CREATION OF OZONE MAPS TO REGISTER THE EVOLUTION OF THIS POLLUTANT USING GEOSTATISTICAL TECHNIQUES

López-Rodríguez, F.

Moral, F.J.

Valiente, P.

Pinilla, E.

Universidad de Extremadura

Abstract

Ground-level tropospheric ozone is one of the air pollutants of most concern. It is mainly produced by photochemical processes involving nitrogen oxides and volatile organic compounds in the lower parts of the atmosphere. Ozone levels become particularly high in regions close to high ozone precursor emissions and during summer, when stagnant meteorological conditions with high insolation and high temperatures are common. Tropospheric ozone levels in Europe continue to exceed both target values and the long-term objectives established in EU legislation to protect human health and prevent damage to ecosystems, agricultural crops and materials. In this work, some results of urban ozone distribution patterns in the city of Badajoz , which is the largest (140.000 inhabitants) and most industrialized city in Extremadura region (southwest Spain) are shown.

Keywords: *Ozone, Geostatistics, Geographic Information System, Variogram, Kriging, Map.*

Introduction

The environmental policy is an issue which attracts an important attention in the European Union and, particularly, in Spain, due to the increasing alarm that economic development causes on human health and security, and the worrying events as the Chernobyl disaster, acid rain, and at worldwide level, greenhouse effect or destruction of the ozone layer.

Today´s society worry for nature and its progressive degradation, due to pollution, has as a consequence that people are demanding a less aggressive way of life for the environment, claiming clean industries, ecological produces, etc. Citizens also demand to their governing class different measures and facts to benefit the environment where they live, favoring a better life quality. Moreover, from a planning point of view, future proceedings in urban areas should consider distribution patterns of pollution.

For the particular case in the city of Badajoz, southwestern Spain, the incidence and the spatial distribution of the atmospheric pollution was studied. Therefore, some geostatistical techniques were used to properly analyze and characterize the spatial distribution of the tropospheric ozone, which is taken as indicative of the pollution level, that is, an excessive level of ozone is indicative of high pollution. Later, this information was incorporated in a geographical information system (GIS) to produce accurate ozone maps, where a continuous graphic representation of the pollutant is shown.

1. Materials and Methods

The first stage of the work was to take samples in the city. Thus, 138 urban locations were chosen as sample points, covering the majority part of the city and taking into account its different characteristic, as inhabitants density, type of streets or roads, etc. Sampling interval was not uniform, ranging from, approximately, 75 to 500 m. This favours the geostatistical study and, in consequence, the interpolation and mapping process. An automatic portable analyzer, based on UV absorption, was used to obtain air ozone concentration, in parts per billion by volume (ppbV).

14 sampling campaigns, one per month (two in August), were carried out between May 2007 and August 2007. All noise measurements were made on working days and under suitable meteorological conditions (no cloudy days). Ground ozone levels were measured between 4 p.m. and 10 p.m. for all sampling campaigns, because during that time span maximum ozone concentrations occur.

Therefore, the final data set consists of ground-level ozone measurements from 138 locations situated throughout Badajoz. For each location, its geographic coordinates were ascertained using a GPS device. From this information, using the GIS software ArcGIS® (version 9.2) and its extension Geostatistical Analyst®, the geostatistical study was carried out and different maps were generated to visualize the spatial distribution of the variable in the experimental area, that is, the city of Badajoz.

2. Geostatistical Estimates

Geostatistics can be defined as the set of tools and techniques to analyze the spatial patterns and predict at unsampled locations the values of a continuous variable distributed in space or in time. It is also denominated spatial statistics, due to its direct application to GIS.

All geostatistical study has to fulfill 3 stages (e.g. Isaaks y Srivastava, 1989):

1. Exploratory analysis of data. Geographical distribution of data is not taking into account. Statistics was applied to check data consistency, removing outliers and identifying statistical distribution where data came from.
2. Structural analysis of data. Spatial distribution of the variable was analyzed. Spatial correlation or dependence can be quantified with semivariograms (or variograms). These function relate the semivariance, half the expected squared difference between paired data values $Z(x_i)$ and $Z(x_i+h)$, to the lag distance, h, by which sample points are separated. When a experimental variogram is defined, i.e. some points of a variogram plot are determined by calculating variogram at different lags, a model (theorical variogram) should be fitted to the points.
3. Predictions. The main objective of a geostatistical study is to get estimates of values of the studied variable at unsampled locations, considering the spatial distribution pattern and integrating information from sample points and observed or known trends, if they

exist. Simulations can also be carried out, taking into account the chosen spatial patterns.

Usually, the variogram is the function used to model the spatial variability. For discrete sampling locations, the function is estimated as:

$$\gamma(h) = \frac{1}{2\,N(h)} \sum_{i=1}^{N(h)} \{Z(x_i) - Z(x_i + h)\}^2 \tag{1}$$

where $\gamma(h)$ is the experimental semivariance value at distance interval h, $Z(x_i)$ are the measured sample values at sample points xi , in which there are data at x_i and x_i+h; N(h) is the total number of sample pairs within the distance interval h. For irregular sampling, h is represented by a distance band because the distance between the sample pairs to be exactly equal to h is very rare.

Geostatistics offers a great variety of methods that provide estimates for unsampled locations. These methods are known as kriging, in honor of Danie Krige, who first formulated this form of interpolation in 1951. Kriging is regarded as the best linear unbiased estimator (BLUE), which is a process of a theoretical weighted moving average:

$$\hat{Z}(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{2}$$

where $\hat{Z}(x_0)$ is the value to be estimated at the location x_0, $Z(x_i)$ is the known value at the sampling place x_i, n is the number of the closest samples used for estimation, and the weights for sample values, λ_i, are calculated based on the parameters of the variogram model. The sum of all weights must be one due to the necessity for ensuring that estimates are unbiased.

The main characteristics that make kriging superior to other traditional interpolation techniques, as inverse distance weighting, triangulation, etc., have been described elsewhere (e.g., Goovaerts, 1997; Moral, 2003).

The more important idea of kriging is based on the theory of regionalized variables (see, for instance, Isaaks y Srivastava, 1989), which show spatial autocorrelation such that samples close together in space are more alike than those that are further apart. All different types of kriging are distinguished depending on the chosen model for the trend of the random function (e.g. Goovaerts, 1997). In this work, the geostatistical interpolation method known as ordinary kriging was used. This procedure considers that the mean fluctuates locally; thus, stationarity is limited to local areas. Deutsch and Journel (1992) described ordinary kriging as the anchor algorithm of geostatistics because of its robustness under different conditions.

3. Results

From measured values of tropospheric ozone, for all sampling locations (Figure 3), a map, where estimated values of that pollutant are shown everywhere in the experimental area with an adequate accuracy, should be generated. A complete geostatistical study was performed to achieve the objective. Thus, during the phase of exploratory analysis of data, histograms and normal QQplots for each sampling campaign indicate normality: the shape of the histograms looks bell shaped and the points of these plots are located close the 45º line (Figure 1 contains the histograms and QQplots for one sampling campaign, which is similar for all others). Moreover, descriptive statistics (Table 1) confirm the normality of the data distribution. The fact that mean and median are very similar and the skewness values are near zero is also indicative of data coming from a normal distribution

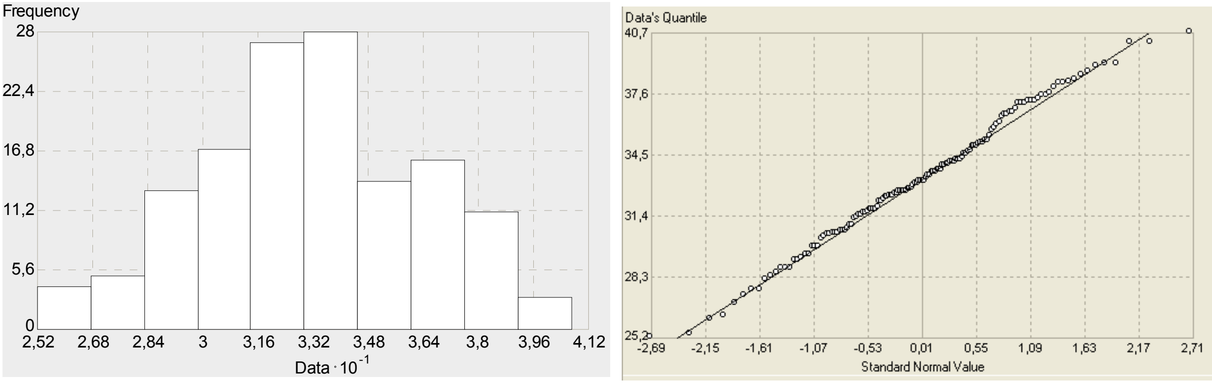


Figure 1: Histogram (left) and QQplot (right) for data, ground-level ozone (ppbV), corresponding to the 17th-19th of July 2007 sampling campaign.

Table 1: Statistics of the ground-level ozone measurements made in 138 points of the city and for all sampling campaigns.

	May 21-24	June 25-28	July 17-19	August 8-10	August 18-20
Mínimum (ppbv)	*25.90*	*26.20*	*25.20*	*35.00*	*27.00*
Maximum (ppbv)	*43.20*	*45.70*	*40.70*	*47.00*	*42.00*
Mean (ppbv)	*33.13*	*36.21*	*33.21*	*39.73*	*36.41*
Median (ppbv)	*33.10*	*36.70*	*33.10*	*39.00*	*37.00*
Standard Deviation (ppbv)	*4.25*	*4.32*	*3.30*	*2.57*	*2.93*
Skewness	*0.04*	*-0.42*	*-0.04*	*0.62*	*-0.67*
Kurtosis	*2.02*	*2.62*	*2.65*	*3.21*	*3.39*

Experimental variograms were determined assuming isotropy conditions because there were no reasons to justify the consideration of anisotropy and, what is more important, with 138 sample points, the influence of different directions in space had supposed the impossibility to define acceptable directional variograms. Therefore, spatial correlation does not depend on directions and experimental variograms were calculated with a directional tolerance of 360º (omnidirectional). When the experimental variogram was calculated, a theoretical variogram was fitted to their points. It is known how the choice of a particular variogram model implies a belief in a certain kind of spatial variability. Possibly, a variable like ground–level ozone is not evenly distributed in reduced distances. In these cases, exponential and spherical models are the most suitable (e,g. Isaaks y Srivastava, 1989); the spherical ones were finally chosen.

Finally, during the most important stage of a geostatistical work, the estimation, the ordinary kriging method was used (e.g. Moral, 2003). A grid, constituted of 40 m side square cells, was designed and superimposed on the city, and estimates were conducted at center of the cells, i.e., ozone concentrations were estimated at a spatial resolution of 40 m throughout Badajoz. The number of observations (neighbours) that were used to estimate the value at each location is at least the closer 15 sample points. From the estimated values, the distribution of ozone levels in the city of Badajoz can be mapped (Figure 3).

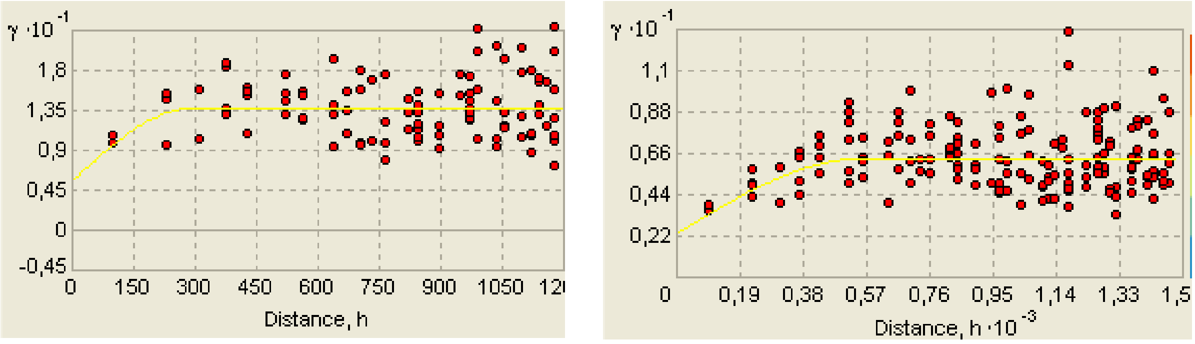


Figure 2: Experimental variograms (red points) and theoretical spherical variograms (yellow lines) for data corresponding to the 25th-28th of June (left) and 8th-10th of August (right) 2007 sampling campaigns.

One important advantage of the geostatistical interpolation techniques with respect to other estimation methods is the possibility of obtaining the reliability of estimates. Thus, together with the estimated value, another output of kriging can be obtained for each location: the kriging variance, or its square root, the kriging standard deviation (KSD). KSD can be mapped similarly to estimates, giving an idea of the quality of the estimates at different places. However, these maps should be used with caution because the reliability of kriging depends on how accurately the variation is represented by the chosen spatial model. Thus, if the nugget effect is overestimated, our estimates could be more reliable than they appear. In general, areas with many sample points or areas where data were sparse but evenly distributed had the most reliable estimates.

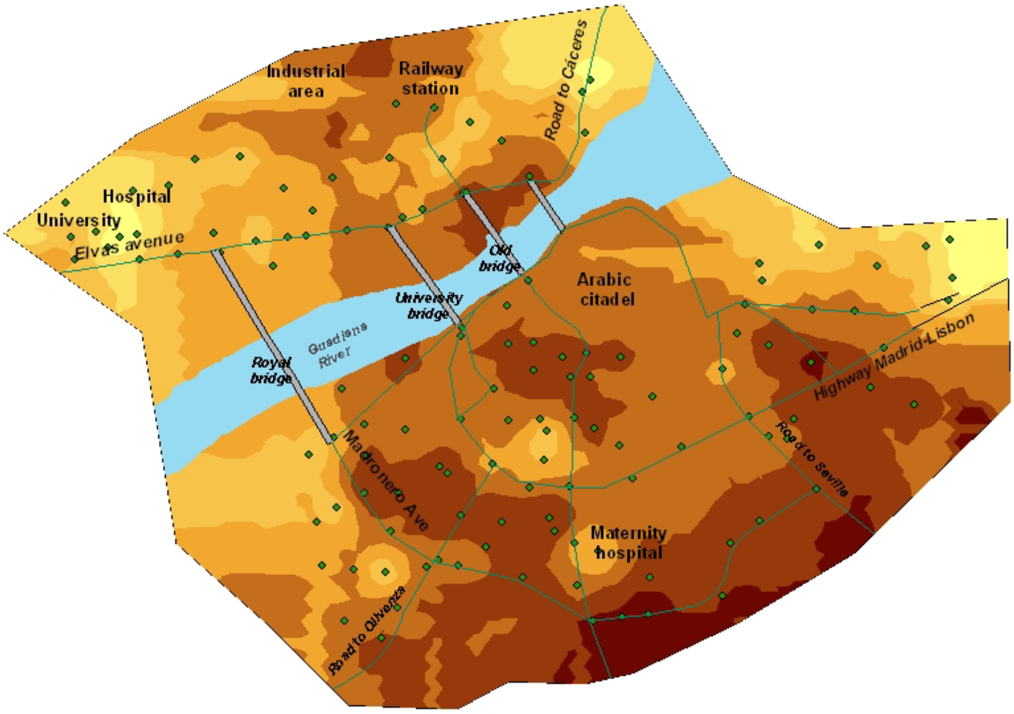
Another interesting application of geostatistics related to ground-level ozone studies is the generation of probability maps (Goovaerts, 1997), which are based on the combination of kriging map and KSD map. For example, if ozone levels higher than 42 ppbV are not considered optimum, according to the regional directive, areas which are likely to surpass that threshold can be delimited. Thus, Figure 4 shows the probability maps corresponding to the 25th-28th of June and 8th-10th of August 2007 sampling campaigns, in which areas with high risk of ground-level ozone exceeding the proposed limit are represented, with the probabilities providing a measurement of confidence for hazard assessment of ozone concentration. Areas with low probabilities, for example <25%, could be regarded as “clean” zones where the ozone level is unlikely to be higher than 42 ppbV and, on the other hand, areas with high probabilities, for example >50%, could be regarded as “dangerous” zones where the noise level is very likely to be higher than 42 ppbV.

4. Conclusions

In this work, using geostatistical techniques, the double objective of characterizing the spatial variability of ground-level ozone and including this information in the interpolation algorithm is achieved; it would not have been possible if other interpolation methods had been applied.

The proposed techniques provide some reliable surfaces at enough spatial resolution to correctly visualize the spatial patterns of this pollutant. Considering different sampling campaigns, the temporal evolution of tropospheric ozone can be analyzed.

Polluted areas in the city have to be delimited. Future actions against ozone should be particularly aimed at reducing the high levels in these zones. Consequently, the ozone maps can influence decisions concerning air-quality policy.



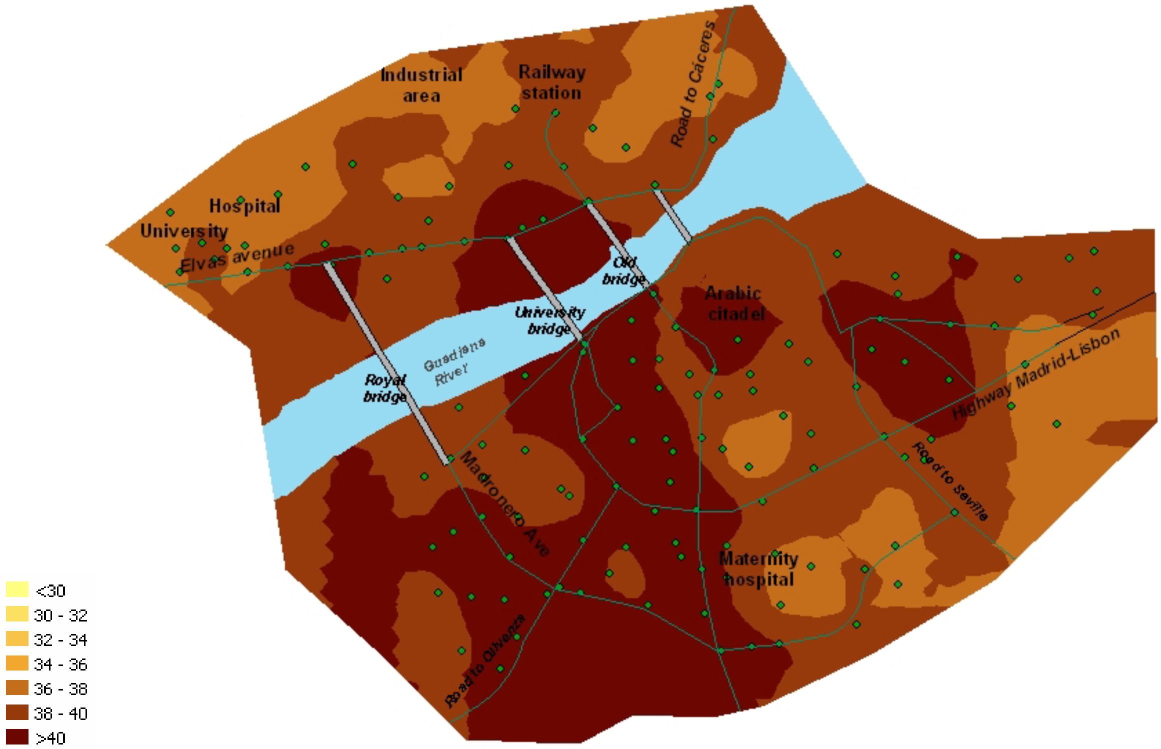


Figure 3: Kriged maps of ground-level ozone in Badajoz city for the 25th-28th of June (up) and 8th-10th of August (down) 2007 sampling campaigns. Scale unity is ppbV.

Green points correspond to sampling locations.

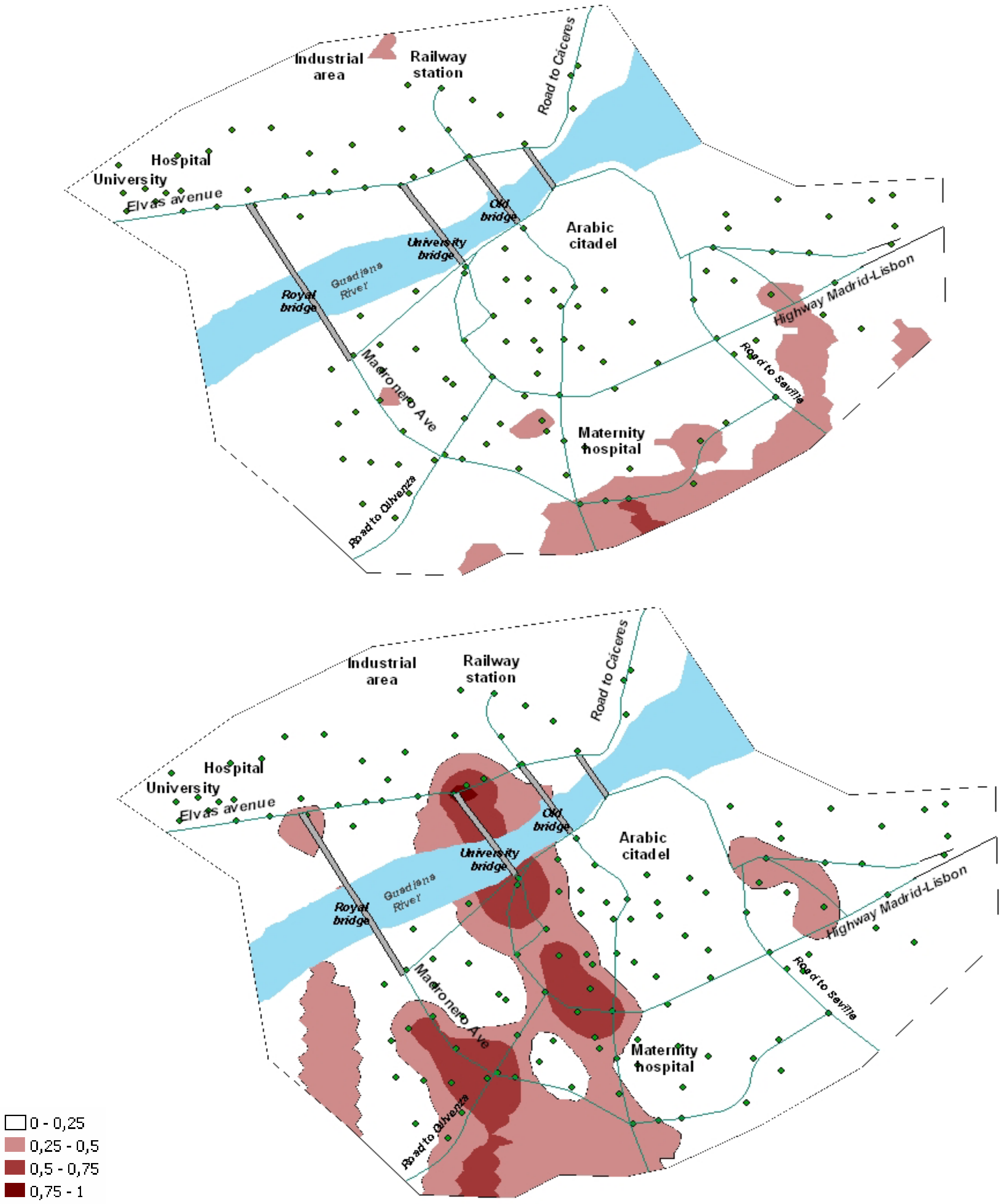


Figure 4: Probability maps of ground-level ozone higher than 42 ppbV in Badajoz city, for the 25th-28th of June (up) and 8th-10th of August (down) 2007 sampling campaigns. Scale unity is probability.

References

Deutsch, C.V., y Journel, A.G. 1992. Gslib: *Geostatistical Software Library and user's guide*, Oxford Univ. Press, New York.

Goovaerts, P., 1997. *Geostatistics for Natural Resources Evaluation*, Oxford Univ. Press, New York. 1997.

Isaaks E.H., and Srivastava, R.M. 1989. *An Introduction to Applied Geostatistics*, Oxford Univ. Press, New York.

Moral, F.J. 2003. *La Representación Gráfica de las Variables Regionalizadas. Geoestadística lineal*. Servicio de Publicaciones de la Universidad de Extremadura, Cáceres.

Correspondence (for further information, please contact):

Fernando López Rodríguez

Phone: (+34) 649833585

E-mail de contacto: ferlopez@unex.es

Empresa / Institución: Universidad de Extremadura