EVALUATION AND ANALYSIS OF THE URBAN AIR POLLUTION BY MEANS OF THE USE OF GEOSTATISTICAL TECHNIQUES AND GEOGRAPHICAL INFORMATION SYSTEMS (SIG)

López-Rodríguez, F
Moral García, F.J.
Departamento de Expresión Gráfica. Universidad de Extremadura

Valiente González , P
Departamento de Química Analítica. Universidad de Extremadura

Abstract

To evaluate and control the pollution in an urban area, diverse graphic representation techniques should be used to map the spatial patterns of the studied pollutants. This is a complex task, which is only feasible if a spatial correlation of the variable of interest is identified. Moreover, the small scale spatial distribution is unusually determined, despite it is fundamental to make a decision from an environmental point of view.

In this work, the link between the Geographic Information System and Geostatistics is proposed to generate pollution maps in urban areas, in which the spatial distribution patterns have high resolution to provide the variability at small distances.

Keywords: *Geostatistics, Geographic Information System, Map, Pollutants.*

1. Introducción

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 [1]. 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 [2], [3]. 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 [4], [5]. 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 [6], [7]. Moreover, they have become useful tools for the study of hazard assessment and spatial uncertainty [8]. 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. [9], but kriging is the only one which incorporates the spatial correlation into its estimation algorithm. Kriging has been used more widely [10] due to its many advantages. Although kriging requires an abundance of sample points to be an accurate spatial interpolation method [11], 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 [12.

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 analyze the temporal evolution and characterize the spatial distribution of ground ozone levels using geostatistical techniques; incorporate this information in a GIS to produce accurate ozone maps; assess the hazard of exceeding some limits with a geostatistical basis.

2. Materials and Methods

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. It is the largest (about 140.000 inhabitants) and most industrialized city in this region.

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. 1). This station is operated by the Department of Environment of the Extremaduran Government. 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). 138 urban locations were chosen as sample points (Fig. 1), covering the majority part of the city.

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. This point of view, when the studied variable (ground-level ozone) is considered random and distributed continuously on the experimental area is adopted to use geostatistics as an estimation technique. It is widely recognized that the statistic approach, geostatistical methods or kriging, has several advantages over the deterministic techniques [4], [5].

3. Results and Discussion

During the first phase of the study, data distribution was described using classical descriptive statistics (Table 1). 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. 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.

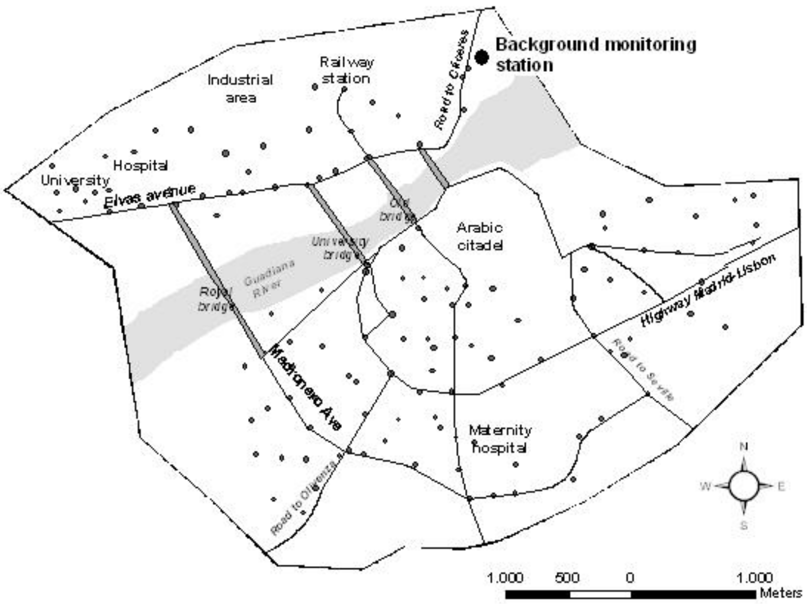

Fig. 1. Map of Badajoz city (urban area) and sampling locations (138).

Experimental variograms were determined assuming isotropy conditions because with the reduced number of sample points, the influence of different directions in space had supposed the impossibility to define acceptable directional variograms.

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 [4]; the spherical ones were finally chosen.

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

In the present study, variograms showed a considerable nugget effect (Table 2), 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 2. 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 31 and 71% (Table 2), 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.

Table 2. 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

Estimated noise levels at unsampled locations were carried out with the ordinary point kriging method. 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 3). Thus: the MPE y MSPE are very low and they suggest that predictions are quite unbiased; 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; 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. 2 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.

Tabla 3. 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

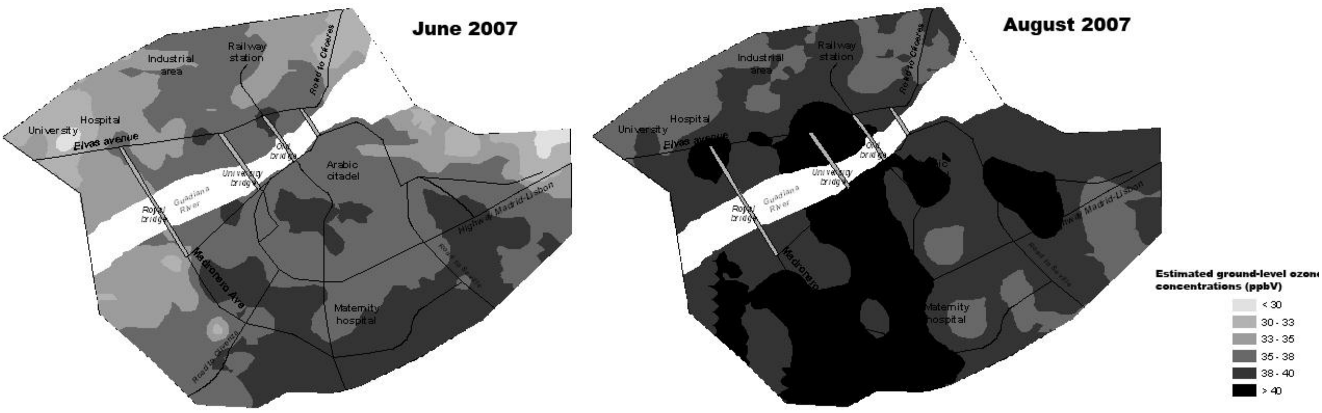

Fig. 2. Spatial distribution of ground-level ozone in Badajoz city for the indicated sampling campaigns.

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 [13], 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 2), 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. 3 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.

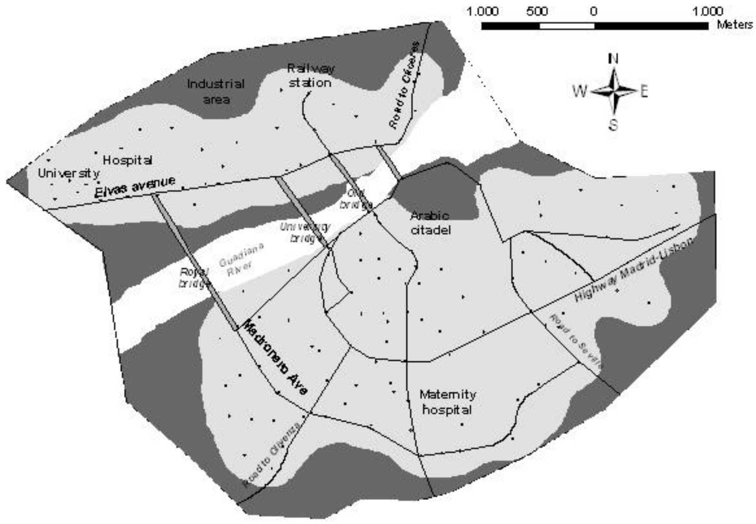

Fig. 3. Map of reliable (kriging standard deviation below 2.5 ppbV, light area) and unreliable (kriging standard deviation above 2.5 ppbV, dark area) estimates of ground-level ozone in Badajoz city.

4. Conclusions

The pollution in an urban environment must be studied by means of high-resolution 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.

Since a strong spatial dependence between ozone data is observed, in this work spatial correlation is properly characterized using omnidirectional spherical variograms, the geostatistical algorithms, particularly the ordinary kriging, provide accurate estimates, as cross validation confirmed. Kriged estimates and their associated kriging standard deviations were incorporated in a GIS to generate ozone and uncertainty maps, which inform about the reliability in predictions.

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 pollution maps can influence decisions concerning air-quality policy, which, in turn, affect the attitudes and behaviors of the general public.

References

[1] 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.

[2] 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.

[3] 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.

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

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

[6] 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.

[7] 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.

[8] P. Goovaerts, “Geostatistical modeling of uncertainty in soil science,” Geoderma, vol. 103, pp. 3-26, 2001.

[9] 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.

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

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

[12] 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.

[13] R. Webster, and M.A. Oliver, Geostatistics for Environmental Scientists. Brisbane, Australia: John Wiley & Sons Ltd, 2001.

Correspondence

Fernando López Rodríguez
Teléfono de contacto: 649833585
E-mail de contacto: ferlopez@unex.es
Empresa / Institución: Universidad de Extremadura