Using locally weighted regressions to model social inequalities in exposure to urban road traffic noise

Paulo Rui Anciaes

Research developed at: London School of Economics, United Kingdom
Current affiliation: University College London, United Kingdom

To present at the City Futures 2014 conference, Paris 18-20 June 2014.

Abstract

Despite the growing body of evidence on environmental inequalities related to urban transport, methodological issues continue to fuel debate among researchers. An unresolved issue regards the scope of analysis, as patterns of inequality may not be stable across an urban area. This issue is policy-relevant, as inequalities may have different understandings in terms of social justice, depending on the subsets of the city where they appear. This paper investigates this question by testing the spatial variability of the relationships between noise exposures and the characteristics of the population living in each place, testing the effect of using alternative definitions of the set of places within the metropolitan area that are included in the estimation of those relationships. The analysis uses a modified version of geographically weighted regression (GWR), weighting observations not in geographical space but in the space defined by variables such as the functional relationship of each neighbourhood in the metropolitan space and its position in relation to the main urban centre. The models are applied to the case of the Lisbon Metropolitan Area, analysing the effect of the radical expansion in the motorway network and road traffic during a decade. It is found that the areas where the disadvantages of low qualified populations are visible have extended from inner to outer suburbs. The changes in the transport system also reduced the advantage of elderly populations within the central areas.

Keywords: urban transport, equity, environmental justice, traffic noise, spatial variability GWR

1. Introduction

Are some social groups disproportionately exposed to the local environmental costs of urban transport? This question has been extensively researched during the last decade, due to the increased political interest in the topic and ease of access to spatially-referenced small-area census data. The hypothesis of environmental inequality has been confirmed in several countries [Braubach and Fairburn 2010, Deguen and Zmirou-Navier 2010], with most studies finding a link between deprivation and exposure to road traffic air pollution [Brainard *et al.* 2002, Mitchell and Dorling 2003, Havard *et al.* 2009] and traffic noise [Brainard *et al.* 2003, Hoffmann *et al.* 2003]. However, methodological questions are still object of intense discussion, limiting the applicability of this type of research in transport policy design and project assessment.

One of the most pressing questions is the definition of the areas over which statistical models are applied. If these areas are not carefully chosen, inequalities may simply reflect variations in confounding factors among the different neighbourhoods [Mitchell and Walker 2007, p.461]. The body of literature on transport-based environmental justice usually considers cities or metropolitan areas as the scope of analysis. However, the borders of these areas do not always correspond to discontinuities in the forces affecting the location of transport infrastructure and the spatial distribution of the population. When those forces operate over a wider area, there are distortions arising from not including all the relevant areas in the analysis ("edge effects"). When those forces operate at a smaller scale, the analysis may hide relevant detail, due to the multiple distributive patterns that arise when the processes behind those patterns are non-stationary in space, that is, vary among the different parts of the study area.

The issue is relevant because the political relevance of the inequalities found depends on their interpretation in terms of principles of social justice. According to authors such as Young (1990) and Elster (1992), these principles are contextual, as in practice, different "goods" are distributed following different principles, by different institutions and at particular times and places. In that sense, the assessment of the social distribution of environmental quality is only relevant for the correction of inequalities that emerged in the light of specific processes [Cutter 1995, Feitelson 2002]. Different policy priorities then apply to the achievement of equality along different socio-economic lines, in specific locations and at specific scales. In addition, the correction of inequalities over space depends on the degree to which they overlap with administrative units and on the availability of the relevant policy instruments to the government of those units.

The hypothesis of this paper is that in urban contexts, and especially in heterogeneous metropolitan areas, households and policy-makers make decisions over space by comparing neighbourhoods according to their functional role. The relationships between levels of pollution and socio-economic variables may then vary with the local relevance of those factors, as these factors have a variable influence in the locations of population and transport infrastructure. In addition, individuals also assess the fairness of transport and environmental policies by comparing the conditions of neighbourhoods that they believe to be geographically or demographically similar. In other words, they construct "communities of justice" defining the set of areas among which they consider that the policy-maker should equalize environmental quality. For these reasons, it is possible to derive different inequality patterns from the same scenario, depending on the attributes defining comparable areas.

The objective of the analysis is then to assess the influence of the definition of different comparison spaces in the relationships between socio-economic variables and road pollution. The analysis focuses on the distribution of a measure of potential noise exposure and uses the Lisbon Metropolitan Area as a case study, evaluating the effects of a large set of motorway projects in the patterns of inequality in the distribution of exposures.

The next section reviews some of the methods used in previous literature to address the issue of spatial stationarity in environmental justice studies, and proposes a new method based on local regressions weighted in attribute space. Section 3 presents the study area, and the variables and model formulation. The results are presented in Section 4. Section 5 discusses the limitations posed to policy intervention by the existence of inequalities in different locations and at different scales and addresses some methodological questions left open. Section 6 concludes the paper.

2. Spatial non-stationarity and "communities of justice"

The concept of spatial non-stationarity describes the variation of a statistical relationship estimated in different places. This issue is especially relevant in relationships measured in large or heterogeneous regions. The relationships between transport pollution and socio-economic variables tend to show some degree of spatial non-stationarity, as they depend on contextual factors that influence both the transport and land use policies applied at each place and the individuals' preferences and choices in the housing and transport markets. Associations may also occur at different spatial scales, as they depend on forces that are active only at those scales: the scale of policy-makers' decisions on the location of transport infrastructure and traffic policies (i.e., administrative areas) and the scale considered by households in their residence location decisions.

Spatial non-stationarity can be modelled by estimating separate regressions on different partitions of the data. These regressions can be related across more than one scale. For example, multilevel models estimate relationships between pollution and socio-economic variables at the level of individuals and places aggregated at different levels [McLeod *et al*. 2000, Briggs *et al*. 2008]. They assume, however, that the processes analysed are discontinuous at the borders between areal units, a hypothesis that collides with the fact that most spatial processes are unbounded [Leung 1987].

Geographically weighted regression (GWR) [Brunsdon *et al*. 1996, Fotheringham *et al*. 2002] is a solution to model the cases when there are no assumptions about the scale at which non-stationary processes operate or about the trends they follow over space. This method assumes that relationships are specific to each place and can be estimated by using all data points weighted by a decreasing function of their distance to that place. The output is a continuous surface of regression parameters across the study area. This procedure is flexible as it relies on user-defined options regarding the number of data points included in each regression and the specification of the weight function. GWR has been applied to show the non-stationary character of social differences in the health risks of pollution exposure [Gilbert and Chakraborty 2011]. Similar results were found for relationships that depend on processes similar to the location of transport infrastructure, such as the location of point sources of pollution [Mennis and Jordan 2005].

While these studies have focused on very broad areas, the spatial variability of environmental inequalities within an urban area presents distinctive characteristics, which remain under-researched. Urban space is not usually perceived as isotropic and continuous and as such, the non-stationarity in statistical relationships is not necessarily linked to geographic location. In fact, the associations between pollution and the social structure of each neighbourhood depend on the set of alternative places considered by households when searching for residences, and on the places considered by the policy-maker when deciding on the location of transport infrastructure.

The assumption implicit in the GWR method is that the degree to which a given place is an alternative for another place depends on the geographic locations of both. If weights are a decreasing function of distance, in practice, each local regression measures the relationships occurring in the areas around a given place. However, households and policy-makers usually consider a broad set of places located in different parts of the metropolitan area but having similar characteristics, for example in terms of centrality or available transport options. The spatial variability of transport inequalities may then be relative not to distances in geographic space but to differences in attributes related to the functional role of each area, as these are the attributes limiting the options and shaping the perceptions of the actors behind the processes leading to the inequalities.

It is also possible to argue that the associations between socio-economic variables and local environmental quality are also socially and politically relevant when they occur in subsets of the metropolitan region that correspond to individuals' judgements of the set of comparable areas in the distribution of pollution. These areas correspond to what [Dobson 1998, pt.3] calls "communities of justice", that is, the group of individuals one judge to be entitled to receive justice as part of public policy. This concept is especially relevant in the case of policies that affect different neighbourhoods in an urban area. Some authors suggest for example that public perceptions of equity in issues such as environmental policy are based on the individuals' delimitation of a "community of justice" that they believe is attached to each policy [Dietz and Atkinson 2005]. In urban contexts, this delimitation is less based on geographic proximity than on the perceived similarity between places. The higher the similarity between two places in terms of these variables, the higher they are judged to be comparable and included in the same community of justice.

This paper assumes therefore that relationships between noise exposures and socio-economic variables are specific to all places sharing certain geographic attributes. In other words, inter-neighbourhood differences in exposures are compared within subsets of the metropolitan area which are similar regarding those attributes. The results are compared with those of the global regression model (which assumes that location is irrelevant in the comparison of inter-neighbourhood differences in exposures) and the GWR model (which assumes that the comparison spaces for each place are the areas around that place).

3. Methods

The objective of the models is to explain the spatial distribution of noise exposures as a function of socio-economic variables. A separate regression is estimated for each place, using the data for all places weighted by their distances measured not in geographic space (as in the GWR approach) but in the space defined by those attributes.

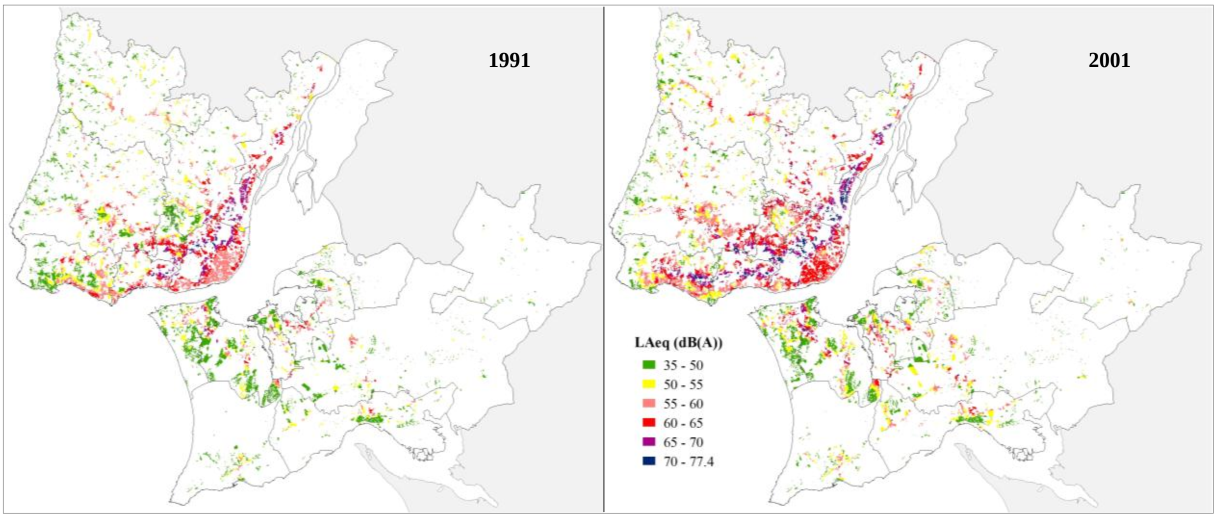
This formulation has methodological advantages when comparing with alternatives. Spatial variability could be inferred, for example, by modelling regression parameters as functions of the weighting variables, using interaction terms or more complex formulations. However, this method imposes a rigid structure to the variability of the relationships. Besides being more flexible, regressions weighted in attribute space also correct for spatial autocorrelation. This issue tends to affect transport equity studies, as census variables and noise levels usually assume similar values in nearby places, leading to biased and unreliable parameter estimates [Buzzelli and Jerrett 2004, Havard *et al*. 2009, Chakraborty 2009]. The use of locally weighted regressions reduces this effect [Fotheringham *et al*. 2002, pt.5].

The Lisbon Metropolitan Area provides an interesting case study for the study of spatial non-stationarity, since it is a highly heterogeneous region in terms of land use, job density and population density. The analysis focuses on two moments in time: 1991 and 2001. This period was chosen as it corresponds to the period with the fastest changes in the transport network in the Lisbon Metropolitan Area.

The unit of analysis is the census enumeration district. An indicator of exposure to noise is defined taking into account the conditions within the whole district. This is achieved by considering that people are exposed when walking in their neighbourhood, and that the probability that they walk to one place in the neighbourhood depends on its population density. A set of representative pedestrian routes used by the population of each district in 1991 and 2001 can then be derived, and overlaid with noise surfaces for these years, yielding indicators of average noise exposures for each district and year. The noise surfaces used are obtained by modelling road and rail traffic, industrial noise, flight path and other noise sources. The distribution of the exposures in 1991 and 2001 is shown in the maps below.

Short description of the map and identify where is Lisbon.

Figure 1: Noise exposures in the Metropolitan Area of Lisbon



The explanatory variables are composite variables extracted from the Portuguese census by a factor analysis. The two main variables of concern, as they measure demographic characteristics usually included in environmental equity studies, are Age and Qualifications. The Age factor is related to both age of people and buildings. The Qualifications factor is an indicator of socio-economic status, including variables such as schooling, skilled jobs, large and owner-occupied dwellings, and value of rents and mortgage payments. A third variable measures Urbanization is included as a control variable. This variable is related to the number of dwellings per building and proportion of non exclusively-residential buildings.

Six alternative geographic attributes are considered for weighting the regressions. *Accessibility* is a gravity-type measure of job accessibility, which takes into account the shortest travel times to a series of major employment centres. *Local accessibility* is defined as the times to the nearest centre. *Distance to Lisbon* and *Time to Lisbon* are indicators of position of each district in relation to the main centre in the metropolitan area (Lisbon) and are estimated by calculating the shortest paths to this city's CBD on the transport network. The remaining two attributes are indicators of the functional role of each place within the metropolitan space: *Time to work* is the weighted average of travel times of the employed population, estimated by modelling walking, private transport and public transport trips at different times of day and considering commuting data from the census and congestion levels.

In each year, the regression parameters vary in the space defined by one of the variables defined above. The model is then formed by a set of local regressions, one for each enumeration district. In the regression for district i, y_i is the dependent variable, x_{ik} are the explanatory variables and b_0 and b_k are respectively the intercept and the regression parameters. These parameters are a function of S_i, the value(s) of the weighting variable(s). The regressions are linear, as it was found in preliminary analysis that non-linear transformations of dependent or independent variables tend to produce poorer models. As the explanatory variables are factor scores, which are standardised and largely uncorrelated, multicollinearity is also not an issue. The specification of the error terms ε_i follows the assumptions of the standard linear regression models: independence and identical distribution of the residuals, which have zero mean and constant variance.

$$y_i = b_0(S_i) + \sum_k b_k(S_i)x_{i,k} + \varepsilon_i$$

The estimation of the regression for each district includes data for all districts weighted by a decreasing function of their distance to the first district measured in the space defined by the weighting variables. The vector of parameters for the regression at district *i* is estimated as

$$b = \left[X^T W(i) X\right]^{-1} X^T W(i) Y$$

where *W(i)* is a diagonal matrix whose elements in the main diagonal give the weights assigned to each district in the data set (See Fotheringham *et al*. (2002, chap.2.7) for the derivation of this formula.)

The distance is the difference in the values of the weighting variable in the two districts. For example, when the variable is accessibility, the regression for a district includes all districts with the same level of accessibility. Districts with different accessibility levels are only partly included, weighted by the difference in accessibility between the two districts. Districts with very different accessibility levels will have a weight close to zero and are virtually excluded from the estimation. It is assumed that weights decrease according to a Gaussian function. The weight of district *j* in the regression for district *i* is given by the function below, where d_{ij} is the distance between the two districts.

$$w_{ij} = \exp\left[-(d_{ij}/b)^2\right]$$

This function depends on a positive parameter *b* called the bandwidth. The lower the bandwidth, the steeper is the decay of the weights with distance. This parameter can be assigned a-priori or estimated by minimizing indices of the models' goodness of fit. The value of the bandwidth in the analysis that follow is the one that minimizes the Akaike Information Criteria (AIC), an index that takes into account the estimated standard deviation of the error term and the number of degrees of freedom in the model, which depend on the bandwidth [Hurvich *et al*. 1998]. The estimation used the GWR3 software, which can be adapted to locations defined in non-geographical space while using the Cartesian system.

4. Results

This section explores the different types of information provided by the estimation of the models defined above.

In a first stage, the results of the models (that is, the sets of local regressions) based on the six weighting schemes are compared with those of the global regression and the GWR model. Each local regression within a model is independent of the others (since it uses a specific set of weights) and yields separate parameter estimates and regression diagnostics, including measures of goodness of fit. This means that it is possible to study the trends that these estimates follow for different values of the weighting variable.

In a second stage, the statistical and spatial distribution of the estimates of one of the models (where the weighting is distance to Lisbon) is studied in more detail, and compared with the distributions obtained with the GWR approach.

Overview

The results are synthesized in **Table 1**. For each year, each line in the table represents the model estimation based on one comparison space. This estimation is composed of a set of regressions, centred at each enumeration district. The first two lines give the results of the global model and of the GWR model. The goodness of fit of each model is given by the AKAIKE index. The local regressions in each model produce separate parameter estimates and coefficients of determination, which can be matched with the values of the weighting variable at the district where the regression is centred. For each year, the second group of columns then gives the Pearson correlations between the weighting variable and the regression R^2 and coefficient estimates.

Table 1: Regressions of noise: Goodness of fit and parameter summary

Weighting scheme	1991					2001				
	AIC	Correlations with the weighting variable				AIC	Correlations with the weighting variable			
		R^2	F_1	F_2	F_3		R^2	F_1	F_2	F_3
(Global model)	101460					142702				
GWR	91799					128652				
Accessibility	97721	0.52	-0.75	-0.14	-0.92	133651	0.48	-0.80	-0.23	-0.74
Local accessibility	100760	0.81	0.09	-0.45	0.17	140603	0.52	0.09	-0.72	-0.52
Distance to Lisbon	97976	-0.18	0.44	-0.30	0.61	136150	0.02	0.61	0.43	0.60
Time to Lisbon	97294	-0.24	0.75	0.24	0.74	135213	-0.40	0.75	0.45	0.69
Time to work	100829	-0.89	-0.10	-0.86	-0.07	141563	-0.51	-0.92	-0.93	0.22
Commuting to Lisbon	98639	-0.24	-0.16	0.25	0.06	137835	-0.08	-0.02	-0.02	-0.46

The column with the AKAIKE index (AIC) shows that all the models weighted in attribute space have better overall goodness of fit than the global model, but lower than the model weighted in geographic space (GWR).

The analysis of the trends followed by the coefficient of determination of the local regressions (R^2) show that the regressions fit better as we move towards areas that have higher accessibility, shorter times to work and are nearer to the main centre. What about commuting to Lisbon. In the first two cases, this occurs alongside an increasing negative influence of the three parameters. In contrast, when the weighting variables are the position in relation to Lisbon, the improvement of the model fit occurs alongside an increasing influence of the three parameters, with the exception of the qualification parameter in the model weighted on travel distance to Lisbon in 1991.

The statistical distribution of the parameter estimates is presented in **Fig.2**, which shows box plots of the four quartiles of the estimates' statistical distribution. The figure shows that…

The hypothesis of non-stationarity of coefficients can be assessed by comparing the standard error of the coefficients estimated in all local regressions (s_{loc}) with the standard error of the coefficients in the global regression model (s_{glo}). Although the theoretical statistical distribution of s_{loc} is an unknown, the hypothesis of non-stationarity of the local coefficients can be formally tested using simulation techniques such as the Monte Carlo test. In the case of locally weighted regressions, a possible test is to compare the position of the observed s_{loc} in a rank list of values obtained re-estimating the model with a large number of different random rearrangements of data in space [Brunsdon *et al.* 1996]. The test reveals that the probabilities of obtaining the observed s_{loc} if the process was stationary tend to zero, suggesting that the parameter variability obtained by the model is the result of a non-stationary process.

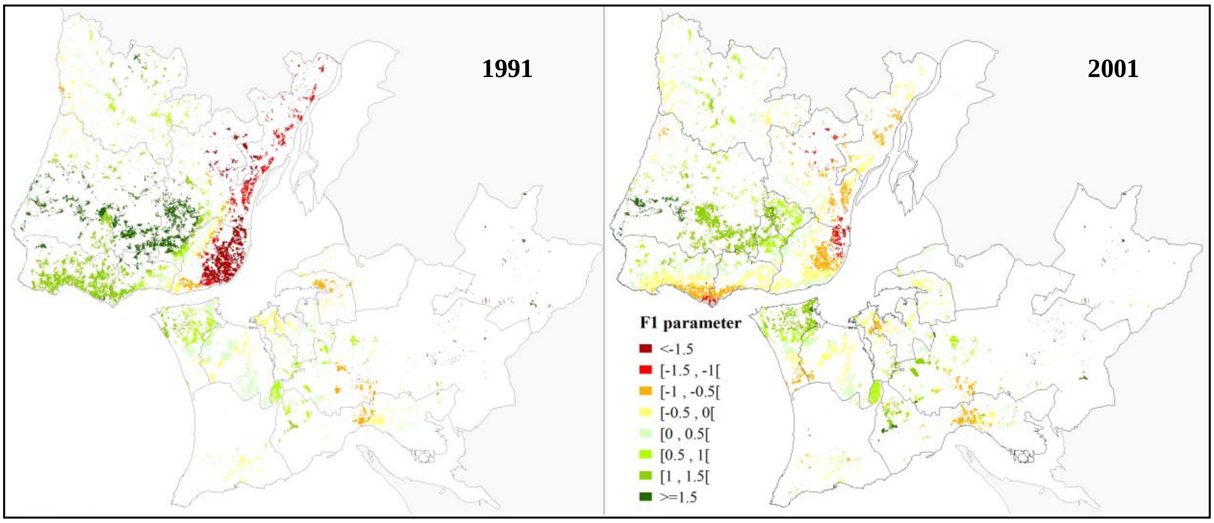
GWR model

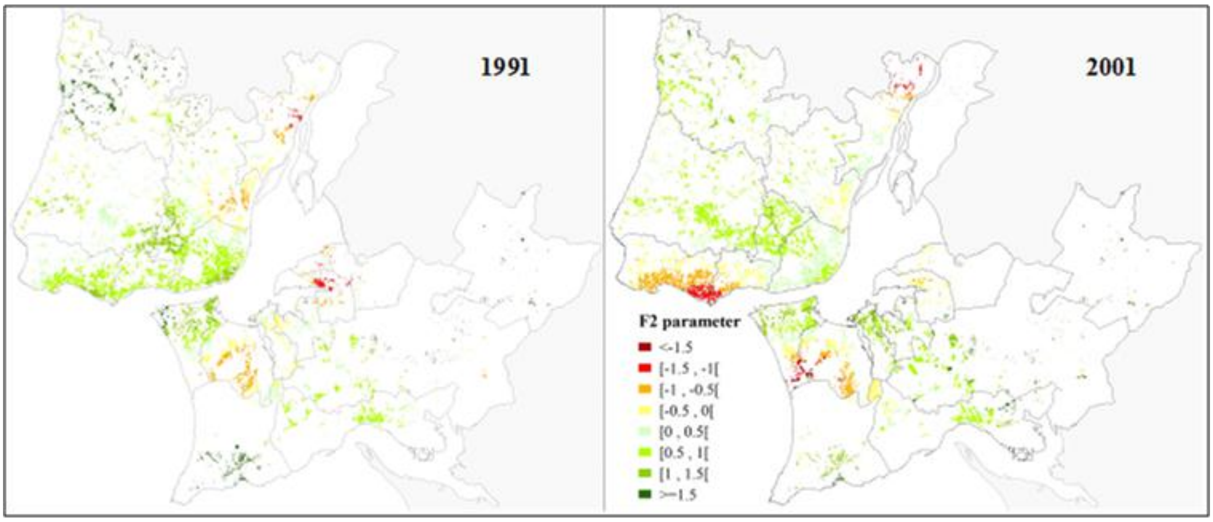
The results of the GWR estimation (where regressions in each district weight other districts based on geographic distance) are depicted in **Fig.3**. The two maps on top show the distribution of the F1 (Age) parameter and the two maps on the bottom show the distribution of the F2 (Qualifications) parameter.

In the case of the Age parameter, the highest negative estimates in 1991 are located in the Lisbon municipality. There is also a clear opposition between the negative estimates found in the Northeast corridor and the positive estimates found in the Northwest and West corridors. This means that although the regressions in the GWR estimation are, in practice, based on areas around each district, the estimates tend to follow patterns that correspond to the main corridors of access to Lisbon. Similar patterns occur in 2001, although the differences between corridors become less marked, as the estimates decreased in absolute values in most districts in the metropolitan area.

The distribution of the Qualifications parameter in 1991 reveals a series of local patterns, but in 2001 patterns at the level of access corridors to Lisbon are also discernible. In particular, there was a marked change in the west corridor of access to Lisbon, with the parameters changing from positive to negative in all districts.

Figure 3: GWR of noise exposures (F1 and F2 parameter)





Model weighted on distance to Lisbon

The next set of figures present the results of the model weighted on distance to Lisbon. This model was chosen due to the ease of understanding of the sub-sets of areas which are directly compared (that is, included in the same model), as these sub-sets are imperfect rings around the city of Lisbon, based on transport network distance.

The charts in **Fig.4** show that noise exposures are negatively associated with age both in the closest and in the most distant areas to Lisbon. The association in the regions close to Lisbon has extended geographically from 1991 to 2001, but has reduced considerably in the areas where it was already present in 1991. In these areas, the urbanization parameter has also become positive. The location of the associations with the Age parameter can be seen in **Fig.5**. The borders of the areas with the highest negative values in 1991 corresponded almost exactly to the Lisbon municipal borders. The associations decreased in the residential areas around the centre, which in 2001 have similar estimates as the suburban areas around, where the parameter changed from positive to negative.

There are also important changes in the spatial patterns of distribution of noise exposures according to qualification levels. The range of the F2 parameter in 1991 is small and there are no intervals in the weighting variable where disadvantage is especially evident, either for the more or less qualified populations. This is not the case in 2001, where the second quartile of distribution of road distances to Lisbon defines a negative peak for the parameter estimates. The map shows that this interval corresponds to the hinterlands of the suburban municipalities surrounding Lisbon, both in the North and the South Banks. These are the areas where the motorway network has expanded the most during the period concerned.

The comparison of **Fig.5** and **Fig.3** reveals some of the differences between this model and the GWR model. In the case of the Age parameter, the changes in the signs of the parameters in some regions from 1991 to 2001 (for example, in the West corridor and in the South bank) are the opposite from the changes shown by the GWR model. In addition, in 2001, the estimates in Lisbon are higher (in absolute value) in the central part of the city, in contrast with the GWR model, where the estimates are higher in the east part.

There are also differences in the Qualification parameter estimates in the two models. In 1991, the GWR model follows similar patterns as the model using distances to Lisbon but the absolute value of the parameters tends to be higher. In addition, the changes from 1991 to 2001 are limited to a small number of areas, unlike the previous case, where there was a structural change in the spatial distribution of the estimates in the suburban areas. Only two zones are "inequality hotspots" using both weighting schemes, in the West corridor and in the two municipalities in the South Bank. In other words, in these areas, lower qualifications are associated with higher noise exposures, either when comparing each district with other districts at the same distance to Lisbon or when comparing with the surrounding districts.

Figure 4: Regressions of noise weighted by distance to Lisbon (chart)

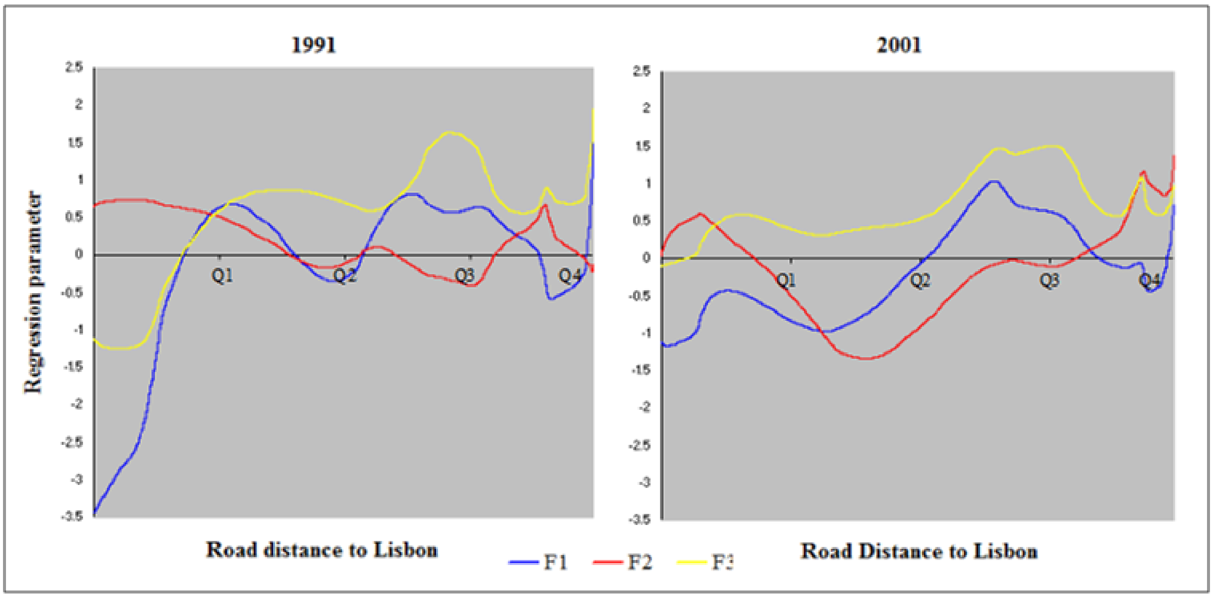
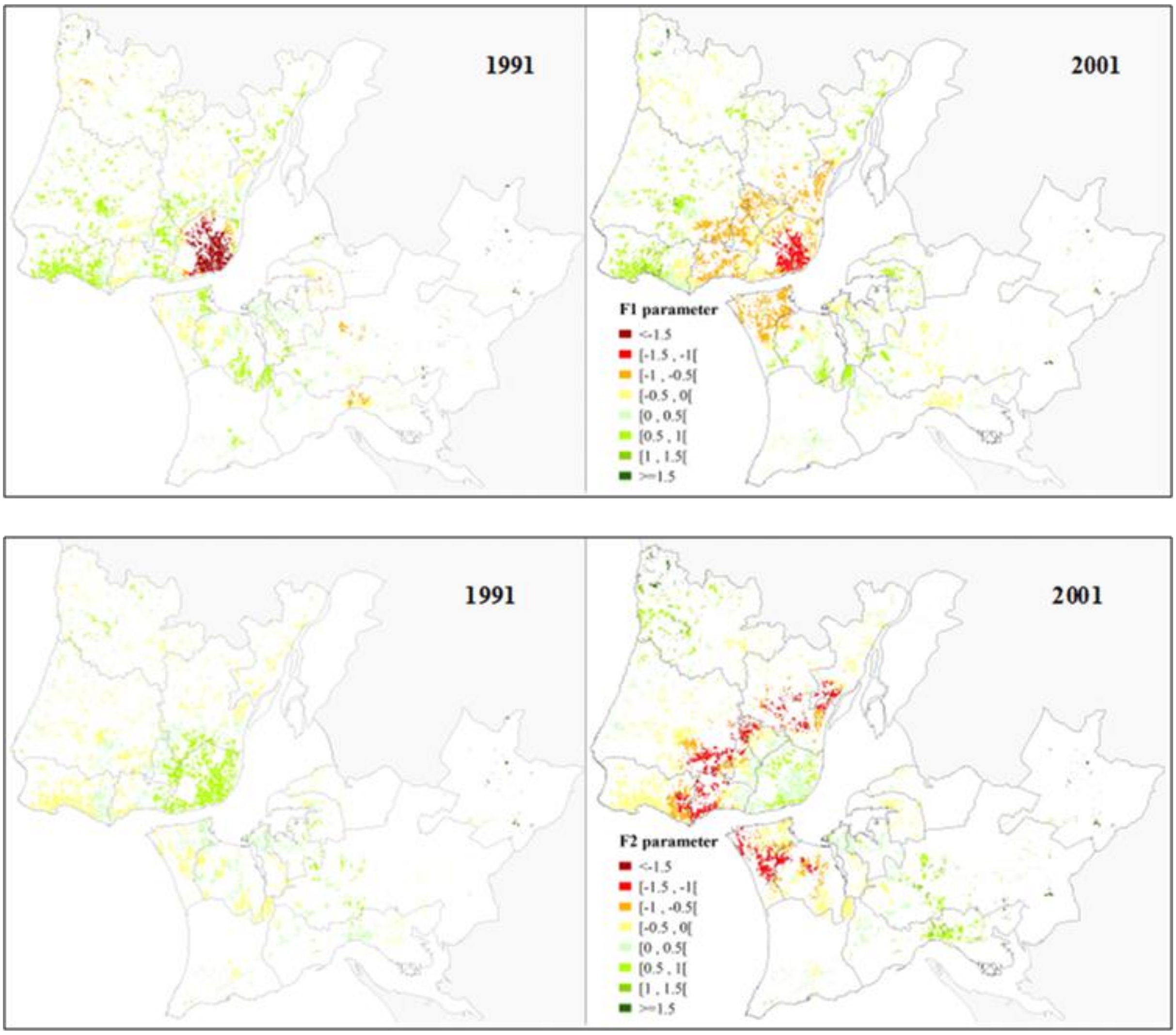


Figure 5: Regressions of noise weighted by distance to Lisbon (F1 and F2 parameter)



In conclusion, the distribution of exposures in the city of Lisbon represents a singular case, with different patterns from the rest of the metropolitan area. Older populations are associated with lower exposures within Lisbon, although this relationships has became weaker from 1991 to 2001. Throughout the metropolitan area, but especially in the suburban areas, the disadvantage of less qualified populations has increased. The extent of this disadvantage depends, however, on the subset of areas we compare directly.

5. Discussion

Policy implications

The existence of different types of inequalities in different parts of the metropolitan area raises questions about which of these inequalities deserves policy priority. The use of several different weighting schemes in the models can provide a guideline. For example, we can attach a higher level of priority to inequalities that appear regardless of the weights used. This is for example the case of the inequalities of populations with different qualifications in suburban regions. Patterns of inequality in other regions are more sensitive to the weights used and would then be attached a lower level of priority.

There are also differences between the types of policies required to address inequalities arising in different parts of the metropolitan area, regardless of their relative priority. This is because the nature of the unfulfilled needs depends on the geographic context in which those needs arise. The improvement of environmental quality in these areas is different from the case of suburban areas, as the main points of concern are not the environment conditions in the central areas of towns but the conditions along pedestrian routes that share the same space as national roads with heavy traffic.

The effectiveness of transport policies in correcting inequalities also depends on the scale of application of those policies, which may not be the same as the spatial scales at which inequalities occur. The social, economic and political processes that lead to social inequalities in the distribution of the benefits and costs of urban transport operate at a given spatial scale, that is, affect only a number of places. The scales implicit in the models in this paper may not overlap with administrative areas. The political structures that decide on the location of transport infrastructure and the regulation of traffic.

The scale of analysis is implicit in the specification of the local regressions. A sub-set of data in each local regression and every observation is assigned a probability of membership in this sub-set, which is defined by the particular weight structure adopted. The estimation of the model with endogenous determination of weights assumes the existence of an “optimal” partition of the data into fuzzy subsets that apply to each regression. As the weights that define this partition are the ones that minimize a measure of the models’ goodness of fit (the AKAIKE index), the size of the subsets of data included in each regression is an indicator of the scale of adjustment between spatial differences in dependent and explanatory variables. This is the scale of operation of the processes that influence the sorting of the different social groups according to the characteristics of the places and the processes that influence the adoption of transport policies by the policy-maker according to the characteristics of the populations affected. The consideration of alternative values for the weighting structure would correspond to different scales of analysis and different assumptions to the scale of operation of those processes.

With the exception of the patterns occurring within the borders of Lisbon municipality, the spatial distribution of the regression parameters obtained with the different weighting schemes suggests that the scale of the relevant processes does not correspond to the scale of the jurisdictions that hold effective power and responsibility over transport issues in the study area (the central and the municipal governments).

The results of this paper suggest that the distribution of noise follows patterns that spread over the set of suburban municipalities around Lisbon. The need to correct those inequalities is an argument in favour for the increase of the power and means available to supra-municipal institutions such as the recently created Lisbon Metropolitan Transport Authority, as the emergence of these inequalities also depends on the dynamics of residence location and commuting patterns that operate at a supra-municipal scale. At the same time, the formulation of transport and urban policies needed to address those inequalities requires the definition of plans and strategies at a scale lower than the national level. The same arguments can be used to support social and political claims for proceeding with abandoned plans for a territorial reorganization of the Portuguese administrative system.

The instruments to address inequalities may also not be available at each scale. For example, traffic restriction policies can promote equity by improving environmental quality in some neighbourhoods. These policies are usually applied at the municipal scale, as they require a level of detail that is only compatible with urban plans and enforcement structures specific to each municipality. If the policies applied in a given municipality follow criteria of equality based only on the comparison of the neighbourhoods within that municipality, then the definition of the areas of application of the policy will depend on the location of the groups considered to be at disadvantage. However, the effect on the distribution of noise exposures based on larger areas is indeterminate, given that the reorganization of traffic will affect the noise levels in places outside that municipality. In addition, there are effects on accessibility of workers living in other municipalities and commuting to the areas in which traffic restriction is applied.

Methods

The analysis in this paper assumed that spatial non-stationarity is a feature of the dataset. However, some of the variation captured by the parameter estimates may derive from the insufficiency of the set of independent variables in explaining differences in the dependent variables, due to omitted variables, model misspecification or measurement errors in dependent and independent variables. The use of more detailed socio-economic data or the inclusion of further variables (such as income levels, socio-professional status and population density) could therefore be beneficial.

There are also limitations on the extrapolation of the results found, as procedures for the statistical inference of models using weighted regressions are still developing. The test of spatial non-stationarity using Monte Carlo randomization distributions is an imperfect solution, as these distributions refer specifically to the data sets tested in each model [Leung *et al*. 2000]. Although the models' residuals do not show relevant spatial patterns, in general it is difficult to disentangle the effects of spatial non-stationarity, spatial autocorrelation and the errors implicit in the definition of dependent and independent variables. In this regard, the use of alternative definitions for the axis along which spatial variability is measured provides an advantage over the measurement of variability only on geographic space, as there is a smaller probability that the variability arose from chance or from insufficiencies of the model in all the different alternatives.

The model can also be improved by adopting local variants of procedures treated as global. The rate of decay of the weighting function was assumed to be constant across the study area, although this function can vary spatially by adopting a weighting system that assigns a gentler bandwidth in regions where data is sparser. The variance of the error term was also assumed to be constant in space. It is possible to address both problems by using other formulations, such as the error variance heterogeneity model [Páez *et al*. 2002].

6. Conclusions

This paper assessed the spatial variability of relationships between potential noise exposures and the characteristics of the population living in each place, testing the effect of using alternative definitions of the sub-set of places within the metropolitan area that are included in the estimation of those relationships. The analysis found multiple spatial patterns in the inequalities in the distribution of exposures. The differences regard both the locations and the scales at which inequalities are more evident. The effects of the expansion of the motorway network on the distribution of the indicators are also different. There was an intensification or reversal of the type of disadvantage in the inner suburbs and also a reduction of the advantage of younger populations.

The measurement of this variability depends however, on assumptions regarding the processes that lead to the adjustments between the different social groups and the conditions of each neighbourhood. The assumption that these processes operate within sets of areas with similar values for a given attribute leads to the identification of a specific patterns of inequalities, arising in certain locations and at certain scales. More generally, the analysis in this paper adds to the idea that there are multiple perspectives to the assessment of environmental inequalities and that these perspectives are often implicit in the statistical methods used.

The regressions in each model represent equilibrium relationships between the location of social groups and exposures at given moments in time, but do not explain the processes leading to these relationships. The analysis in this paper provides a guideline for the definition and the interpretation of these relationships. The confirmation that these relationships are product of an adjustment process of individuals to conditions existing in the different neighbourhoods requires the study of relationships between changes in the transport system, the neighbourhoods' social structure and patterns of urban land use.

Acknowledgements

Thanks go to Giles Atkinson, Steve Gibbons, Helena Titheridge and Andrew Lovett for their comments and suggestions. Financial support from the Portuguese Foundation for Science and Technology is acknowledged.

References

Brainard, J., Jones, A., Bateman, I., Lovett, A. (2002) Modelling environmental equity: access to air quality in Birmingham, England. *Environment and Planning A* **34 (4)**, 695-716.

Brainard, J., Jones, A., Bateman, I., Lovett, A. (2003) Exposure to environmental urban noise pollution in Birmingham, UK. *Urban Studies* **41 (13)**, 2581-2600.

Braubach, M., Fairburn, J. (2010) Social inequities in environmental risks associated with housing and residential location – a review of evidence. *European Journal of Public Health* **20 (1)**, 27-35.

Briggs, D., Abellan, J J., Fecht, D. (2008) Environmental inequity in England: Small area associations between socio-economic status and environmental pollution. *Social Science & Medicine* **67 (10)**, 1612-1629.

Brunsdon, C., Fotheringham, A S., Charlton, M. (1996) Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis* **28 (4)**, 281-298.

Buzzelli, M., Jerrett, M. (2004) Racial gradients of ambient air pollution exposure in Hamilton, Canada. *Environment and Planning A* **36 (10)**, 1855-1876.

Chakraborty, J. (2009) Automobiles, air toxics, and adverse health risks: Environmental inequities in Tampa Bay, Florida. *Annals of the Association of American Geographers* **99 (4)**, 674-697.

Cutter, S. (1995) Race, class and environmental justice. *Progress in Human Geography* **19 (1)**, 111-122.

Deguen, S., Zmirou-Navier, D. (2010) Social inequalities resulting from health risks related to ambient air quality – A European review. *European Journal of Public Health* **20 (1)**, 27-35.

Dietz, S., Atkinson, G. (2005) Public perceptions of equity in environmental policy: Traffic emissions policy in an English urban area. *Local Environment* **10 (4)**, 445-459.

Dobson, A. (1998) *Justice and the Environment: Conceptions of Environmental Sustainability and Social Justice*. Oxford University Press, Oxford.

Elster, J. (1992) *Local Justice*. Russell Sage Foundation, New York.

Feitelson, E. (2002) Introducing environmental equity dimensions into the sustainable transport discourse: issues and pitfalls. *Transportation Research D* 7 (2), 99-118.

Fotheringham, S A., Brunsdon, C., Charlton, M. (2002) *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*. Wiley, Chichester.

Gilbert, A., Chakraborty, J. (2011) Using geographically weighted regression for environmental justice analysis: cumulative cancer risks from air toxics in Florida. *Social Science Research* **40 (1)**, 273-286.

Havard, S., Deguen, S., Zmirou-Navier, D., Schillinger, C., Bard, D. (2009) Traffic-related air pollution and socioeconomic status: A spatial environmental equity study on a small-area scale. *Epidemiology* **20 (2)**, 223-230.

Hoffmann, B., Robra, B P., Swart, E. (2003) Social inequality and noise pollution by traffic in the living environment: an analysis by the German Federal Health Survey. *Gesundheitswesen* **65 (6)**, 393-401. [Article in German]. English abstract available at http://www.ncbi.nlm.nih.gov/pubmed/12836130

Hurvich, C M., Simonoff, J S., Tsai, C-L. (1998) Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society Series B* **60 (2)**, 271-293.

Leung, Y. (1987) On the imprecision of boundaries. *Geographical Analysis* **19 (2)**, 125-151.

Leung Y., Mei, C-L., Zhang, W-X. (2000) Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A* **32 (1)**, 9-32.

McLeod, H., Langford, I., Jones, A., Stedman, J., Day, R., Lorenzoni, I., Bateman, I. (2000) The relationship between socio-economic indicators and air pollution in England and Wales: implications for environmental justice. *Regional Environmental Change* **1 (2)**, 78-85.

Mennis, J L., Jordan, L. (2005) The distribution of environmental equity: Exploring spatial nonstationarity in multivariate models of air toxic releases. *Annals of the Association of American Geographers* **95 (2)**, 249-268.

Mitchell, G., Dorling, D. (2003) An environmental justice analysis of British air quality. *Environment and Planning A* **35 (5)**, 909-29.

Mitchell, G., Walker, G P. (2007) Methodological issues in the assessment of environmental equity and environmental justice., in M Deakin, G Mitchell, P Nijkamp and R Vreeker (Eds.) *Sustainable Urban Development Volume 2: The Environmental Assessment Methods*. Sustainable Urban Development Series. Routledge, London., pp. 447-472.

Páez, A., Uchida, T., Miyamoto, K. (2002) A general framework for estimation and inference of geographically weighted regression models: 1. Location-specific kernel bandwidths and a test for locational heterogeneity. *Environment and Planning A* **34 (4)**, 733-754.

Young, I. M., (1990) *Justice and the Politics of Difference*. Princeton University Press, Princeton.