BUILDING A SPATIO-TEMPORAL ENVIRONMENTAL QUALITY INDEX: THE CASE OF MADRID

J.M. MONTERO

Faculty of Ciencias Juridicas y Sociales, Castilla-La Mancha University, 45071 Toledo, Spain – Jose.MLorenzo@uclm.es

B. LARRAZ

Faculty of Ciencias Juridicas y Sociales, Castilla-La Mancha University, 45071 Toledo, Spain – Beatriz.Larraz@uclm.es

C. CHASCO*

Department of Applied Economics, Autonoma University of Madrid, 28049 Madrid, Spain – Coro.Chasco@uam.es

Abstract

The elaboration of Environmental Quality Indexes (EQI) for metropolitan areas is one of the main topics in environmental economics. This article introduces some methodological and practical novelties building an EQI in Madrid (Spain). Managing with general air quality information, from the point of view of the selection of the variables, we consider noise – joint to air pollution – as a relevant environmental variable. Because the group of environmental variables is only available at a number of environmental monitoring stations, from the point of view of the computation process, we use kriging to match the monitoring stations registers to the Census data in order to map the EQI. In a first step, we krige the environmental variables to the complete surface and finally, we elaborate the environmental index. We follow an inverse process as usual in the literature, since it leads to better estimates. At last, in order to build the final synthetic index, we use the Pena Distance method. We compare EQI results computed in 2001 and 2008 considering the evolution in terms of air pollution and noise individually and as a whole.

Key words: air pollution, distance indicator, environment, kriging, noise.

JEL codes: C21, C43, Q53

---

* *Jose M. Montero is Professor in Statistics; Beatriz Larraz is Lecturer in Statistics and Coro Chasco is Lecturer in Econometrics.*

1 INTRODUCTION

Air pollution is at the top on the list of citizens' environmental concerns. This is particularly true in big cities where more than half the world's population (3.3 billion people) lives. The link between air quality and human health worries many health experts, policy-makers and citizens. The World Health Organization states that almost 2.5 million people die each year from causes directly attributable to air pollution. In this sense, the elaboration of Environmental Quality Indexes (EQIs) for big cities is one of the main topics in regional and environmental economics. Making EQIs can pursuit several objectives. The main one is to report daily air pollution levels to the public in order to prevent from potential health effects of air pollutants and determine specific actions when alert thresholds are exceeded. Environmental variables are also important as determinants of housing prices. In effect, it is reasonable to assume that pollution enters into the utility function of potential house buyers, since consumers are willing to pay for environmental goods, such as air quality, absence of acoustic pollution, etc. In the two last decades, Smith and Kaoru (1995), Smith and Huang (1993, 1995), Kim et al. (2003), Anselin and Le Gallo (2006) and Anselin and Lozano-Gracia (2008) among others, are good examples of the focus on hedonic property-value models for estimating the marginal willingness of people to pay for a reduction in the local concentration of specified air pollutants.

For all the abovementioned reasons, in this paper we elaborate an EQI for the municipality of Madrid (Spain), at the spatial level of census tracts, since there are no similar measures for this city and it can be compared with other standard socio-economic statistics. We propose some methodological and practical improvements, which are novel in this kind of analysis. From the point of view of the selection of the variables, first we will consider noise – joint to air pollution – as a relevant environmental variable. Because the group of environmental variables is only available at a number of environmental monitoring stations, from the point of view of the computation process, we will use kriging to match the monitoring stations registers to the much numerous census tracts in order to map the environmental quality surface. We follow an inverse process as usual: in a first step, we krige the environmental variables to the complete surface and finally, we elaborate the environmental index. It can be demonstrated that this process leads to better estimates (less MSE). At last, in order to build the final synthetic index, we use the Pena Distance method (DP2), which allows comparing space-time data.

In addition, we analyze the evolution of such EQI in the last decade. The comparison is developed in terms of air pollution and noise individually and in common. Thus, we obtain the impact of noise in global air quality concept.

The paper is organized as follows. In the following section, we present the methodological aspects used in the paper. In the third section, we describe the complete construction process of a Environmental Quality Index (EQI) for the city of Madrid. The article concludes with a summary of key findings and future research.

2. METHODOLOGICAL QUESTIONS

2.1 NOISE POLLUTION

As stated before, in order to build a more complete environmental index, we propose the consideration of noise to the group of air pollution environmental variables. In fact, though noise policies have been implemented in several developed countries in the recent decades, the proportion of the population that is exposed to noise levels above legal limits is still relatively important. For this reason, in the urban contexts, noise levels have an economic value (e.g. on housing prices) that has been quantified in the empirical literature using different methodologies. The hedonic approach is the more dominant. It infers individual preferences as revealed in the markets (Baranzini and Ramírez, 2005). For example, housing market data can be analyzed in order to assess whether and how much of the house selling price differentials can be explained by different noise levels.

In the specialized literature on hedonic house price models, where these kind of environmental indexes haven been built as explanatory variables (see Escobar 2006), it is not frequent to find applications using so many pollution variables including noise. Hedonic specifications typically include air pollutants such as ground-level ozone (Banzhaf 2005, Hartley *et al.* 2005 and Anselin and Le Gallo 2006), or particle matter (Chay and Greenstone 2005, Murthy *et al.* 2003), since these are more visible (like smog) and have the greatest impact on health. Sometimes, they include two pollutants, such as carbon monoxide and particle matter or ground level ozone (Neill at al 2007, Anselin and Lozano-Gracia 2008, respectively). Moreover, as far as we know Baranzini and Ramirez (2005) is the only case that considers jointly air and acoustic pollutants.

2.2 THE COMBINATION OF POINT-DATA AND AREA-DATA WITH KRIGING

Mapping an EQI implies the combination of different kind of data available at different spatial supports. While the variables are registered in a small number of monitoring stations, which produces point-data, we want to provide information for area-data at the level of much numerous census tracts. We also find that the location of the air quality monitoring stations rarely coincides with the acoustic ones. In effect, the location of environmental monitoring stations is based on regular sampling and unfortunately, they are certainly scarce due to both physical and economic constraints. This is the case of many other similar applications as De Iaco *et al*. (2002), which work with an air pollution data set available at 30 locations in Milan district or Anselin and Le Gallo (2006), Anselin and Lozano-Gracia (2008), which consider 27 and 28 stations in California, respectively.

Matching all these heterogeneous data can lead to a well-known situation called the "change of support problem" (COSP). Kriging is very often the solution to overcome this mismatch of spatial support (Gotway and Young 2002), particularly when dealing with socioeconomic data, since it takes into account spatial dependence. In the specialized literature, the usual solution to the abovementioned problem is to interpolate the environmental variable(s) to obtain their interpolated values in the locations where socioeconomic data are available (Census data, housing prices, etc.). Several interpolative alternatives have been considered in recent research: Thiessen polygons, inverse distance method, splines, kriging and cokriging, though the last two ones are more appropriate when dealing with environmental variables (Anselin and Le Gallo, 2006). When dealing with an only spatial environmental variable, kriging is a good option to get optimal estimates, since it considers its spatial dependence. In a multivariate approach, cokriging can also be a good option but it is more complex than kriging and in the isotopic case does not provide added benefits (Subramanyam and Pandalai, 2004).

Kriging is a univariate procedure, which interpolates the values of the target variable at unobserved locations using the available observations of the same variable. This interpolation procedure, which is a minimum mean-squared-error method of spatial estimation, produces the best linear unbiased estimator. In order to obtain the interpolative estimates, it uses the covariance or variogram function, which is the spatial equivalent of the autocorrelation function in time series analysis. Kriging strategy is based on the idea that variables follow a stochastic process over space. It takes into account the multidirectional feature of space in a similar fashion as time series in the unidirectional stochastic process.

Depending on the nature of stochastic processes, there are different kinds of kriging: simple kriging (SK), ordinary kriging (OK) and universal kriging (UK). In this work, we will use OK since the stochastic processes are intrinsically stationary with unknown constant means.

The usual procedure in the literature of synthetic indexes consists of building first an environmental synthetic index that will be kriged afterwards to the whole map, arguing that it is a way to transform a multivariate problem in an univariate one (Preisendorfer 1988; De Iaco *et al.* 2001, 2002). Nevertheless, we think that our option —kriging the environmental variables first and elaborating the synthetic index afterwards— is a better option because it leads to a lower error variance (Myers, 1983). Another alternative could be the direct estimation of the environmental index including the correction factor and the conditions proposed by Matheron (1979), but it is in our opinion much more difficult to implement than our proposal.

2.3 THE USE OF DP2 TO BUILD ENVIRONMENTAL QUALITY INDEXES

Finally, in order to build the global synthetic index, we opt to use a distance indicator, the Pena Distance or DP2, instead of the more commonly used PCA. DP2 is an iterative procedure that weights partial indicators depending on their correlation with a global index. Its most attractive feature is that it uses all the valuable information contained in the partial indicators eliminating all the redundant variance present in these variables (i.e. avoiding multicollinearity). This method has mainly been used to compute quality of life and other social indicators (Pena 1977, Zarzosa 1996, Royuela et al. 2003). However, we propose its use in other fields – like environmental indexes – due to its good statistical properties; i.e. multidimensionality, comparability and comprehensibility. Note that PCA and DP2 are complementary no substitute methods (see Zarzosa 1996, p. 194 or Cancelo and Uriz 1994, pp. 177-178). The first is capable of reducing the information of a group of variables eliminating redundant information. Nevertheless, DP2 also allows relative comparisons between different spatial units and/or time periods. The synthetic index derived from PCA cannot measure disparities between spatial units and/or periods of time, since it is an ordinary-type indicator. It is only capable of determining whether the state of the environment quality in unit 'A', time '*t*' is better/worse than in unit 'B', time '*t*+*k*' (for *k*=1,…). However, DP2 is a cardinal measure. It means, it is also capable of determining how much better/worse the state of the environment quality in unit 'A', time '*t*' is with respect to unit 'B', time '*t*+*k*'. This is because the partial indicators in DP2 are not the raw variables (as in PCA) but simple distance indicators for each raw variable. Therefore, unlike PCA, the numerical results of a synthetic index obtained with DP2 are quantitatively meaningful, allowing comparisons between units across space and/or time.

The main characteristics of this distance indicator are, apart from the abovementioned, first, that it is a multidimensional indicator, which is able to aggregate different environmental quality variables expressed in different measurement units. Second, it is a quantitative distance indicator, which allows comparing the environmental quality in several spatial units, since it is referred to a same base or 'ideal state'. Third, it is an exhaustive indicator, which is not based on a mere reduction of information as PCA. It uses all the 'valuable information' contained in the partial indicators; i.e. it gets the statistical information that is not either false or duplicate, which can be interpreted using ordinal or – better – cardinal scales. This property allows including a great number of variables since all useless redundant variance will be removed by the own process, avoiding multicollinearity. Following Ivanovic (1974), the more data are included in the partial indicators (related to the subject matter) the more complete will be the final synthetic index, since each variable always contain unique and proper information not present in the others. DP2 can eliminate all the superfluous common variance selecting only the part of the information, which is original.

These characteristics allow including several sources of pollution in the same synthetic index, such as air and noise. Although these data are measured in different units and can contain more or less repeated information, DP2 distance method will express all them in abstract comparable units, taking into account only the useful variance, excluding the rest.

DP2 is a relatively complex method, which implies several iterations or matrix rearrangements. The point of departure of the whole process is a matrix V of order (K,m), in which m is the number of census tracts and K is the number of partial indicators (which includes both the interpolated objective variables and the subjective ones). Each element of this matrix, v_{kj}, represents the state of the partial indicator k in the census tract j. In this matrix, those partial indicators negatively connected with environmental quality must change their sign (i.e. all their data must be multiplied by -1). On their side, variables positively linked with environmental quality do not suffer any change. As a result, an increase/decrease of the values of any partial indicator will correspond with an improvement/worsening in environmental quality.

In a second stage, we compute a distance matrix D such that each element, d_k, is defined as follows:

$$d_{kj} = \left| v_{kj} - v_{kj}^{*} \right| \tag{1}$$

where v_{kj}^{*} is the k^{th} component of the reference base vector $v_{j}^{*} = \left\{ v_{1j}^{*} \quad v_{2j}^{*} \quad \ldots \quad v_{Kj}^{*} \right\}$. It is necessary to define a reference value for each partial indicator in order to make comparisons in terms of environmental quality between different spatial units (census tracts). In quality-of-life applications, it is quite common to consider the minimum value as the reference (Vicéns and Chasco 2001, Sánchez and Rodríguez 2003, CES Murcia 2003). As a result, a higher value in DP2 (which will always adopt positive values) will imply a higher environment quality level, since it implies a longer distance respect to a theoretical 'non-desired' situation. Some indicators have clear reference values (e.g. those legally established by national or international organizations). This is the case of most air quality variables (SO_2, CO, etc.), for which the EU has fixed limit levels for the protection of human health (Official Journal of the European Union, 2008). However, we have opted not to use them due to the complexity and diversity of the measurements, which do not match with the average monthly data available for the city of Madrid. In addition, this property allows making a ranking between the spatial units in terms of environmental quality. Therefore, d_k measures the distance between the partial indicator *k* in the census tract *j* and its reference value.

In a third stage, in order to express all the indicators in abstract comparable units, we compute a first global index, the Frechet Distance (DF), which is defined as:

$$DF\left(j\right)=\sum_{k=1}^{K}\frac{d_{kj}}{\sigma_k}=\sum_{k=1}^{K}\frac{\left|v_{kj}-v_{kj}^{*}\right|}{\sigma_k} \quad ; \quad j=1,2,\dots,m \tag{2}$$

where σ_k is the standard deviation of partial indicator *k*. For each partial indicator, the distance between two spatial units d_k is weighted by the inverse of σ_k. That is to say, the contribution of each d_k to the global indicator is inversely proportional to their corresponding indicator standard deviation. This weighting scheme, which is similar to those used in heteroskedastic models, gives less importance to those distances with more variability, and vice versa.

DF is a valid concept of distance only in a theoretical situation of uncorrelated indicators. When there is a direct relationship between the partial indicators (as it is usual), DF will include some duplicated information. Therefore, DF must be corrected in order to eliminate this dependence effect (i.e. the redundant information existent in other variables), which is supposed to be linear. This is why, for each spatial unit *j*, DF is the maximum value that can reach DP2, which is defined as follows:

$$DP2\left(j\right)=\sum_{k=1}^{K}\frac{d_{kj}}{\sigma_k}\left(1-R^2_{k\cdot k-1,k-2,\dots,1}\right) \quad ; \quad j=1,2,\dots,m \tag{3}$$

where $R^2_{k\cdot k-1,k-2,\dots,1}$ is the determination coefficient of the regression of each partial indicator *k* on the others (*k*–1, *k*–2,…,1). It expresses the part of the variance of *k* that is linearly explained by the rest of partial indicators. As a result, the correction factor $\left(1-R^2_{k\cdot k-1,k-2,\dots,1}\right)$ deducts the part of the variation of the observed values that is explained by the linear dependence. Note that R^2 is an abstract concept, which is unrelated with the measurement units of the indicators.

DP2 implies a decision about the entrance order of the partial indicators in the computation process. That is to say, it must be decided which partial indicator

k is the first in contributing its variance to the global index, which one will be the second, etc. In this process, the first indicator (k=1) will contribute all its information to the global index (d_1/σ_1). However, the second indicator (k=2) will only add the part of its variance that is not correlated with the first one: $\left(d_2/\sigma_2\right)\left(1-R^2_{2\cdot1}\right)$. Regarding the third indicator, it will contribute to DP2 the part of its variance that is not correlated with the first and the second one: $\left(d_3/\sigma_3\right)\left(1-R^2_{3\cdot2,1}\right)$. And so forth.

Obviously, depending on the decision DP2 will adopt different values. Thus, it is important to find an objective hierarchical method that leads to a unique entrance order of the partial indicators. If DF is a compendium of all the partial indicators, it seems logical to make the selection taking into account the correlation between each partial indicator and DF. The indicator with the highest correlation with DF will be the leader given that it is the most informative; i.e. the indicator that contributes more variance to the global index.

The whole process is a four-step procedure that can be summarized as follows:

- First, we compute the DF values for each spatial unit using expression (9); i.e. taking into account the reference base vector v^* of minimum values.
- Second, we calculate the correlation coefficients of the partial indicators and DF to ordering the former in accordance with their degree of dependence with the later.
- Third, we compute DP (expression 10) considering the previously determined entrance order of the partial indicators. This first global index is called DP-1.
- Forth, we make a new ranking with the partial indicators in accordance with their correlation degree with DP-1 with the aim of re-computing DP. We call this second global index as DP-2.
- We repeat this iterative process until a convergence is reached; i.e. the difference between two DP contiguous indexes is null. In the case of non-convergent DP values, we can choose the first DP index (or even the average of the two final ones).

The numeric value of DP index has no real sense but it is useful to compare the state of different spatial units (census tracts) about environmental quality. We can rank census tracts according to this criterion. If we use the same variables and method, we can compare our results for Madrid with those obtained in other cities or even in other moments of time. DP2 lets comparing changes in relative positions and even detecting their causes.

3. DATA ANALYSIS

There are several types of air pollutants. These include the primary pollutants, which are directly emitted from a process, and the secondary ones, which are formed in the air when primary pollutants react or interact together to produce harmful chemicals. Primary pollutants are the ones that cause most damage to ecosystems and human health. They are, among others, sulphur dioxide (SO_2), oxides of nitrogen (NO_x), nitrogen dioxide (NO_2), carbon monoxide (CO) and particulate matter (PM). Regarding secondary pollutants, ground-level ozone (O_3) is considered, joint with PM, the most dangerous pollutant for human health.

'Noise pollution' is the name given to the unwanted sound. Noise is the most pervasive environmental pollutant of the modern world. The excessive noise induce imbalance in a person's mental state, affecting its psychological health. It can cause annoyance, high stress levels as well as noise-induced hearing loss. The source of most acoustic pollution worldwide is transportation systems (motor vehicles, aircrafts, rails), as well as machinery and construction works. It is measured in decibels *(dB(A)).*

These seven environmental variables provide all the necessary scientific information about air and sound pollution in a specific neighborhood of a metropolitan area.

Tab. 1: Description of the environmental variables.

Variables		Unit	Spatial level	Reference
1. Objective indicators				
1.1. Air quality indicators				
SO_2	Sulphur dioxide	$\mu g/m^3$	27 stations	Annual average (years 2001, 2008)
CO	Carbon monoxide	$\mu g/m^3$	27 stations	Annual average (years 2001, 2008)
NO_x	Oxides of nitrogen	$\mu g/m^3$	27 stations	Annual average (years 2001, 2008)
NO_2	Nitrogen dioxide	$\mu g/m^3$	27 stations	Annual average (years 2001, 2008)
PM	PM_{10} particulate matter (fraction of suspended particles < 10 $\mu g/m^3$ in diameter)	$\mu g/m^3$	27 stations	Annual average (years 2001, 2008)
O_3	Ground-level ozone	$\mu g/m^3$	27 stations	Annual average (years 2001, 2008)
1.2. Noise pollution indicators:				
L_{Aeq}	Equivalent continuous noise, dB(A)	dB(A)	27 stations	Annual average (years 2001, 2008)

Note: Own elaboration with data from Madrid Council

The data used in this paper come from two different sources (Table 1). On the one hand, the ‘Atmosphere Pollution Monitoring System’ publishes the air pollution measures, which can be downloaded from the Municipality of Madrid’s web page (www.munimadrid.es). The six air pollution variables are measured at 25 fixed operative monitoring stations as annual averages of daily readings in 2001 and 2008. The noise measure comes from 28 fixed operative monitoring stations, which include the above mentioned. It indicates the equivalent continuous noise level in 2001 and 2008.

Figure 1 shows the locations of the operative air quality and noise monitoring stations. As can be seen, most of them are located in the central districts and only a relatively small number can be found in the periphery. Note the reasonable coverage of the domain under study by the monitoring stations since every district has one o more stations or, in the case of the peripheral less densely populated ones, share a station with their neighbors.

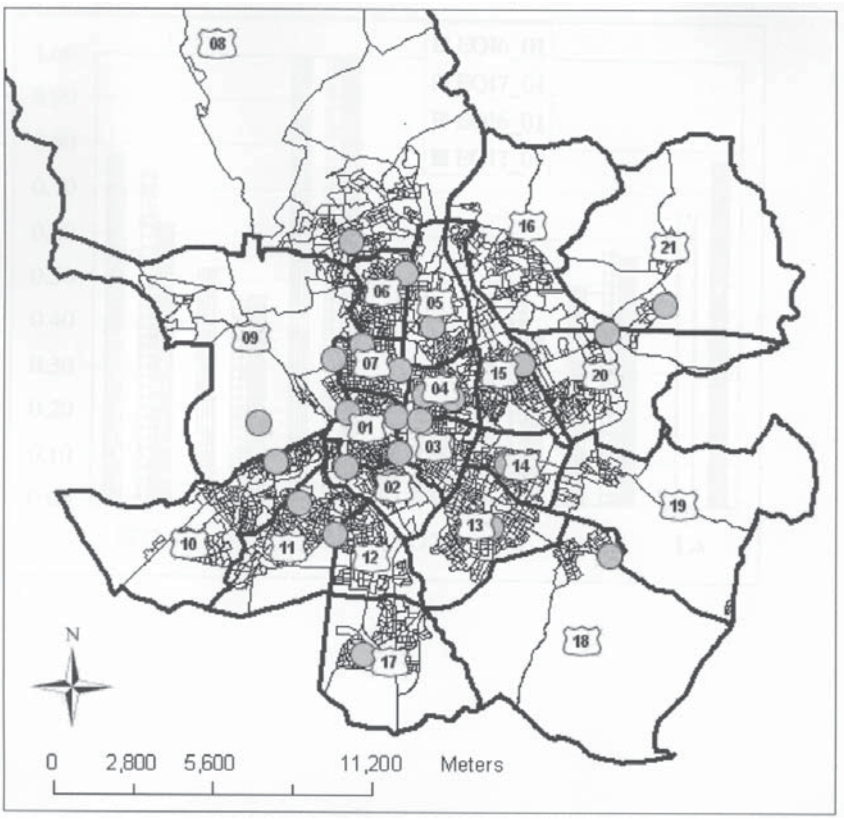

Fig. 1: Location of the active monitoring stations in the districts of Madrid

As it was pointed out in the introduction, there is a mismatch between the spatial level of the environmental measured variables and the census tract level used to map. This disparity lead us to interpolate the values at the monitoring stations to the locations of every 2,358 census tracts using kriging.

As a first step, in order to show the structure of the spatial dependence of each variable, we have computed their experimental and fitted theoretical variograms (Table 2). In order to calculate the experimental variograms, we have considered 10 lags with a lag size of 600 meters. Once the variogram models have been chosen, we next proceed to the kriged estimation. (ordinary kriging, OK) of the annual averages of each pollutant in the total area of Madrid (and thus, in the 2,358 census tracts of this city) using 2001 and 2008 data.

Tab. 2: Variogram fitting.

Environmental pollution variables		**Variogram model**		**Sill**	**Range**
Sulphur dioxide (·g/m^3)	SO_2	2001	Gaussian	43.45	4,600m
			Nugget	0.43	
		2008	Spherical	28.87	
Carbon monoxide (mg/m^3)	CO	2001	Spherical	0.0805	5,250m
		2008	Gaussian	0.0195	3,600m
			Nugget	0.0005	
Oxides of nitrogen (·g/m^3)	NO_x	2001	Quadratic	2,254.7	8,550m
		2008	Gaussian	1,130.4	600m
Nitrogen dioxide	NO_2 (·g/m^3)	2001	Quadratic	220.3	8,700m
		2008	Spherical	174.0	950m
Particulate matter (·g/m^3)	PM	2001	Spherical	21.47	5,550m
		2008	Gaussian	41.27	650m
Ground-level ozone (·g/m^3)	O_3	2001	Power	Coeff:0.02	Slope:0.64
			Nugget	28.66	
		2008	Gaussian	10	2,500m
			Nugget	0.4000	
Noise (dB(A))	L_{Aeq}	2001	Spherical	20.5417	8,550m
		2008	Spherical	16.6664	5,000m

2. MAIN RESULTS

In the previous sections, we have presented our research variables and we have kriged the objective indicators with the intention of building the matrix V of partial indicators. Note that we have not followed the same process as usual in this kind of applications. Firstly, we have kriged each of the indicators to secondly building a synthetic index for both 2001 and 2008.

In order to compare the variation of pollution in the period 2001-2008, we estimate EQIs for both air pollution and air-noise pollution as a whole. Firstly, we have applied DP2 to two matrices: V_1, which is of order 6 air-pollution indicators by 2,358 census tracts, and V_2, which is of order 7 indicators (air-pollution variables plus noise) by 2,358 census tracts.

Before starting the DP2 computation process, we must determine how the partial indicators contribute to the global one; i.e. if they have a positive or negative impact on environmental quality. As it was stated before, all the indicators must have a positive contribution to the phenomenon we are measuring, which is, in our case, environmental quality. In our case, the whole set of variables is negatively related to environmental quality. Hence, we decided not to change the sign of all the variables so as instead of measuring environmental quality, we will measure pollution; i.e. an increase/decrease of the values of any partial indicator will correspond with an improvement/worsening in pollution.

Next, for each matrix V_1 and V_2, we have computed their corresponding DF considering the minimum value of every partial indicator as the reference base $\left(v^* = \min\left\{v_{kj}\right\}\right)$. As a result, a higher value in the global indicator will imply a higher pollution level, since it implies a longer distance respect to a theoretical 'most-desired' situation. After that, we have calculated the correlation coefficients of each variable and their corresponding DF what lead to a first arrangement of the variables, in order to compute a first estimation of DP2. The final DP2 indexes reached the convergence after 2 and 3 iterations for EQI6 and EQI6, respectively (either for 2001 or 2008). In Table 3, we show the main results of DP2 computations for the two environmental indexes each year.

Concerning DP2 results, NO_x registers the highest correlation with the final DP2 in all the cases. This is why in the computation of DP2, it enters the first contributing all its variance to the final DP2 (correction factor=1). While a primary gaseous pollutant, NO_x, is the most important variable, the second contributor to DP2 is – in most of the cases– NO_2 (it enters the second'– in 2008 – and the third – in 2001 – in the DP2 index). Nevertheless, its information is mostly included in NO_x; i.e. it only adds 18-21% of its variance to the final indicator (correction factor=0.18 – in 2001 – and 0.21 – in 2008). In the case of EQI7, noise (L_{Aeq}) is the third contributor, donating a 66-77% of its variance in 2008 and 2001, respectively. For this reason, noise will have a relevant role in those indexes that include this variable. It must be noted that though O_3 is the least important indicator in both global indexes, the rest of partial indicators collects less than the 50% of its variance. This is why it gives to the final DP2 a range of values from 48% to 57% of its information.

Tab. 3: Main results of DP2 computations for the EQIs in 2001 and 2008.

		Rank				Correction factor			
		EQI6		EQI7		EQI6		EQI7	
		2001	2008	2001	2008	2001	2008	2001	2008
SO_2	Sulphur dioxide ($\mu g/m^3$)	4	5	4	6	0.77	0.74	0.56	0.62
CO	Carbon monoxide ($\mu g/m^3$)	2	4	2	4	0.52	0.45	0.52	0.46
NO_x	Oxides of nitrogen ($\mu g/m^3$)	1	1	1	1	1.00	1.00	1.00	1.00
NO_2	Nitrogen dioxide ($\mu g/m^3$)	3	2	3	2	0.18	0.21	0.18	0.21
PM	PM_{10} particulate matter ($\mu g/m^3$)	5	3	5	5	0.47	0.49	0.47	0.38
O_3	Ground-level ozone ($\mu g/m^3$)	6	6	7	7	0.49	0.57	0.48	0.56
L_{Aeq}	Equivalent continuous noise, dB(A)	-	-	3	3	-	-	0.66	0.77

Note*:* EQI6*: Environmental Quality Index for air-pollution, EQI7: Environmental Quality Index for air-pollution and noise, Rank: entrance order of partial indicators in the final DP2, Correlation coefficient: Pearson correlation coefficient of each indicator with the final DP2, Correction factor:* $\left(1-R^2_{k\cdot k-1,k-2,\ldots,1}\right)$ *the part of the variance that is not explained by the previously introduced indicators.*

The correction factor can be interpreted as the "weight" that corrects the standardized distance measure ($\boldsymbol{d}_k j/\sigma_k$). This weight does not depend of the importance of a partial variable on the synthetic indicator. It depends of the 'original' variance that a partial variable contains with respect of the rest of the variables (in fact, it is the part of the variance that is not explained by the previously introduced indicators). In Figure 2, we show a complete depiction of the correction factors for the two EQIs each year.

The computation results are apparently quite similar for the four indexes (EQI6 and EQI7 in 2001 and 2008), though some interesting differences can be detected in their spatial distribution. In Figure 3, we have represented these indexes. However, we have previously standardized the DP2 variables to facilitate their interpretation. In effect, though the original DP2 values are nonsensical in real terms, it is possible to compute the deviation to the mean value (multiplied by 100). Therefore, a value of 100 will correspond to area/time base; i.e. the DP2 city average in 2001. Values of DP2 above/below 100 mean pollution levels better/worse than the city average in 2001.

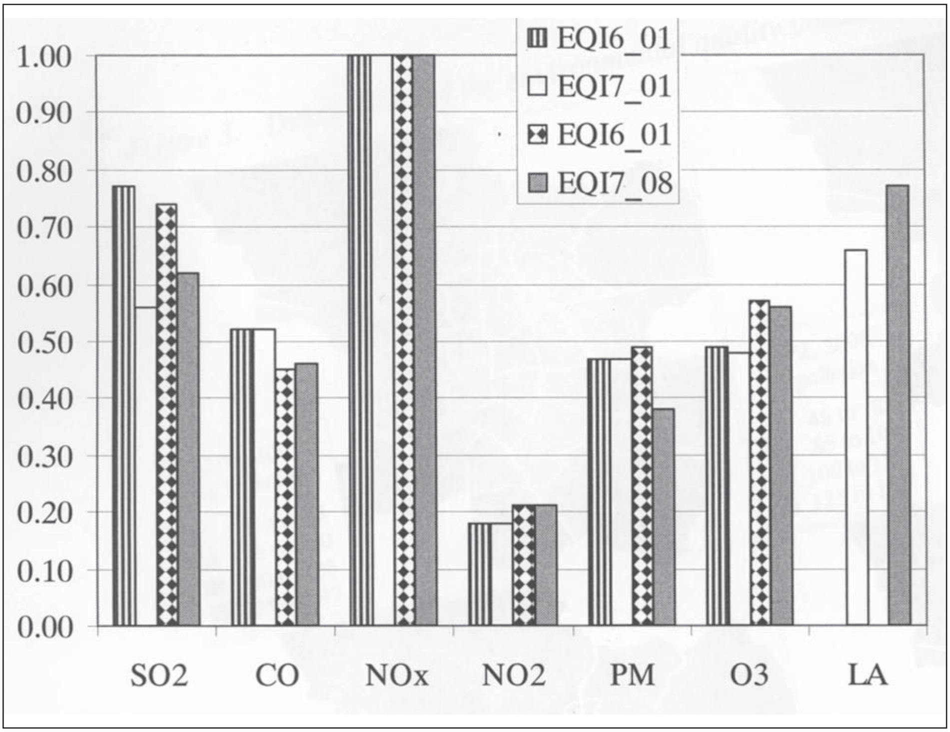

Fig. 2: Correction factor for air pollutants and noise in 2001 and 2008.

An initial analysis of those maps could result in the same conclusion: the highest levels of pollution are concentrated in the 'Central Almond' (the 7 central districts surrounded by the M-30 first belt) and the industrial southeastern periphery. Regarding the lowest levels of pollution, they seem to be located in some eastern/ western districts. Nevertheless, some interesting differences can be appreciated when comparing both indexes: air-pollution and air plus noise pollution. Actually, the introduction of noise in the EQI produces a more or less similar map in 2001, though in 2008, it penalizes some peripheral neighborhoods affected by the main radial highways, the M-30 and M-40 belts. In fact, more census tracts are in the upper intervals of DP2 when noise is included, especially in the peripheral areas of the city. This is particularly true when computing the evolution of the EQIs in the period 2001-2004 (Figure 4). Therefore, noise matters and must be considered as a pollutant when evaluating environmental quality.

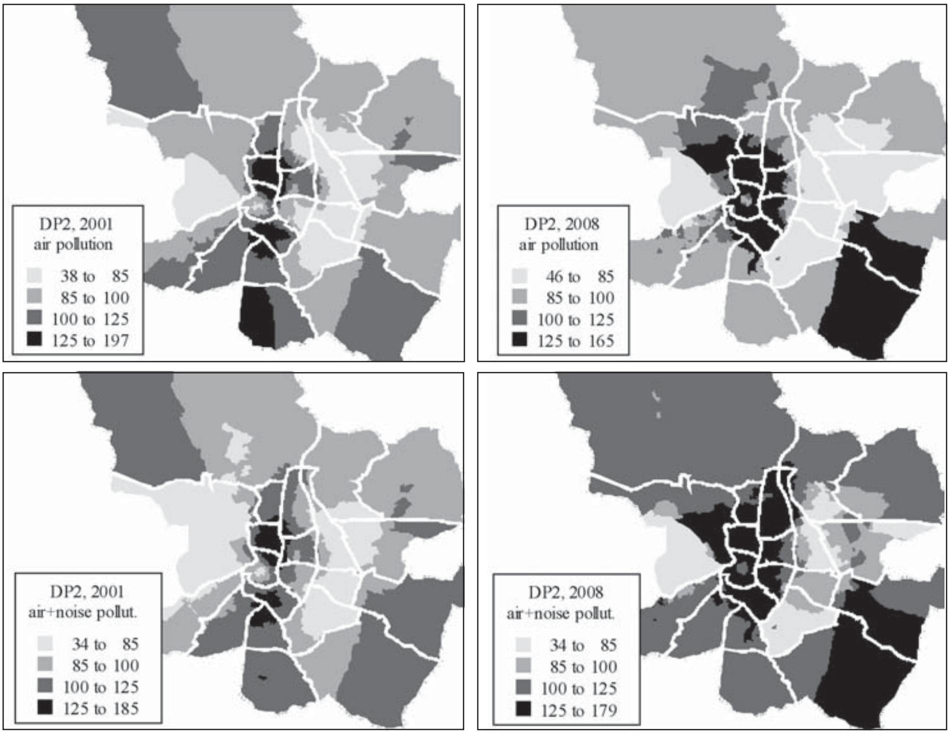

Fig. 3. Distribution of the environmental quality indexes in the city of Madrid.

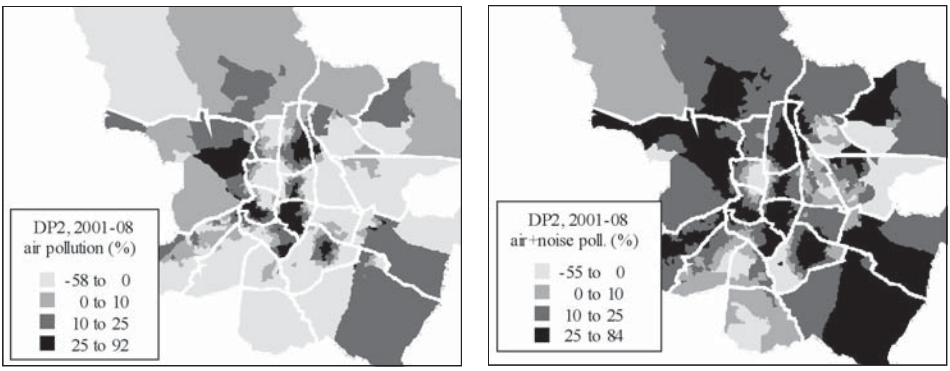

Fig. 4: Variation rates of air and noise pollution in Madrid.

4. CONCLUSIONS

As it is well known, the elaboration of Environmental Quality Indexes (EQIs) for metropolitan areas is one of the main topics in regional and environmental economics. In this paper, we have contributed to the development of the topic with some empirical and methodological novelties. Concerning the first, we propose an EQI with seven indicators: air pollutants (SO_2, CO, NO_x, NO_2, PM and O_3), jointly with noise while in the literature it is difficult to find environmental indexes with more than three partial indicators. Even the consideration of noise as a pollutant is a rather novel issue.

From a methodological point of view, mapping an EQIs can lead to the well-known ‘change of support’ problem. Kriging is the solution we propose to interpolate since it takes into account spatial dependence, which is a usual effect in the environmental variables. Although this scope is not new in the literature, we propose, as an innovation, a change of order in the procedure, since it leads to lower estimation errors. Firstly, we obtain the kriged estimates of the partial indicators for the desired locations, and secondly we compute the global index. Furthermore, we also recommend using a distance indicator the Pena Distance or DP2 instead of other synthesis methods, such as PCA. On the one hand, PCA is based on a mere reduction of information, while DP2 uses all the valuable information contained in the partial indicators, eliminating all the redundant variance present in these variables. On the other hand, DP2 has good statistical properties (i.e multidimensionality, comparability and comprehensibility) and it allows comparing values though space and time.

The abovementioned practical and methodological novelties have empirically been developed in a study case: the elaboration of an Environmental Quality Index for the city of Madrid in a spatio-temporal framework. Results have been certainly satisfactory and some interesting differences can be detected in their spatio-temporal distribution. For example, when including noise as a pollutant, the EQI penalizes some peripheral neighborhoods affected by the main radial highways and city-belts.

Once shown the main concluding remarks, new future lines of research immediately arise. For instance, in certain situations, cokriging could overcome better than kriging the ‘change of support’ problem. Besides, the use of observation networks could reduce the estimation errors in the interpolative stage of the elaboration of the index. At last, in other empirical context, Environmental Quality Indexes could be used as explanatory variables in hedonic housing price models.

AKNOWLEDGEMENTS

Jose-Maria Montero and Beatriz Larraz acknowledge financial support from the FEDER PAI-05-021 Project of the *Junta de Comunidades de Castilla-La Mancha*. Coro Chasco also acknowledges financial support from the CCG08-UAM/HUM-4173 of the *Consejería de Educación* (*Comunidad de Madrid*) and the Spanish Ministry of Education and Science SEJ2006-02328/ECON.

REFERENCES

ANSELIN L., LE GALLO J. (2006) "Interpolation of air quality measures in hedonic house price models: Spatial Aspects", *Spatial Economic Analysis* 1-1, pp. 31-52.

ANSELIN L., LOZANO-GRACIA N. (2008) "Errors in variables and spatial effects in hedonic house price models of ambient air quality", *Empirical Economics* 34, pp. 5-34.

BANZHAF H.S. (2005) "***Green price indices"***, ***Journal of Environmental Economics** and Management* 49(2), pp. 262-280.

BARANZINI A., RAMÍREZ J.V. (2005) "Paying for quietness: the impact of noise on Geneva rents", *Urban Studies* 42-4, pp. 633-646.

CANCELO J.R., URIZ P. (1994) "Una metodología general para la elaboración de índices complejos de dotación de infraestructuras", *Estudios Regionales* 40, pp. 167-188.

CES Murcia (2003), "La renta familiar disponible bruta y el índice de bienestar de los municipios de la Región de Murcia durante el periodo 1995-2000: estimación y análisis". Available from www.cesmurcia.org.

CHAY K.Y., GREENSTONE M. (2005) "Does air quality matter? Evidence from the housing market", *Journal of Political Economy*, 113(2), pp. 376-424.

DE IACO S., MYERS D.E., POSA D. (2001) "Total air pollution and space–time modelling". In: Monestiez P, Allard D., Froidevaux R. (eds.), *GeoEnv III—Geostatistics for Environmental Applications*, Kluwer Academic Publishers, Dordrecht, pp. 45–56.

DE IACO S., MYERS D.E, POSA D. (2002) "Space-time variograms and a functional form for total air pollution measurements", *Computational statistics & Data Analysis* 41, pp. 311-328.

EMERY X. (2000) *Geoestadística lineal*, Departamento de Ingeniería de Minas, Facultad de CC. Físicas y Matemáticas, Universidad de Chile.

ESCOBAR L. (2006) "Indicadores sintéticos de calidad ambiental: un modelo general para grandes zonas urbanas", *Eure* XXXII-96, pp. 73-98.

GOOVAERTS P. (1997) *Geostatistics for natural resources evaluation*, Oxford University Press, New York.

GOTWAY C.A., YOUNG L. (2002) "Combining incompatible spatial data", *Journal of the American Statistical Association* 97 (458), pp. 632-648.

Hartley P.R., Hendrix M.E., Osherson D (2005) *Real estate values and air pollution: measured levels and subjective expectations*, Discussion Paper, Rice University.

ISATIS (2001) *Isatis v4.1.1. Software Manual*, Avon, France: Geovariances and Ecole des Mines de Paris.

IVANOVIC B. (1963) "Classification of underdeveloped areas according to level of economic development", *Eastern European Economics* II, pp. 1-2.

IVANOVIC B. (1974) "Comment ètablir une liste des indicateurs de development", *Revue de Statistique Apliquée* 22-2, pp. 37-50.

Jenks G.F., Caspall F.C. (1971) "Error on choroplethic maps: definition, measurement, reduction", *Annals of the Association of American Geographers* 61-2, pp. 217-244.

KIM C.-W., PHIPPS T.T., ANSELIN L. (2003) "Measuring the benefits of air quality improvement: a spatial hedonic approach", *Journal of Environmental Economics and Management*, 45, pp. 24-39.

LARK R.M., PAPRITZ A. (2003) "Fitting a linear model of corregionalization for soil properties using simulate annealing", *Geoderma* 115, pp. 245-260.

MATHERON G. (1979) *Recherche de simplification dans un problème de cokrigeage*. Publication nº 628, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau.

MONTERO J.M., LARRAZ B. (2006) "Estimación espacial del precio de la vivienda mediante métodos de krigeado", *Estadística Española* 48, pp. 62-108.

MURTHY M.N., GULATI S.C., BANERJEE A. (2003) *Hedonic property prices and valuation of benefits from reducing urban air pollution in India*, Delhi Discussion Papers 61, Institute of Economic Growth, Delhi, India. Available from http://ideas.repec.org/p/ind/iegddp/61.html.

MYERS D.E. (1983) "Estimation of linear combinations and cokriging", *Mathematical Geology* 15, pp. 633-637.

NEILL H.R., HASSENZAHL D.M., ASSANE D.D. (2007) "Estimating the effect of air quality: spatial versus traditional hedonic price models", *Southern Economic Journal* 73 (4), pp. 1088-1111.

Official Journal of the European Union (2008) *Directive 2008/50/Ec of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe*. Available from http://eur-lex.europa.eu/JOIndex.do.

PENA J.B. (1977) *Problemas de la medición del bienestar y conceptos afines (Una aplicación al caso español)*, Presidencia del Gobierno, Instituto Nacional de Estadística, Madrid.

PREISENDORFER R.W. (1988) *Principal Component Analysis in meteorology and oceanography*. Elsevier, Amsterdam.

ROYUELA V., SURIÑACH J., REYES M. (2003) "Measuring quality of life in small areas over different periods of time. Analysis of the province of Barcelona", *Social Indicators Research* 64, pp. 51-74.

SÁNCHEZ M.A., RODRÍGUEZ N. (2003) "El bienestar social en los municipios andaluces en 1999", *Revista Asturiana de Economía* 27, pp. 99-119.

SMITH V.K., HUANG, J.C. (1993) "Hedonic models and air pollution: twenty-five years and counting", *Environmental and Resource Economics* 36-1, pp. 23-36.

SMITH V.K., HUANG J.C. (1995) "Can markets value air quality? A meta-analysis of hedonic property value models", *Journal of Political Economy* 103, pp. 209-227.

SMITH V.K., KAORU Y. (1995) "Signals or noise-explaining the variation in recreation benefit estimates", *American Journal of Agricultural Economics* **72,** pp. 419-433.

SPENCE J.S., CARMACK P.S., GUNST R.F., SCHUCANY W.R., WOODWARD W.A., HALEY W. (2007) "Accounting for spatial dependence in the analysis of SPECT Brain Imaging Data", *Journal of the American Statistical Association* 102, pp. 464-473.

SUBRAMANYAM A., PANDALAI H.S. (2004) "On the equivalence of the cokriging and kriging systems", *Mathematical Geology* 36-4, pp. 507-523.

TZENG S., HUANG H-C., CRESSIE N. (2005) "A fast, optimal spatial-prediction method for massive datasets", *Journal of the American Statistical Association* 100, pp. 1343-1357.

VICÉNS J., CHASCO C. (2001) *Estimación del indicador sintético de bienestar social publicado en el Anuario Social 2000,* Colección Estudios Sociales. Working Papers 1.

ZARZOSA P. (1996) *Aproximación a la medición del bienestar social*, Secretariado de Publicaciones e Intercambio Científico, Universidad de Valladolid.

LA COSTRUZIONE DI UN INDICE DI QUALITÀ AMBIENTALE SPAZIO-TEMPORALE: IL CASO DI MADRID

Riassunto

L'elaborazione di indici di qualità ambientale (EQIs, Environmental Quality Indexes) per le aree metropolitane è uno dei temi maggiormente dibattuti in ambito di economia ambientale. Il presente lavoro introduce alcune novità metodologiche e pratiche per la costruzione di un EQI, con riferimento alla città di Madrid (Spagna). Nell'ambito della gestione delle informazioni generali sulla qualità dell'aria, consideriamo il rumore, insieme all'inquinamento atmosferico, come una variabile ambientale rilevante. Poiché la disponibilità di tali variabili ambientali è limitata ad un sotto-insieme di stazioni di monitoraggio ambientale, dal punto di vista del processo di computazione utilizziamo il kriging method per abbinare i registri delle stazioni di monitoraggio ai dati censuari, al fine di mappare l'EQI. In una prima fase, le variabili ambientali sono abbinate all'intera superficie, procedendo successivamente all'elaborazione dell'indice ambientale. L'approccio seguito è di tipo inverso rispetto a quanto usualmente compiuto in letteratura, poiché esso consente di giungere a stime migliori. Da ultimo, per la costruzione dell'indice sintetico finale, utilizziamo il metodo basato sulla misura di distanza di Pena. I valori dell'indice EQI così ottenuti per gli anni 2001 e 2008 vengono confrontati, considerando l'evoluzione in termini di inquinamento atmosferico e acustico presi singolarmente ed insieme.