Unemployment estimation
Spatio-temporal point process approach

Soraia Pereira *, **Feridun Turkman** *, **Luís Correia** **
* CEAUL, Faculty of Sciences, University of Lisbon, Portugal
** National Statistical Institute of Portugal
Corresponding author. E-mail: sapereira@fc.ul.pt

Introduction

The official labor market estimates, published quarterly by the INE (National Institute of Statistics) are based on a direct method using the Portuguese Labor Force Survey sample. These estimates are available both at national level and for NUTS II regions of Portugal. Currently, knowledge of the labor market requires reliable estimates at a more disaggregated level, particularly at NUTS III level. However, due to the small size of these areas, there is insufficient information to obtain estimates with acceptable accuracy using the direct method. See figure 1 for the relationship between the NUTS in Continental Portugal.

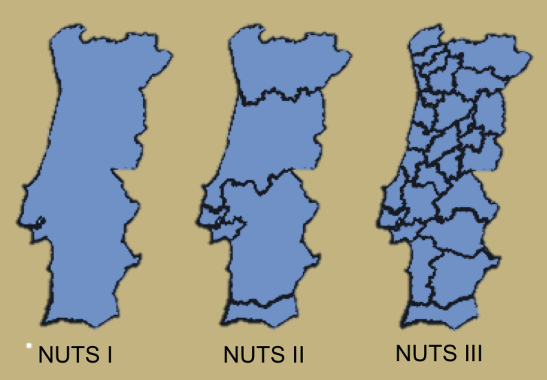

Figure 1: NUTS I, NUTS II and NUTS III in Continental Portugal

Main Objectives

Recently, there have been considerable methodological developments to solve small area estimation problems. The majority of these methods are based on generalized linear models applied to areal data. In this study, we propose the application of a spatio-temporal point process approach, thereby taking into account the individual spatial location and information specific to the individuals, unlike the aggregate data models that do not take into account this information. For modeling the relative intensity of unemployment, we used a Cox log-Gaussian model, using the SPDE and INLA methodologies.

Data and Methods

In this study, we analyzed the quarterly data of the Labour Force Survey from the 1st quarter of 2011 to the 4th quarter of 2013. The locations of the sampled unemployed people in this period are illustrated in Figure 2. As we might expect, the point pattern shows a greater intensity in the regions with more population density. For this study we are interested in modeling the relative intensity of unemployment, so we must also take into account the location of individuals sampled who are not unemployed.

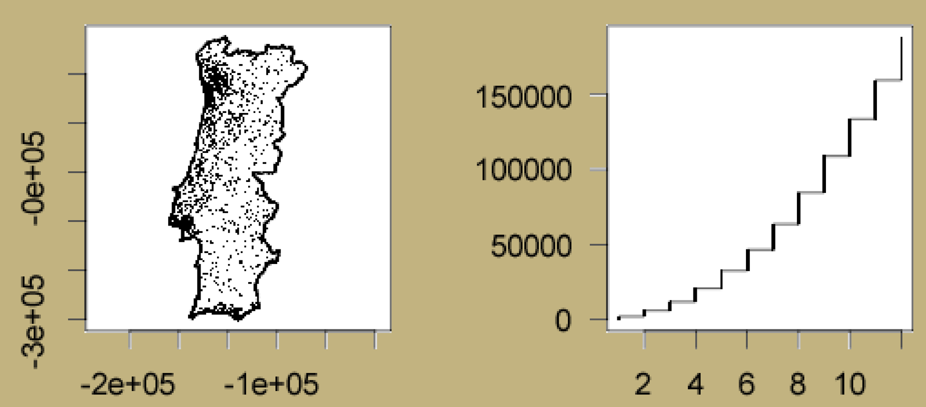

Figure 2: Static display of the data consisting of locations in the left-hand panel and the cumulative distribution of the times in the right-hand panel

Point referenced data are realizations of a spatial point process $(y(s), s \in D)$ caracterized by a spatial index s which varies continuously in the fixed domain D. Blangiardo et al (2015) provide some very useful tools for implementing methods that can be used when modeling these processes.
A simple model widely used for modeling of pontual processes is the inhomogeneous Poisson model. The temporal extension assumes that the number of points within a region D at time t has Poisson distribution with mean $\Lambda(D) = \int_D \lambda(s,t)ds$ where $\lambda(s,t)$ is the intensity surface of the process at time t. If the intensity surface is treated as a realization of a random field, we call it by Cox processes.
In this study, we are interested in the case of log-Gaussian Cox processes, where the intensity surface is modeled as

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

where $Z(s,t)$ is a gaussian random field.
This field can be modeled using the linear predictor

$$Z(s,t) = \mu(t) + \xi(s) + \phi(s,t) \tag{2}$$

where $\mu(t)$ is the temporal dependence, $\xi(s)$ is the spatial dependence and $\phi(s,t)$ is the spatio-temporal dependence. Each of these components can be decomposed into deterministic components (covariates) and stochastic components (random effects).
These processes fit naturally within a Bayesian hierarchical modeling approach. The inference made was based on the method proposed by Simpson et al (2011), using for this purpose the INLA and SPDE approaches.

The covariates considered were: age, gender, education level (available at individual and quarter levels), and the proportion of registered unemployed in the employment centers (available at NUTS III and quarter levels).
An autoregressive process of order 1 (AR1) was assumed for the random effects, and a Gaussian process stationary and isotropic on space for the spatial random effects. For the spatio-temporal effects a Gaussian process stationary and isotropic on space and autoregressive and stationary on time, was assumed.
The temporal and spatio-temporal random effects were not considered significant, probably due to the small size of the study period (12 quarters).

Results

The maps of the posterior mean from the 1st quarter of 2011 to the 4th quarter of 2013 are reported in figure 3. We can see some seasonality in Algarve and Porto as we would expect. Unemployment in Algarve is lower in the summer probably due to tourism and unemployment in Porto is higher during this season probably due to agriculture, which is a common activity in the regions near Porto, and adversely affected in the summer months.

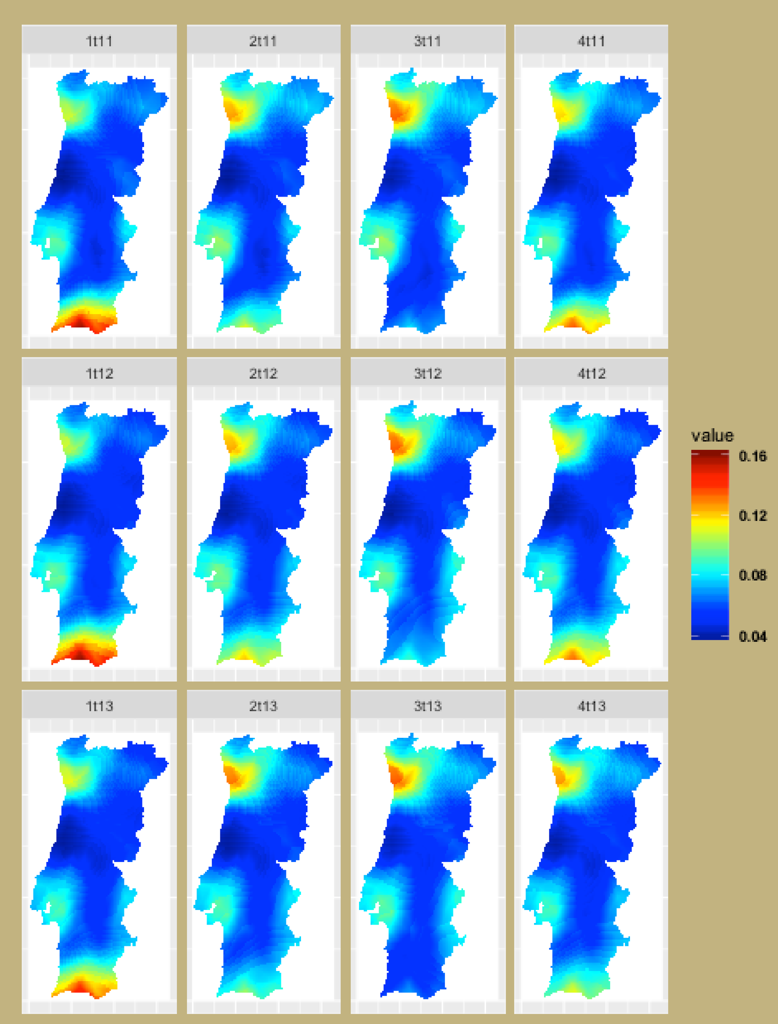

Figure 3: Maps of the posterior mean of the relative intensity from 1st quarter of 2011 to the 4th quarter of 2013.

In addition to taking into account the locations of the individuals, this approach has the advantage of providing estimates for any area (NUTS, district, municipality, or even non-administrative regions). Figure 4 shows the posterior mean by NUTS III for the 4th quarter of 2013.

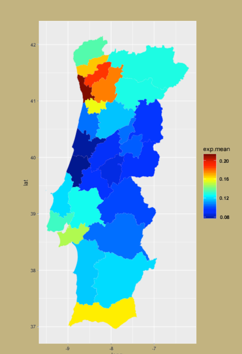

Figure 4: Map of the posterior mean at the NUTS III level for the 4th quarter of 2013

Acknowledgements

We would like to thank CEAUL for the funding through the project UDI/MAT/00006/2013 and FCT through the PhD grant SFRH/BD/92728/2013.

Note

The information above is part of an ongoing study and does not reflect the official opinion of INE.

References

Blangiardo, M., Cameletti, M. (2015) Spatial and Spatio-temporal Bayesian Models with R-INLA. Wiley. pp. 193-258.
Simpson, D., Illian, J.B., Lindgren, F., Sørbye, S.H., Rue, H. (2011) Going off grid: Computationally efficient inference for log-Gaussian Cox processes. https://arxiv.org/pdf/1111.0641v3.pdf.

Encontro Ciência 2016