Environmental Engineering and Management Journal October 2014, Vol.13, No. 10, 2507-2515
http://omicron.ch.tuiasi.ro/EEMJ/

"Gheorghe Asachi" Technical University of Iasi, Romania

GEOSTATISTICAL ANALYSIS OF THE CAUSES OF ENVIRONMENTAL NOISE IN SPAIN

María-Dolores Huete-Morales, José-Manuel Quesada-Rubio*, Esteban Navarrete-Álvarez, María-Jesús Rosales-Moreno, María-José Del-Moral-Ávila

University of Granada, Faculty of Science, Department of Statistics and Operations Research, Avda. Fuentenueva s/n 18071 Granada, Spain

Abstract

The problem of noise levels in the home is of increasing interest for reasons of health and psychological well-being, among others. This study analyses in detail the meta-data compiled in the Spanish Census of Population and Housing, which provided data on a large number of environmental variables, including noise pollution in the home. A geostatistical study is conducted from the data provided by the latter survey, in which spatial autocorrelation is measured by analytical techniques by using R statistical software and the ESRI geographic information system. We study the empirical variogram, analize different theoretical models (gaussian, exponential and spherical) and estimate the parameters from different perspectives: constant or linear trend, weighted least squares, maximum likelihood and restricted maximum likelihood. The noise level is estimated through the kriging interpolation technique, using the different parameters of each model.

Key words: environment, Kriging interpolation technique, noise, spatial statistical analysis, variogram

Received: May, 2014; Revised final: October, 2014; Accepted: October, 2014

1. Introduction

Environmental noise has become one of the most annoying contaminants of modern society that directly affects the quality of life for people, among other environmental contaminants (Huete et al., 2014). At present, the noise is considered inherent to the development of any activity such us transport (Nicolae and Anda, 2014), industry, commerce or entertainment, and although it is assumed as a problem in developed countries, the population exposure to noise is superior in many occasions in emerging countries due to major deficiencies in planning, construction and control.

The consequences of noise pollution on health occur cumulatively from medium to long term. Hearing loss is one of the most common effects of continuous exposure to high noise levels or very prolonged exposure to noise levels of average intensity However there are multiple effects mainly affecting the vegetative and neuroendocrine systems. The psychological effects include changes in sleep, with subsequent consequences such as lack of sleep, lack of concentration, headaches, stress, or behavioral disorders such as aggression, irritability, decreased memory performance or concentration ability. The effects of noise began to be considered as a polluting agent from the last quarter of the twentieth century. The first international declaration which contemplated the consequences of noise dates back to 1972, when the World Health Organization (WHO) decided to generically catalog it as a kind of pollution

In 1976 the World Medical Association adopted its Declaration on Pollution which maintains that noise pollution is made up of "excessively high levels of sounds produced by industrial plants, transportation systems, audio systems and other

* Author to whom all correspondence should be addressed: e-mail: quesada@ugr.es

means that can produce a permanent hearing loss, other pathophysiological effects and emotional problems". In 1979 the noise is classified as a specific pollutant by the Stockholm Conference, and in 1990 WHO creates the "Inter Health" program, warning about diseases related to modern lifestyle, among which are the ones derived from the noise.

In 1995 WHO produces international broadcast guides in order to strengthen the scientific knowledge on the effects of urban noise on health and guide authorities and environmental health professionals trying to protect people from the effects of noise. The "Guidelines for Community Noise" were prepared as a practical response to the need to develop measures against noise as well as the need to improve the legislation, management and guidance to national and regional levels.

The European Union became aware of the problem of noise pollution from the Green Paper of the European Commission on "Future Policy Anti-Noise" (Brussels, 04.11.1996) which states the need to clarify and standardize the regulatory environment noise, implement adequate controls over noise generating products and the coordinated action of the states in other areas in order to undertake labors to prevent and reduce noise. Following this line in Spain the Noise Law (37/2003, of 17 November) was established. To counteract noise there are basically two general methods that competent authorities should apply: the technical solutions, which generally do not act on the root cause, but that mitigate the effects and prevention, which should be promoted by different environments (public institutions, industries etc.) with the aim of raising public awareness and actually decrease noise levels.

It is essential that the actions developed to limit and control exposure to environmental noise are supported by adequate scientific evaluation of available data on location of risk areas, effects of noise or pollution levels. These studies provide the basis of the assessment process and risk management.

Among other goals, analyses of environmental quality are aimed at protecting the population from the discomfort caused by excessive noise. Researchers seek to identify where noise is most evident and how it is produced. There is a strong relationship between economic and social activity in the study area and noise pollution (Iosub et al., 2009). It is also notable that the most populated cities are environmentally noise polluted (Trombetta et al., 2002; Petrescu and Borza, 2013). The issue of noise levels in the home is of increasing interest for reasons of health and psychological well-being, among others. It is true that noise sensitivity is different for each person (Soames, 1999) but the evidence for effects of environmental noise on health is strongest for annoyance, sleep and cognitive performance in adults and children; noise interferes in complex task performance and modifies social behaviour (Stansfeld and Matheson, 2003) and has a direct impact on quality of life (Donath, 2006). According to the World Health Organization, noise pollution has a strong adverse effect on health. Various agencies are currently performing surveys into the question of environmental noise. The Ministry of Agriculture, Food and the Environment has established a Noise Pollution Information System but local control of noise has not been successful in most places (Goines and Hagler, 2007).

In this paper, we propose a geostatistical analysis to determine the noisiest areas in Spain; to achieve this goal, the study has been developed in the following phases:

- Phase 1: Exploratory analysis of the information with the percentage of households affected by noise problems and obtaining spatial autocorrelation indicator.
- Phase 2: Structural analysis by estimating the experimental semivariogram and fitting the theoretical model.
- Phase 3: Prediction of the study variable by conventional kriging.

2. Material and methods

Information on the external noises affecting households was obtained from the Census of Population and Housing, compiled by the Spanish National Statistics Institute (INE, 2001a). This census provides abundant data on the attitudes of Spanish households toward the environment, broken down by municipalities, provinces and autonomous regions, and classified according to numerous sociodemographic variables regarding individuals resident in Spain. The response variable is the percentage of households who have reported problems with noise in their province of residence. This percentage, expressed as $Y(s_i), i = 1,....,N$, was measured for the 52 Spanish provinces, whose coordinates $s_i = (x_i, y_i), i = 1,.....,N$ represent the longitude and latitude of the centroids of the polygon for each province. The study was carried out using geostatistical analysis methods with spatially referenced data, and the geospatial units were limited to the peninsular provinces. This was done because the study goal is to determine spatial dependence in terms of the distance between data locations, and so the inclusion of insular regions might have distorted the results obtained. The methodology used is characterized by its consideration of variations that are spatially dependent (i.e., there is a spatial component). Only by such methods is it really possible to reflect the inherent variability of geographically-located data and thus discover, among the erratic fluctuations of data and the long-range trend component, the underlying spatial process that characterizes the phenomenon in question. The squared differences $1/2(Y(s_i) - Y(s_j))^2$ are termed empirical variogram ordinates, such that the behaviour of the expectations of the ordinates of the semivariogram is constant and equal to σ^2 if the correlation is zero, i.e., if there is no spatial correlation. In general, if correlation exists, it is expected to decrease with increasing distance between locations and will tend to disappear at large distances, such that the values of the expectations of the ordinates of the empirical semivariogram will tend toward σ^2 as the distance increases. The experimental semivariogram for the residuals of the model was obtained using the method described by Hawkins and Cressie (Cressie, 1993). The covariance structure can be defined within the variogram in terms of sill, range or nugget correlation parameters. The estimation of the variogram (Barry et al., 1997) or the covariance parameters of a series of plausible models forms part of the analysis that can also be used to make predictions, as well as contributing to our understanding of the structure that models the dependence itself.

Other authors who have contributed to the estimation of the variogram are Zimmerman and Zimmerman (1991). The reliability of the variograms obtained can be seriously affected by outliers or highly asymmetric distributions, and in such cases robust estimators of the variogram are recommended, as defined in Cressie and Hawkins (1980) and in Cressie (1985).

Two common approaches to fitting the model are to use unweighted or ordinary least squares (OLS) or the weighted least squares (WLS) criterion, either by numbers of pairs or by Cressie's method. Parameter estimation can also be performed by the maximum-likelihood (ML) or restricted maximum-likelihood (REML) procedure. The empirical variogram and the theoretical curve (the estimated parametric variogram) that gives the best fit can be shown together on a single graph. The theoretical form adopted (exponential, Gaussian, etc.) will approximately represent the trajectory of the points of the empirical variogram, assuming randomness. In the geostatistical context, references for this estimation method include Pardo-Igúzquiza (1997), Stein (1999) and Webster et al. (2006).

In this paper, we study the goodness of fit of different theoretical models (Gaussian, exponential and spherical) and estimate the parameters from different perspectives: constant or linear trend, weighted least squares, maximum likelihood and restricted maximum likelihood. For this spatial analysis it was necessary to obtain maps or shapefiles for Spain (DICES, 12013) for the corresponding peninsular provinces, using the polygons and their centroids, with the ETRS89_30N system (ESRI, 2013). Geostatistical analysis was performed using R software (R Project, 2013), together with SPSS 20.0 for data cleaning.

3. Exploratory analysis of the information

An essential prior analysis was conducted of the spatial information concerning the percentage of households affected by noise problems. Fig. 1 shows the data recorded for each province, where the mean value was 24.06%, although in one province it was just 10.0% (well below the first quartile value of 17.95) while in another it rose to 41.43%. These results show, first, that the spatial behaviour for the variable under study is extremely heterogeneous. With respect to the spatial classification of the variable (Fig. 2), in which it was divided into five equal parts, by quantiles, it can be seen that the periphery (especially the south and east) of the peninsula presented much higher values, apart from those for the capital, Madrid.

Visually, there is a clear relationship between the spatial coordinates and the study variable; the value of the Moran index is 0.067 (p<0.0001), which confirms the existence of spatial autocorrelation between the observations (Moran, 1950; Li et al., 2007; Quesada et al., 2011). An exploratory study of the data, to analyse the trend and variability, shows that in principle there are no anomalous data. The range is (10.00 - 41.43), with an interquartile range of (17.95 - 29.82), a coefficient of variation of 34.30%, a slight asymmetry to the right and negative kurtosis. Therefore, the data variability is considerable. A Box-Cox transformation (Box and Cox, 1964) was performed to reduce the possible variability of the data (lambda = 0.6405). The interpolation maps (Fig. 3) clearly show that the south-eastern provinces, together with Madrid and some northern areas, present a different pattern of behaviour from the rest of the peninsula. The values observed indicate that this does not arise from a stationary process, as the variability is high and there exists a spatial trend. Therefore, a model is required to eliminate this trend.

The covariable was taken as the total number of companies per province in 2001 (Fig. 4), according to the Central Business Directory (INE, 2001b). The Pearson correlation coefficient between the two variables is 0.575, which indicates the presence of a relationship between the noise affecting the population and the number of companies in the province. To eliminate the trend from the spatial information analysed, in order to obtain a plausible constant trend hypothesis, we fitted a model that accounts for the percentage of households, based on the reference coordinates for each region and the covariable for the number of companies *Z(s)* included in the spatial model (Eq. 1).

$$Y(s_i) = f(s_i) + Z(s_i) + \varepsilon \qquad (1)$$

The model fitted to estimate the spatial trend (R-squared = 0.7052, p-value <0.0001) is shown in Table 1. A trend function was fitted by least squares and the residuals $r(s_i) = Y(s_i) - \mu(s_i)$ were analysed to determine the covariance structure.

The residuals presented a normal behaviour pattern and the scattergrams with respect to the geographical coordinates show that the trend has been successfully eliminated from the model (Fig. 5).

4. Structure of the spatial correlation

A structural analysis was performed by estimating the experimental semivariogram and fitting the theoretical model. The procedures employed for this were ordinary least squares (OLS), the OLS-npairs method and Cressie's weighted least squares method (Table 2, Fig. 6).

Other methods considered were fitting by maximum likelihood (ML) and restricted maximum likelihood (REML).

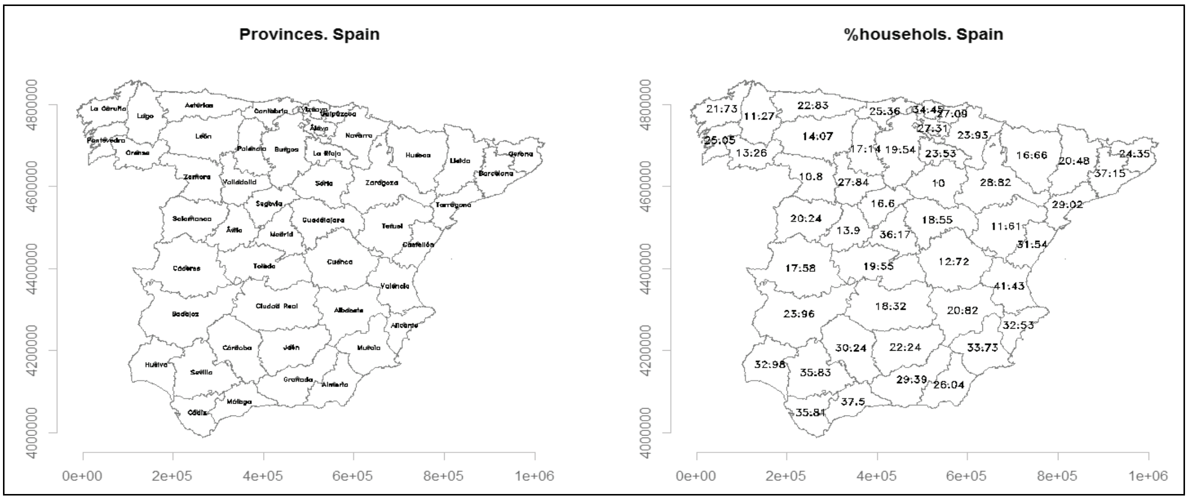

Fig. 1. Percentage of households affected by noise

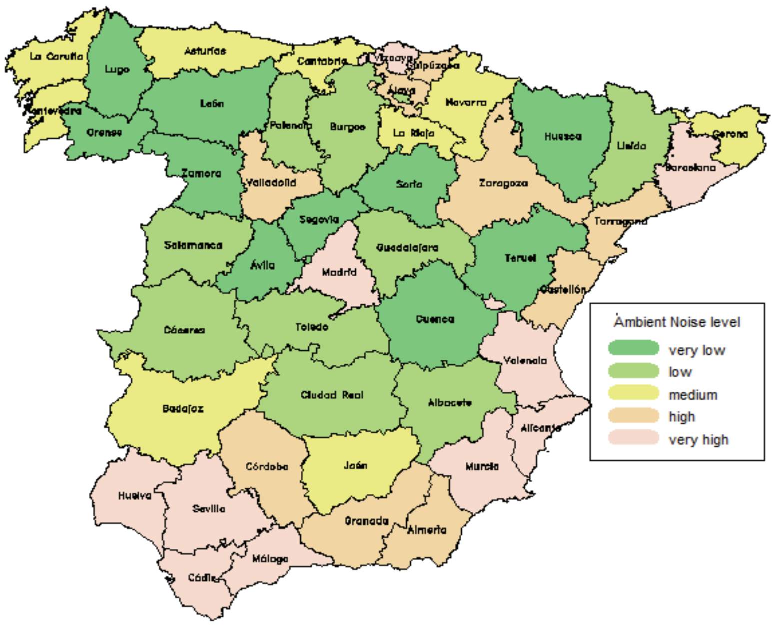

Fig. 2. Classification of percentage of households affected by noise

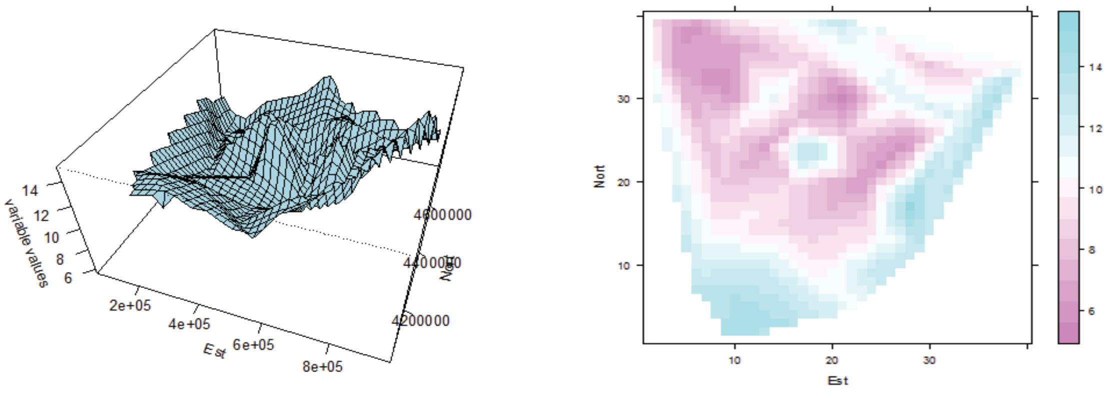

Fig. 3. Perspective map and level plot

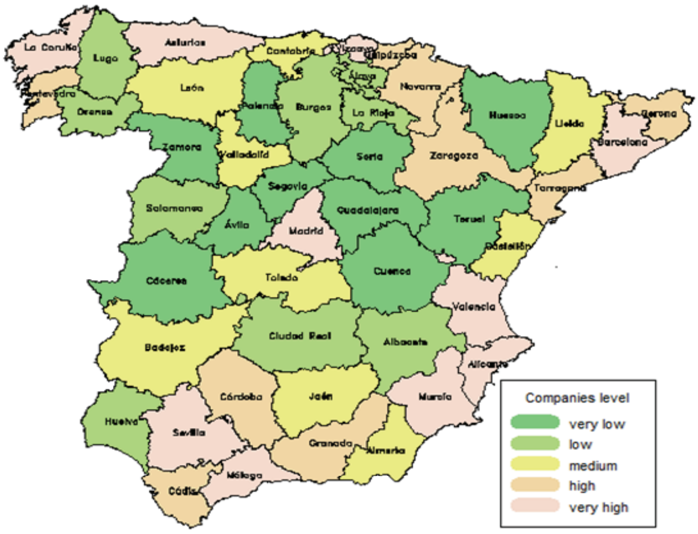

Fig. 4. Concentration of companies per province, by quantiles

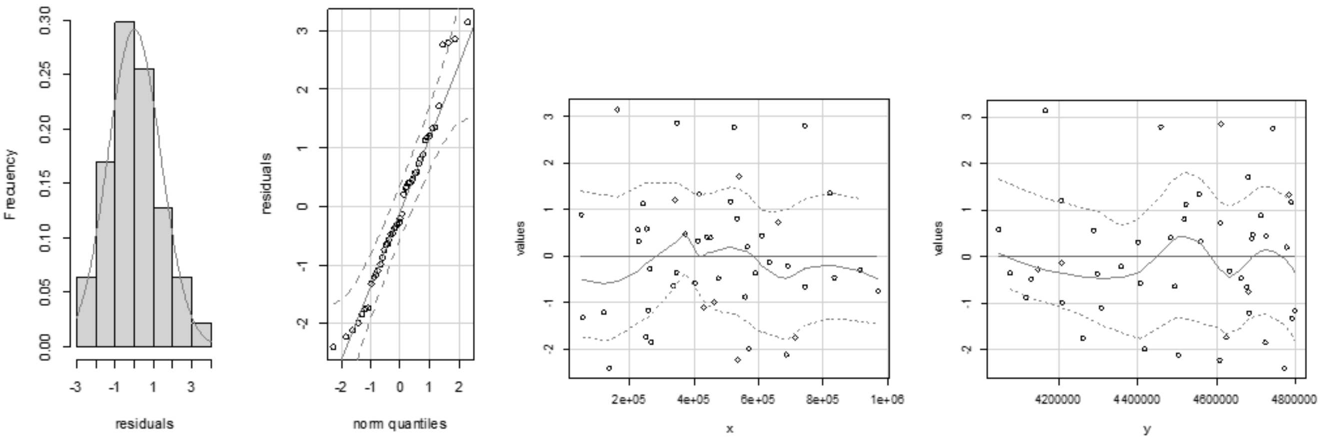

Fig. 5. Exploratory Graphics of residuals

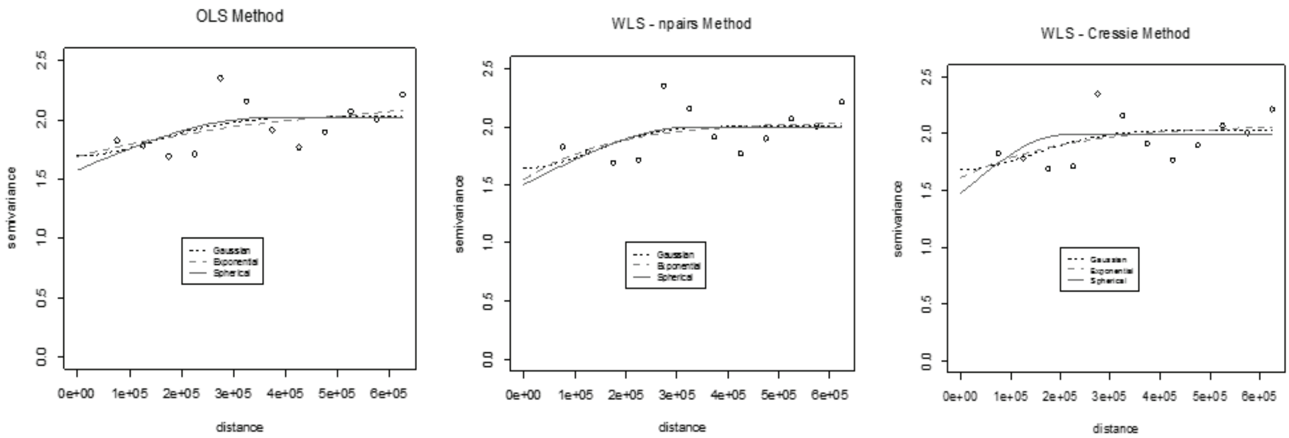

Fig. 6. OLS and WLS fit to the experimental variogram

Table 1. Ajusted model for spatial trend

Coefficients	***Estimate***	***p-value***
Constant	$2.570{\cdot}10^{2}$	0.0331
x	$2.032{\cdot}10^{-6}$	0.0592
y	$-1.093{\cdot}10^{-4}$	0.0431
y^2	$1.191{\cdot}10^{-11}$	0.0492
z	$7.312{\cdot}10^{-5}$	$5.13{\cdot}10^{-7}$
z^2	$1.477{\cdot}10^{-10}$	$3.98{\cdot}10^{-5}$

In all cases, Gaussian, exponential and spherical models were considered. The results are shown in Table 3 and Fig. 7. The prediction errors are shown in Table 4; the fewest such errors were obtained by the exponential model estimating by ordinary least squares and weighted least squares, by the pair number method.

5. Interpolation with kriging techniques

By obtaining predictions over a grid of locations on the Spanish peninsula, it is possible to graphically visualize the behaviour throughout the surface of the region. Figs. 8 and 9 show the prediction of the study variable for a grid of points on the peninsula, obtained by conventional kriging and using the estimated parameters of each model as if they were known (ordinary kriging).

6. Conclusions

Spain is one of the noisiest countries in the European Union and its inhabitants consider the problem one that is very difficult to overcome. Some residents have made improvements to their homes to reduce noise levels, and many demands have been made and actions taken in this respect.

Table 2. Estimated parameters. OLS and WLS method

Parameter	*Tausq (nugget)*	*Sigma^2*	*Phi (range)*	*Sum Squared*	*Efective Range*
	OLS Method				
Gaussian	1.698	0.328	215118.5	0.3808	372331
Exponential	1.691	0.519	462271.3	0.3843	1384841
Spherical	1.575	0.438	353507.9	0.3803	353507
	WLS – npairs Method				
Gaussian	1.643	0.364	189680.4	31.8997	328302
Exponential	1.543	0.499	172747.5	32.8138	517505
Spherical	1.501	0.501	337934.7	31.2850	337935
	WLS – Cressie Method				
Gaussian	1.686	0.346	208147.8	8.2705	360266
Exponential	1.609	0.469	208147.8	8.5817	623555
Spherical	1.469	0.520	208147.8	9.2749	208147

Table 3. Estimated parameters. ML and REML method

Parameter	*Beta*	*Tausq (nugget)*	*Sigma^2*	*Phi (range)*	*Log likelihood*	*Efective Range*
	ML Method					
Gaussian	0.021	1.567	0.289	208100	-80.45	360266
Exponential	0.023	1.662	0.193	208100	-80.92	623555
Spherical	0.025	1.506	0.317	208100	-80.65	208147
	REML Method					
Gaussian	0.197	1.531	0.415	208100	-78.81	
Exponential	0.035	1.463	0.539	208100	-79.17	
Spherical	0.029	1.476	0.393	208100	-79.31	

Table 4. Prediction Errors

Model	*Method*	*Error*
Gaussian	OLS	0.02820802
Gaussian	WLS – npairs	**0.01972972**
Gaussian	WLS – cressie	0.02912018
Gaussian	ML	0.06318817
Gaussian	REML	0.04555753
Exponential	OLS	**0.01645304**
Exponential	WLS – npairs	**0.01817376**
Exponential	WLS – cressie	0.02438761
Exponential	ML	0.07284642
Exponential	REML	0.04228689
Spherical	OLS	0.03069649
Spherical	WLS – npairs	**0.01873634**
Spherical	WLS – cressie	0.02221045
Spherical	ML	0.08759555
Spherical	REML	0.06667787

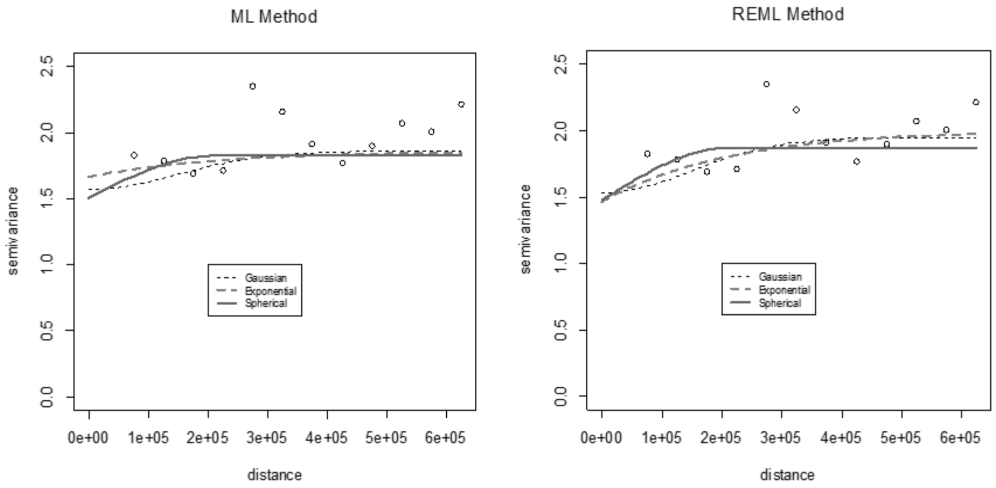

Fig. 7. ML and REML fit to the experimental variogram

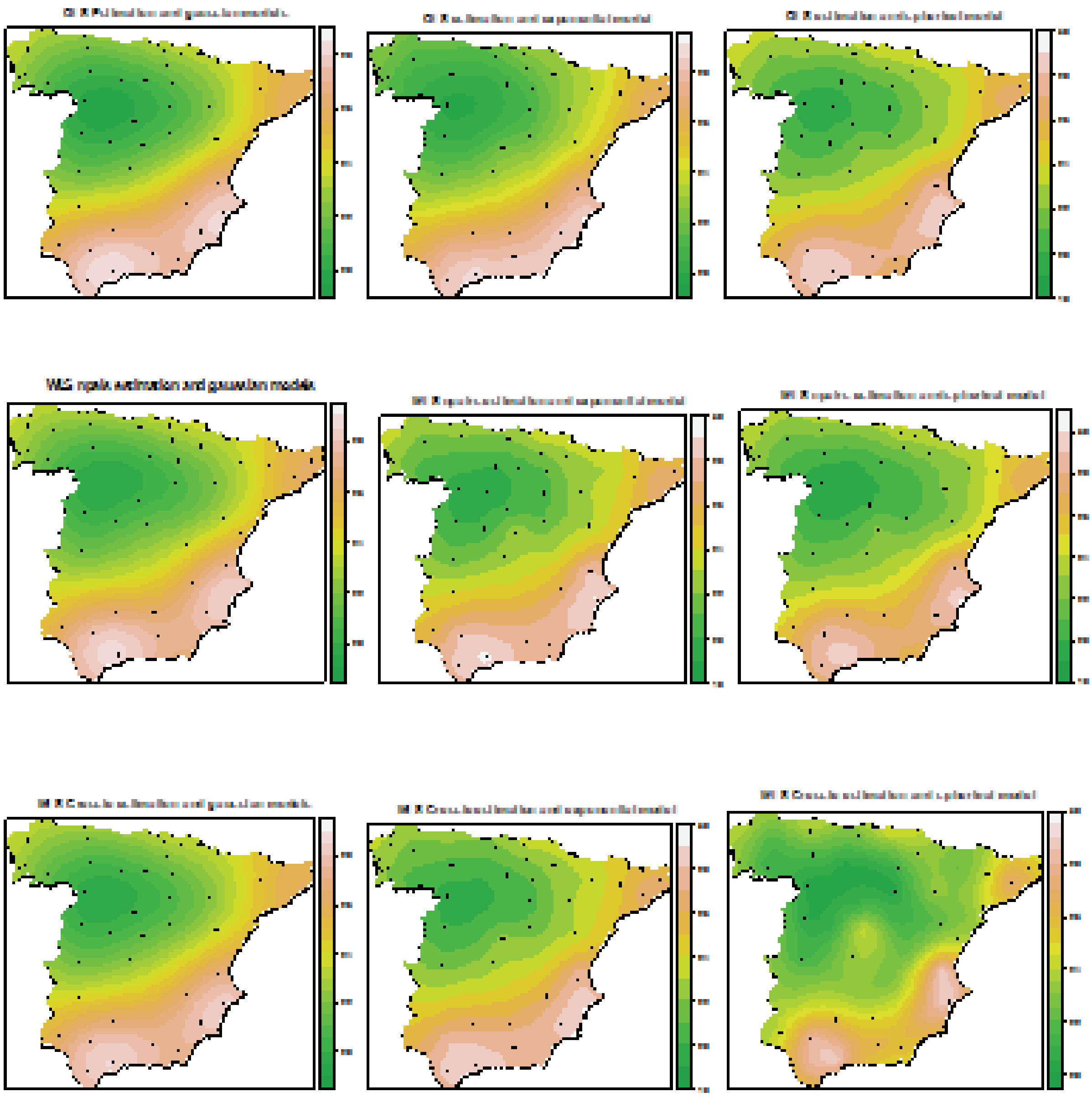

Fig. 8. Prediction by Least squares (gaussian, exponential, spherical models)

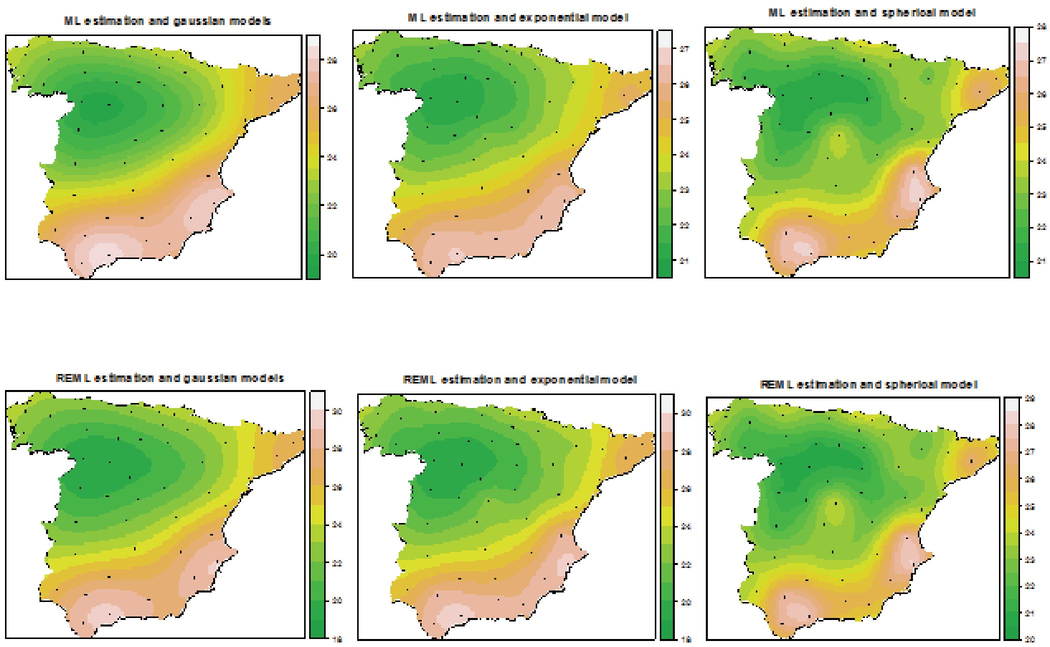

Fig. 9. Prediction by maximum likelihood (gaussian, exponential, spherical models)

Accordingly, the authorities and other bodies with responsibilities in this area need to be aware of where actions are most needed to prevent this type of pollution. Spatial correlation indicators reflect the geographic distribution of areas with similar patterns of noise behaviour, showing some areas to be significantly noisier than others, depending on the causes of the environmental noise.

It has been used geostatistical analysis methods; the aim of these techniques is to determine spatial dependence in terms of the distance between locations with spatially referenced data. By such methods is it possible to reflect the inherent variability of this type of data. Through the empirical variogram, it has been estimated the theoretical form adopted (exponential, Gaussian y spherical).

We have studied the goodness of fit of these theoretical models, using estimated parameters from different perspectives (constant or linear trend, weighted least squares, maximum likelihood and restricted maximum likelihood). Finally, the study variable has been predicted for a grid of points on the peninsula, by kriging method (using the estimated parameters of each model). This methodology is useful when there is no data in any of the study areas, in which the spatial information is interpolated.

The present study is intended to provide useful information for stakeholders in this field, for decision-making on spatial magnitudes. Using these techniques, they can understand the relationships between different areas of study and predict where necessary.

References

Barry J., Crowder M., Diggle P., (1997), *Parametric estimation of the variogram*, Technical Report ST-97-06, Dept. Maths and Stats, Lancaster University, Lancaster, UK.

Box G.E.P., Cox D.R., (1964), An analysis of transformations, *Journal of the Royal Statisistical Society, Series B*, **26**, 211-246.

Cressie N.A.C, (1993), *Statistics for Spatial Data*, Wiley, New York.

Cressie N., Hawkins D.M., (1980), Robust estimation of the variogram, *Journal of the International Association of Mathematical Geology*, **12**, 115–125.

Cressie N., (1985), Fitting variogram models by weighted least squares, *Journal of the International Association of Mathematical Geology*, **17**, 563–586.

DICES, (2013), Cartographic directory of Spain (in Spanish), On line at: http://www.dices.net/poblaciones/4292.html.

Donath T., (2006), Effects of environmental noises and sounds on the quality of life controlled by the central nervous system, *Orvosi Hetilap*, **147**, 737-739.

Esri, (2013), Map Service, On line at: http://www.arcgis.com/home/gallery.html.

Goines R.N., Hagler M.D., (2007), Noise pollution: a modern plague, *Southern Medical Journal*, **100**, 287-294.

Huete-Morales M.D., Quesada-Rubio J.M., Navarrete-Álvarez E., Rosales-Moreno M.J., Del-Moral-Ávila M.J. (2014), Influence of environmental and economic factors on endemic flora in the European Union, *Environmental Engineering and Management Journal*, **13**, 751-758.

INE, (2001a), Census of Population and Housing (in Spanish), On line at: http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft25%2Fp500&file=inebase&L=0.

INE, (2001b), Central Business Directory (in Spanish), On line at: http://www.ine.es/jaxiBD/menu.do?type=db&divi=dir&his=0&L=0.

Iosub I., Iosub S., Grigorie V., Meghea A., (2009), Noise pollution and health impact in urban Arges area, *Environmental Engineering and Management Journal*, **8**, 855-858.

ISO 1999:1990, (1990), Acoustics-Determination of occupational noise exposure and estimation of noise-

Induced hearing Impairment, International Organization for Standardization, Genéve, Suiza.

Li H., Calder C.A., Cressie N., (2007), Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model, *Geographical Analysis*, **39**, 357–375.

Moran P.A.P., (1950), Notes on Continuous Stochastic Phenomena, *Biometrika*, **37**, 17–23.

Nicolae S., Hionis C., (2014), Preventing risk of noise exposure in working environment using noise mapping, *Environmental Engineering and Management Journal*, **13**, 1349-1354.

Pardo-Igúzquiza E., (1997), MLREML: a computer program for the inference of spatial covariance parameters by maximum likelihood and restricted maximum likelihood, *Computers and Geosciences*, **23**, 153–162.

Petrescu V., Borza S., (2013), Research on noise pollution in densely populated urban areas, *Environmental Engineering and Management Journal*, **12**, 409-415.

Quesada-Rubio J.M., Huete-Morales M.D., Del-Moral-Ávila M.J., Navarrete-Álvarez E., Rosales-Moreno M.J. (2011), Impact of environmental and human factors on rural fire occurrence, *Environmental Engineering and Management Journal*, **10**, 1857-1863.

R-Project, (2013), The R-project for statistical computing, On line at: http://www.r-project.org/.

Soames R.F., (1999), Noise sensitivity as a factor influencing human reaction to noise, *Noise & Health*, **1**, 57-68.

Stein M.L., (1999), *Interpolation of Spatial Data*: Some Theory for Kriging. Springer-Verlag, New York.

Stansfeld S., Matheson M., (2003), Noise pollution: non-auditory effects on health, *British Medical Bulletin,* **68**, 243–257.

Trombetta P.H., Belisário F., Alves W., (2002), Environmental noise pollution in the city of Curitiba, Brazil, *Applied Acoustics*, **63**, 351–358.

Webster R., Welham S.J., Potts J.M., Oliver M.A., (2006), Estimating the spatial scale of regionalized variables by nested sampling, hierarchical analysis of variance and residual maximum likelihood, *Computers and Geosciences*, **32**, 1320–1333.

WHO, (1995), Guidelines for Community Noise-document guides the Department of the Protection of the Human Environment, Occupational and Environmental Health, World Health Organization, Geneva, Switzerland.

Zimmerman D., Zimmerman M., (1991), A comparison of spatial semivariogram estimators and corresponding kriging predictors, *Technometrics,* **33**, 77–91.