Changes in the Spatial Distribution of Pollutant Releases in Europe

Marta R. Casanova
Department of Applied Economics
University of Valencia

Vicente Orts
Department de Economics
University Jaume I

José M. Albert
Department of Economics
University Jaume I

Last Version: October 2016

In this paper we estimate the intensity function of a Cox process modelling the point pattern representing the spatial distribution of pollutant releases in Europe, in two different moments in time. The intensity of a Cox process is variable, depending on the location; therefore we can work with inhomogeneous spatial point patterns. We obtain a non-parametric spatial distribution function, which allows us to detect a clear tendency toward clustering in the location of polluting facilities in some areas of the European Union and the study of the spatial variations of pollutant releases in Europe.

Keywords: *pollutant releases, Cox process, inhomogeneous point process, agglomeration*

JEL codes: *C15, C21, L16, Q21, R30*

1. Introduction

Climate change is one of the most important challenges on the agenda of governments, politicians and institutions worldwide. Representatives of the world's governments meet regularly in order to reach an international agreement to combat climate change.

The United Nations Framework Convention on Climate Change (UNFCCC) was created, in 1992, by the United Nations General Assembly to establish a framework for intergovernmental negotiations, in order to agree an international response to climate change. This year, the UNFCCC tried to combat climate change by stabilizing global emissions and so limiting average global temperature increases. This non-binding agreement soon had to be improved to strengthen the global response to climate change and, in 1997, these negotiations concluded with the Kyoto Protocol.

The Kyoto Protocol adopted binding emission reduction targets for all developed and transition-economy countries, but did not set emissions targets for developing countries. However, due to a difficult process of ratification, it did not enter into force until 2005. This protocol affected to 37 industrialized countries and the European community in its first commitment period. The Protocol's first commitment period started in 2008 and ended in 2012 and only binded to developed countries because it recognized that these countries are the main responsible for the current high levels of greenhouse gases (GHG) emissions in the atmosphere as a result of more than 150 years of industrial activity. The second commitment period began in 2013 and will end in 2020.

At the Paris climate conference, celebrated in 2015, parties launched new negotiations to agree a climate deal for the period beyond 2020. Therefore, the Paris Agreement seeks to accelerate and intensify the actions to strengthen the ability of countries to deal with the impacts of climate change. This agreement will try to keep a global temperature rise, this century, below 2 degrees Celsius above pre-industrial levels and will pursue efforts to limit the temperature increase even further to 1.5 degrees Celsius. Moreover, the Paris Agreement requires all Parties to put forward their best efforts through "nationally determined contributions" (NDCs) and to strengthen these efforts in the years ahead. This includes requirements that all Parties report regularly on their emissions and on their implementation efforts.

In this way, by means of this brief summary about climate change politics, we have realized that from finals of last century many measures have been adopted to decelerate the negative impacts of the climate change; however, do the adopted measures really get the desired aims initially?

To answer this question, we are going to compare the pollutant releases emitted by European facilities in two different years (2007 and 2013). To be more precise, we are going to analyse the existing changes in the spatial distribution of pollutant releases during these two years. In this way, we are not going only to know whether the emissions of pollutant releases increase or decrease during the six years studied, but we will also know where these emissions suffer the changes.

Spatial agglomeration or clustering is one of the key stylized facts of the geographical location of individual firms and industries. Consequently, the spatial distribution of economic activity is basically inhomogeneous, and the spatial inequalities become the norm more than the exception at many economic relevant levels. The topics of economic activity location and spatial concentration of firms have attracted the attention of economists for long time ago. Thus, from the pioneering works of Von Thünen (1826), Marshall, (1890), Weber (1909), Lösch, (1940), Hoover (1948), among others, to more recent contributions of the "new economic geography" (for a survey of this literature see Duranton, (1997), Ottaviano and Puga, (1998), Fujita, Krugman and Venables, (1999), Neary, (2001), and Jovanovic, (2007)), the interest of economists has been to characterize the patterns of geographic concentration of firms and industries, and analyze the forces that can allow to explain these inequalities across space. However, in contrast to the recent advances in our understanding of economic forces that can allow to explain spatial disparities, the advances in the empirical characterization of the spatial patterns of economic activity have been comparatively small, and a lot of work must be done to provide both accurate feedbacks to theory, and strong stylized facts in this field.

The work on the characterization of the spatial distribution of economic activity has focused mainly on the calculation of indices, more or less complex, and more recently, estimating homogeneous spatial distribution functions (Sweeney and Feser (1998), Marcon and Puech (2003), Duranton and Overman (2005)). These new tools use distance based methods, avoiding one of the more annoying problems of traditional indices, the modifiable areal unit problem (MAUP), that is, their dependence of the particular administrative scale chosen (see Figure 1). Nevertheless, these methods face at least one important shortcoming: they suppose homogeneous spatial processes[1], becoming inappropriate to analyze non-homogeneous point sets. In this paper we take a step more forward and propose a distance-based method[2], but able to detect spatial structure of inhomogeneous process, the estimation of an intensity function of a Cox process.

Finally, we illustrate the use of this tool with an application to an interesting database – one that deals with information on the polluting facilities[3] and their spatial location within the Europe territory. Thus, the second objective of the paper is to test statistically whether there is some regularity in the selection of the spatial location of polluting facilities in Europe, with significant differential agglomeration in specific locations, which would open the debate on the possible existence of polluting havens into Europe.

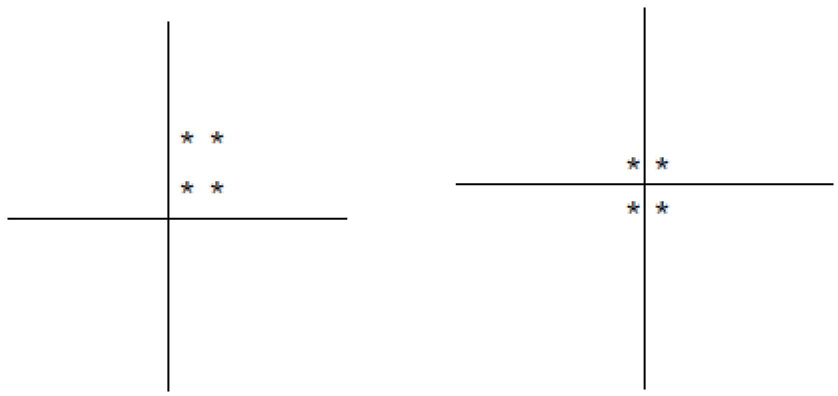


Figure 1: The cluster level of a certain points set can depend on the regional borders

The rest of de paper is organized as follows. Section 2 shows the statistical framework, section 3 outlines the methodology used for the intensity function estimation, in section 4 we describe the data and develop a practical application analysing the spatial

[1] These authors used as method of analysis the *K*-function (Ripley, 1977), that means that the measures only yield valid results when they are used to asses spatial patterns where there is no large scale variation in the mean of the process (homogeneous point process).
[2] This method is also used by Quah and Simpson (2003).
[3] Data obtained from the European Pollutant Release and Transfer Register (E-PRTR).

distribution of the pollutant releases in Europe, and finally, in the section 5 some final considerations are made.

2. Statistical framework

Definition 1. *A point process is a stochastic model governing the locations of events* $\{s_i\}$ *in some set X, a bounded region in* R^d*, Cressie (1993).* A spatial point pattern is a partial realisation of a stochastic point process (Cox and Isham (1980)). Such a realization may be visualized very easily as a cloud of points in a bounded region.

Definition 2. *A homogeneous Poisson process with intensity* λ *is defined by:*

1. *The number of events* $s_1, s_2, ..., s_n$ *in any bounded region* $A \in X$ *follows a Poisson distribution with mean* ¸ $\lambda|A|$
(being $|A|$ = surface of A)

2. *The n events are uniform and independently distributed in A.*

In accordance with condition (2.1), the intensity λ is constant. The condition (2.2) avoids the existence of interactions among the events. In economics the presence of agglomeration phenomena is usual, and this kind of phenomena cannot be treated with homogeneous processes, we must work with inhomogeneous point processes. Any inhomogeneous process defines a variation of the inhomogeneous Poisson process in the following way.

Definition 3. *Let* $\lambda(s): X \rightarrow R_+$ *be a non-constant function on X. The set of events* $s_1, s_2, ..., s_n$ *in X is an inhomogeneous Poisson process if:*

1. *For each bounded region* $A \in X$ *the number of events N(A) = n follows a Poisson distribution with mean* $\int_A \lambda(s)ds$.

2. *Given n events in A, the localizations* $s_1, s_2, ..., s_n$ *in X form a random and independent sample of this distribution on A with a density that depends on the intensity function* $\lambda(s)$, $s \in A$*, and whose values depend on the different localizations.*

In some applications it is useful to allow that the intensity function $\lambda(s)$ of a Poisson process can vary in space, to transform itself into a stochastic process $\Lambda(s)$.

Definition 4.

1. *Let* $\Lambda(s)$, $s \in X$ *denote a non-negative stochastic process on X.*

2. *A realization of* $\Lambda(s)$ *will be a Cox process driven by* Λ *if this realization is an inhomogeneous Poisson process with intensity function* $\Lambda = \lambda$

The result of the realization inherits the properties of the process $\Lambda(s)$ in a natural way.

3. Estimation of the intensity function

In this paper we use the estimation method proposed by Berman and Diggle (1989). Our aim is to estimate the inhomogeneous version of the intensity function λ and $\lambda(s)$ respectively, starting from a data set, $\{s_i \in A : i = 1,2,...,n\}$ being *A* a planar region.

We can estimate λ by the usual estimator

$$\tilde{\lambda} = n / |A|$$

since λ is defined as the expected number of events by area unit.

The usual estimator for $\lambda(s)$ is

$$\tilde{\lambda}(s,h) = (\pi h^2)^{-1} N(\pi h^2),$$

the observed number of events by area unit in a disk with radius *h*, centred in a point $s_i \in s$.

Using kernel method for smoothing point process data,

$$\tilde{\lambda}(s;h) = h^{-2} \sum_{i=1}^{n} f\{h^{-1}(s - s_i)\}$$

where $(s - s_i)$ denotes distance among points, and

$$f(u) = \{\pi^{-1} : \|u\| \le 1$$

$$0 : \|u\| > 1$$

The border effects are kept in mind modifying the denominator and obtaining

$$\tilde{\lambda}(s;h) = N\left(\pi h^2\right) / A(s;h)$$

where $A(s;h)$ is the area of the intersection among the observed region *A* and the disk with radius *h* centred in *s*.

The final form of the estimator is:

$$\tilde{\lambda}(s;h) = h^{-2} \sum_{i=1}^{n} f\left\{h^{-1}(s - s_i)\right\} / A(s;h)$$

The method chooses the smoothing parameter *h* to be that value minimizing the mean square error, for $MSE(h) = E\left\{\left[\tilde{\lambda}_h(x) - \Lambda(x)\right]^2\right\}$.

4. Application

4.1. Data

Some of the problems that have persisted along years in environmental data sets were large measurement errors, insufficient information or missing data values. However, this can cease being a drawback with the European Pollutant Release and Transfer Register, henceforth E-PRTR. This database brings together information about which pollutants are being released, where, how much and by whom.

OECD has played an important role in developing the concept of the E-PRTR and supported the development and implementation of an E-PRTR in member countries.

Our study is based in more than 35,000 pollutant releases reported by industrial facilities included in the E-PRTR. These industrial facilities[4] should report information about specific pollutants when they exceed the applicable capacity thresholds[5]. European Pollutant Release and Transfer Register provides a rich source of data about pollutants released to air, water and soil.

The E-PRTR covers the 27 EU Member States[6] as well as Iceland, Liechtenstein, Norway, Serbia and Switzerland. The data is reported annually by each facility for which the applicable thresholds are exceeded. This information is entirely reliable,

[4] Set out in Annex I of E-PRTR Regulation.
[5] Listed in Annex I and II of E-PRTR Regulation.
[6] Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and United Kingdom.

transparent, consistent and comparable along the years and the different countries. In this way, although a little fraction of the actual releases are not captured by E-PRTR because they are below the required reporting threshold, the objective of this European Register is to cover 95% of the emissions of each selected pollutant.

The industrial facilities that report information cover 65 economic activities and our database provides 91 different pollutants, classified in 7 groups: greenhouse gases, other gases, heavy metals, pesticides, chlorinated organic substances, other organic substances and inorganic substances.

We use geographical coordinates (longitude and latitude) in order to locate all the polluting industrial facilities. These coordinates give a precision of at least ± 500 meters and refer to the geographical centre of the site of the facility. Our analysis takes into consideration the whole of the countries which we have information, except for Iceland because its location.

In this way, we analyse the changes that appear in the spatial distribution of pollutant releases, by taking into account two different years (2007 and 2013). This analysis is carried out by means of the geographical coordinates of the industrial facilities that emit pollutants. The first reporting year of the Register, 2007, we have 38,926 pollutant releases, and the last reporting year, 2013, we have 37,690 pollutant releases.

We should highlight that the European Pollutant Emission Register (EPER), the previous register to the E-PRTR, provided more longitudinal information, but its database had less than half of countries, and also the pollutants taken into account were much scarcer. Therefore, as the comparison between both registers was not possible, we decided to use the E-PRTR in order to have comparable and homogeneous information.

4.2. Empirical Results

Figure 1 shows the two point patterns which represent the location of the pollutant facilities taken into account. The left hand side of this figure represents year 2007 and in the right hand side we find the distribution of pollutant releases in 2013.

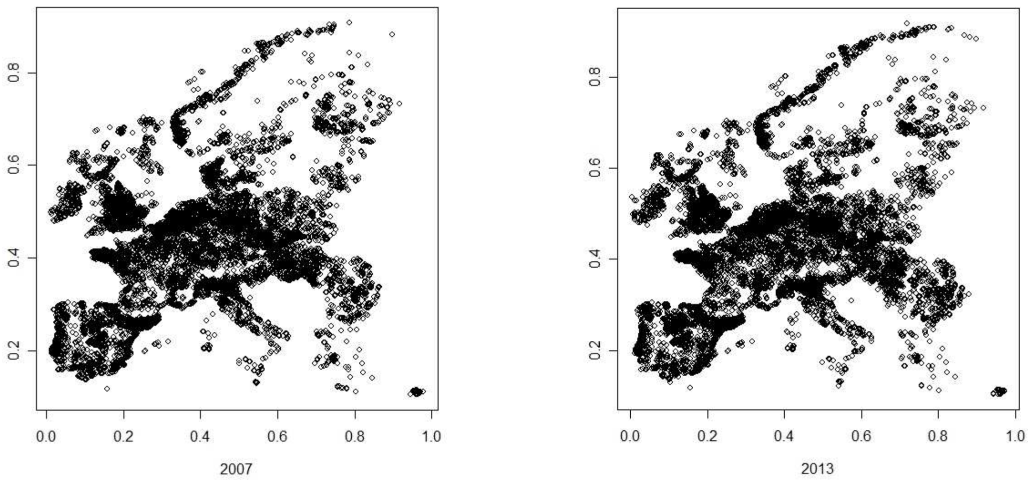


Figure 1. Spatial location of pollutant releases in Europe

A simple visual inspection of this graph makes us think that we are faced with a point pattern which is a realization of an inhomogeneous spatial point process.

Thus, next, we should estimate the intensity function that detects the spatial structure of this inhomogeneous process using the method of Berman and Diggle, by means of the free software *R* and *CRAN* (2004). This software allows us to calculate our intensity function for a certain number of sub-regions inside the study region. The study region of our analysis is divided in 400 equal square sub-regions, i.e. 20x20. Given that the sides of this region measure 3,500 km, approximately, the side of each sub-region will measure 175 km of longitude. The calculation of the intensity has been performed taking an interaction distance of $h = 0.01$ (35 km). This value of h minimizes the mean square error and is the optimum smoothing parameter. That is to say, we considered the distance between any polluting facility and all those located in their surroundings in a radius of 35 km.

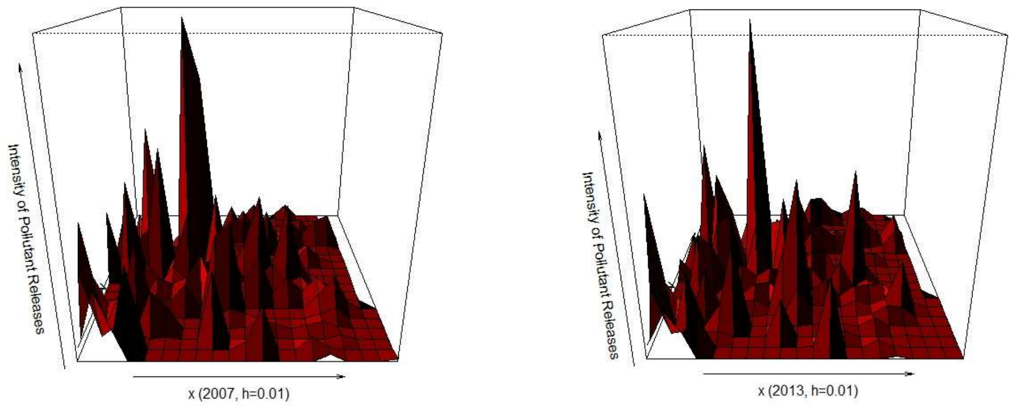


Figure 2. Intensity of pollutant facilities in Europe

The main results are shown graphically in *Figure 2*. The surface variations on the graphs of this figure represent the values of the 400 estimated local intensities, whose magnitude[7] are shown in the chart of the Figure 3 and Figure 4, respectively. Each one of these 400 sub-regions is identified by their geographical coordinates; therefore, for each sub-region we know both, the location and the intensity in pollutant industrial facilities.

	20	0	0	0	0	0	0	0	0	0	0	0	0	0	14547	62661	0	0	0	0	0
	19	0	0	0	0	0	0	0	0	0	0	24528	92050	0	12406	0	0	0	0	0	0
	18	0	0	0	0	0	0	0	0	0	0	85259	0	0	0	0	2525	0	0	0	0
	17	0	0	0	0	0	0	0	18394	981	136362	0	0	0	2554	0	66308	0	0	0	0
	16	0	0	0	0	0	0	0	184419	20037	0	0	0	137328	0	30769	0	0	3461	4055	0
	15	0	0	0	0	0	72654	219880	0	0	0	0	35313	31493	0	112942	55090	117840	146205	0	0
	14	0	0	0	0	0	82578	1016	0	11327	155101	154947	85844	172796	0	0	0	12053	13949	0	0
	13	0	83543	53530	215257	37711	50440	0	0	10941	0	26662	2065	0	0	1917	52763	0	0	0	0
	12	0	3159	233759	46616	0	0	0	0	137144	99326	99533	0	0	0	6328	78518	21799	0	0	0
	11	29894	249769	13179	638529	178115	24744	0	112815	131873	65119	74242	16276	9411	15237	20544	0	0	0	0	0
	10	296892	305	161850	226042	578440	0	1073269	850337	174824	194486	75639	100721	56238	77263	42485	0	0	0	0	0
	9	0	0	7925	11130	8272	137668	326702	42104	406275	64163	188022	396169	86012	359374	120799	0	0	0	0	0
	8	0	0	480518	145717	40685	119344	36220	131961	143975	170574	61316	77345	127749	43609	71714	1503	4884	0	0	0
	7	0	0	0	15898	104082	75806	102811	210520	0	40860	93167	204674	35756	104863	41627	108182	34914	0	0	0
	6	0	0	0	0	34563	43434	111592	0	315900	376234	0	0	0	0	0	67003	145204	35269	0	0
	5	200190	72511	93310	326304	52453	5856	0	0	1557	43479	24336	0	0	0	0	17231	52975	0	0	0
	4	432402	9628	139460	5168	157606	97447	0	0	1672	0	3260	34478	19892	0	13586	31686	8750	0	0	0
	3	273742	11627	60244	33691	200	7209	0	0	278387	0	0	0	0	0	0	0	19727	0	0	0
	2	49481	242281	71132	338843	0	0	0	0	0	0	36160	185155	0	0	56093	15555	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32869	0	0	0	12928
Y		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	X																				

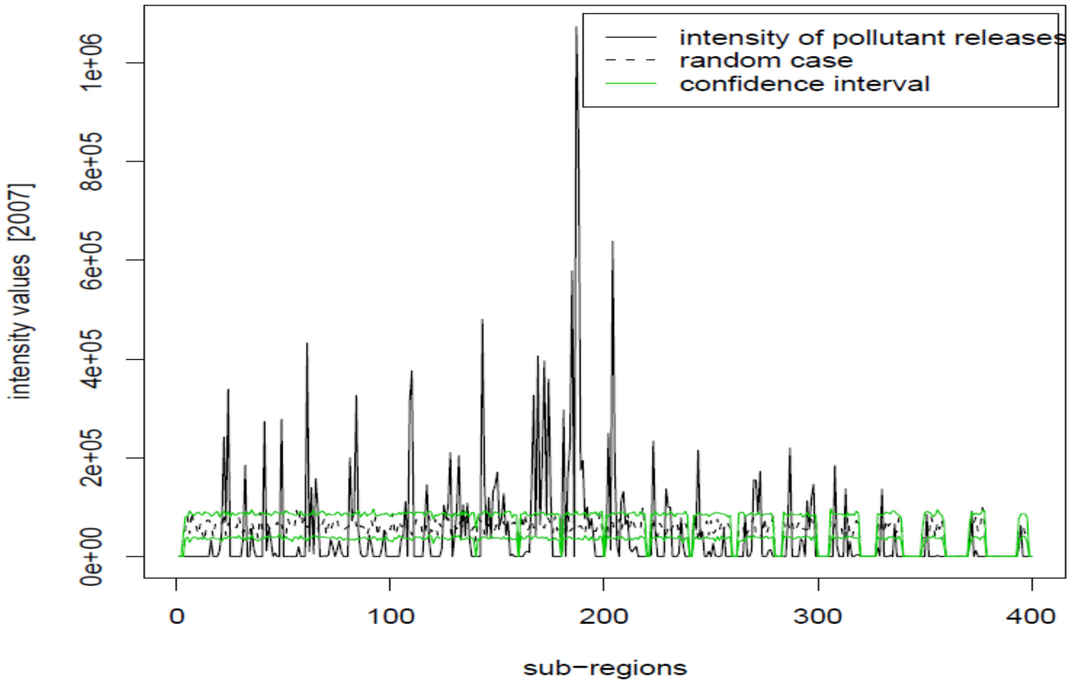
7 This magnitude is function of both, the number of pollutant facilities inside the sub-region, and the number of pollutant facilities located around this sub-region in a distance equal or smaller than 35 kilometres.

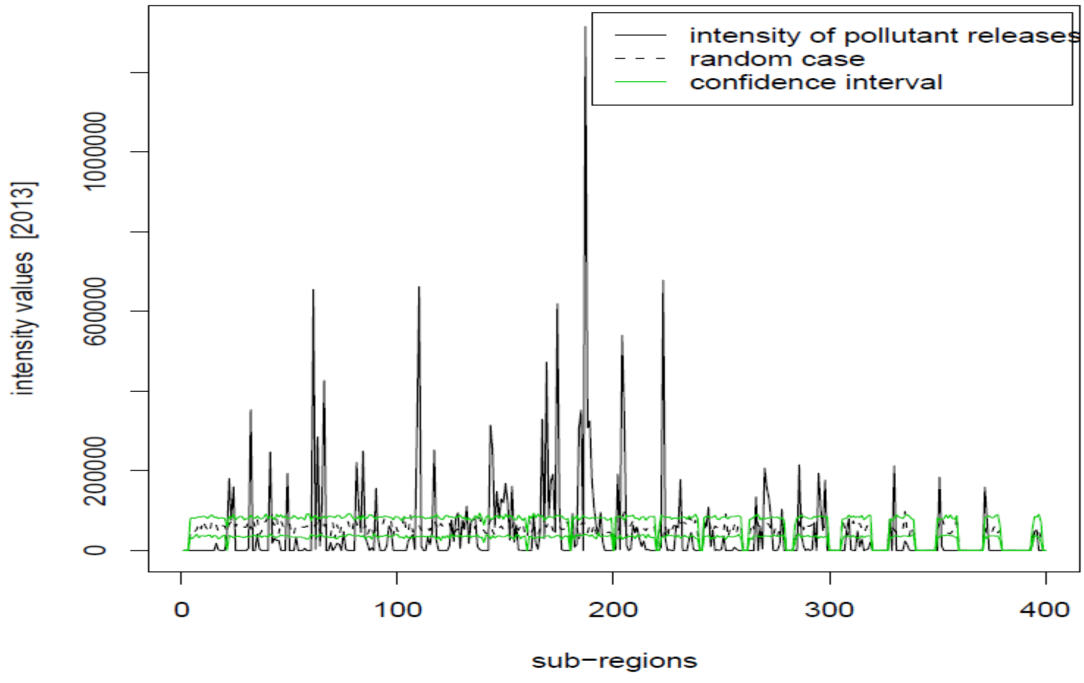
Figure 3. Intensity of pollutant releases in Europe (2007)

	20	0	0	0	0	0	0	0	0	0	0	0	0	0	14791	47498	52168	0	0	0	0
	19	0	0	0	0	0	0	0	0	0	0	0	159168	75630	0	0	0	0	0	0	0
	18	0	0	0	0	0	0	0	0	0	0	183894	11080	0	0	0	0	0	0	0	0
	17	0	0	0	0	0	0	0	77383	0	212628	0	0	0	0	23892	14881	1139	0	0	0
	16	0	0	0	0	0	0	0	51663	79800	1970	0	0	49265	9360	29311	0	2039	4350	21256	0
	15	0	0	0	0	0	215147	40522	10681	0	3962	0	2689	70343	0	193828	102876	50513	176761	0	0
	14	0	0	0	0	0	134576	0	58830	4073	207028	156903	130313	77485	0	0	0	16393	103844	0	0
	13	0	72522	61107	109058	9406	52024	0	2150	0	3140	36971	0	0	0	0	8208	3220	0	0	0
	12	0	8331	678028	28857	0	2173	0	0	62467	57043	178250	0	0	0	27948	45677	3580	0	0	0
	11	3313	192016	2338	540008	335102	14470	0	4162	64059	44771	56114	29701	9964	24134	1301	3247	0	0	0	0
	10	93194	8255	15617	307923	352556	0	1314383	307902	325281	176223	107086	60680	37727	96129	21875	0	0	0	0	0
	9	0	0	91075	18731	3269	44382	329133	29923	472754	83901	174747	191734	45880	619296	105583	2538	0	0	0	0
	8	0	0	314603	249714	58561	148511	80243	129280	120176	168291	121156	26187	161249	21081	64954	1166	0	0	0	0
	7	0	0	0	11657	76783	22446	82051	94925	5627	76730	53540	111195	18104	72034	77581	80351	10124	2909	0	0
	6	0	0	0	0	25365	31246	40239	2566	302671	662265	10972	0	0	34049	14584	45666	252180	24766	0	0
	5	221134	108069	45187	249468	34448	18287	0	7109	0	156239	29199	0	0	0	12139	62283	52079	0	0	0
	4	654971	3358	285198	11001	116407	426472	0	0	20223	0	7551	18466	15499	0	47665	3716	0	0	0	0
	3	246967	18467	30926	25315	25824	0	0	0	193591	0	0	0	33613	0	0	0	5476	0	0	0
	2	38001	180846	76031	159714	0	0	0	0	0	0	6360	353166	0	0	41252	4234	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18005	1943	0	0	0
Y		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	X																				

Figure 4. Intensity of pollutant releases in Europe (2013)

In this paper we take a further step more and present the Cox process as a useful instrument to measure the cluster or concentration of inhomogeneous point processes. Since the results obtained with the method of Berman and Diggle are sensitive both to the number of points and to the distance among them, to make comparable the estimates obtained on points patterns with different number of points, we act in the following way: for each different number of points – in year 2007, 38,926 points and in year 2013, 37,690 points – and, through Monte Carlo simulations, we obtain 100 random points patterns to build confidence intervals with the values obtained when fitting the intensity of a Cox process. Using the measures of intensity that surpass those of the confidence interval, we can make comparisons of concentration measures among different point patterns.





In these Figures, the axis of ordinates measures the fitted values of local intensity, and we use these values as measures of pollutant industrial facilities concentration in EU regions. The confidence interval (95% of confidence) is represented by the distance among the two parallel lines in the low side of the graph. These values above the confidence limit indicate cluster, and the more far away they are from the interval, the larger is the cluster. The axis of abscissas represents each one of the 400 sub-regions in which we have divided our area of study.

The global pollutant releases have increased from year 2007 to year 2013.

We detect a tendency towards concentration of pollutant facilities, in both years, and this tendency is not fortuitous but statistically significant.

We do not find great differences in the spatial distribution of pollutant facilities, although it seems that these ones tend to reallocate in a lower number of locations. So, it is worth noting that those locations with higher levels of pollutant releases tend to get even more elevated levels of pollution.

In this way, should we consider that the adopted agreements do not get the desired aims initially, making that the global emissions reduce their level? Or it depends on regional environmental policies more than on international environmental policies?

References

Berman, M. and Diggle, P.J. (1989) Estimating weighted integrals of the second order in-tensity of a spatial point process. *Journal of the Royal Statistical Society*, B: 81-92.

European Commission (2001) European pollutant emission register. Electronic publication, European Commission, http://www.eper.cec.eu.int

Cox, D.R. and Isham, V. (1980) *Point processes*. Chapman & Hall, London.

Cressie, N. (1993) *Statistics for Spatial Data*. John Wiley and Sons, New York.

Duranton, G. (1997) La nouvelle economie géographique: agglomération et dispersion.

Economie et Prévision, 131(5): 1-24.

Duranton, G. and Overman, H.G. (2005) Testing for localization using micro-geographic data. *Review of Economics Studies*, 72: 1077-1106.

Krugman, P. Fujita, M. and Venables, A.J. (1999) *The Spatial Economy*. MIT Press, Cambridge, Massachusetts.

Hoover, E.M. (1948) *The Location of Economic Activity*. New York: MacGraw-Hill.

Jovanovic, M.N. (2007) *Economic Integration and Spatial Location of Firms and Industries*, volume I, II of *The International Library of Critical Writings in Econonomics*. Edward Elgar Publishing Limited, Cheltenham, UK, Northampton.

Leisch, F. (2004) Precompiled binary distributions. Technical report, http://cran.r-project.org/, June.

Lösch, A. (1940) *Die RaÄumliche Ordnung der Wirtschaft*. Iena, Gustav Fisher. Englishtraslation: The Economics of Location, Yale University Press (1954).

Marcon, E. and Puech, F. (2003) Evaluating the geographic concentration of industries using distance-based methods. *Journal of Economic Geography*, 3: 409-428.

Marshall, A. (1890) *Principles of Economics*. London: Macmillan.

Neary, P. (2001) Of hype and hyperbolas: Introducing the new economic geography. *Journal of Economic Literature*, 39: 536-61, June.

Puga, D. and Ottaviano, G.I.P. (1998) Agglomeration in the global economy: a survey of the new economic geography. *The World Economy*, 21(6): 707-731.

Quah, D.T. and Simpson, H. (2003) Spatial cluster empirics. Working Paper, Economics Department, LSE, London, U.K., June.

Sweeney, S.H. and Feser, E.J. (1998) Plant size and clustering of manufacturing activity. *Geographical Analysis*, 30(1): 45-64.

Von Thünen, J.H. (1826) *Der Isolierte Staad in Beziehung auf Landwirtschaft un NationalÄokonomie*. Hamburg: Perthes. English traslation: The Isolated state, Oxford, Pergammon Press (1966).

Weber, A. (1909) *Ueber den Standort der Industrien*. TÄubingen: J.C.B. Mohr. English traslation: The Theory of the Location of Industries, Chicago University Press (1929).