, the National Statistical Institute of Portugal presently uses direct estimation methods to estimate the national unemployment figures. Recently however, there has been increased interest in estimating these figures in smaller areas. Due to reduced sampling sizes in small areas, direct estimation methods tend to produce fairly large sampling variations. Therefore, model based methods should be favoured as these tend to borrow strength from area to area by making use of the areal dependence. These model based methods tend to use areal counting processes as models and typically introduce spatial dependence through the model parameters using a latent random effect.

In this paper, we suggest using point referenced models as an alternative to the traditional small area estimation methods for unemployment estimation. Specifically, we model the spatial distribution of residential buildings across

---

\*Corresponding author

*Email address:* [soraia.gpereira@gmail.com](mailto:soraia.gpereira@gmail.com) (Soraia Pereira )

Portugal using a log Gaussian Cox process, and the number of unemployed people per residential unit as a mark attached to these random points. Thus, the main focus of the study is to model the spatial intensity function of this marked point process. The number of unemployed people in any region can then be estimated using a proper functional of this marked point process. The principal objective of this point referenced method for unemployment estimation is to produce reliable estimates at higher spatial resolutions, and at the same time to incorporate into the model any available auxiliary information of the residential units, such as mean age or education level as compared to areal unit averages used in small area estimation.

*Keywords:* Unemployment estimation, spatial marked point processes, Log-Gaussian Cox model, INLA, SPDE, small area estimation

---

## 1. Introduction

The knowledge and understanding of unemployment rates at a regional level has been increasingly used to make political decisions. In Portugal, the official unemployment figures are published quarterly by the National Statistical Institute of Portugal (INE) at the national level as well as for NUTS II regions. NUTS is the classification of territorial units for statistics (see Figure 1 for a better understanding of NUTS regions and the 278 counties in mainland Portugal). The calculation of the official numbers is based on a direct method from the sample of the Portuguese Labour Force Survey (PLFS). This method is an extension of the Horvitz-Thompson estimator (Horvitz and Thompson, 1952), with a correction for non-response and a calibration for the known population totals.

Currently, there is a need to obtain reliable estimates at a more disaggregated level, particularly for the NUTS III regions, or even for the individual counties themselves. However, due to the relatively small size of these areas, there is insufficient information available to obtain estimates with an acceptable level of accuracy using the direct method without altering the dimension of the PLFS.

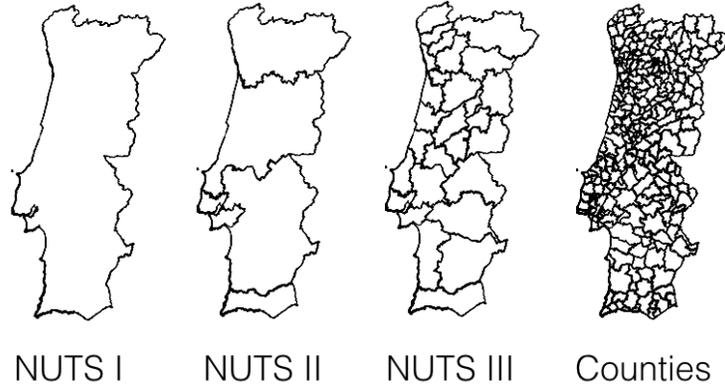


Figure 1: NUTS and counties in mainland Portugal

Small area estimation methods (SAE) provide useful tools for such studies (Rao and Molina, 2015).

The fundamental objective of these methods is to estimate the quantity of interest, such as total unemployed numbers, unemployment proportions or rates of areal units by linking them through a regression model to improve upon the direct estimators. By borrowing strength from covariates and neighbouring units these methods give more precise estimates, particularly for areal units with smaller sample sizes. The basic Fay-Herriot model (1979) achieves this objective by relating the direct estimates of the quantity of interest to relevant auxiliary variables available at areal level in a classical state space representation in two stages. In the observation equation, the direct estimate  $\hat{\theta}_i$  of areal unit  $i$  is expressed as

$$\hat{\theta}_i = \theta_i + e_i, \quad i = 1, 2, \dots, m$$

where the state  $\theta_i$ , or the true latent quantity of interest of the areal unit, is modelled by

$$\theta_i = \mathbf{x}'\boldsymbol{\beta} + v_i,$$

20 where,  $e_i$ ,  $i = 1, 2, \dots, m$  are zero-mean independent random variables often assumed to have normal distribution with known variances, whereas  $v_i$  are assumed to be independent zero-mean Gaussian random variables with unknown

variances, independent of  $e_i$ . On the account that it implicitly includes the sampling mechanism by using the direct estimators in the observation equation, this basic model is bias free and significantly reduces variance, provided that the auxiliary information is informative.

Recently, there have been some quite extensive improvements and extensions to this basic model. You and Chapman (2006) extend this basic model taking into account the extra uncertainty of estimating the sampling error variances, and Marhuenda *et al.* (2013) include spatial, as well as temporal, dependence structures in the model by adding spatially and time varying latent random effects into the state equation. The basic Fay-Herriot model implicitly assumes normality assumption and this may not be adequate if the quantity of interest is the unemployment rate or proportions. Da Silva and Migon (2016) give a dynamic beta model for modelling unemployment rates. Several other alternative generalized linear models with spatio-temporal random effects are given by Molina *et al.* (2007), Lopez-Vizcaino *et al.* (2015) and Pereira *et al.* (2018).

From fourth quarter of 2014 onwards, INE started geo-referencing all the sampling units (dwellings) that are chosen as part of the Portuguese labor force survey (PLFS), according to the coordinates of the respective residential buildings, which may be composed by of one or more dwellings. Geo-referencing residential units opens new opportunities in modeling unemployment by using point referenced models as compared to areal models based on SAE methods. Geo-referencing embeds the sampling units in a continuous space, contrary to SAE methods, which embeds the sampling units in predesignated areal units. Surveys, such as the PLFS we specifically study, generate a collection of sampling units randomly distributed in space and due to the fact that the surveys highly depend on population density across space, the random distribution of these sampling units in space are highly non-homogeneous. The model we propose not only models the specific attribute of the sampling units conditional on a specific spatial configuration of the sample, it also models this random configuration of the sampling units by a non-homogeneous point process. We call the former model marks whereas the latter points. Therefore marked point

processes give a joint model for these two type variations inherent in the data.

55 For modelling the intensity of sampled residential unit locations and their associated marks, we suggest a marked log Gaussian Cox processes (LGCP) as model. The LGCP is a class of flexible models widely used in the context of spatial point processes (Moller *et al.*, 1998, Moller and Waagepetersen, 2004, Illian *et al.*, 2008, Baddeley *et al.*, 2016). Typically, in this framework, the log  
60 intensity of the point process is assumed to be a (latent) Gaussian random field. In order to facilitate calculations, often marks are assumed to be independent of point patterns so that marks and spatial patterns can be modeled separately. However, for inference on such models, Illian *et al.* (2012) proposed a flexible framework using INLA (Rue *et al.*, 2009, Martins *et al.*, 2013, Rue *et al.*, 2017),  
65 in which the spatial pattern of points and marks are allowed to be dependent, assuming independence conditional on a common latent spatial Gaussian processes. Inference for these models is not straightforward. Due to computational problems that emerge in this framework, Lindgren *et al.* (2011) proposed a more computationally tractable approach based on stochastic partial differen-  
70 tial equation (SPDE) models, which permit the approximation of a Gaussian field by a Gaussian Markov random field and we follow this method.

As far as we know, there is no literature of spatial point referenced models applied to unemployment estimation. Mostly methods are restricted to direct estimation methods and their extension to areal models. Hence, we believe that  
75 our methods bring methodological novelty to unemployment estimation by employing marked point processes as models for unemployment. It brings together many benefits in this field; First, by making the unemployment intensity as the primary source of interest, it is possible to obtain unemployment estimates in any desired (small) areal units by numerical integration. Second, the auxiliary  
80 information is introduced in smaller units (residential buildings) as compared to areal unit level, thus taking care of spatial heterogeneity in a much efficient manner.

The structure of this paper is as follows: In sections 2 and 3, we explain the sampling design of the PLFS and the consequent data available for the analysis.

85 In section 4, we explain our models as well as inferential methods and give the  
unemployment estimates for the 28 NUTS III regions, as well as for the 278  
counties in mainland Portugal.

A comparison of reported results with the direct and small area estimation  
methods is also given in section 5. Finally, in section 6, we give a brief account  
90 of possible extensions to the models and methods employed.

## 2. Labor Force Surveys

The methodologies proposed in this study are highly dependent on the sam-  
pling design of the PLFS. The calculation of the estimates for the population  
depends on the selection probabilities of the sampling units, which, in return,  
95 are calculated according to the sampling design. Therefore, it is important  
to understand both how the sampling units are drawn and how the inclusion  
probabilities are calculated so that bias free inference on the population can be  
made based on the sample. The PLFS is a continuous survey of the population  
living in private dwellings within the Portuguese national territory. The survey  
100 provides an understanding of the socioeconomic situation of these individuals  
during the week prior to the interview (reference week). The dwellings are the  
sampling units and the inhabitants living in these dwellings are the observation  
units.

During the study period, the sampling of the PLFS was stratified into 30  
105 NUTS III regions (see the regions of mainland Portugal in figure 2). In each  
region, multi-stage sampling is conducted, where the primary sampling units  
are areas consisting of one or more  $1\text{km}^2$  cells of the INSPIRE grid, and the  
secondary units are dwellings. Every selected dwelling and all its inhabitants  
are surveyed. Since all the individuals in each selected dwelling are surveyed,  
110 their selection probabilities are equal to the respective selection probability of  
dwelling. Here, these selection probabilities change from one inspire grid cell  
to another, but within each grid cell, they are identical. Hence, these selection  
probabilities have a resolution of  $1\text{km}^2$ . We refer the reader to INE (2016) for

further details of the PLFS. Based on this sample survey design, our funda-  
 115 mental assumption is that dwellings of the sample are selected independently  
 and identically among the population with constant selection probabilities and  
 that average characteristics of inhabitants of dwellings are constant within each  
 $1\text{km}^2$  inspire grid cells.

The official estimates of the unemployment figures are calculated using a di-  
 120 rect method, based on the Horvitz-Thompson estimator (Horvitz and Thomp-  
 son, 1952) in which the weights attached to each sampling unit are inversely  
 proportional to the selection probability. In section 4.2, we treat the points of  
 sampled residential units as a thinned version of the population residential units,  
 with thinning probabilities equal to the selection probabilities. These selection  
 125 probabilities will play a crucial role when making inference on the number of  
 unemployed people.

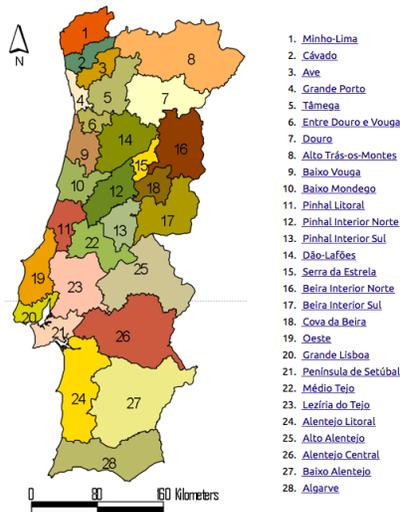


Figure 2: NUTS III regions in mainland Portugal (version 2002)

### 3. Data

In this study, we only model spatial variation of unemployment and specifically we analyse the data of the Portuguese Labour Force Survey (PLFS) regarding the fourth quarter of 2014. In section (6), we mention how such models can be extended to model space-time variation.

The sample size is about 40,000 observations and each individual can be classified into one of the following three categories: employed, unemployed and inactive. An inactive individual is an individual who is not employed neither unemployed. Covariate information about the individuals is available, such as gender, age and education level.

Within a point process modelling scheme, the choice of dwellings as sampling units, creates problems. Since residential units are geo-referenced, multiple dwellings may appear with the same spatial location, making part of the same residential unit. The sampling design and the consequent data are not sufficiently detailed to obtain good information on the multiplicity distribution of dwellings in each residential building, therefore we use residential buildings as design units. Thus, we aggregate the number of unemployed people observed in dwellings with the same spatial location, and we denote by  $Y(s_j)$ , the number of unemployed in the residential building sampled at spatial position  $s_j \in \mathcal{S}$ .

#### 3.1. Covariates

One of the benefits of point referenced models for modeling the spatial variation of unemployment is that detailed auxiliary information can be given at unit level (average attributes of observation units within each residential unit), such as age and average education. At areal level, we have two pieces of auxiliary information, namely: the population density at counties (which we will use as offset in the model), and the number of individuals registered as unemployed at county unemployment offices. Other information at any areal level can be easily incorporated into the model as it becomes available.

The median of the education level in each residential building and the mean age were considered as useful information to model the marks. The use of the

median instead of mean of the education level is due to its asymmetric distribution. Although the education level does not constitute a quantitative variable, it was treated as such due to its ordinal meaning (1-primary level, 2-secondary level, 3-higher level). Higher values of this variable in the Lisbon and Península de Setúbal regions can be clearly seen (figure 3). Here, the extrapolation of these covariates is done taking into account the spatial dependence at a detailed geographical level. These auxiliary information are then extended from sampling units to all domains of study by simple non-parametric kernel spatial smoothing method (Nadaraya, 1964, 1989; Watson, 1964), giving us a reasonably accurate auxiliary information for out-of-sample residential units. The idea behind this method, also known as the Nadaraya-Watson smoother is the following: if the observed values are  $y(s_1), \dots, y(s_n)$  at locations  $s_1, \dots, s_n$  respectively, then the smoothed value at a location  $u$  is given by

$$g(u) = \frac{\sum k(u - s_i)y(s_i)}{\sum k(u - s_i)} \quad (1)$$

where  $k$  is a probability density. A common choice for this is the isotropic Gaussian probability density and in this case it is called a Gaussian kernel. The standard deviation of the kernel is the smoothing bandwidth. A larger bandwidth gives more smoothing. As Baddeley et al (2016) explain, for very large values of smoothing bandwidth, the result will be approximately constant and equal to the average mark value in the entire dataset. For small values of bandwidth, the result becomes closer to the nearest data point. The choice of the bandwidth involves a balance between the bias and variance: as bandwidth increases, the bias increases and variance decreases. The proportion of unemployed people registered in the regional employment centers is highly correlated to the variable of interest. Notice that here we are interested in the concept of unemployment for a National Statistical Institute. The unemployment registered is used as a regional covariate. Figure 3 shows the spatial variation of the covariates. It is interesting to see the spatial distribution of the mean age, with more younger people near the country's coastline than in its interior.

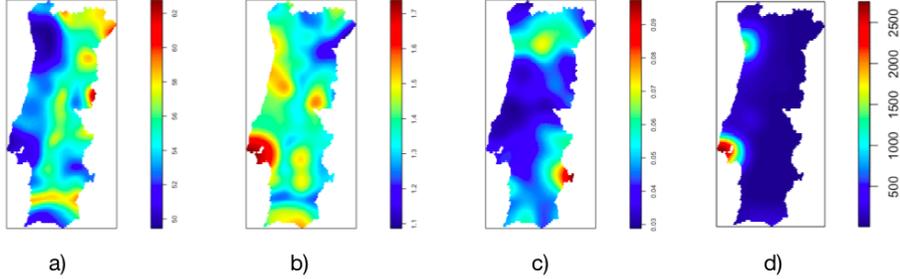


Figure 3: Kernel estimates of: a) the mean age per residential building; b) the median of the education level per residential building; c) the proportion of registered unemployed people in centers of employment; d) the population density (inhabitants/ $km^2$ ).

#### 185 4. A spatial point patterns approach

We model the spatial variation of unemployed people by a marked point process. Here we give a brief description of point processes (detailed study can be found in Illian *et al.* (2008)).

A point process is a stochastic model describing the random configuration of point patterns in some abstract space. A spatial point process  $P_1(s), s \in \mathcal{A} \subset \mathbb{R}^2$  describes the probability structure of random point patterns in space. Related concept is the counting measure, such that for some  $A \subset \mathcal{A}$ ,  $N(A)$  is the random variable that counts the number of such points in  $A$ . When this counting measure is Poisson, then for some intensity function  $\lambda(s), s \in S$ , the probability mass function of the random variable  $N(A)$  is given by

$$P(N(A) = x) = e^{-\mu(A)} (\mu(A))^x / x! \quad , \quad x = 0, 1, 2, \dots,$$

where  $\mu(A) = \int_{s \in A} \lambda(s) ds$ . A spatial log Gaussian Cox process is a doubly stochastic point process for point patterns in space, in which the intensity function  $\lambda(s), s \in S$  is itself a stochastic process where, typically

$$\log \lambda(s) = \mu(s) + W(s).$$

Here,  $\mu(s)$  is a deterministic function defined over  $s \in S$  and  $W(s)$  is a stationary isotropic, zero-mean Gaussian field, typically assumed to have a Matérn

covariance function (Banerjee *et al.*, 2004).

Marked point processes are generalizations of point processes where every point  $s \in S$  of the random point patterns is assigned a further quantity  $Y(s)$  which provides additional information on the object represented by the point  $s$ . Thus we represent the marked point process by

$$M(s) = (P_1(s), Y(s), s \in S).$$

In this particular context, when counting people,  $Y(s)$  is a discrete counting random variable defined over  $\mathbb{N} \cup \{0\}$  with probability mass function  $P(Y(s) = y)$ . When the point process  $P_1(s), s \in S$ , is a Poisson point process, then the marked point process itself is a Poisson point process where for any  $A \subset S$  and  $B \subset \mathbb{N} \cup \{0\}$ ,  $M(A, B)$  represents the number of points in  $A$  with its mark in  $B$  and is given in terms of an intensity function  $\lambda(s, y), s \in S, y \in \mathbb{N} \cup \{0\}$  and

$$P(M(A, B) = x) = e^{-\mu(A, B)} (\mu(A, B))^x / x!, x = 0, 1, 2, \dots,$$

where  $\mu(A, B) = \int_{s \in A, y \in B} \lambda(s, y) ds dy$ . Most often, it is assumed that point patterns are independent of the marks, in which case this intensity function takes a more manageable form

$$\lambda(s, y) = \lambda(s)P(Y(s) = x),$$

so that point patterns and marks can be modeled separately.

In our context,  $M(s) = (P_1(s), Y(s))$ , represents the marked point process of residential units in space and  $Y(s)$  is the corresponding mark of a residential unit at spatial location  $s$ , which is the number of unemployed residing in that unit.

We model the spatial point process  $P_1(s), s \in S$ , of residential units using a log Gaussian Cox process with intensity  $\lambda_1(s)$ , with

$$\log \lambda_1(s|W_1(s)) = \alpha_1 + \mathbf{z}'_1(s)\boldsymbol{\theta}_1 + W_1(s), \quad (2)$$

and assume that for every  $s$ , the mark  $Y(s)$  is a Poisson random variable with

200 probability mass function

$$P_{Y(s)|W_2(s)}(y) \sim \text{Poisson}(\lambda_2(s|W_2(s))), \quad (3)$$

where

$$\log \lambda_2(s|W_2(s)) = \alpha_2 + \mathbf{z}'_2(s)\boldsymbol{\theta}_2 + W_2(s). \quad (4)$$

Here,  $W_1(s)$  and  $W_2(s)$ ,  $s \in S$  are independent latent zero-mean isotropic Gaussian Markov fields belonging to the Matérn class (Banerjee *et al.*, 2004),  $\mathbf{z}_1(s)$  and  $\mathbf{z}_2(s)$  are generic auxiliary information which may help in understanding the spatial patterns of points as well as the marks, and  $\boldsymbol{\theta} = (\alpha_1, \alpha_2, \alpha_3, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$  are the model parameters, which we assume to be random as part of our Bayesian hierarchical model. 205

We follow Simpson *et al.* (2017) and Fuglstad *et al.* (2017) to construct a joint penalising complexity (PC) prior density for the model parameters of each field. Here, the PC priors for the spatial range,  $\rho$ , and the marginal standard deviation,  $\sigma$ , are given by 210

$$p(\rho, \sigma) = RS\rho^{-2}e^{-R\rho^{-1}-S\sigma} \quad (5)$$

where  $R$  and  $S$  are hyperparameters determined by  $R = -\log(\alpha_1)\rho_0$  and  $S = \frac{-\log(\alpha_2)}{\sigma_0}$ . The practical approach for this in INLA is to require the user to indirectly specify these hyperparameters through  $P(\rho < \rho_0) = \alpha_1$  and  $P(\sigma < \sigma_0) = \alpha_2$ . Here, we consider  $\rho_0 = 400, \alpha_1 = 0.5, \sigma_0 = 1, \alpha_2 = 0.5$ . 215

With the independence assumption, the corresponding marked point process  $M(s, y), s \in S, y \in \mathbb{N} \cup \{0\}$  has the following structure:

$$M(s, y)|\lambda(s, y) \sim \text{Poisson}(\lambda(s, y)), \quad (6)$$

with

$$\lambda(s, y) = \lambda_1(s)P_{Y(s)}(y) \quad (7)$$

4.1. Target quantities for inference

Our objective is to make inference on the number of unemployed people of the population in any given region  $A \subset \mathcal{S}$ , based on the sample survey. Thus, our target quantity is given by

$$U(A) = \sum_{j=1}^{U_1(A)} Y(s_j),$$

220 where  $U_1(A)$  is the number of residential buildings in region  $A$  of the population,  $Y(s_j)$ , the number of unemployed in these residential units. When detailed census information is available,  $U_1(A)$  is not a random number, but can be considered as a population constant. However, in this work we treat spatial point patterns of population residential units, and consequently  $U_1(A)$  as random.

225 We will base our inference on the respective sample quantities

$$N(A) = \sum_{j=1}^{N_1(A)} Y(s_j), \quad (8)$$

where  $N_1(A)$  represents the number of residential units of the sample in  $A$ .

Let  $\mathbf{s} = (s_1, \dots, s_n)$  be the location of the sampling units chosen in the sampling survey,  $\mathbf{y}(s)$  be the number of unemployed people in residential units at location  $s$  and  $\mathbf{z}_1$  and  $\mathbf{z}_2$  be the auxiliary information specific to residential units and marks respectively observed at these locations. In this study, we do not know the exact geo-referenced positions of all the residential buildings in the population and therefore we treat these locations in space as a non-homogeneous random point patterns.

235 We denote by  $\mathbf{x} = (n, \mathbf{s}, \mathbf{y}(s), \mathbf{z}_1, \mathbf{z}_2)$  the observed data obtained from the sampling survey. Randomly chosen point patterns  $\mathbf{s}$  of this sample is a thinned version of the true point patterns of the population and this location specific thinning operator which we denote by  $p(s), s \in \mathcal{S}$  is defined precisely by the sample survey design, which we will explain in section 4.2.

Our specific target parameters are the posterior predictive mean and variance

240 of the random variable  $N(A)$  given respectively by

$$\mathbb{E}(N(A)|\mathbf{x}) = \mathbb{E}_{(W_1(s), W_2(s), \boldsymbol{\theta}|\mathbf{x})}[\mathbb{E}(N(A)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta})] \quad (9)$$

and

$$\begin{aligned} \text{Var}(N(A)|\mathbf{x}) &= \text{Var}_{(W_1(s), W_2(s), \boldsymbol{\theta}|\mathbf{x})}[\mathbb{E}(N(A)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta})] \\ &+ \mathbb{E}_{(W_1(s), W_2(s), \boldsymbol{\theta}|\mathbf{x})}[\text{Var}(N(A)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta})]. \end{aligned} \quad (10)$$

here, the expectations are taken with respect to the joint posterior density of the parameters  $(W_1(s), W_2(s), \boldsymbol{\theta})$ .

Calculation of

$$\mathbb{E}(N(A)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}) = \mathbb{E}\left(\sum_{j=1}^{N_1(A)} Y(s_j)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}\right) \quad (11)$$

245 and

$$\text{Var}(N(A)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}) = \text{Var}\left(\sum_{j=1}^{N_1(A)} Y(s_j)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}\right), \quad (12)$$

require certain assumptions. Assume that  $I_j$  represent the INSPIRE grid cells with  $||I_j|| = 1km^2$ .

1. Conditional on  $W_1(s)$ , the point patterns of the Cox process over disjointed regions are independent. Consequently, conditional on  $W_1(s)$ , the point patterns over the design grid cells  $I_j$  are also independent. We also assume that within each grid cell the intensity function of the Cox process is homogeneous so that  $\lambda_1(s) = \lambda_1(I_j)$  for every  $s \in I_j$ .
2. We assume that the marks  $Y(s)$  are independent of the point patterns so that the conditional intensity function of the marked point process is given by

$$\lambda(s, y|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}) = \lambda_1(s|\mathbf{x}, W_1(s), \boldsymbol{\theta})P_{Y(s)|\mathbf{x}, W_2(s), \boldsymbol{\theta}}(y),$$

3. We assume that conditional on  $W_2(s)$ , marks observed on disjointed sets are independent.

255 4. Finally, we assume, in line with the sampling survey design that, in each  
each  $1\text{km}^2$  grid cells  $I_j$  that the marks  $Y(s_i)$  are identical for every  $s_i \in I_j$ ,  
and that the number of unemployed in every residential unit in grid cell  
 $I_j$  are identical. Hence, in each grid cell we replace  $E(Y(s_i))$  by  $E(Y(I_j))$ .  
We also assume that average auxiliary covariate information are identical  
260 for residential units which are in the same grid cell. Hence, within each  
grid cell, residential units are identical in all aspects.

Let  $N_1(I_j)$  be the number of residential units in each grid cell  $I_j$ . Then with  
assumptions (1)-(4),

$$\begin{aligned} E(N(A)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}) &= E\left(\sum_{j=1}^{N_1(A)} Y(s_j)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}\right) \\ &= E\left(\sum_{I_j \in A} \sum_{i \in I_j} Y(s_i)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}\right) \\ &= \sum_{I_j \in A} E\left(\sum_{i=1}^{N_1(I_j)} Y(s_i)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}\right) \quad (13) \end{aligned}$$

$$= \sum_{I_j \in A} E(N_1(I_j)|W_1(I_j))E(Y(I_j)|\mathbf{x}, W_2(I_j), \boldsymbol{\theta}) \quad (14)$$

$$= \sum_{I_j \in A} ||I_j||\lambda_1(I_j|\mathbf{x}, W_1(I_j), \boldsymbol{\theta})\lambda_2(I_j|\mathbf{x}, W_2(I_j), \boldsymbol{\theta}) \quad (15)$$

$$\approx \int_{s \in A} \lambda_1(s|\mathbf{x}, W_1(s), \boldsymbol{\theta})\lambda_2(s|\mathbf{x}, W_2(s), \boldsymbol{\theta})ds \quad (16)$$

Here,  $W_1(I_j)$  and  $W_2(I_j)$  represent the independent latent Gaussian Markov  
265 random fields respectively approximating the latent Gaussian random fields  
 $W_1(s)$  and  $W_2(s)$  obtained by the SPDE method. (15) follows on from the  
conditional independence and homogeneity of the point patterns as well as the  
marks within each  $1\text{km}^2$  grid cell, whereas (16) proceeds from the approximation  
of integrals by sums over the design grid cells. Thus, the  $1\text{km}^2$  design grid cells  
270 are the smallest units over which we approximate the point referenced process.

We can calculate, with similar arguments

$$\begin{aligned} \text{Var}(N(A)|\mathbf{x}, W_1(s), W_2(s), \boldsymbol{\theta}) &\approx \int_{s \in A} \lambda_1(s|\mathbf{x}, W_1(s), \boldsymbol{\theta})\lambda_2(s|\mathbf{x}, W_2(s), \boldsymbol{\theta})ds \\ &+ \int_{s \in A} \lambda_1(s|\mathbf{x}, W_1(s), \boldsymbol{\theta})\lambda_2^2(s|\mathbf{x}, W_2(s), \boldsymbol{\theta})ds \end{aligned} \quad (17)$$

The mean and the variance of the predictive distribution given in (9) and (10) can be calculated numerically. The INLA package permits the calculation of the intensity function  $\lambda_1(s|\mathbf{x}, W_1(s), \boldsymbol{\theta})$  as well as the mean mark  $\lambda_2(s|\mathbf{x}, W_2(s), \boldsymbol{\theta})$ . INLA also simulates from the marginal and joint posterior densities of the latent process, as well as the model parameters, and thus target quantities (9), (10) as well as posterior intensities

$$\lambda_1(s|\mathbf{x}) = \text{E}_{(W_1(s), \boldsymbol{\theta}|\mathbf{x})}[\lambda_1(s|\mathbf{x}, \boldsymbol{\theta}, W_1(s))], \quad (18)$$

$$\lambda_2(s|\mathbf{x}) = \text{E}_{(W_2(s), \boldsymbol{\theta}|\mathbf{x})}[\lambda_2(s|\mathbf{x}, \boldsymbol{\theta}, W_2(s))] \quad (19)$$

and therefore the posterior intensity of the marked point process for the sample

$$\lambda(s|\mathbf{x}) = \lambda_1(s|\mathbf{x}) * \lambda_2(s|\mathbf{x}), \quad (20)$$

can be efficiently calculated within the INLA platform. Our methodology and the consequent inference on  $M(A)$ , the population totals of unemployed can be described in the following steps:

- From the sample, estimate posterior intensity function  $\lambda(s|\mathbf{x})$ , given in (20). This intensity, through expression (9) will count the unemployed people in the sample in any region.
- Use design based inclusion probabilities to estimate the thinning probabilities  $p(s)$ ,  $s \in \mathcal{S}$ , and the above intensity function to calculate the intensity function of the marked point process for the population units. How this

thinning probability is calculated will be given in section 4.2. We note that  
 290 the role of  $p(s)$  is somewhat similar to the detection functions often used  
 in distance sampling methods in estimation of animal abundance in eco-  
 logical studies (Yuan et al., 2017), which connects the observed numbers  
 to true population.

- The marked point process posterior intensity of the population unit thus  
 295 are given by

$$\lambda^*(s|\mathbf{x}) = \lambda(s|\mathbf{x})/p(s) \quad (21)$$

This will count the unemployed people in the sample over any region  
 $A \subset \mathcal{S}$  through the expression (9) substituting  $\lambda(s|\mathbf{x})$  by  $\lambda^*(s|\mathbf{x})$ .

#### 4.1.1. Model selection

Several alternative spatial random effects were used in modelling the inten-  
 300 sities, namely (i) two independent latent processes  $W_1$  and  $W_2$  for the points  
 and their marks respectively, as described in 3-4, (ii) one random effect  $W$  for  
 points, (iii) one random effect  $W$  for points and its scaled version  $\alpha_3 W$  for the  
 marks.

To evaluate the significance of each covariate and random effect, we consid-  
 305 ered different models and compared the results of two model selection criteria:  
 deviance information criterion (*DIC*), proposed by Spiegelhalter *et al.* (2002),  
 and Watanabe-Akaike information criterion (*WAIC*), proposed by Watanabe  
 (2010).

Table 1 shows the values of these two criteria for the models considered  
 310 for the marked point process. In this case, the model with the best performance  
 was the one that took into account the following factors:

- the offset term given by the population density to model the intensity of  
 the point process ( $\text{offset}_1$ );
- the covariates to model the mark intensity: number of individuals per  
 315 residential building ( $\text{nind}_2$ ), the median of the education level ( $\text{edu}_2$ ),

Table 1: DIC, WAIC and the effective number of parameters

log intensity of points and marks	$DIC$	$WAIC$	$p_{DIC}$	$p_{WAIC}$
$\alpha_1 \parallel \alpha_2$	110707.09	110724.00	2.18	19.04
$\alpha_1 + \text{offset}_1 \parallel \alpha_2$	84826.76	84853.04	2.198	28.41
$\alpha_1 + \text{offset}_1 \parallel \alpha_2 + \text{offset}_2$	110707.09	110724.00	2.18	19.04
$\alpha_1 + \text{offset}_1 \parallel \alpha_2 + \text{nind}_2$	84791.91	84818.07	3.199	29.29
$\alpha_1 + \text{offset}_1 \parallel \alpha_2 + \text{nind}_2 + \text{edu}_2$	84790.11	84816.27	4.198	30.28
$\alpha_1 + \text{offset}_1 \parallel \alpha_2 + \text{nind}_2 + \text{edu}_2 + \text{age}_2$	84714.83	84741.04	5.198	31.34
$\alpha_1 + \text{offset}_1 \parallel \alpha_2 + \text{nind}_2 + \text{edu}_2 + \text{age}_2 + \text{iefp}_2$	84706.69	84733.00	6.197	32.444
$\alpha_1 + \text{offset}_1 + W \parallel \alpha_2 + \text{nind}_2 + \text{edu}_2 + \text{age}_2 + \text{iefp}_2$	48538.24	67474.12	1817.38	15031.22
$\alpha_1 + \text{offset}_1 + W \parallel \alpha_2 + \text{nind}_2 + \text{edu}_2 + \text{age}_2 + \text{iefp}_2 + \alpha_3 W$	48656.62	67562.53	1796.70	14998.79
$\alpha_1 + \text{offset}_1 + W_1 \parallel \alpha_2 + \text{nind}_2 + \text{edu}_2 + \text{age}_2 + \text{iefp}_2 + W_2$	48505.25	67443.79	1842.59	15058.96

the mean age ( $\text{age}_2$ ), and the proportion of registered unemployed people ( $\text{iefp}_2$ ). Here, subscripts 1 and 2 indicate that the corresponding covariate is used in modelling intensity  $\lambda_1(s)$  and  $\lambda_2(s)$  respectively;

- two independent latent processes  $W_1$  and  $W_2$  used for points and their marks.

320

It is clear from the table that the model that employs all the auxiliary information and two independent latent processes, one for points the other for marks (model (i)), seems to give the best fit, and that the model with a single common latent process for the points and marks (model (iii)) comes second place. We note here that model (iii) implies preferential sampling (Diggle *et al.*, 2010). Since INE does not use preferential sampling in the survey design, this model is not compatible.

325

We also considered a negative binomial distribution for the marks as an alternative to the Poisson marks, but these models did not bring significant gains in terms of DIC.

330

To perform the spatial prediction, we created a regular grid of  $1\text{km}^2$  in the domain. A projection from the mesh to the grid was performed and the resulting maps of the posterior mean of the logarithmic transformation of the intensity of the residential units  $\log(\lambda_1(s))$  and the logarithmic of the marks mean are shown in figure 4. The plot of the logarithmic transformation of the intensity

335

provides a clearer picture of the spatial variation of the residential buildings. As we expected, the highest values are concentrated in the Lisboa, Porto, and Algarve regions. The intensity is clearly higher near the coast and lower in the interior of the country.

340 With these estimates, we can conclude that the average number of unemployed people per residential building is higher in the Grande Porto, Península de Setúbal and Alentejo Central regions.

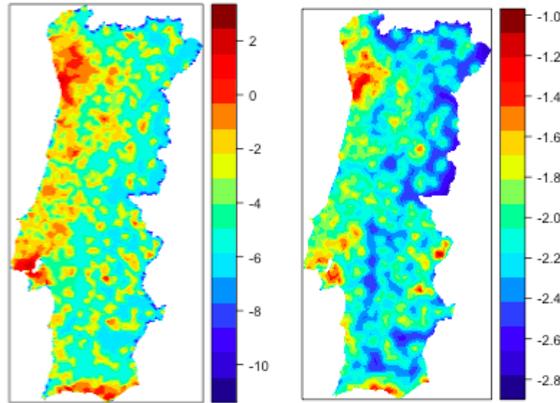


Figure 4: Posterior mean of: log intensity (left), and log mean marks (right)

#### 4.2. Unemployment estimation for NUTS III regions and counties

The marked point process model explained in the previous section, projects  
 345 the sampling survey in space. This point process is a thinned version of the true point patterns of the residential units across Portugal, together with their marks.

Let  $N_1^*(s)$  represent the true point patterns of the residential units with intensity  $\lambda_1^*(s)$ . Then

$$\lambda_1^*(s) = \frac{\lambda_1(s)}{P(RU(s))}, \quad (22)$$

350 where,  $P(RU(s))$  is the probability that a residential unit at  $s$  is included in the survey.  $P(RU(s))$  should be interpreted as the proportion of the residential

units in any infinitesimal area which is included in the sampling survey.

If we assume also that  $N_2^*(s)$  represents the true intensity of the number of unemployed people observed in residential unit at location  $s$ , then the intensity  $\lambda_2^*(s)$  of this counting process is given by

$$\lambda_2^*(s) = \frac{\lambda_2(s)}{P(D(s)|RU(s))}, \quad (23)$$

where the probability  $P(D(s)|RU(s))$  should be interpreted as the proportion of dwellings in a residential unit which are included in the sampling survey.

Target quantities (9) and (10) depend on the multiplicative intensity

$$\lambda(s) = \lambda_1(s)\lambda_2(s), \quad (24)$$

which is a thinned version of

$$\lambda^*(s) = \lambda_1^*(s)\lambda_2^*(s), \quad (25)$$

and this relationship is given by

$$\lambda^*(s) = \frac{\lambda(s)}{p(s)}, \quad (26)$$

where

$$\begin{aligned} p(s) &= P(RU(s))P(D(s)|RU(s)) \\ &= P(D(s)), \end{aligned}$$

since  $P(D(s)|RU^c(s)) = 0$ .

Here,  $p(s)$  should be interpreted as the proportion of dwellings in any infinitesimal area that are chosen in the sampling survey.

To define the intensity of the full version of the spatial point process, the knowledge of the sampling probabilities  $p(s)$  for the whole domain is required. These probabilities are calculated as part of the PLFS at  $1\text{km}^2$  inspire grid cells and we smooth these probabilities over the domain of study using a non-parametric kernel method with a resolution of  $1\text{km}^2$  grid cells. Thus,  $p(s)$  is

370 non-homogeneous across the domain of study, but this probability is constant  
 within each grid cell.

We simulated 1000 values of the predictive posterior distribution of  $\lambda_1(s|\mathbf{x})$   
 and  $\lambda_2(s|\mathbf{x})$  as described in (18) and (19) for each cell  $I_j$  to estimate the target  
 quantities. This was achieved by simulating samples from the posterior distri-  
 375 butions of the model parameters and the latent Gaussian Markov fields used in  
 the model.

The multiplicative (predictive) intensity for the population  $\lambda^*(s|\mathbf{x})$  will form  
 the basis for calculating the unemployment in any region  $A$ , expressed in terms  
 of  $E(N(A)|\mathbf{x})$  as given in (9).

380 Figure 5 gives the predictive means of the multiplicative intensity functions  
 $\lambda(s|\mathbf{x})$  and  $\lambda^*(s|\mathbf{x})$ , as well as the selection probability  $p(s)$ .

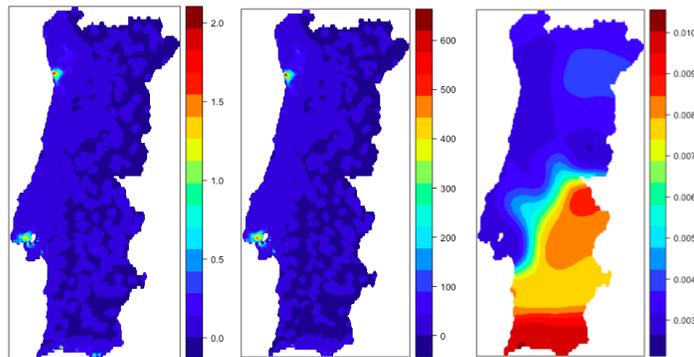


Figure 5: Predictive means of multiplicative intensity functions for the sample (left) and the  
 population (middle), as well as the selection probability (right)

We calculated the credible intervals for the posterior mean by NUTS III  
 region and compared those with the direct estimates (figure 6). The x-scale  
 in these graphs represents the NUTS III regions accordingly to the codes in  
 figure 2. We note that the direct estimates are inside the intervals for almost all  
 385 regions. The highest points correspond to Porto and Lisboa. The figure reveals  
 that there is an underestimation in these regions and an overestimation in the  
 others. This behaviour is probably due to the large differences in the intensity

of these regions and consequent smoothing of intensities across the space.

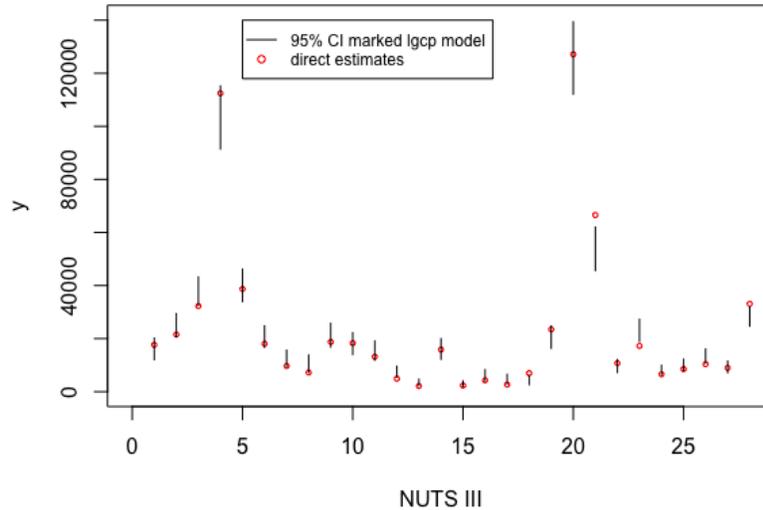


Figure 6: 95% credible intervals for the posterior mean (black) and direct estimates (red) by NUTS III regions

390 **5. Comparison between the results of the point-referenced data model  
and the traditional small area models**

We conducted a comparison between the marked LGCP model described in the previous sections, and some traditional small area models. For a first comparison, we chose the hierarchical Bayesian model versions of the standard  
395 Fay and Herriot model (FH model) proposed by Rao and Molina (2015) and its extension with a latent CAR model (FH-CAR model) to borrow strength from adjacent areal units given by You and Zhou (2011). In both of these models we use the same set of auxiliary information, aggregated to NUTS III regions and we report the total unemployment estimates for each of these 28 regions  
400 obtained from the 3 alternative methods. We also report the performance of

these 3 models at a higher resolution, namely the estimates of the total number of unemployed people in 278 counties.

Figure 7 shows the estimates obtained by the direct method, the FH model, the FH-CAR model and the LGCP model. We can see that the FH models (FH  
405 and FH-CAR) disagree with the direct method in the Lisbon region. This result, in the most densely populated region, may indicate some fragility. Moreover, coefficients of variation (CVs), given by the ratio of the standard deviation to the mean, obtained by the FH model for the Serra da Estrela region are quite high (figure 8).

410 As we expected, for the majority of the regions the FH model presents lower CVs in comparison with the FH-CAR and LGCP models. This result can be explained by the difference in the number of parameters in these models. Moreover, the FH models use the direct estimates as the data in the modelling process and assume that the variances are fixed. The LGCP model uses the data  
415 from the sample of the Labour Force Survey, and do not assume the variances as fixed. In any case, the CVs of the estimates obtained by the FH-CAR model, as well as the CVs of the estimates obtained by the LGCP model do not achieve 25% in any NUTS III region.

In general, the FH-CAR and LGCP models present similar CVs. Since the  
420 estimates and the CVs of these two models are close, we think that the LGCP model proposed here brings many advantages to this problem, as we explained in the introduction section.

Note that if the direct method is applied to more disaggregated geographical levels, such as counties, it is not possible to provide estimates for the regions  
425 without observations in the LFS or without any unemployment observation (figure 9). Consequently, the basic FH models cannot provide estimates for these regions since they use direct estimators with known variances as data. The FH-CAR model, on the other hand, can do this using the spatial effects. However, in some counties there are no observations even in the neighbours. Therefore, the  
430 estimates obtained by the FH models in these counties are questionable. Figure 9 shows that in the majority of regions with observations in the LFS, the LGCP

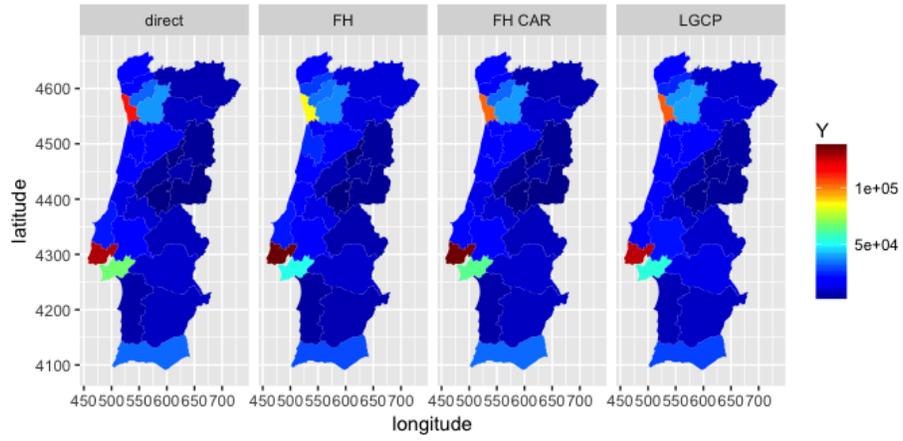


Figure 7: Estimates of total unemployed from the direct method, FH model, FH-CAR model and LGCP model, by NUTS III for the fourth quarter of 2014

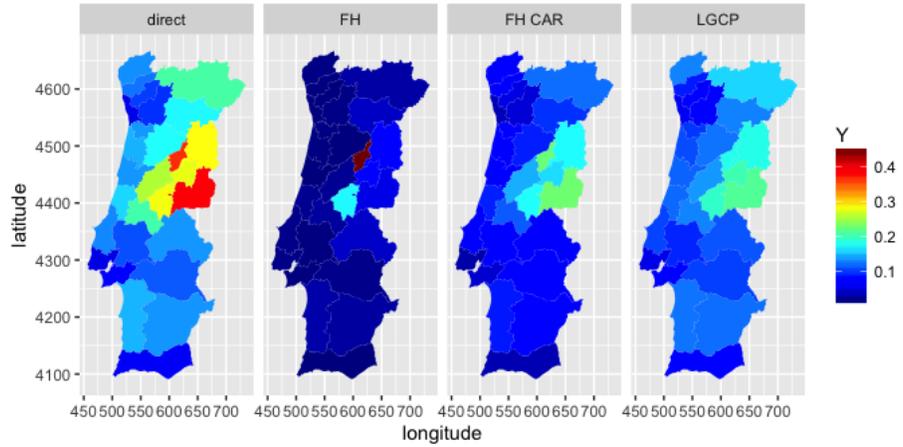


Figure 8: Coefficients of variation of the estimates obtained by the direct method, FH model, FH-CAR model and LGCP model

model is the one with the lowest CVs. For a clear comparison, we show the boxplots of the CVs obtained from each model in figure 10. As we expected, the higher the level of disaggregation, the greater the difference between the  
435 LGCP and the traditional SAE models in terms of variability.

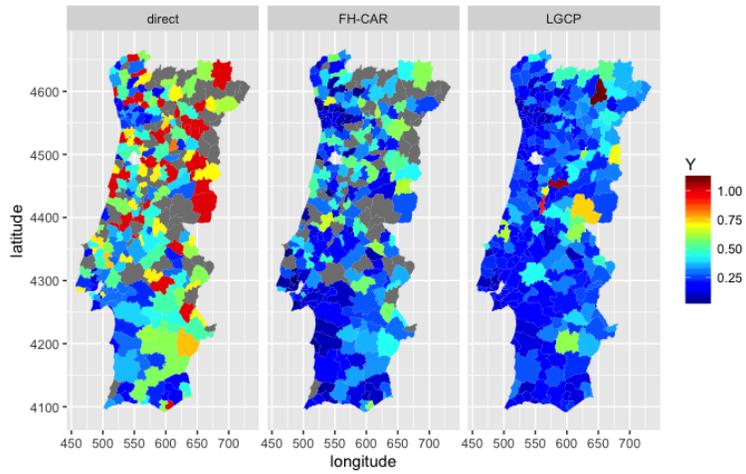


Figure 9: CVs of the estimates obtained by the direct method (left), FH-CAR model (middle) and LGCP model (right)

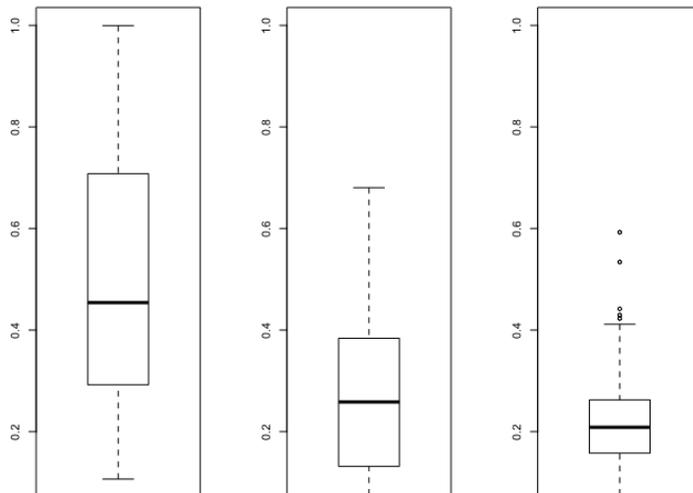


Figure 10: Boxplots with the CVs of the estimates obtained by the direct method (left), FH-CAR model (middle) and LGCP model (right)

## 6. Discussion, conclusions and further work

In this study, we employed point referenced spatial models to generate unemployment estimates by making use of the newly available geo-referenced locations of sampling units. Unlike the direct method and area level models, our approach takes into account this important new information and permits spatial smoothing at higher spatial resolutions, by making use of the disaggregated data.

Although for large areal units such as NUTS III regions, unemployment figures estimated from our method do not differ significantly from the estimates obtained by the standard areal models employed within the SAE methodology, the main benefits of the proposed method can clearly be seen when one looks at the unemployment estimates at smaller areal units, such as counties. By comparing the estimates obtained from the proposed methodology with the gold standard, namely direct estimation, we see that even in fairly large areas where

450 direct estimates are available, our estimates compare fairly well with the gold standards. Despite the fact that we did not carry out a full sensitivity analysis, this examination, performed in the manner suggested by Roos *et al.*(2015) may suggest how the proposed models can be improved.

National Statistical Institutes usually require a great deal of consistency  
455 between the estimates obtained at different areal resolutions, but this requirement is not easy to satisfy using the standard small area models. Since this new methodology operates independently from the administrative limits of geographical units, it can provide us with the necessary means to meet this requirement both in a consistent and timely fashion.

460 Despite the relative complexity of this methodology, the computational costs are not high due to the availability of the R-INLA package in the software R.

In this work, we do not report on the time dynamics of unemployment. At present, there are only 14 quarterly sample surveys with geo-referenced sampling units. As these further quarterly data sets become available, it may be  
465 possible to investigate, with a certain degree of precision, how spatial variation of unemployment changes over time. This can be done by considering space-time marked point processes, in which the latent process becomes a space time Gaussian process. By adding time varying covariates into the model, such as a linear or quadratic trend function in time, it may be possible to capture, in  
470 detail, the time dynamics of the unemployment across the domain of study. It is possible to infer on such models within the INLA platform with only a moderate increase in computational time.

Recently, INE has begun a much more ambitious geo-referencing method by identifying and geo-referencing not only the sampled residential units, but all  
475 the residential units in the whole of Portugal. Use of such geo-referenced out of sample information on the population units will inevitably result in estimates with much better precision, as the variation and uncertainty due to random configuration of residential units which are present in our model, will no longer exist. Point referenced methods and models that are adequate in the presence of  
480 this extra information will be discussed elsewhere as new data become available.

## Acknowledgements

This work was supported by the project *UID/MAT/00006/2013* and the PhD scholarship *SFRH/BD/92728/2013* from Fundação para a Ciência e Tecnologia. Instituto Nacional de Estatística and Centro de Estatística e Aplicações da Universidade de Lisboa are the reception institutions. We would like to thank professor Antónia Turkman, Elias Krainski and Paula Pereira for their help.

## Note

This study is the responsibility of the authors and does not reflect the official opinions of Instituto Nacional de Estatística.

## References

- [1] Baddeley, A., Rubak, E., Turner, R. (2016) *Spatial Point Patterns - Methodology and Applications with R*. Chapman and Hall/CRC.
- [2] Banerjee, S., Carlin, B. P., Gelfand, A. E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Chapman and Hall/CRC.
- [3] Chambers, R., Salvati, N., Tzavidis, N. (2012) M-Quantile Regression for Binary Data with Application to Small Area Estimation. *Centre for Statistical and Survey Methodology, University of Wollongong. Working paper, 2012*, <http://ro.uow.edu.au/cssmwp/101>
- [4] Chandra, H., Salvati, N., Chambers, R., Tzavidis, N. (2012) Small Area Estimation under Spatial Nonstationarity. *Computational Statistics and Data Analysis*, **56**, 2875-2888.
- [6] Da-Silva, C.Q., and Migon,H.S.(2016) Hierarchical Dynamic Beta model. *REVSTAT*, **14**, 49-73.
- [6] Diggle, P. J., Menezes, R., Su, T. (2010) Geostatistical inference under preferential sampling. *Appl. Statist.*, **59**, 191-232.

- [7] Fay, R.E, Herriot, R.A.(1979) Estimates of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association*, **74**, 269-277.
- [8] Fuglstad, G.-A., Simpson, D., Lindgren, F., and Rue, H. (2017) Constructing Priors that Penalize the Complexity of Gaussian Random Fields. arXiv:1503.00256
- [9] Gelfand, A. E., Schmidt, A.M., Banerjee, S. and Sirmans, C.F. (2004) Non-stationary multivariate process modeling through spatially varying coregionalization. *Test*, **13**, 263-312.
- [10] Gelman, A., Hwang, J., and Vehtari, A. (2014) Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, **24**, 997-1016.
- [11] Hobza, T., Morales, D. (2016) Empirical best prediction under unit-level logit mixed models. *Journal of Official Statistics*. **32**, 661-692.
- [12] Horvitz, D. G., Thompson, D. J. (1952) A Generalization of Sampling Without Replacement From a Finite Universe. *Journal of the American Statistical Association*, **47**, 663-685.
- [13] Illian, J., Penttinen, A., Stoyan, H., Stoyan, D. (2008) *Statistical analysis and modelling of spatial point patterns*. Wiley. Statistics in Practice.
- [14] Illian, J. B., Sørbye, S. H., Rue, H. and Hendrichsen, D. K. (2012) Using INLA to fit a complex point process model with temporally varying effects – a case study. *Journal of Environmental Statistics*, **3**.
- [15] Illian, J. B., Sorbye, S. H., Rue, H., (2012) A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). *The Annals of Applied Statistics*, **6**, 1499-1530.
- [16] INE (2016) Documento metodológico do Inquérito ao Emprego. <http://smi.ine.pt/DocumentacaoMetodologicaPorTema?clear=True>

- [17] Lindgren, F., Rue, H., Lindstrom, J. (2011) An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach (with discussion). *Journal of Royal Statistical Society Series B*, **73**, 423-498.
- 535
- [18] Lopez-Vizcaino, E., Lombardia, M. J., Morales, D. (2015) Small area estimation of labour force indicators under a multinomial model with correlated time and area effects. *Journal of Royal Statistical Society Series A*, **178**, 535-565.
- [19] Marhuenda, Y., Molina, I., Morales, D. (2013) Small area estimation with spatio-temporal Fay-Herriot models. *Computational Statistics and Data Analysis*, **58**, 308-325.
- 540
- [20] Martins, T., G., Simpson, D., Lindgren, F., Rue, H. (2013) Bayesian computing with INLA: New features. *Computational Statistics and Data Analysis*, **67**, 68-83.
- 545
- [21] Molina, I., Saei, A. and Lombardia, M. J. (2007) Small area estimates of labour force participation under a multinomial logit mixed model. *J.R. Statist. Soc. A*, **170**, 975-1000.
- [22] Moller, J., Waagepetersen, R. P. (2004) *Statistical Inference and Simulation for Spatial Point Processes*. Wiley.
- 550
- [23] Moller, J., Syversveen, A. R., Waagepetersen, R. P. (2004) Log gaussian cox processes. *Scandinavian journal of statistics*. **25**, 451-482.
- [24] Nadaraya, E. (1989) *Nonparametric estimation of probability densities and regression curves*. Vol. 20 of Mathematics and its Applications (Soviet Series), Kluwer Academic Publishers Group, Dordrecht. Translated from the Russian by Samuel Kotz.
- 555
- [25] Nadaraya, E. (1964) On estimating regression. *Theory of Probability and its Applications*, **9**, 157-159.

- [26] Opsomer, J.D., Claeskens, G., Ranalli, M.G., Kauermann, G. and Breidt, F.J. (2008) Non-parametric Small area estimation using penalized Spline Regression. *J.R. Statist. Soc. B*, **70**, 265-286.
- [27] Pereira, S., Turkman, F., Correia, L. (2018) Spatio-temporal analysis of regional unemployment rates: A comparison of model based approaches. *Rev-stat.* **16**, 515-536.
- [28] Rao, J.N.K., Molina, I. (2015) *Small Area Estimation* (second edition). New York: Wiley.
- [29] Roos, M., Martins, T., Held, L., Rue, H. (2015) Sensitivity Analysis for Bayesian Hierarchical Models. *Bayesian Analysis*, **10**, pp. 321-349.
- [30] Rue, H., Martino, S., Chopin, N. (2009) Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with discussion). *Journal of the Royal Statistical Society Series B*, **71**, 319-392.
- [31] Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., Lindgren, F. K. (2017) Bayesian computing with INLA: A review. *Annual Reviews of Statistics and Its Applications*, **4**, 395-421.
- [32] Simpson, D. , Illian, J. , Lindgren, F. , Sørbye, S. , and Rue, H. . (2016) Going off grid: Computational efficient inference for log-Gaussian Cox processes. *Biometrika*, **103** ,1-22, 2016. (doi: 10.1093/biomet/asv064).
- [33] Simpson, D. P., Rue, H. , Riebler, A. , Martins, T. G. , and Sørbye, S. H. (2017) Penalising model component complexity: A principled, practical approach to constructing priors (with discussion). *Statistical Science*, **32** ,1-28.
- [34] Spiegelhalter, D. J., Best, N.G., Carlin, B.R., van der Linde, A. (2002) Bayesian measures of model complexity and fit (with discussion). *Journal of Royal Statistical Society Series B*, **64**, 583-639.

- 585 [35] Watson, G. S. (1964) Smooth regression analysis. *Sankhya Ser. A*, **26**,  
359-372.
- [36] You, Y., Chapman, B. (2006) Small area estimation using area level models  
and estimated sampling variances *Survey methodology*, **32**, 97-103
- [37] You, Y. and Zhou, Q. M. (2011) Hierarchical Bayes small area estimation  
590 under a spatial model with application to health survey data. *Survey method-*  
*ology*, **37**, 25-237.
- [38] Y. Yuan, F. E. Bachl, F. Lindgren, D. L. Brochers, J. B. Illian, S. T.  
Buckland, H. Rue, T. Gerrodette. (2017) Point process models for spatio-  
temporal distance sampling data from a large-scale survey of blue whales  
595 *Annals of Applied Statistics*. **11**, 2270-2297.