World Academy of Science, Engineering and Technology
International Journal of Civil and Architectural Engineering
Vol:4, No:1, 2010

Geostatistical Analysis and Mapping of Ground-level Ozone in a Medium Sized Urban Area

F. J. Moral García, P. Valiente González, and F. López Rodríguez

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

***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.

In this work, some results of a study about urban ozone distribution patterns in the city of Badajoz, which is the largest and most industrialized city in Extremadura region (southwest Spain) are shown. Fourteen sampling campaigns, at least one per month, were carried out to measure ambient air ozone concentrations, during periods that were selected according to favourable conditions to ozone production, using an automatic portable analyzer.

Later, to evaluate the ozone distribution at the city, the measured ozone data were analyzed using geostatistical techniques. Thus, first, during the exploratory analysis of data, it was revealed that they were distributed normally, which is a desirable property for the subsequent stages of the geostatistical study. Secondly, during the structural analysis of data, theoretical spherical models provided the best fit for all monthly experimental variograms. The parameters of these variograms (sill, range and nugget) revealed that the maximum distance of spatial dependence is between 302-790 m and the variable, air ozone concentration, is not evenly distributed in reduced distances. Finally, predictive ozone maps were derived for all points of the experimental study area, by use of geostatistical algorithms (kriging). High prediction accuracy was obtained in all cases as cross-validation showed. Useful information for hazard assessment was also provided when probability maps, based on kriging interpolation and kriging standard deviation, were produced.

***Keywords*—**Kriging, map, tropospheric ozone, variogram.

I. Introduction

An increasing ozone concentration has been observed in many European countries, from Scandinavia and Britain to Spain [1]-[3]. Some measurements in Spain, particularly in the Mediterranean region, have shown high tropospheric ozone concentrations [4], [5]. Ground ozone levels is a topic of considerable environmental concern, since excessive level of ozone are taken as indicative of high pollution.

Ozone at ground level – in the air we breathe – is not to be confused with the ozone layer in the upper atmosphere, which shields the Earth from the sun's ultraviolet rays. Up there ozone is “good”; at ground level it is “bad”.

Human activites have led to much higher ground-level ozone concentrations in Europe today than in the pre-industrial era. In the troposphere, near the Earth's surface, human activities lead to ozone concentrations several times higher than the natural background level. Too much of this ground-level ozone is “bad” as it is harmful to breathe and also damages the environment. When ozone mixes with other air pollutants, especially nitrogen oxides and particulate matter, it can form a harmful smog. This smog occasionally takes place in polluted city areas.

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 [6]. In addition, ozone is a greenhouse gas which may have important global climatic consequences.

Ozone is a natural component of the troposphere, produced by photochemical reactions of nitrogen oxides and volatile organic compounds (VOC), collectively called ozone precursors, enhanced by temperature and sunlight. Emissions from car exhausts, power plants and industrial facilities are the major sources of nitrogen oxides and VOC. Since 1990 approximately, emissions of ozone precursors in the European Union have declined about 30%, due mainly to the widespread introduction of catalytic converters in cars and new laws to reduce air pollution. However, this has not resulted in comparable decrease in ozone levels, particularly in city centres, which could be explained by the dependence of ozone generation on climatic conditions and complex chemical processes.

Lelieveld and Detener [7] suggest that around 70% of the tropospheric ozone is photochemically produced within the troposphere itself, and about half of this ozone is anthropogenic; only the rest, 30%, is originated from the stratosphere. In the Mediterranean region, these percentages are higher: 90% of the tropospheric ozone is generated in situ and the anthropogenic fraction is around 75% [2].

The reactive nature and the photochemical origin of ozone generate important temporal (hourly, daily, seasonally and annually) and spatial variations in its concentrations. Chemical reactions involving ozone formation (and removal)

F. J. Moral García is with the Graphic Representation Department, University of Extremadura, Badajoz, 06071, Spain (corresponding author phone: +34 924 28 96 00; fax: +34 924 28 96 01; e-mail: fjmoral@unex.es).

P. Valiente González is with the Analytical Chemistry Department, University of Extremadura, Badajoz, 06071, Spain (e-mail: valiente@unex.es).

F. López Rodríguez is with the Graphic Representation Department, University of Extremadura, Badajoz, 06071, Spain (e-mail: ferlopez@unex.es).

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

occur within a time interval of a few hours and over tens of kilometers, and during the next weeks after its generation it can be transported by the wind to hundreds of kilometers.

Although annual variation in tropospheric ozone depends on different factors, as proximity to large source areas of NO_x and VOC, meteorological conditions and geographical location [8], a clear annual cycle over mid-latitudes, with a broad summer maximum in populated and industrialized areas or a spring maximum in rural zones, is apparent. The summer maximum is related to local photochemical productions [9], when stagnant meteorological conditions with high insolation and high temperatures take place, whereas the spring maximum is attributed to enhanced photochemical reactions due to the more important solar radiation which acts on the precursors, NO_x and VOC, accumulated during the winter [10]. According to Monks [11], the maximum ozone levels has increased over the last two decades.

The daily pattern of tropospheric ozone is driven mainly by the cycle of NO_x and hydrocarbon emissions as well as by the solar radiation [12]. During the day, ozone concentrations will usually build up and peak in the afternoon. At nighttime ozone concentrations decline.

Generally, ozone concentrations peak in summertime and in the afternoon, because the formation of ozone needs sunlight.

Ground level ozone has been studied extensively in recent years (e.g. Cocchi and Trevisano [13] provide a critical review of the different approaches) with the aim of getting more accurate predictive models and the need to take confounding effects into account when ozone trends and the health effects of ozone are investigated. However, studies about the temporal y/o spatial variability of ground level ozone are very limited [1], [14], and, as far as we know, there is no any work where a detailed, small-scale variability of this pollutant is analyzed in a city, probably due to the difficulty of obtaining enough measurements.

At the present time, ozone is measured at thousands of stationary monitors scattered across Europe, most of which are located in urban areas. In Spain, there are also an important number of monitoring stations, with uneven densities depending on the different regions. But, in all cases, they are insufficient to perform a precise study about the spatial distribution of tropospheric ozone. Only the tendencies at regional levels, at the very best, can be analyzed, as tens of kilometers of spatial resolution is available.

It is known how monitoring atmospheric pollution in urban areas involves mapping techniques that assist the decision-maker to describe and quantify the pollution at locations where no measurements were available. The preparation of pollution maps is a complex task, which is only feasible if a spatial correlation of the variable of interest is identified [15]. The existence of a spatial correlation of atmospheric pollutants is not only a condition for an optimum interpolation of the data in space in order to generate a map of pollution, but it also provides very useful insights on the structure of the air quality patterns. Some studies have identified a strong spatial variability of air pollutants [16], [17]. The main goal of interpolation is to discern the spatial patterns of atmospheric pollution concentrations by estimating values at unsampled locations based on measurements at sample points. Geostatistics provides an advanced methodology to quantify the spatial features of the studied variables and enables spatial interpolation, kriging [18], [19]. In addition, geographical information systems (GIS) and geostatistics have opened up new ways to study and analyze spatial distributions of regionalized variables, i.e. distributed continuously on space [20], [21]. Moreover, they have become useful tools for the study of hazard assessment and spatial uncertainty [22]. Without a GIS, analysis and management of large spatial data bases may not be possible.

Many air pollution studies have employed distance-weighting methods, e.g. [23], but kriging is the only one which incorporates the spatial correlation into its estimation algorithm. Kriging has been used more widely [20], [24] due to its many advantages [19]. Although kriging requires an abundance of sample points to be an accurate spatial interpolation method [25], even when relative small data sets and not exhaustive samplings are available it is a reliable technique for investigating the distribution and sources of pollutants [26].

To inform decisions regarding, for instance, the protection of public health from elevated ozone levels in a urban area, high-resolution maps are necessary. Therefore, the main objectives of this paper were to: (1) analyze the temporal evolution and characterize the spatial distribution of ground ozone levels using geostatistical techniques; (2) incorporate this information in a GIS to produce accurate ozone maps; (3) assess the hazard of exceeding some limits with a geostatistical basis.

II. Materials and Methods

A. Survey Area. Selection of Sampling Points and Ozone Field Measurements

In this work, some results of urban ozone distribution patterns in the city of Badajoz (38º 53' 12" N, 6º 58' 15" W, 170 meters above mean sea level), a medium-sized ancient town which belongs to the Autonomous Community of Extremadura, southwestern Spain, are shown (Fig. 1). It is the largest (about 140.000 inhabitants) and most industrialized city in this region. The town has a strategic situation on the Portuguese border and on the Madrid-Lisboa highway.

In Badajoz, there is only one monitoring station, which is continuously measuring ozone levels and other pollutants, situated in the northeast of the city (Fig. 2). This station is operated by the Department of Environment of the Extremaduran Government as part of a network (5 monitoring stations) to monitor background air pollution. Thus, the information provided by this monitoring station is indicative of a "mean" ozone level over the town. Since we are interesting in studying the spatial distribution in the town, ground-level ozone measurements at different locations have to be obtained. Therefore, an automatic portable analyzer,

based on UV absorption, was used to obtain air ozone concentration, in parts per billion by volume (ppbV).

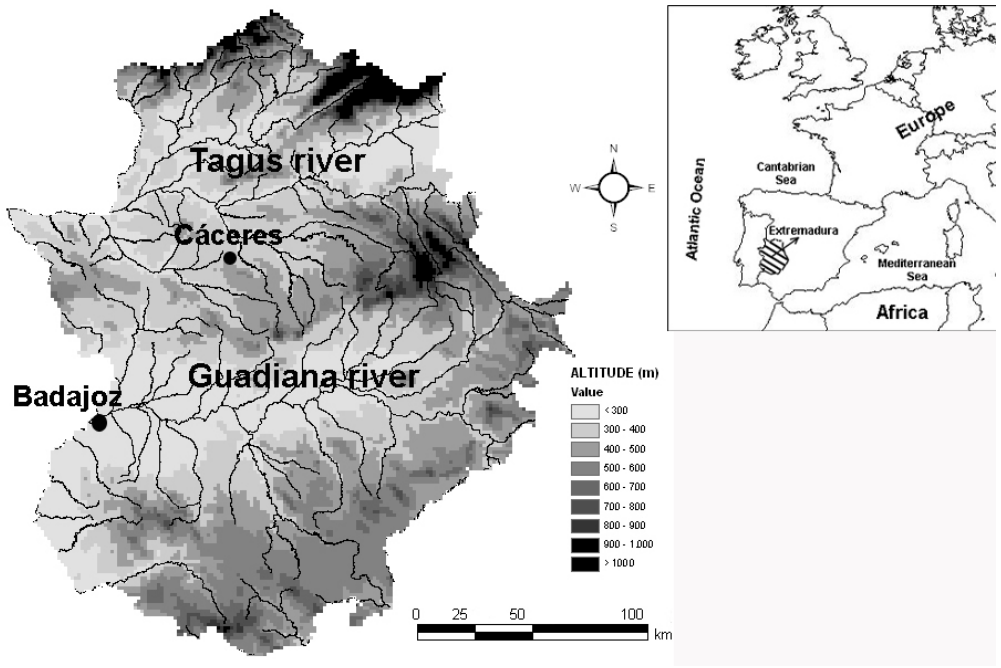


Fig. 1 Location map of Badajoz, with reference to the Autonomous Community of Extremadura and Spain

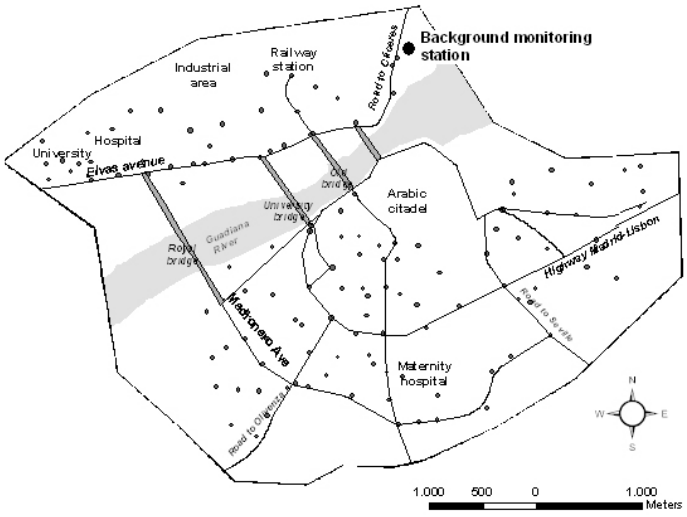


Fig. 2 Map of Badajoz city (urban area) and sampling locations (138)

Another important issue is the selection of all locations where a sample has to be taken. In the context of ozone mapping, a small sampling interval is preferable because it facilitates accurate spatial interpolation and it is essential for assessing operational scale [27]. Operational scale is the spatial extend at which a particular phenomenon operates, being inversely related to spatial complexity, i.e., a phenomenon with small-scale variations has a small operational scale.

Finally, 138 urban locations were chosen as sample points (Fig. 2), 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 (see section III).

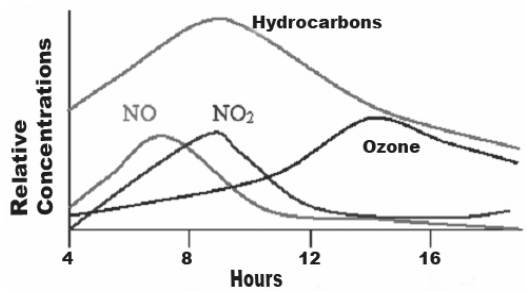


Fig. 3 Typical daily evolution of ozone and its main precursors

An additional problem related to ground-level ozone measurements is the time interval to be considered. If a portable analyzer is used and we are interested in analyzing both the spatial patterns and temporal evolution of ground ozone levels, measurements have to be carried out when the maximum concentrations are apparent. Fig. 3 shows the typical daily evolution of ozone and its main precursors. The maximum ozone levels are reached at early evening. During the morning, traffic flows and industrial activities are very important and, consequently, high concentrations of primary pollutants, precursors, are produced. Later, temperature and

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

sunlight increases and, in consequence, through photochemical reactions, the secondary pollutants (ozone) are generated. In the late evening, when temperature and sunlight decrease, the ozone generation decreases too.

To establish the daily time span where maximum ozone concentration levels are expected and, consequently, the time interval in which samples have to be taken, hourly data from the monitoring station corresponding to some typical days are analyzed. Fig. 4 shows the temporal evolution of ozone levels during three typical spring and summer days. A constant daily pattern is apparent: maximum ozone concentrations occur from midday to late evening. To define more precisely the time interval in which measurements have to be conducted, ozone concentrations were measured simultaneously from 10 a.m. to 10 p.m. approximately during four consecutive days at the same locations. As Fig. 5 shows, the patterns are very similar for each location and maximum was always obtained between 1 p.m. and 8 p.m. approximately. In consequence, ground ozone levels will be measured during that time span for all sampling campaigns.

Another interesting observation from Fig. 5 is the difference, for the same day, of the maximum ozone values for each location. This denotes the existence of a small-scale variability in the city.

When the optimum time interval to measure maximum ozone levels was established, successive sampling campaigns, at least one per month, were conducted. Thus, data were collected at 138 locations in the city of Badajoz during 14 sampling campaigns, one per month (two in August), which were carried out between May 2007 and May 2008. All noise measurements were made on working days and under suitable meteorological conditions (no cloudy days). 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.

A. Geostatistics

After obtaining all ground-level ozone measurements, the spatial distribution of this pollutant in the city was analyzed for each month and later it was necessary to estimate the ozone level at other locations where direct measurements were not carried out.

Since the factors that determine the values of environmental variables are numerous, largely unknown in detail, and interact with a complexity that we can not unravel, we can regard their outcomes as random. If a stochastic point of view is adopted, then there is not just one value for a property but a whole set of values at each point in space. We regard the observed value there as one drawn at random according to some law, from some probability distribution.

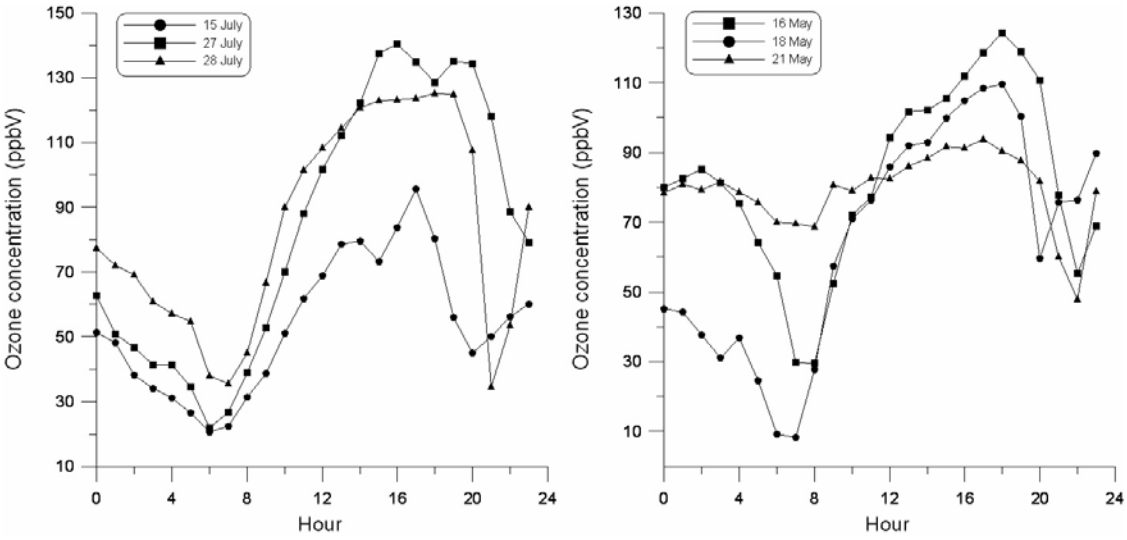


Fig. 4 Temporal evolution of ground-level ozone during three summer (left) and spring (right) days. Data are from the background monitoring station located in Badajoz

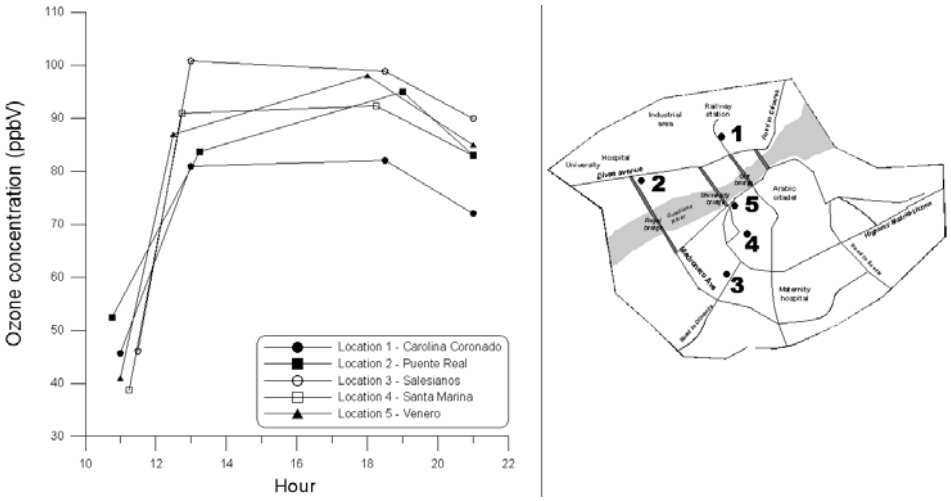


Fig. 5 Temporal evolution of ground-level ozone at five locations (right) in the city of Badajoz. Data are mean values for four consecutive days (April 2007)

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

This point of view, when the studied variable (maximum ground-level ozone) is considered random and distributed continuously on the experimental area (regionalized variable), is adopted to use geostatistics as an estimation technique.

It is recognized that the statistic approach, geostatistical methods or kriging, has several advantages over the deterministic techniques [18], [19]. The fact of giving unbiased predictions with minimum variance and taking into account the spatial correlation between the data recorded at different locations is an important advantage of kriging. Moreover, besides interpolation, kriging provides information on interpolation errors. Such values can be mapped to generate error surfaces which inform about the reliability of 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 [18], [19]. Geostatistics is based on the theory of regionalized variables [28], which show spatial autocorrelation such that samples close together in space are more alike than those that are further apart.

In this study, 3 phases were completed to conduct the geostatistical work [18]:

1º) Exploratory analysis of data. Statistics was applied to check data consistency, removing outliers and identifying statistical distribution where data came from.

Certain kriging methods work best if the data are approximately normally distributed [19]. The histograms and normal QQplots can be used to check the normality.

Furthermore, statistics provide additional numerical information to confirm the graphical tendencies shown with the histograms and QQplots [29].

Outlier is a measured sample point that has a very high or low value relative to the values in the dataset. It is important to detect outliers because they may be values that were measured or recorded incorrectly and, in this case, their effects on subsequent stages of the geostatistical study are very negative [20]. One simple way to detect them is selecting points on the tails of the distribution and checking whether the extreme values are isolated locations, surrounded by very different values (then they may require further research and, if necessary, be removed), or not [29]. Moreover, high skewness values can indicate the existence of outliers.

2º) 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(xi) and Z(xi+h), to the lag distance, h, by which sample points are separated [18], [30]. For discrete sampling locations, the function is estimated as:

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

where γ(h) is the experimental semivariance value at distance interval h, Z(xi) are the measured sample values at sample points xi , in which there are data at xi and xi+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.

The variogram shows the degradation of spatial correlation between two points of space when the separation distance increases. This function has two components [18], [19], [30]; the first is the nugget effect, which characterizes the discontinuity jump observed at the origin of distances, quantifies the short-term, erratic variations of the studied phenomenon plus measurements and data errors. The second is the increasing part of the variogram, which may reach the sill (theoretical sample variance), leveling off the curve, for a distance called range, or keep on increasing continuosly with distance. The non-nugget part of the variogram measures the nonrandom part of the phenomenon and models its average medium-scale behaviour in space.

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

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 [18], [19], [30]. Although there are some statistical techniques to justify the choice of a theoretical variogram [31], subjective criteria and previous experiences are the main tools to choose one. The fitted model provides information about the spatial structure as well as the input paramenters for the last phase, kriging interpolation [18], [30].

3º) 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. 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.

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). This statistical measure indicates the reliability of estimates (it is an additional advantage of kriging over other interpolation techniques, as it was previously stated). KSD depends on the sample distribution and variogram structure.

All different types of kriging are distinguished depending on the chosen model for the trend of the random function. In this work, the geostatistical interpolation method known as ordinary kriging was used [18], [19]. This procedure considers that the mean fluctuates locally; thus, stationarity is limited to local areas. Deutsch and Journel [32] described ordinary kriging as the anchor algorithm of geostatistics because of its robustness under different conditions. Point ordinary kriging is defined when estimates are related to a point.

The geostatistical analysis, including all 3 phases previously described, was carried out with the extension Geostatistical Analyst® of the GIS software ArcGIS® (version 9.2).

Before producing the final maps, cross validations were used to validate the accuracy of all interpolations [18], [19], [29], [30]. Cross validation withholds one observation at a time, estimating the value at that location with the remaining data. Later, for all data points, the difference between the actual and the estimated value is calculated. Finally, some statistics are computed to examine how well the model predicts the values at unknown locations. In this work, the mean prediction error, MPE, the mean standardized prediction error, MSPE, the root-mean-square prediction error, RMSPE, the average kriging standard error, AKSE, and the root-mean-square standardized prediction error, RMSSPE, were used [29]. These are defined as:

$$\mathrm{MPE} = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{Z}(x_i) - Z(x_i)\right] \tag{3}$$

$$\mathrm{MSPE} = \frac{1}{n}\sum_{i=1}^{n}\left[\hat{Z}(x_i) - Z(x_i)\right] / \hat{\sigma}(x_i) \tag{4}$$

$$\mathrm{RMSPE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\hat{Z}(x_i) - Z(x_i)\right]^2} \tag{5}$$

$$\mathrm{AKSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\hat{\sigma}^2(x_i)} \tag{6}$$

$$\mathrm{RMSSPE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\left(\hat{Z}(x_i) - Z(x_i)\right) / \hat{\sigma}(x_i)\right]^2} \tag{7}$$

where $\hat{\sigma}^2(x_i)$ is the kriging variance for location x_i [18], [19], [29], [30].

If the predictions are unbiased, the ME should be near zero. However, this statistic has some important drawbacks: it depends on the scale of the data and is insensitive to inaccuracies in the variogram. So, usually the MSPE is used to standardize the ME, being ideally zero, i.e., an accurate model would have a MSPE close to zero. The RMSPE should be as small as possible, denoting how closely the model predicts the measurement values. Besides making predictions, each of the kriging techniques gives the kriging variances which estimate the variability of the predictions from the known values. The kriging variances must be accurately calculated because they have an important influence on some applications of kriging, e.g., the probability kriging. If the RMSPE is close to the AKSE, the prediction errors are correctly assessed. If the RMSPE is smaller than the AKSE, then the variability of the predictions is overestimated; conversely, if the RMSPE is greater than the AKSE, then the variability of the predictions is underestimated. The same could be deduced from the RMSSPE statistic. It should be close to one. If the RMSSPE is greater than one, the variability of the predictions is underestimated; likewise if it is less than one, the variability is overestimated [29].

After conducting the cross validation process, maps of kriged estimates were generated which provided a visual representation of the distribution of the maximum ground-level ozone in Badajoz. These maps were produced with the ArcMap® module of the ArcGIS®.

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

III. Results and Discussion

A. Exploratory Analysis of Data

TABLE I
STATISTICS OF THE GROUND-LEVEL OZONE MEASUREMENTS MADE IN 138 POINTS OF THE CITY AND 14 SAMPLING CAMPAIGNS (SD = STANDARD DEVIATION)

	May 07	June 07	July 07	August 07	August -2 07	Sept 07	Oct 07
Mean (ppbV)	33.13	36.21	33.21	39.73	36.41	29.41	25.94
Median (ppbV)	33.10	36.70	33.10	39.00	37.00	29.10	25.70
SD (ppbV)	4.25	4.32	3.30	2.57	2.93	3.86	3.65
Minimum (ppbV)	25.9	26.3	25.3	35.0	27.0	20.9	18.9
Maximum (ppbV)	43.2	45.7	40.7	47.0	42.0	38.0	33.4
Skewness	0.03	-0.41	-0.04	0.63	-0.67	0.25	-0.07
Kurtosis	2.02	2.62	2.65	3.21	3.39	2.23	1.99

	Nov 07	Dec 07	Jan 08	Feb 08	March 08	April 08	May 08
Mean (ppbV)	21.39	18.97	20.61	30.18	31.89	35.54	36.7
Median (ppbV)	21.40	19.15	20.35	30.00	32.00	35.40	37.00
SD (ppbV)	1.90	2.08	1.54	1.08	1.43	1.52	1.79
Minimum (ppbV)	17.5	15.1	16.8	27.0	28.0	32.0	32.0
Maximum (ppbV)	24.7	22.5	24.7	32.0	35.0	39.2	41.0
Skewness	-0.22	-0.23	0.21	-0.42	-0.46	0.01	-0.29
Kurtosis	2.01	1.91	2.67	3.14	2.56	2.33	2.75

During the first phase of the study, data distribution was described using classical descriptive statistics (Table I). For each sampling campaign, the mean and median are very similar which is indicative of data coming from a normal distribution. This is ratified by the fact that skewness values near zero are obtained. The skewness value is based on the size of the tails of a distribution and provides a measure of how likely the distribution will produce outliers. Thus, in this work, outliers should be scarce, if they exist, which is important to obtain accurate estimates.

Histograms and normal QQplots for each sampling campaign also indicate normality. The shape of the histograms looks bell shaped and the points of these plots are located close the 45º line [29]. Fig. 6 contains the histograms and QQplots for two sampling campaigns, which are similar for all others.

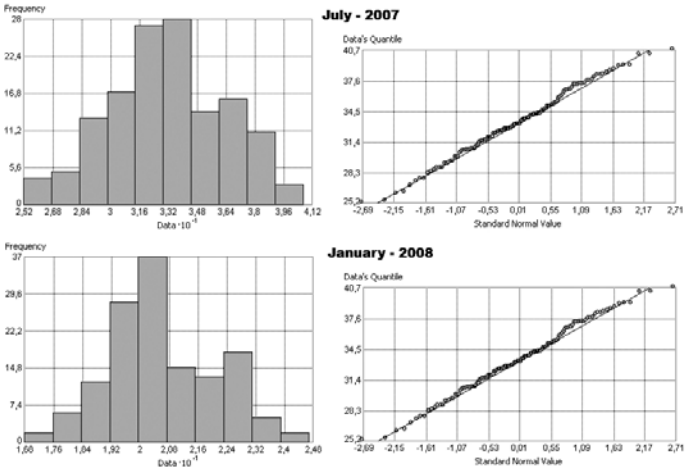


Fig. 6 Histograms and normal QQplots for data, ground-level ozone (ppbV), corresponding to two sampling campaigns

Although normality is not a prerequisite for kriging, it is a desirable property. Kriging will only generate the best absolute estimate if the random function fits a normal distribution.

Fig. 7 shows the temporal evolution of the mean ozone level for each sampling campaign. The typical annual cycle, with a distinct maximum in late spring-summer, is apparent.

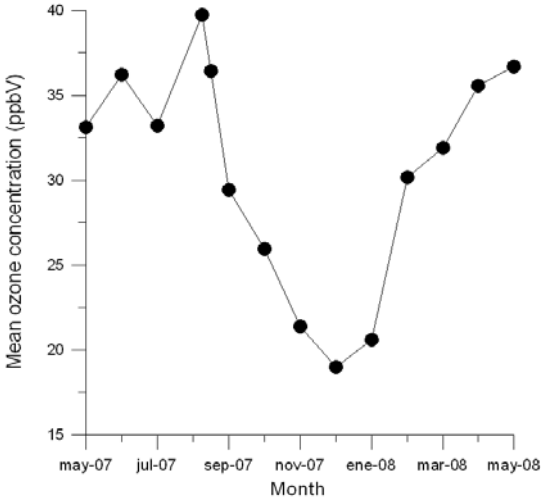


Fig. 7 Temporal evolution of the mean ground-level ozone measured during each sampling campaign

B. Structural Analysis of Data

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 [18], [19]. 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 [19]; the spherical ones were finally chosen.

TABLE II
THEORETICAL SPHERICAL VARIOGRAMS FITTED TO EXPERIMENTAL OMNIDIRECTIONAL VARIOGRAMS FOR ALL SAMPLING CAMPAIGNS

	Range (m)	Nugget	Sill	Nugget/Sill (%)
May 07	473	10,53	19,33	54,47
June 07	302	4,55	13,86	32,83
July 07	594	2,81	8,94	31,43
August 07	617	2,09	6,74	31,01
August-2 07	690	3,83	7,74	49,48
Sept. 07	534	7,24	14,03	51,60
Oct 07	790	9,06	12,79	70,84
Nov. 07	601	1,32	3,4	38,82
Dec. 07	509	1,48	4,46	33,18
Jan. 08	406	0,81	2,38	34,03
Feb. 08	412	0,48	1,19	40,34
March 08	447	0,61	2	30,50
April 08	512	0,3	1,61	18,63
May 08	486	0,42	2,45	17,14

In the present study, variograms showed a considerable nugget effect (Table II, Fig. 8), which indicates that ground ozone level variability can occur at a scale smaller than the minimum lag distance (around 75 m). All characteristics of the variograms for each sampling campaign are shown in Table II. The maximum distance of spatial dependence, the range, varies between 302 m for June, 2007, and 790 m for May and October, 2007. This means, for example if June is considered, that sample points 302 m or more distant from each other, are spatially independent. This information could be also taken into account for future studies on the same topic, if an optimal sampling design is desired. Furthermore, the fact that the nugget-sill ratio is moderately high, between 17 and 71% (Table II), and a mean value around 37 %, indicates a moderate-strong spatial dependence between data, because the part of the variance due to the nugget effect is not very important, as well as the necessity for considering some close samples to properly calculate the nugget effect.

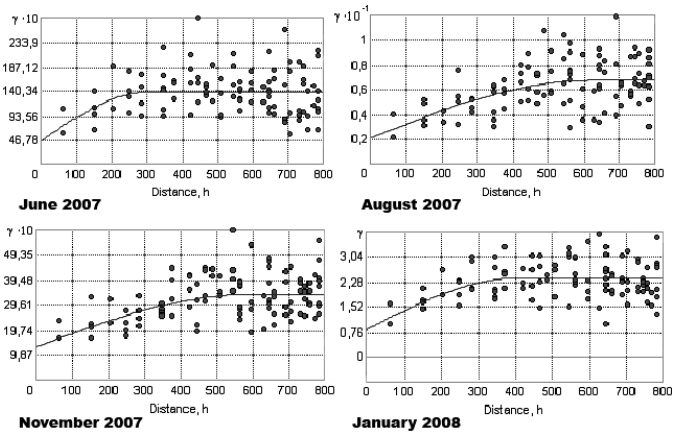


Fig. 8 Omnidirectional experimental variograms (points) and theoretical spherical variograms (lines) for data corresponding to the indicated sampling campaigns

As Diem [27] indicates, the operational scale is represented by the range of the variogram. Thus, in this study we obtained operational scales ranging from 302 m to 790 m (values of all ranges), depending on the month, with a mean value of 517 m., that is, small-scale geographic variations within these mean values dominates the ground-level ozone surfaces. This information is only available if a dense network of sampling points is considered to enable the construction of stable variograms, which is only possible when samples are at different distances, including some closer locations to better estimate the nugget effect.

Other previous studies, e.g. [23], [27], have reported operational scales values around 5 km, always considering wide sampling intervals. It seems that a finer operational scale exists and, in any case, the spatial complexity depends on local sources of ozone precursors, especially nitrogen oxides, and the urban configuration, which can generate urban canyon effects.

C. Geostatistical Estimation. Spatial Distribution Maps and Hazard Assessment of Ozone

Estimated noise levels at unsampled locations were carried out with the ordinary point kriging method, integrating the spatial correlation structures described with the variograms.

It is known that the success of kriging depends mainly on the estimation and modeling of the variogram. Strong spatial correlation, i.e., low semivariances at small spatial lags and the existence of a range (operational scale) are necessary to perform geostatistical estimates. Only spatial lags showing spatial dependence should be taken into account for kriging; in consequence, the appropriate search radius to be used (neighbourhood) is the operational scale value.

A grid, constituted of 80 m side square cells, was designed and superimposed on the city, and estimates were conducted

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

at center of the cells, i.e., ozone concentrations were estimated at a spatial resolution of 80 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. Previously, the accuracy of estimates and the validity of the prediction errors, the uncertainty, were assessed by means of cross validations (Table III). Thus:

1) the MPE y MSPE are very low and they suggest that predictions are quite unbiased.

2) For all sampling campaigns the RMSPE are less than the AKSE; this indicates that the variability of the predictions are overestimated, i.e., the predictions are, at least, as reliable as the value of the kriging variance indicates. The RMSSPE also suggest the same since they are less than one.

3) In general, the kriging variances are fair indicators of the variability in the predictions for all cases because the differences between RMSPE and AKSE are very small.

Fig. 9 shows some kriged ground-level ozone maps. Areas with higher ozone levels are usually those where traffic flow is more intense (near the main avenues, crossroads and access to the bridges).

The highest concentrations of ozone have not to be found in those urban areas where the pollutants that form ozone are emitted. In Badajoz city, where industrial activity is not excessively important, traffic is the main source of ozone precursors, so it is expected that nitrogen oxides and VOC are more abundant in areas in which traffic flow is more intense. But if there is an abundance of nitrogen oxide, ozone formation is suppressed. In consequence, ozone concentration is sometimes low in those areas. This is not the case in Badajoz, as it was previously stated. Maybe the fact that the avenues which support more traffic are wide, allowing the movement of precursors, prevents nitrogen oxide accumulating excessively.

Although spatial variability seems to be more important during spring-summer, when ozone levels are higher, it is also evident during autumn-winter, with lower ground-level ozone, and always around the main roads.

KSD can be mapped similarly to estimates, giving an idea of the quality of the estimates at different places. However, according to Webster and Oliver [30], 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 this work, the nugget effect was high (Table II), so we can consider that predictions are, at least, as reliable as the value of the KSD indicates. Cross validation statistics also confirm the reliability of estimates as previously was discussed. KSD maps are similar for all sampling campaigns because the sample locations are always the same and the variogram structures are alike. Fig. 10 shows, for instance, the KSD map for the sampling campaign of August 2007 to illustrate that the periphery of the town, where samples are sparse, has more doubtful estimates. In general, areas with many sample points or areas where data were sparse but evenly distributed had the most reliable estimates.

TABLE III
CROSS VALIDATION STATISTICS FOR THE ESTIMATES FOR ALL SAMPLING CAMPAIGNS USING THE ORDINARY KRIGING APPROACH (MPE = MEAN PREDICTION ERROR; MSPE = MEAN STANDARDIZED PREDICTION ERROR; RMSPE = ROOT-MEAN-SQUARE PREDICTION ERROR; AKSE = AVERAGE KRIGING STANDARD ERROR; RMSSPE = ROOT-MEAN-SQUARE STANDARDIZED PREDICTION ERROR)

	MPE	MSPE	RMSPE	AKSE	RMSSPE
May 07	0.056	0.012	4.24	4.33	0.98
June 07	-0.069	-0.016	3.69	3.74	0.99
July 07	0.015	0.005	2.61	2.64	0.99
August 07	0.064	0.019	2.14	2.27	0.96
August-2 07	0.058	0.021	2.40	2.55	0.95
Sept. 07	0.005	0.003	3.48	3.61	0.97
Oct 07	-0.009	-0.001	3.41	3.47	0.98
Nov. 07	0.030	0.014	1.56	1.67	0.94
Dec. 07	0.029	0.012	1.80	1.95	0.93
Jan. 08	0.019	0.009	1.43	1.49	0.96
Feb. 08	0.019	0.016	1.04	1.07	0.97
March 08	0.015	0.008	1.26	1.33	0.96
April 08	0.035	0.022	1.04	1.11	0.95
May 08	0.034	0.018	1.34	1.39	0.97

Another interesting application of geostatistics related to ground-level ozone studies is the generation of probability maps [19], 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, Fig. 11 shows the probability map corresponding to the August 2007 sampling campaign, 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.

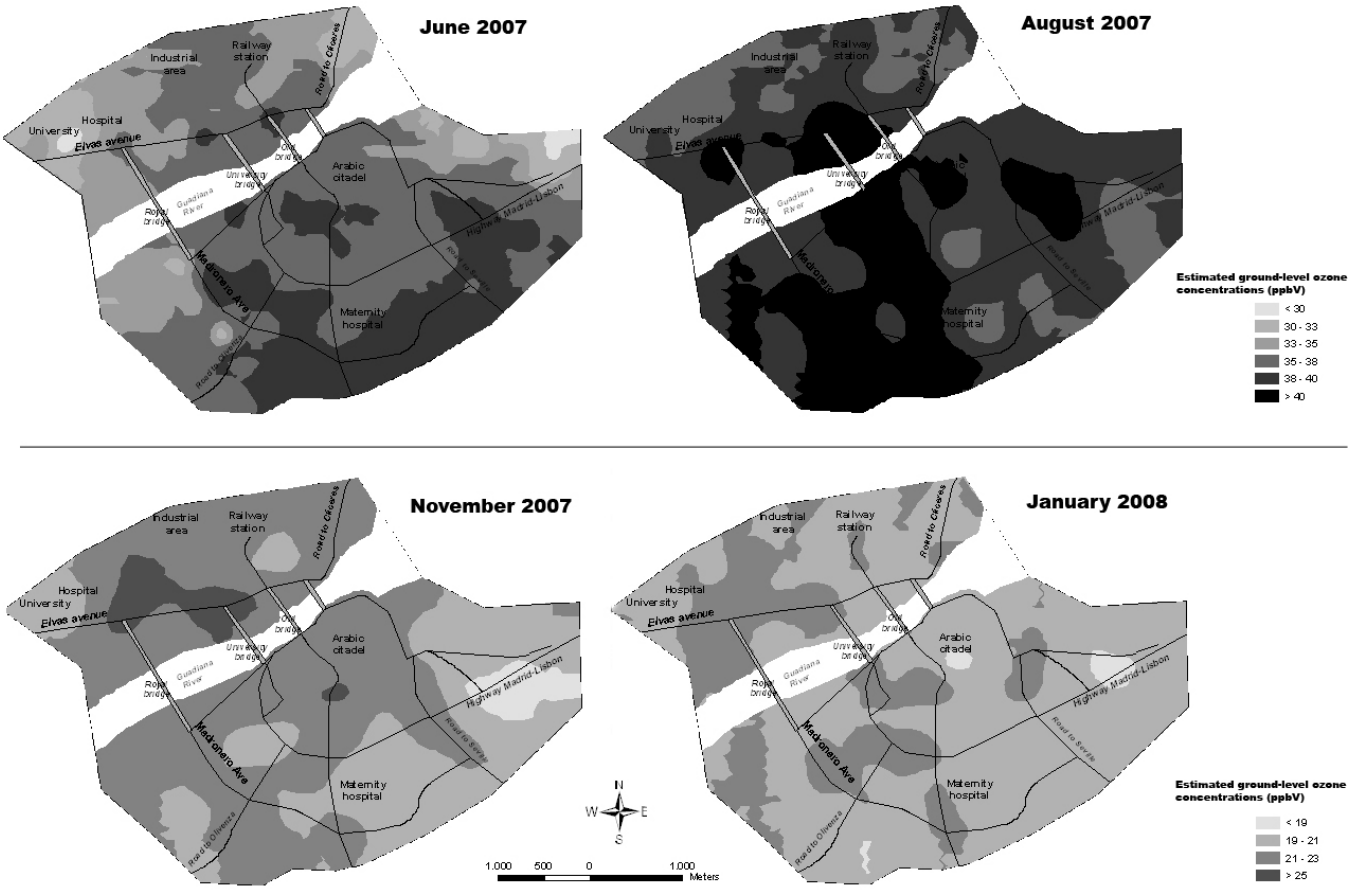


Fig. 9 Spatial distribution of ground-level ozone in Badajoz city for the indicated sampling campaigns (see corresponding variograms for each case in Fig. 8)

IV. Conclusion

The ground-level ozone in an urban environment must be studied by means of high-resolution ozone maps, which are essential tools to properly diagnose and propose control measures with the aim of minimizing its effects. In this work, geostatistical techniques are considered to model the ambient air ozone distribution over the experimental area. For this task, field measurements have to be sufficient to characterize the small-scale variability. This is only possible if a portable analyzer is used and measurements are taken during the daily time span where maximum ground-level ozone concentrations are expected (in this case, between 1 p.m. and 8 p.m.).

In this work, spatial correlation is properly characterized using omnidirectional spherical variograms, revealing that spatial dependence between data has a mean value around 500 m, considering all months, and important variations at close locations. Since a strong spatial dependence between ozone data is observed, the geostatistical algorithms, particularly the ordinary kriging, provide accurate estimates, as cross validation confirmed.

Later, kriged estimates and their associated kriging standard deviations are incorporated in a GIS to generate ozone and uncertainty maps, which inform about the reliability in predictions.

As predictions errors are fair indicators of the variability in predictions, as cross validation revealed, and normality is apparent, probability maps were finally generated. They are very useful tools for hazard assessment and decision support.

Although the real spatial complexity of ozone surfaces can not be captured, the proposed techniques provide some reliable surfaces at enough spatial resolution to correctly visualize the spatial patterns of this pollutant.

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, which, in turn, affect the attitudes and behaviors of the general public.

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

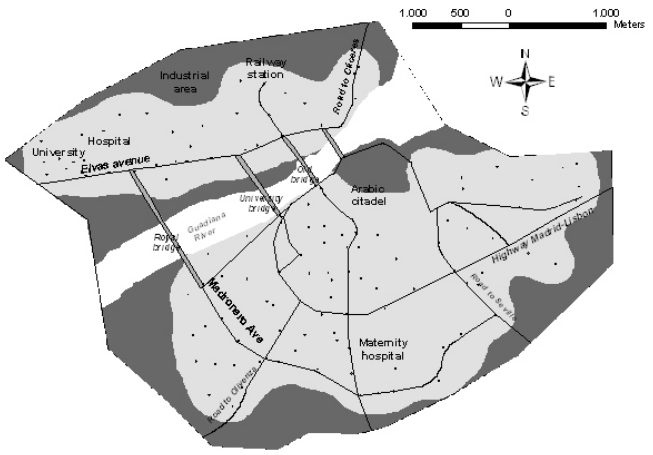


Fig. 10 Map of reliable (kriging standard deviation below 2.5 ppbV, light area) and unreliable (kriging standard deviation above 2.5 ppbV, dark area) estimates of ground-level ozone in Badajoz city. Circles represent sample points distribution

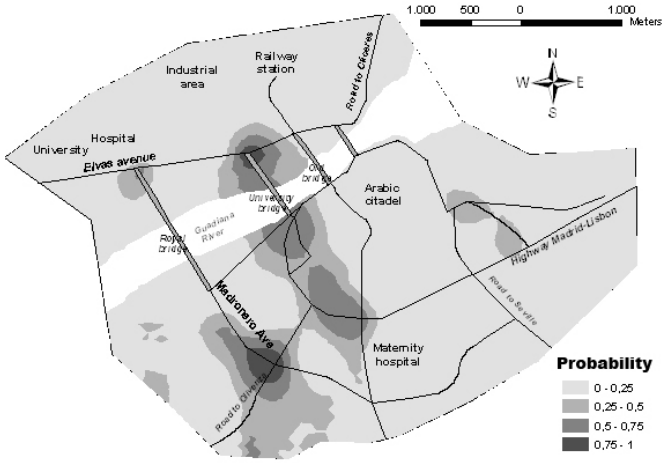


Fig. 11 Probability map of ground-level ozone higher than 42 ppbV in Badajoz city, for the August 2007 sampling campaign

ACKNOWLEDGMENT

The authors are grateful to the Consejería de Economía, Comercio e Innovación of the Junta de Extremadura, Spain, for the financial support to conduct the study through the grants for the research groups of the Autonomous Community. Moreover, the authors are also grateful to the Red Extremeña de Protección e Investigación de la Calidad del Aire of the Consejería de Industria, Energía y Medio Ambiente of the Junta de Extremadura for providing the ozone data from the background monitoring stations.

REFERENCES

[1] A. Ribas and J. Peñuelas, "Temporal patterns of surface ozone levels in different habitats of the North Western Mediterranean basin," *Atmospheric Environment*, vol. 38, pp. 985-992, 2004.

[2] J. Lelieveld, H. Berresheim, S. Borrmann, P.J. Crutzen, F.J. Dentener, H. Fisher, J. Feichter, P.J. Flatau, J. Heland, R. Holzinger, R. Korrmann, M. Lawrence, Z. Levin, K.M. Markowicz, N. Mihalopoulos, A. Minikin, V. Ramanathan, M. de Reus, G.J. Roelofs, H.A. Scheeren, J. Sciare, H. Schlager, M. Schultz, P. Siegmund, B. Steil, E.G. Stephnou, P. Stier, M. Traub, C. Warneke, J. Williams, and H. Ziereis, "Global air pollution crossroads over the Mediterranean," *Science*, vol. 298, pp. 794-799, 2002.

[3] J. Beck and P. Grennfeld, "Distribution of ozone over Europe," in *The Proccedings of the EUROTRAC Symposium 92*, The Hague, The Netherlands, 1993, pp. 43-58.

[4] M.J. Sanz, V. Calatayud and E. Calvo, "Spatial patternof ozone injury in Aleppo pine related to air pollution dynamics in a coastal-mountain region of eastern Spain," *Environmental Pollution*, vol. 108, pp. 239-247, 2000.

[5] A. Ribas and J. Peñuelas, "Effects of ethylene diurea as a protective antiozonant on beans (*Phaseolus vulgaris* cv Lit) exposed to different tropospheric ozone doses in Catalonia (NE Spain)," *Water, Air and Soil Pollution*, vol. 117, pp. 263-271, 2000.

[6] European Environmental Agency, *The European Environment. State and Outlook 2005*, Copenhagen, 2005.

[7] J. Lelieveld and F.J. Denterner, "What controls tropospheric ozone?," *Journal of Geophysical Research*, vol. 105, pp. 3531-3551, 2000.

[8] J.A. Logan, "Tropospheric ozone: seasonal behaviour, trends, and anthropogenic influence," *Journal of Geophysical Research*, vol. 90 (D6), pp. 10463-10482, 1985.

[9] A.S. Lefohn, *Surface Ozone Exposures and their Effects on Vegetation.* Chelsea: Lewis publishers, 1992.

[10] D. Simpson, "Biogenic emissions in Europe, 2, Implications for ozone control strategies," *Journal of Geophysical Research*, vol. 100, pp. 891-906, 1995.

[11] P.S. Monks, "A review of the observations and origins of the spring ozone maximum," *Atmospheric Environment*, vol. 34, pp. 3545-3561, 2000.

[12] M. Peleg, M. Luria, G. Sharf, A. Vanger, G. Kallos, V, Kotroni, K. Lagouvardos, and M. Varinou, "Observational evidence of an ozone episode over Grater Athens Area," *Atmospheric Environment*, vol. 31, pp. 3969-3983, 1997.

[13] D. Cocchi and C. Trivisano, "Ozone," in *Encyclopedia of environmetrics*, New York, 2002, pp. 1518-1523.

[14] C. Dueñas, M.C. Fernández, S. Cañete, J. Carretero, and E. Liger, "Assessment of ozone variations and meteorological effects in an urban area in the Mediterranean coast," *The Science of the Total Environment*, vol. 299, pp. 97-113.

[15] L.P. Hopkins, K.B. Ensor, and H.S. Rifai, "Empirical evaluation of ambient ozone interpolation procedures to support exposure models," *J. Air & Waste Management Association*, vol. 49, pp. 839-846, 1999.

[16] S. Vardoulakis, N. Gonzalez-Flesca, B.E.A. Fisher, and K. Pericleous, "Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris," *Atmospheric Environment*, vol. 39, pp. 2725-2736, 2005.

[17] A. Coppalle, V. Delmas, and M. Bobbia, "Variability of NOx and NO2 concentrations observed at pedestrian level in the city centre of a medium sized urban area," *Atmospheric Environment*, vol. 35, pp. 5361-5369, 2001.

[18] E.H. Isaaks, and R.M. Srivastava, *An Introduction to Applied Geostatistics*. New York: Oxford University Press, 1989.

[19] P. Goovaerts, *Geostatistics for Natural Resources Evaluation*. New York: Oxford University Press, 1997.

[20] D. McGrath, C.S. Zhang, and O.T. Carton, "Geostatistical analyses and hazard assessment on soil lead in Silvermines area, Ireland," *Environmental Pollution*, vol. 127, pp. 239-248, 2004.

[21] A. Korre, S. Durucan, and A. Koutroumani, "Quantitative-spatial assessment of the risks associated with high Pb loads in soils around Lavrio, Greece," *Applied Geochemistry*, vol. 17, pp. 1029-1045, 2002.

[22] P. Goovaerts, "Geostatistical modeling of uncertainty in soil science," *Geoderma*, vol. 103, pp. 3-26, 2001.

[23] D.L: Phillips, D.T., Tingey, E.H., Lee, A.A., Herstrom, and W.E. Hogsett, "Use of auxiliary data for spatial interpolation of ozone exposure in southeastern forests," *Environmetrics* vol. 8, pp. 43-61, 1997.

[24] M. Tayanc, "An assessment of spatial and temporal variation of sulfur dioxide levels over Istanbul, Turkey," *Environmental Pollution*, vol. 107, pp. 61-69, 2000.

[25] D.E. Myers, "Interpolation and estimation with spatially located data," *Chemometrics and Intelligent Laboratory Systems*, vol. 11, pp. 209-228, 1991.

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf

[26] C. Carlon, A. Critto, A. Marcomini, and P. Nathanail, "Risk based characterisation of contaminated industrial site using multivariate and geostatistical tools," *Environmental Pollution*, vol. 111, pp. 417-427, 2001.
[27] J.E. Diem, "A critical examination of ozone mapping from a spatial-scale perspective," *Environmental Pollution*, vol. 125, pp. 369-383, 2003.
[28] P.A. Burrough, and R.A. McDonnell, *Principles of Geographical Information Systems*. Oxford: Oxford University Press, 1998.
[29] K. Johnston, J. M. Ver Hoef, K. Krivoruchko, and N. Lucas, *Using ArcGIS Geostatistical Analyst*. Redlands, USA: Environmental Systems Research, 2001.
[30] R. Webster, and M.A. Oliver, *Geostatistics for Environmental Scientists*. Brisbane, Australia: John Wiley & Sons Ltd, 2001.
[31] N. Cressie, "Fitting variogram models by weighted least squares," *Mathematical Geology*, vol.17, pp. 563-586, 1985.
[32] C. Deutsch, and A. Journel, *Geostatistical Software Library and User's Guide*. New York: Oxford University Press, 1992.
[33] L.A. McNair, R.A. Harley, and A.G. Russell, "Spatial inhomogeneity in pollutant concentrations, and their implications for air quality model evaluation," *Atmospheric Environment*, vol. 30, pp. 4291-4301, 1996.

Open Science Index, Civil and Architectural Engineering Vol:4, No:1, 2010 publications.waset.org/9050/pdf