Measuring environmental inequalities: insights from the residential segregation literature

Yves Schaeffer, Mihaï Tivadar

Measuring environmental inequalities: insights from the residential segregation literature ‡

Schaeffer[1,*] Y., Tivadar[1] M.

[1] Univ. Grenoble Alpes, Irstea, LESSEM, 38000 Grenoble, France

[*] Corresponding author: yves.schaeffer@irstea.fr

‡Authorship is equally shared among co-authors.

Abstract:

Inequalities in exposure to environmental hazards and access to environmental amenities have been documented in many cities, in relation to residential segregation of low-income or minority groups. The literature on residential segregation measurement, however, has not yet been considered a source of insights for the measurement of environmental inequalities. Here, we propose two segregation-based indices – the Environmental Dissimilarity gap index (ΔED_K) and the Environmental Centralization index (EC_d) – and a randomization method to make robust environmental inequality assessments. In addition, to help policy-makers target local policies better, we developed an original approach to identify and map hotspots that have a large influence on environmental inequalities. These methods are applied in Grenoble, France, to study the distribution of green spaces and industrial risks between poor and non-poor households.

Keywords: Environmental Justice; Environmental Equity; Spatial Segregation; Monte Carlo Simulations, Jackknife Simulations; France

1. Introduction

The overarching goal of environmental justice (EJ) studies is to inform EJ policies, whose objective according to Harner et al. (2002, p. 318) is "to create environmental equity: the concept that all people should bear a proportionate share of environmental pollution and health risk and enjoy equal access to environmental amenities". More precisely, Boyce et al. (2016) underline the policy-relevance of studying both inter-individual (i.e. vertical) and between-group (i.e. horizontal) environmental inequalities. Many national constitutions as well as environmental statutes and regulations endorse the normative principle that every person has the right to clean and safe environment. And – in the United States at least – the requirement of equity across groups defined on the basis of race, ethnicity and economic status is explicitly inscribed in environmental policy.

In practice, a large number of EJ studies are conducted at a city or metropolitan area scale. They focus on interactions between population groups unequally distributed in space and spatialized environmental amenities/disamenities. As an example, many studies consider the uneven distribution of urban green spaces (e.g., Apparicio et al., 2016; Frey, 2016; Schwarz et al., 2015; Shanahan et al., 2014; Wen et al., 2013; Zhou and Kim, 2013; Pham et al., 2012; Landry and Chakraborty, 2009) or unequal exposures to urban air pollutants (e.g., Carrier et al., 2014; Zwickl et al. 2014; Harner et al., 2002; Sheppard et al., 1999). They are concerned primarily with between-group inequalities, in accordance with the compelling idea that already disadvantaged groups – such as low-income and racial minorities – should not in addition face environmental disadvantages (Boyce et al., 2016). On a methodological level, these analyses are mainly based on between-group comparisons of means or medians, bivariate correlations and multivariate regressions (Mitchell and Walker, 2005).

In an urban context, such environmental inequalities are likely to be linked to residential segregation, i.e. the spatial separation of population groups between urban neighbourhoods. This multidimensional phenomenon has been well conceptualized by Massey and Denton (1988) and its measurement has been abundantly and thoroughly discussed by sociologists, demographers, geographers and economists for the past half-century and more (e.g., Tivadar, *forthcoming*; Reardon and O'Sullivan, 2004; Wong, 1993; Morrill, 1991; White, 1983; Duncan and Duncan, 1955a). Surprisingly enough, however, this rich literature has not yet been considered a source of insights for conceptualizing and measuring environmental inequalities.

In this paper, we take advantage of this opportunity and suggest 'segregation-based' environmental inequality indices, measuring spatial distributions of social groups and environmental phenomena relative to one another. We propose the Environmental Dissimilarity gap index (ΔED_K) to analyse areal-level environmental data (such as vegetation cover or pollution loads in census blocks) and the Environmental Centralization (EC_d) index to analyse multiple-points environmental data (such as geocoded hazardous sites or urban parks). In addition, following a recommendation made by Sheppard et al. (1999), we suggest a randomization strategy based on Monte Carlo experiments to make robust distribution-free environmental inequality assessments at a city-wide scale.

The recent residential segregation literature points out that more attention needs to be paid to local areas if the ultimate goal is to contribute to public policies development (Brown and Chung, 2006; Folch and Rey, 2016). Accordingly, we also propose a procedure that allows identifying and mapping hotspots that have a large influence on environmental inequalities. This method uses Jackknife simulations to identify spatial units whose removal would result in significant decreases/increases in the values of ΔED_K or EC_d. This approach may help urban policy-makers target EJ policies better. It could inform public decision about, for instance, where to implement greening policies, or priority actions to protect people from hazards or pollutants, or social housing policies, etc.

To illustrate our proposals, we provide a case study with respect to spatial distributions of vegetation (i.e., areal-level data) and dangerous industrial sites (i.e., point data) in Grenoble-Alpes Métropole, France. We examine segregation-based environmental inequalities between low-income households and other households, mobilizing gridded residential data provided at a very fine spatial scale.

The next section presents the related literature and our conceptual framework. Section 3 introduces methods for global and local environmental inequality analyses. Section 4 presents our case study. Section 5 concludes.

2. Related Literature and Conceptual Framework

2.1. Environmental Justice Studies

From its beginnings in the late 1970s to the present, Environmental Justice (EJ) has been both an area of academic research and the banner of a civil movement calling for policies to address inequalities in environmental conditions. Originally focused on racial and social inequalities in the distribution of toxics and hazardous waste in the United States, it has since continuously expanded its thematic and geographical scope (Schlosberg, 2013). Inequalities with respect to environmental 'bads' (nuisances or risks) have remained a prominent topic (see, for example, Hajat et al., 2015, for a review on air pollutants), but attention has also been paid to environmental 'goods' (see, e.g., Jennings et al., 2012, on green spaces). In addition, the field has shifted from documenting inequalities to analysing the underlying reasons for these inequalities (Timmons et al., 2018; Mohai et al., 2009). It has also moved from a conception of distributive equity to a more pluralistic conception of justice, including issues of recognition, participation, capabilities, community justice and justice beyond the humans (Schlosberg, 2013). In parallel, EJ literature and movement have spread out in various countries, with transnational links and consideration of global issues, such as climate justice (Mohai et al., 2009; Schlosberg and Collins, 2014).

Today, we can consider that EJ refers to four different types of environmental inequalities (Laurent, 2011): (i) exposure and access inequalities, i.e. unequal distributions of environmental bads and goods between individuals and groups; (ii) policy-effect inequalities, i.e. unequal effects of environmental policies; (iii) policy-making inequalities, i.e. unequal access to environmental policy-making; and (iv) impact inequalities, i.e. unequal environmental impacts of different individuals and social groups. In this respect, our study falls within the first and oldest strand of the EJ literature, about exposure and access inequalities, and is underpinned by a distributive conception of justice such as the above-mentioned one of Harner et al. (2002). More specifically, it focuses on measuring inequalities between social groups in urban contexts. We present here relevant examples of urban EJ studies, which provide an overview of the most standard methods used for this analysis.

A first example is Carrier et al. (2014) on air pollutants in Montreal, Canada. They consider three statistical methods widely used in the EJ literature to analyse environmental inequalities with respect to visible minorities, low-income individuals, young and elderly people. First, for each social group, they compute weighted averages of several pollutant indicators at a small dissemination subdivision (i.e. city block level, the weight of a city block being its share of the total group population), and compare them with t-tests to similar averages obtained for the rest of the population. Secondly, they compute Spearman's correlation coefficients to examine statistical dependencies between rankings of proportions of groups and of pollutant indicators across city blocks. Finally, they perform multivariate regressions with each of the pollution indicators as the dependent variable in each case and proportions of groups as independent variables, controlling for spatial dependencies. Their results for low-income population and visible minorities are consistent with the bulk of EJ studies: they tend to reside in more polluted areas.

Another example is Schwartz et al. (2015), who examine potential inequities in relationship to race/ethnicity and income associated with distributions of urban tree canopy (UTC) in seven U.S. cities. Data are analysed at the Census Block Group level using Spearman's correlations and multivariate regressions. Schwartz et al. (2015, p. 11) stress that the key question – "is UTC cover distributed equally in the cities examined?" – is answered relying on Spearman's correlations. A significant coefficient implies that: "regardless of what drives the pattern, the pattern exists". This method provides a baseline picture that is comparable across all cities. Multivariate regressions help to get closer to causation and answer supplementary questions: "what other variables drive the distribution of UTC cover?" and "do the data have significant spatial structure?" With respect to the baseline diagnostic, Schwartz et al. (2015) find a significant positive correlation with income across all cities, but less striking results for races.

The two examples presented above analyse areal-level data, but EJ studies also deal with geocoded environmental data. In an older but still relevant study, Sheppard et al. (1999) examine associations between toxic sites and minority/poor populations in Minneapolis, MN, comparing commonly used spatial coincidence and analytical buffering methods. In the first method, proximate populations are defined as those residing in census enumeration units containing toxic sites. In the second, a GIS-based buffer analysis is performed, in which proximate populations are defined as those residing within a predefined distance from a toxic site. These proximity measures being defined, poverty rates of proximate and non-proximate subpopulations – or other socio-demographic or racial characteristics – can be compared.

Sheppard and colleagues examine the sensitivity of their results to proximity measures and buffer distances.[1]

In addition, Sheppard et al. (1999) suggest a methodology for evaluating the statistical significance of their results. The hypothesis investigated is whether the ratios of poverty percentages between proximate and non-proximate subpopulations are large by comparison to what would have been observed if the toxic sites had been placed randomly within the city. They carried out a series of randomization experiments to simulate a hypothetical set of toxic sites distributions. Simulation results indicate that, broadly speaking, observed toxic site locations in Minneapolis were associated with unusually high poverty rates. However, as far we know, Sheppard et al. (1999) have not been followed (except by Chakraborty and Armstrong, 2001) in their recommendation to use randomization strategies for the generation of robust distribution-free environmental inequality assessments.

On exposure to hazards in New York City, Jacobson et al. (2005) review advanced geostatistical techniques and discuss several inequity measures. As an alternative to analytical buffering, they adopt a distance-decay modelling approach (or gravity model) to measure exposure gradients to highways within ethnic groups at the block group level. They then define and advocate a class of inequity measures that equates statistical (or conditional) independence between exposure and demographics as 'perfect equity', with degree of inequity computed as degree of departure from independence. They underline that standard regression approaches (either bivariate or multivariate) look only at departures from zero correlations, which is different from independence. On the other hand, the *Theil* index, commonly used to study income inequality, is shown to be relevant to comparing exposure inequality under different grouping frameworks (e.g. race vs class). But again, this index is not grounded on the independence assumption. Jacobson et al. (2005) thus propose a graphical approach allowing a direct comparison of the empirical joint distribution of exposure and ethnic group and the one implied by the conditional independence model (conditioned on block groups' median income). These plots show that Hispanics and Asians are more exposed than average to highways, whatever the income level.

---

[1] On GIS-based analysis, see also, e.g., Harner et al. (2002), who test several methods for comparing subpopulations in at-risks and not-at-risks zones, or conversely the concentration of risks in disadvantaged areas and elsewhere; or Maantay (2002) and Maantay (2007), who consider issues of optimal study area, spatial resolution, data aggregation, data deficiencies, areal extent of exposure and areal interpolation; or Chakraborty et al. (2011), on similar topics and on emerging geostatistical techniques.

The above-mentioned studies provide an overview of methods used to analyse between-group environmental inequalities. In their review of statistical techniques used in EJ studies, Mitchell and Walker (2007) note that linear regression is by far the most popular method.[2] Although there is a long tradition of using inequality indices to study income and wealth distributions, few EJ studies have explored their relevance for analysing horizontal environmental inequalities, Jacobson et al. (2005) and Lopez (2002) being rare exceptions in this line.[3] On the opposite, inequality indices have been widely mobilized and adapted for studying residential segregation, another kind of between-group (spatial) inequality.

2.2. From Residential Segregation to Segregation-based Environmental Inequalities

Residential segregation refers to the geographic separation of social groups, usually in an urban context. According to Park and Kwan (2017), it was only in the early 21st century that some scholars have sought to understand the association between segregation and environmental inequalities. This association has been investigated primarily with respect to air pollution in U.S. metropolitan areas. Several studies have confirmed that increased segregation tends to be associated with increased racial inequality in exposure to health risks (e.g., Morello-Frosh and Lopez, 2006; Lopez, 2002), but more ambiguous results have sometimes been obtained (Downey al., 2008). More recently, Saporito and Casey (2015) have investigated residential segregation and differences in exposure to green space in U.S. metropolitan areas. Findings show that lower-income people and members of minority groups live in neighbourhoods with much less vegetation than their wealthier, white counterparts, and these differences are exacerbated in racially and economically segregated cities.

These studies have examined statistical links between residential segregation and environmental inequalities. Yet they have not reconsidered how horizontal environmental inequalities are defined and measured and the conceptual and methodological perspectives that the segregation literature could bring to this issue.

The conceptual framework of our methodological proposals is as follows:

- Environmental (dis-)amenities are defined as place-based environmental attributes that provide *local* (dis-)services to people, where geographical proximity between

[2] But referring to Bowen (2002), they underline that few studies adequately report on diagnostic tests (non-linearity, multi-collinearity, heteroscedasticity, etc.) so that the adequacy of this approach is questionable.

[3] Lopez (2002) adapts the net difference score, based upon cumulative frequency distributions, to measure Black/White inequality in exposure to air toxics. Inequality indices have also been suggested to measure vertical (i.e., inter-individual) environmental inequalities (Boyce et al., 2016).

attributes and people enhances the amounts of (dis-)services delivered (Schaeffer and Dissart, 2018).

- We name 'Environmental Segregation' the geographic separation between a social group and an environmental (dis-)amenity. The more this group is segregated from an amenity, the less likely it is to benefit from it. The more it is segregated from a disamenity, the less likely it is to be harmed by it.
- A 'Segregation-based Environmental Inequality' is then a difference between two social groups as for their respective degrees of environmental segregation. A social group is disadvantaged, relative to another group, when it is more segregated from an environmental amenity or less segregated from an environmental disamenity.

Residential segregation is classically conceptualized and measured along five dimensions identified by Massey and Denton (1988): evenness, exposure, concentration, centralization, and clustering. Evenness refers to inequalities in the distribution of population groups between neighbourhoods. Exposure (or isolation) captures opportunities for contacts between members of different (or similar) groups within neighbourhoods. Concentration refers to inequalities in regard to the physical space occupied by groups. Centralization describes the distribution of groups around a city centre. Finally, clustering consider the proximity of groups within and across neighbourhoods.

Some of these dimensions are interrelated and several simplifications of this typology have been suggested (Reardon and O'Sullivan, 2004; Brown and Chung, 2006; Wong, 2008). To study segregation-based environmental inequalities, we believe two dimensions are particularly relevant: evenness (incorporating clustering and concentration notions) and centralization.

Evenness is the long-standing dominant dimension in segregation analysis. To date, the most standard measures of segregation still are Duncan and Duncan's (1955a, 1955b) Dissimilarity (*D*) and Segregation (*IS*) indices. Yet, in the early 80s, these indices have been criticized for not taking account of local spatial interactions between social groups (White, 1983, Morrill, 1991, Wong, 1993). In an urban residential context, it is indeed obvious that people interact with one another across neighbourhoods' boundaries, so that the geographical configuration of the neighbourhoods matters. Morrill (1991), Wong (1993) and other scholars have thus proposed 'spatial evenness' indices (i.e., adjusted *D* indices) that incorporate the clustering dimension in an evenness framework. Concentration is also strongly connected to evenness.

Duncan and Duncan (1961) have proposed the *Delta* index, which is mathematically equivalent to the *D* index: the latter measures the extent to which two population groups' distributions among spatial units differ, while the former measures the extent to which the distribution of a population group differs from the one of land.

Centralization has been especially studied in relation to the 'white flight' hypothesis (Crowder, 2000): the idea that large and growing populations of Blacks spurred Whites to leave urban neighbourhoods in which they would have otherwise remained. But this dimension has been deemed less and less important due to the emergence of sprawled and polycentric cities (Brown and Chung, 2006, Wong, 2008), and the gentrification of many central districts (Hwang and Lin, 2016). Centralization measures have also been considered weaker as they require defining a city 'centre', which is not a straightforward geographical feature (Folch and Rey, 2016). Recently, however, scholars have shown that the standard Relative Centralization index (*RCE*) – originally proposed by Duncan and Duncan (1955b) to study centralization at a city-wide scale – could easily be generalized to a polycentric context (Tivadar, *forthcoming*), or recast into a local centralization index useful for exploring local segregation around any relevant reference location (Folch and Rey, 2016).

Hence, we propose: (i) to adapt *D* and its spatialized versions to the measurement of inequalities related to areal-level environmental data (such as vegetation cover or pollution loads), and (ii) to adapt *RCE* and its local version to the measurement of inequalities related to multiple points environmental data (such as urban parks or hazardous sites). The following section presents these proposals. It also suggests a method based on randomization experiments to perform global (i.e. city-wide) environmental inequality assessments, and a procedure that allows identifying and mapping hotspots that have a large influence on segregation-based environmental inequalities.

3. Measuring Environmental Inequalities

3.1. Segregation-based Environmental Inequality Indices

3.1.1. Environmental Dissimilarity

Duncan and Duncan's (1955a) Dissimilarity index *D* is one of the most widely used measures of segregation. It measures – for two social groups – departure from the perfectly unsegregated situation, where relative distributions of minority and majority groups across spatial units are similar. It ranges theoretically between 0 and 1 and corresponds to the share of the minority group that would have to change its place of residence – moving from one spatial unit to another – to make the unsegregated situation occur.

$$D^{x,y} = \frac{1}{2}\sum_{i=1}^{n}\left|\frac{x_i}{X} - \frac{y_i}{Y}\right| \quad (1)$$

where n is the number of spatial units, x_i and y_i the population of each group in unit i, and $X = \sum_{i=1}^{n} x_i$ and $Y = \sum_{i=1}^{n} y_i$ are each group total population.

As mentioned above, this framework has already been adapted by Duncan and Duncan (1961) to analyse the relation between the people and an environmental data (the amount of land). The *Delta* index measures the concentration of a given social group, that is, the dissimilarity between its distribution and the one of the land among spatial units.

Following them, we propose that the Environmental Dissimilarity index *ED* measures the dissimilarity between the distribution of a population group and the one of an environmental (dis-)amenity among spatial units. The *Delta* index then corresponds to a particular *ED* index, where land area is supposed to be an environmental amenity.

Formally, *ED* is given by:

$$ED^{x,a} = \frac{1}{2}\sum_{i=1}^{n}\left|\frac{x_i}{X} - \frac{a_i}{A}\right| \quad (2)$$

where a_i is the environmental value in spatial unit i and $A = \sum_{i=1}^{n} a_i$ the total value of the environmental variable.

ED measures – for a social group – the departure from the 'environmentally unsegregated' situation, where the distribution of this group among spatial units is similar to the one of the environmental (dis-)amenity. It ranges between 0 and 1 and can be interpreted as the share of the population group that would have to move to reach the 'environmentally unsegregated' state.

ED has a geometrical interpretation based on an adaptation of the Lorenz curve, standard in economic inequality analysis. Figure 1 presents a Lorenz-like curve, where the horizontal axis shows the cumulative proportion of population x in spatial units ordered by the density of the environmental variable per inhabitant, and the vertical axis shows the cumulative proportion of a. *ED* is the maximum distance between this 'environmental segregation curve' and the diagonal corresponding to the environmentally unsegregated situation. A spatialized *Gini* index applied to environmental segregation could also be defined based on this curve: it would correspond to the grey surface between the curve and the diagonal.

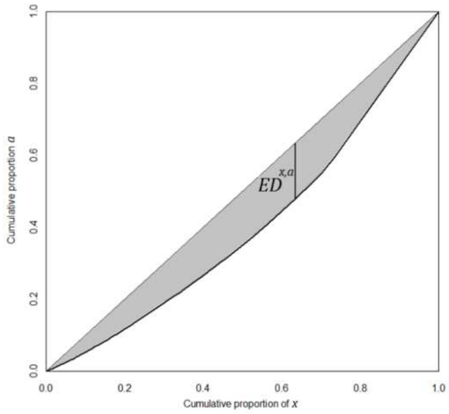

Fig. 1.: Environmental Segregation curve
Notes: The curve indicates, e.g., that 50% of the population X located in the spatial units where the density of the amenity A is the lowest benefits from less than 40% of the total amount of amenity; The *Environmental Dissimilarity* index corresponds to the maximal distance between the diagonal (i.e. the perfectly unsegregated state) and the environmental segregation curve; An environmental *Gini* index would correspond to the grey area between the diagonal and the curve.

The first strength of *ED* is its ease of interpretation. It has a more intuitive meaning than the corresponding *Gini* index, or than the *Theil* index used by Jacobson et al. (2005). The second strength of the *ED* index is that it can be adjusted to account for local spatial interactions.

Morrill (1991) has first developed an adjusted *D* index, where *D* is scaled down when opportunities to interact across adjacent spatial units are present. As it is possible that relevant local interactions go beyond the first order contiguity, the Morrill's adjusted *D* index has been generalized to the k-th order contiguity (Tivadar, *forthcoming*). We can adapt this approach to define an adjusted *ED* index, given by:

$$ED_K^{x,a} = ED_0^{x,a} - \sum_{k=1}^{K} f(k) \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}^{k} ED_{ij}^{x,a}}{\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}^{k}} \quad (3)$$

where $ED_0^{x,a} = ED^{x,a}$, $f(k)$ is a distance-decay function defined by contiguity order *k*, with $f'(k)<0$, $f(1)=1$ and $f(k)_{k\to\infty}=0$. In the empirical case below, we use the usual negative exponential function: $f(k)=\exp(1-k)$. c_{ij}^{k} are the elements of the spatial weights matrix defined as the *k* order contiguity matrix. For each level of contiguity, these elements equal 1 when spatial units *i* and *j* are contiguous of order *k*, 0 otherwise. $ED_{ij}^{x,a}$ captures the potential of interactions across borders between the population and the (dis-)amenity located in two contiguous (of order *k*) spatial units *i* and *j*: the more dissimilar their spatial distributions are, the more local interactions across borders are likely to occur, in the spirit of Morrill's original expression of local interactions between two population groups.[4]

After Morrill, Wong (1993) has developed more refined spatial weights matrices: one where interactions between two contiguous spatial units are proportional to the length of their shared boundary, and another that takes account of spatial units' shapes (i.e., their perimeter/area ratios). These methods can also be transposed to environmental dissimilarity measurement and would be particularly relevant when spatial units have irregular forms (e.g., Census Block Groups), which is not the case in our empirical case.

So far we have presented environmental segregation measures, but we are more specifically interested in between-group environmental inequalities. Thus we define the Environmental Dissimilarity Gap (*ΔED*) to measure – for two social groups – the difference in their degrees of environmental segregation. *ΔED* ranges between -1 and 1 and is given by:

$$\Delta ED^{x,y} = ED^{x,a} - ED^{y,a} = \frac{1}{2}\sum_{i=1}^{n}\left(\left|\frac{x_i}{X}-\frac{a_i}{A}\right| - \left|\frac{y_i}{Y}-\frac{a_i}{A}\right|\right) \quad (4)$$

We also define an adjusted version of the Environmental Dissimilarity gap:

---

[4] Morrill's original interaction term for adjusting the *D* index is based on group proportions in contiguous spatial units: between-group interactions across borders are more likely to occur when spatial units have very distinct shares for the two groups (e.g. one is dominated by black people and the other by white people).

$$\Delta ED_K^{x,y} = ED_K^{x,a} - ED_K^{y,a} \tag{5}$$

which can be rewritten as:

$$\Delta ED_K^{x,y} = \Delta ED_0^{x,y} - \sum_{k=1}^{K} f(k) \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}^k \Delta ED_{ij}^{x,y}}{\sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}^k} \tag{6}$$

where $\Delta ED_0^{x,y} = \Delta ED^{x,y}$.

3.1.2. Environmental Centralization

The Relative Centralization (*RCE*) index has been proposed by Duncan and Duncan (1955b). It allows comparing the locations of two social groups around a point, typically the Central District of a city. It equals 0 when the two groups have similar locations relative to the centre and ranges between -1 and +1 otherwise, the sign indicating which group is closer to this centre. It is given by:

$$RCE^{x,y} = \left(\sum_{i=2}^{n} x_{i-1} y_i\right) - \left(\sum_{i=2}^{n} x_i y_{i-1}\right) \tag{7}$$

where x_i and y_i are ordered by the distance to the city centre. If $RCE^{x,y} > 0$, population *x* is located closer to the centre than population *y*, and conversely if $RCE^{x,y} < 0$.

This index is a particular form of the spatialized Gini index, based on a Lorenz-like curve similar to the one presented on Fig. 1, but where the vertical axis is the cumulative proportion of y and the spatial units are ordered according to their distance to the centre: *RCE* is the surface between the curve and the diagonal.

RCE can be used to compare locations around a specific environmental (dis-)amenity (Folch and Rey, 2016). But EJ studies are generally interested in (dis-)amenities present in multiple locations (e.g., U.S. toxic release inventory sites, as in Harner et al., 2002 or Sheppard et al., 1999). The easiest way of generalizing the *RCE* index to multiple locations is to consider, for each spatial unit, its distance to the closest (dis-)amenity (Tivadar, *forthcoming*). Another option – advantageous when all (dis-)amenities do not have the same importance and/or their impacts are cumulative – could be to consider weighted distances to multiple (dis-)amenities.[5]

Folch and Rey (2016) have defined a local version of *RCE*, by selecting either the k nearest neighbours to the reference centre, or the spatial units that fall within a set distance band around this centre. We propose to recast this method in a polycentric context to obtain a spatially constrained global index. This Environmental Centralization (EC_d) index is thus formally equivalent to *RCE*, but generalized to polycentrism and applied to spatial units located at a distance less than d from their closest environmental (dis-)amenity:

$$EC_d^{x,y} = \left(\sum_{i=2}^{k} x_{i-1} y_i\right) - \left(\sum_{i=2}^{k} x_i y_{i-1}\right) \tag{8}$$

where x_i and y_i are ordered by the distance to the closest environmental (dis-)amenity, and k is the rank of the last spatial unit who respect the spatial constraint: $d_i = \min_a \{d_i^a\} \leq d$. If $EC_d^{x,y} > 0$ population x is located closer to environmental (dis-)amenities than population y, and conversely if $EC_a^{x,y} < 0$.

If we take into account all spatial units, $k = n$, we obtain the unconstrained form of the index $EC_{d_{\max}}^{x,y} = \left(\sum_{i=2}^{n} x_{i-1} y_i\right) - \left(\sum_{i=2}^{n} x_i y_{i-1}\right)$, where $d_{\max} = \max_i \{d_i\}$ is the maximal distance to the closest (dis-)amenity in the study zone.

3.2. Statistical Approaches

3.2.1. Method for Global Analysis

To make a robust global (i.e., city-wide) environmental inequality analysis, one also needs a statistical approach appropriate to spatial data. Following Sheppard et al. (1999), we propose a randomization strategy. The general idea is to test whether the *empirical value* of an environmental inequality index (which measures differences in locations between groups in relation to spatialized environmental attributes) is statistically different from its *expected*

[5] The population-weighted distance approach (Zhang et al., 2015) would be particularly relevant for implementing this option since it uses a measure of spatial interaction (a gravity model incorporating a distance decay function and a factor of importance for amenities) as weights to calculate a weighted distance measure.

value under the null hypothesis that all groups come from the same population (i.e. are distributed in space in a similar way).[6]

In practice, we developed a Monte Carlo permutation test. In a simulation, all households in both groups (poor and non-poor) are randomly assigned to residential locations, and their probability of being assigned to a given location is proportional to the residential capacity of that location (i.e. the actual number of households observed in it).[7] Thus a simulation generates a counterfactual spatial distribution of the population on which an inequality index (ΔED_K or EC_d) can be calculated. We make 499 simulations to obtain a distribution of *simulated values* for this index. Adding the empirical value in this distribution and looking at its position allows computing a (pseudo) p-value and test the null hypothesis.

This statistical approach is consistent with our overall conceptual framework, inspired by the residential segregation literature. Randomizing the spatial distribution of households eliminates every kind of between-group inequalities related to places of residence. It thus provides counterfactual simulations for testing the significance of both residential segregation measures and segregation-based environmental inequalities indices.

3.2.2. Method for Local Analysis

The global environmental inequality analysis provides answers to environmental inequality questions at the study zone level. But policy-makers, to implement EJ policies, must target their interventions on the ground. They have to decide for instance where to implement greening policies, or priority actions to protect people from hazards or pollutants, or social

[6] That is, we randomize the spatial distribution of households rather than the one of amenities or dis-amenities. Sheppard and colleagues test whether their empirical measures of unequal exposition to industrial toxics are statistically different from those obtained when randomizing the locations of toxic installations (i.e. selecting Cartesian coordinates within municipal boundaries). In other words, they test whether the observed spatial distribution of toxic sites is more socially unequal than a random one. One obvious advantage of this strategy is computational since toxic sites are much less numerous than households. However, it is far from realistic to consider that every point in space could have been a possible candidate for a facility (e.g. topographical reasons make some points unsuitable). Instead, toxic facilities could be randomly assigned only to sites that could have been elected (but these latter are generally unknown to the researcher). Randomizing households is an alternative which we believe is much more compelling.

[7] We test whether population groups are distributed unequally *in the existing housing stock* in relation to environmental (dis-)amenities. So the spatial distributions of dwellings and residential areas are considered exogenous. Alternatively, we could have considered counterfactual spatial distributions of dwellings and/or non-developed areas. But we don't know the constraints on the spatial distribution of dwellings and whether non-developed areas are so for choice or because they are improper to be developed. Physical geography, density regulations and growth control policies would have to be taken into account. In that respect, an analysis making endogenous all the space would be much more difficult to justify and implement. We thank an anonymous reviewer for raising this issue.

housing policies, etc. Thus we also provide a statistical approach that allows identifying – and mapping – the local hotspots that have the largest influences on environmental inequalities. This method is based on Jackknife simulations, where the spatial units of the study zone are successively withdrawn.[8] For each simulation, the inequality index (ΔED_K or EC_d) is computed, and we obtain at the end of the process a distribution of simulated index values. Then we look at outliers in the tails of this distribution.[9] Each of these outlier values is attached to a particular simulation, where one spatial unit has been removed from the whole set of spatial units. So we can identify the 'outlier' spatial units whose removal has the greatest impacts on the inequality index. We can map them and thus visualize hotspots for potential policy interventions.

4. Empirical example

This section is not intended to be a comprehensive analysis of environmental inequalities in Grenoble-Alpes Métropole, but aims to illustrate our methodological proposals.

4.1. Background and Data

In France, a Métropole is an administrative entity organized around a large city, in which several municipalities co-operate to plan their development. Grenoble-Alpes Métropole is located at the foot of the French Alps in the South-East of France. It comprises the city of Grenoble and 48 neighbouring municipalities, hosting altogether a population of about 450,000 inhabitants on an area of about 550 km^2. It is an attractive place for students and highly-skilled workers, thanks to its renowned university and advantageous labor market. The unemployment rate is low and job opportunities are numerous in international high-tech companies and more generally in a very dynamic industrial sector. Households are thus on average wealthier and more educated than in most French urban areas of comparable size. However, this generally favourable situation does not come without strong income inequalities and residential segregation dynamics. Indeed, Grenoble-Alpes Métropole is also known for its deprived urban districts, with a high concentration of poor, unemployed, low-skilled and immigrant people, and its hilly suburbs, where affluent households flock.

[8] The number of simulations is thus equal to the number of spatial units.

[9] Several approaches can be used to define outlier values. In the example presented below, outliers are obtained using the boxplot method: they are values that deviate from the mean of the simulated distribution by two standard deviations or more. Other standard techniques that could be employed are the different scores methods (normal, t Student and chi-squared scores) and the median absolute deviation method.

In this respect, it is a place where issues of environmental equity could be quite acute. Specifically, some industrial sites pose risks to the people living in their vicinity, and the exposition to these risks might be socially unequal due to residential segregation. Another potential problem concerns access to vegetation and the resulting health benefits. Although there are underprivileged neighborhoods with large urban parks, disadvantaged populations tend to live in the central part of the Métropole, which has little vegetation in general, while richer populations tend to live in the green urban periphery.

The study uses gridded socio-economic data provided by the National Institute of Statistics and Economic Studies (Insee), based on administrative fiscal data for 2010. For each inhabited cell of a rectangular 200x200m grid, we have estimations of both the total number of households and the number of low-income households. Low-income households are those whose income per consumption unit (before taxes and benefits) is less than 60% of the median of Metropolitan France distribution (i.e. 11,249.40 euros). In what follows, for the sake of simplicity, low-income households will often be referred to as 'poor' and the other households as 'non-poor'. The right panel of Figure 2 presents the share of low-income households in the 4,342 inhabited cells of the study area. The 'holes' in the map, covering about two thirds of the total area of the Métropole, correspond to uninhabited areas occupied either by infrastructures, industrial and commercial zones, water bodies and agricultural fields, or – mainly – by forests and natural areas located on mountains.

As an example of areal environmental data, we consider vegetation cover. Each pixel of 5x5m located in the study zone is classified as vegetated or not, depending on the value of a Normalized Difference Vegetation Index (NDVI) obtained by remote sensing (RapidEye images, 2010). A pixel is considered vegetated for NDVI values greater than 0.35. Then we count the number of vegetated pixels in each inhabited cell and thus obtain a proxy for the vegetation cover. In this example, we focus on the distribution of vegetation across inhabited cells. The vegetation of the uninhabited part of the study zone also provides services to the population and would be important to consider in further research.

As for point data, we consider hazardous industrial sites located in the metropolitan area. This geocoded data is provided by the French Ministry of Ecology 2016, as an application of the European Seveso-III Directive (Directive 2012/18/EU) on Technological Disaster Risk Reduction. It provides the location of industrial establishments where dangerous substances are used or stored in large quantities, with either a low or a high threshold of risk.

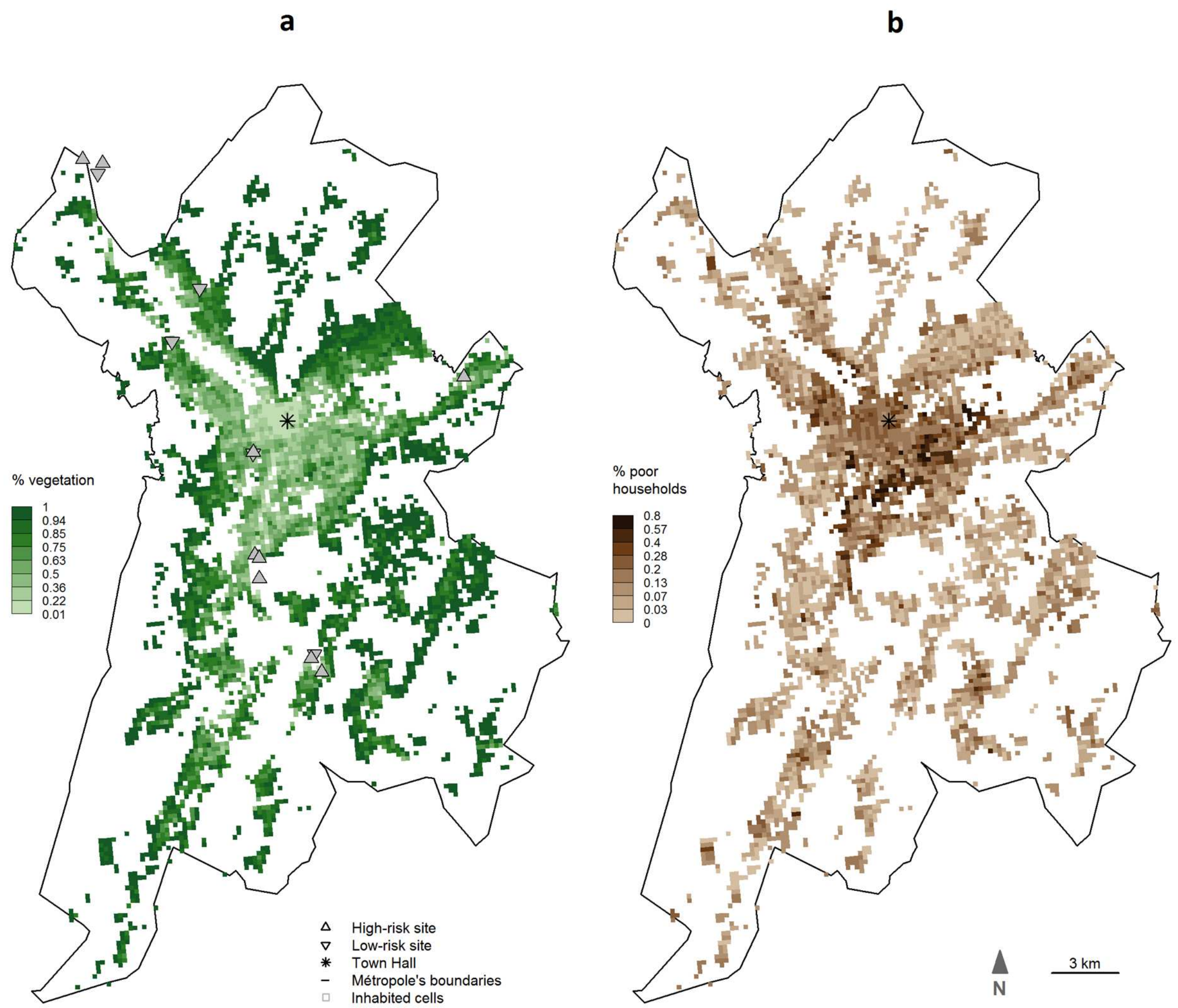

Fig. 2. Vegetation cover, industrial hazards and low-income households
Sources: Own treatments based on French tax database Insee RFL, 2010, RapidEye, 2010, French Ministry of Ecology Seveso database 2016.
Notes: Variables are classified based on Jenks natural breaks.

Table 1 presents basic descriptive statistics for socio-economic and environmental data. Overall, there are 184,485 households and 18% of low-income households in our study zone, which comprises 4,342 inhabited cells. The mean value of the number of households is quite low (42.5) and the one of the vegetation cover is quite high (3.1 ha) due to the high number of cells located in outer suburbs or peri-urban areas. The standard deviation of the number of poor households is very high (21.9) relative to the mean (7.7), as a result of their strong concentration in cells of the urban core and the inner suburbs (see Figure 2).

Table 1
Basic statistics for environmental and socio-economic data

	Grid cell level				Study zone level
	Mean	SD	Min	Max	
Households [a,d]	42.5	10.9	0.3	684	184,485
Low-income households [a,d]	7.7	21.9	0	385	33,537
Low-income households (%) [a,d]	10.5	10.9	0	0.8	18.2
Vegetation (ha) [b,e]	3.1	0.9	0.03	4	13,271
Vegetation (%) [b,e]	76.5	23.4	0.01	100	76.5
Number of dangerous industrial sites [c]					14
Number of high-risk industrial sites [c]					9
Distance to the closest industrial site (km) [c]	4	2.6	0	17.8	
Distance to the closest high-risk site (km) [c]	4.5	2.6	0	17.8	

Sources: Own calculations based on: [a] French tax database Insee RFL, 2010; [b] RapidEye, 2010 (treatments L. Martinez); [c] French Ministry of Ecology Seveso database 2016.
Notes: [d] These proxies provided by Insee accept values below 1; shares of low-income households are bounded to 0.8 for confidentiality reason (see insee.fr/fr/statistiques/2520034#documentation); [e] A pixel is considered vegetated for NDVI values greater than 0.35.

4.2. Results

4.2.1. Global Dissimilarity Analysis for the Vegetation Cover

The first empirical case considers the environmental segregation of poor and non-poor households with respect to vegetation. Figure 3 shows the environmental segregation curves corresponding to ED_0 indices for poor and non-poor households, and Table 2 (and Figure 4) displays values of ED_K and ΔED_K for distinct K. Obviously, as they are always many more people and much less vegetation in the centre of an urban region than in its fringe, obtaining low values for ED_K indices would be rather odd. Indeed, both poor and non-poor households appear to be highly segregated from green spaces: ED_0 values (respectively 0.79 and 0.64) indicate that more than a half of them would have to be displaced to reach the unsegregated state (the diagonal line on Fig. 3). Regarding inequalities, and as might be expected given the local context (see 4.1), these values show that the poor are more segregated from green spaces than the non-poor, with a difference of 15 points in the shares of households to be displaced.

Incorporating local interactions, the environmental segregation of poor and non-poor households decrease strongly (from K=0 to 4) and then stabilize as the spatial scale of the adjustment widens (for K ≥ 5).[10,11] Since poor people are more segregated from green spaces than non-poor people at the residence cell scale (without spatial adjustment), we expect that the former benefit more than the latter from local interactions with neighbouring cells. Indeed, the absolute decrease of ED_K with K is larger for poor than for non-poor households: ΔED_K decreases strongly and then stabilize around 0.05. In sum, there are opportunities of local interactions with green spaces in our study zone, which contribute to reducing the environmental inequality, but that do not make it disappear.

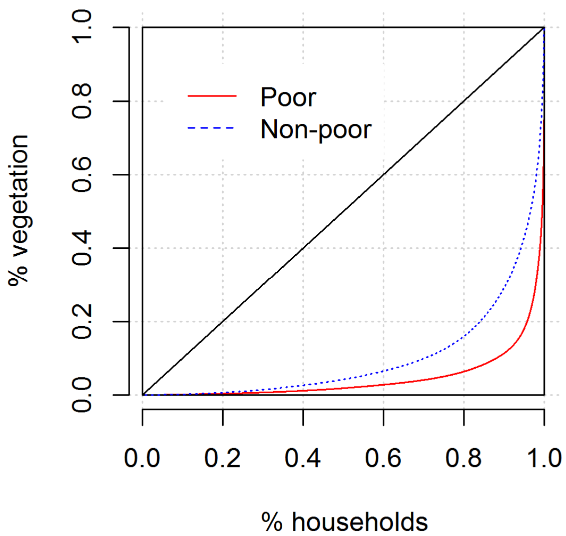

Fig. 3. Environmental Segregation curve for vegetation cover
Notes: e.g., the 80% of the poor households living in the less vegetated grid cells benefits from less than 10% of the total amount of vegetation (as indicated by the red curve), whereas the same proportion of non-poor households benefits from nearly 20% (as indicated by the blue curve).

[10] Adjusted-*ED* indices equal the *ED* index less the mean value of the local spatial interactions. The fact that ED_1 is much lower than ED_0 means there are strong local interactions: indeed, many densely inhabited cells with low greenery are neighbouring low density cells with high greenery, especially along the urban fringe.
[11] Interactions are weighted by a distance decay function, which explains the stabilization of ED_K values when K is increased from 5 to 10. Interactions at the firth, second, third, fourth and fifth order contiguity are weighted respectively by a factor of 1, 0.37, 0.14, 0.05, 0.02 and 0.01, and weights beyond the fifth contiguity are very close to zero. The choice of a negative exponential in this empirical example reflects the assumption that proximity is of great importance for people's interactions with the greenery. In a comprehensive analysis, the choice of the distance decay function and its implications should be considered more in depth.

Table 2

Global dissimilarity analysis for vegetation cover: observed and simulated indices

		Empirical ED_K	Simulated ED_K (mean)	Rank [a]	Pseudo P-value
K=0	Poor	0.789	0.67	500	0.002
	Non-poor	0.642	0.664	1	0.002
	Difference	0. 147	0.006	499	0.004
K=1	Poor	0.498	0.401	500	0.002
	Non-poor	0.411	0.411	340	0.322
	Difference	0.087	-0.01	500	0.002
K=2	Poor	0.377	0.292	500	0.002
	Non-poor	0.316	0.308	500	0.002
	Difference	0.061	-0.016	500	0.002
K=3	Poor	0.330	0.251	500	0.002
	Non-poor	0.278	0.269	500	0.002
	Difference	0.052	-0.017	500	0.002
K=4	Poor	0.313	0.236	500	0.002
	Non-poor	0.265	0.254	500	0.002
	Difference	0.048	-0.018	500	0.002
K=5	Poor	0.306	0.23	500	0.002
	Non-poor	0.259	0.248	500	0.002
	Difference	0.047	-0.018	500	0.002

Notes: [a] The rank is the position of the observed index value in the distribution of simulated values.

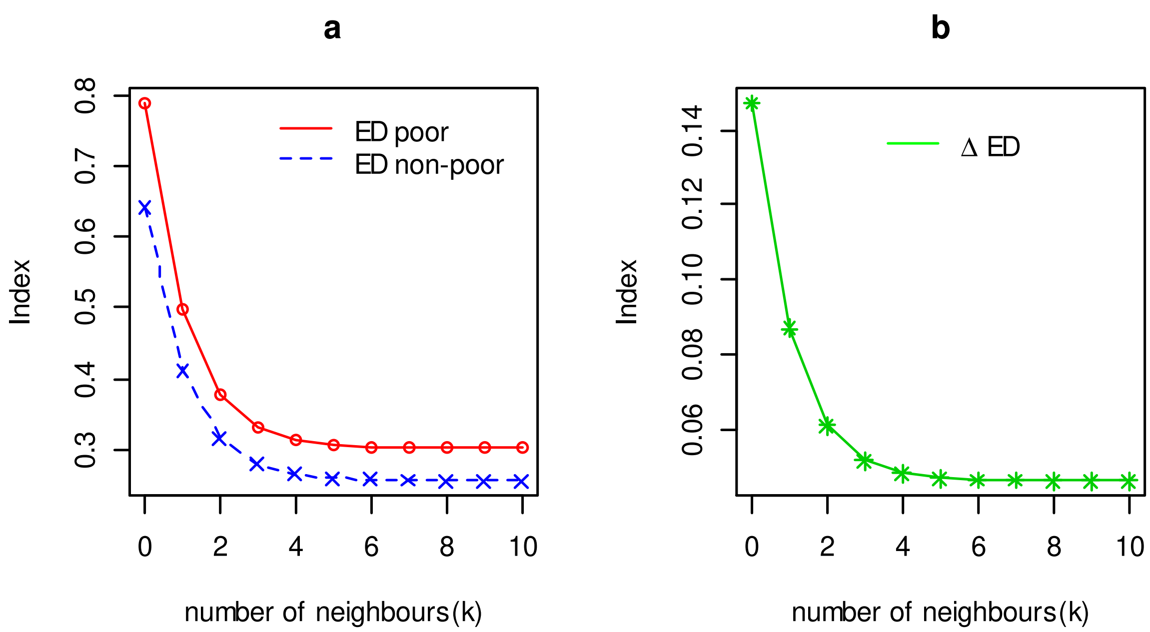

Fig. 4. Environmental Dissimilarity indices for vegetation cover

To reach robust conclusions, we must check that the observed differences between the ED_K values of the poor and the non-poor do not reflect random differences in the spatial distributions of the two groups. Table 2 provides pseudo p-values for ED_K and ΔED_K indices, computed based on 499 Monte Carlo spatially constrained random draws (see section 3.2.1 for details). It shows that all but one value is highly statistically significant. Most importantly, the environmental segregation of poor households is always significantly higher in the empirical case than in the random scenario, as well as the degree of environmental inequality between poor and non-poor households. The only non-significant value is the ED_1 index computed for the non-poor (see below for more explanations).

Since the results presented in Table 2 are qualitatively equivalent for K=2 to 5, we comment more on K=0, K=1 and K=5. Figure 5 presents the position of the empirical value in the distribution of simulated values for the ED_K indices of the poor (first column) and the non-poor (second column), and for ΔED_K (third column), either without adjustments for local spatial interactions (first row) or with adjustments at the 2nd (second row) or 5th contiguity order (third row).

As expected, on subfigures *a, d, g*, empirical ED_K values are much larger than simulated values: poor households always have a highly significant degree of environment segregation: non-random factors push them away from green spaces. Similarly, on subfigures *c, f, i*, environmental inequality, as measured by ΔED_K, is always positive and statistically significant. The poor are always more environmentally segregated than the non-poor.

Results concerning non-poor households require further reflection. The subfigure *b* reveals that non-poor households are significantly *less* environmentally segregated than when allocated randomly: there are non-random factors that make the non-poor choose greener locations. But at the 1st order contiguity (subfigure *e*), the empirical ED_1 index of non-poor is in the middle of the simulated distribution: they are segregated ‘just as expected’ in the random case. Lastly, at the 5th order (subfigure *h*), non-poor households are significantly *more* environmentally segregated than expected (and the graphic would be equivalent for all $K \geq 2$).

Here we should recall that the adjusted index equals the unadjusted index less the mean value of the local interactions. This implies that simulated local interactions for non-poor households are larger than observed local interactions for K=1, and even more for K=5: there are *fewer local interactions* than would be in a random scenario. Our interpretation is as follows: non-poor households have fewer local interactions with the vegetation than when allocated randomly – with probabilities based on housing stocks – because they are more dispersed than the dwellings: they are disproportionately more represented in low-density areas and less in dense urban neighbourhoods. Since local interactions are maximal between high-density highly artificialized neighbourhoods and low-density heavily vegetated ones and minimal between neighbourhoods of similar population density and vegetation cover, a greater dispersion of non-poor households is associated with fewer local interactions.

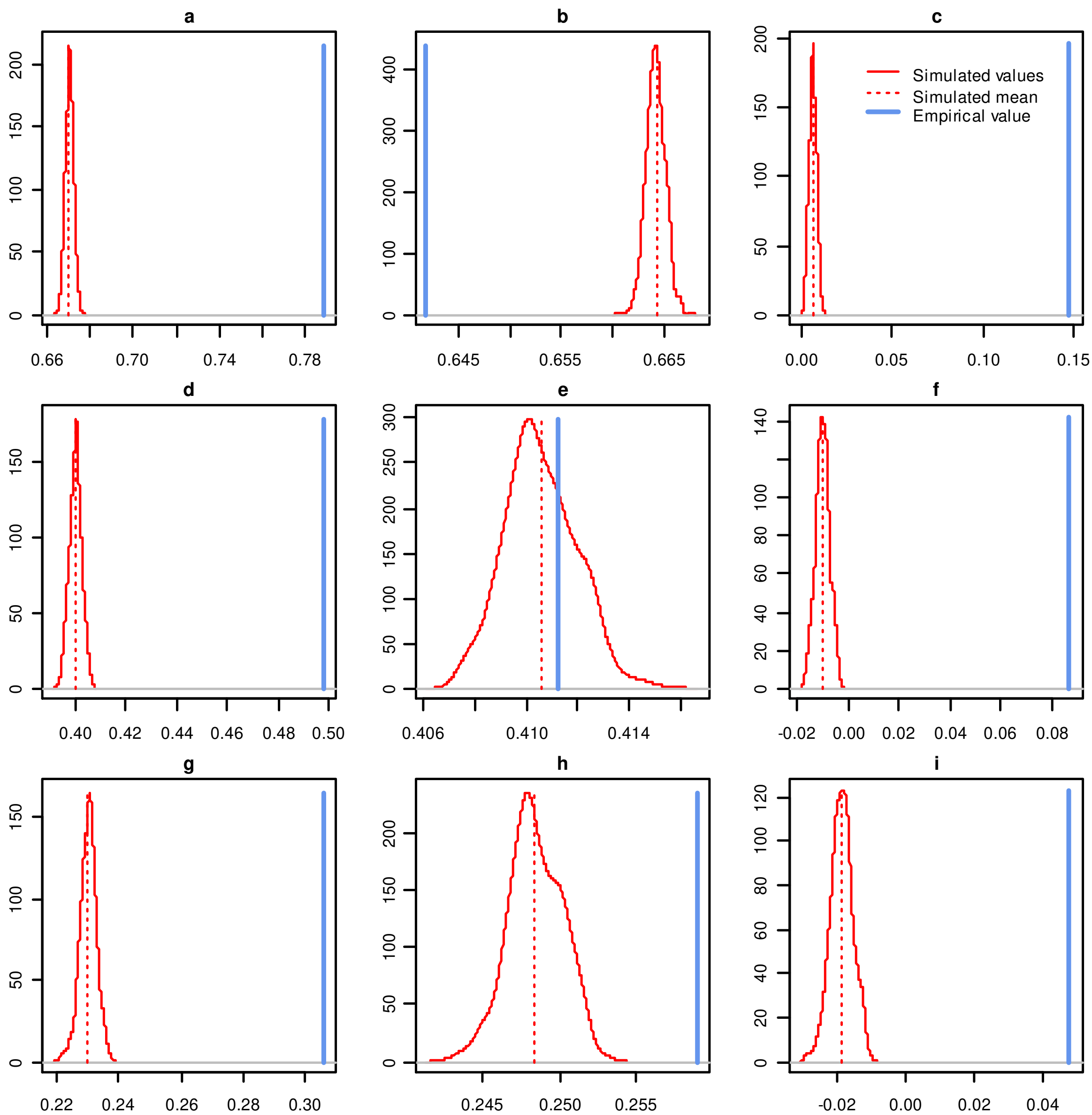

Fig. 5. Global dissimilarity analysis for the vegetation cover: empirical and simulated indices
Notes: Each subfigure provides the probability density function for simulated values, the mean of the simulated values and the empirical value of the index, for *ED* indices computed for the poor (1st column), the non-poor (2nd column), and for ΔED (3rd column), either without adjustments for local spatial interactions (1st row) or with adjustments at the 2nd (2nd row) or 5th contiguity order (3rd row).

Non-poor households are thus segregated 'just as expected' for K=1 and more segregated than in a random scenario for K≥2, but in any case they remain less environmentally segregated than the poor: the significant (non-random) environmental segregation of the poor always generates a significant (non-random) environmental inequality with the non-poor.

4.2.2. Local Dissimilarity Analysis for the Vegetation Cover

The global analysis has confirmed that the poor are more segregated from green spaces than the non-poor. Against this background, we may consider alternatively two political goals: (i) the first focuses on the poor and aims to bring closer together the poor and the vegetation cover; (ii) the second focuses on reducing the environmental inequality between the poor and the non-poor. The former is focused on the absolute situation of the poor, and the second on their relative situation compared to the non-poor. In accordance with these goals, the local analysis may focus on identifying: (i) the hotspots having the largest influences on the ED_0 of the poor; or (ii) the ones that have the largest influences on ΔED_0.[12]

In line with the first goal, the left panel of Figure 6 shows the box plot of the Jackknife simulations of ED_0 for poor households, and maps the cells whose removal is associated with an important decreasing ED_0 (i.e., corresponding to low outliers on the box plot, see 3.2.2). We will call them 'A' cells. Compared to others, they are much less vegetated and much more populated, with higher shares of low-income households (Table 3). On the map (subfigure *a*), we can distinguish a group of cells (say 'A1' cells, coloured in blue) located in the centre of the city, where vegetation is scarce (see Fig. 2a), and a second group of cells (say 'A2' cells, coloured in red) located mainly in the South-East suburb of the city, where shares of poor are very high (see Fig. 2b).[13]

In line with the second goal, the right panel of Figure 6 shows the box plot corresponding to Jackknife simulations of ΔED_0, and maps the hotspots having the largest positive impacts on the environmental inequality between the poor and the non-poor (i.e., corresponding to low outliers on the box plot). Clearly, this second set of cells – say 'B' cells – is quite different from the former. Overall, B cells appear much more vegetated and less inhabited than A cells (see Table 3). The B set comprises cells (say 'B1' cells, coloured in blue on Fig. 6) with very low shares of poor, mainly located in the peri-urban part of the study zone. They also contain

[12] For space reasons, we perform the local analysis only on non-adjusted *ED* values.

[13] The value of 0.3 retained to highlight red cells with high shares of poor households (vs blue cells with low shares) corresponds to the threshold of high outliers in the distribution of this variable (see Table 1).

cells (say 'B2' cells, coloured in red, cf. footnote 14) that were also part of the A2 group, with high shares of poor and located in the suburbs of Grenoble. But the cells of the A1 group – precisely the less vegetated and the most populated – do not pertain to this new set.

These results are interesting as they show that the two political goals stated above should not be confused and may require different spatial targets. If the political goal is focused on the absolute situation of the poor, it may be worth targeting the urban heart of the city (A1 cells), where the poor are numerous, to introduce more vegetation where possible. But the centre of the city is populated by poor as well as non-poor households, so that this targeting may not be the most relevant to reduce the environmental inequality. If the political goal is primarily to reduce this inequality, then targeting more specifically the places where the shares of low-income households are the highest (A2 and B2 cells) may be more effective. Also, the environmental inequality stems from the quasi-absence of poor households in peri-urban communities. Thus, helping more low-income households to settle in these areas (B1 cells), for instance through social housing programmes, may be another relevant political option.

4.2.3. Global Centralization Analysis for Industrial Hazards

The second empirical example concerns disamenities, namely hazardous industrial sites. The values of the unconstrained *EC* index ($EC_{d=dmax}$) are positive (0.036 and 0.99 respectively for all and high risks), showing that the poor are globally more centralized than the non-poor with respect to dangerous sites (Table 4). The Monte Carlo significance tests confirm this environmental inequality: observed values do not reflect random differences in the spatial distributions of the two groups. These results aren't surprising given the local context (see again 4.1): the segregation dynamics that separates poor and non-poor people is likely to be linked to both amenities (e.g., vegetation) and disamenities (e.g. industrial risks) and the presence of environmental inequalities was the expected outcome.

These inequalities can be visualized on Figure 7, which shows the segregation curves for all risks or high risks, where the *EC* index is the area between the curve and the diagonal (see section 3.1.2). For all risks (subfigure *a*), considering spatial units at short distances from dangerous sites (the left side of the graph), either poor or non-poor households can be overrepresented (the curve is successively under or above the diagonal), but considering the whole area, poor households are clearly closer to dangerous sites. For high risks, poor households are unambiguously closer to dangerous sites (the curve is always under the diagonal); the curve deviates more from the diagonal when more distant spatial units are taken into account. We can conclude that the inequality is mainly driven by a strong over-representation of non-poor households in residential locations that are the farthest from industrial hazards.

Table 3
Local dissimilarity analysis for the vegetation cover: descriptive statistics of hotspots

	A cells				B cells			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Households	397.4	135.1	158	684	106.5	126.2	29.7	684
Low-income households	145.7	51	90	385	44.4	79.4	0	385
Low-income households (%)	40.7	17	17	74.2	16.4	25.5	0	79.5
Vegetation (ha)	1.3	0.7	0.1	2.7	3.1	0.7	0.6	3.9
Vegetation (%)	32	17.4	2.5	68.6	71.7	17.7	13.8	98.7

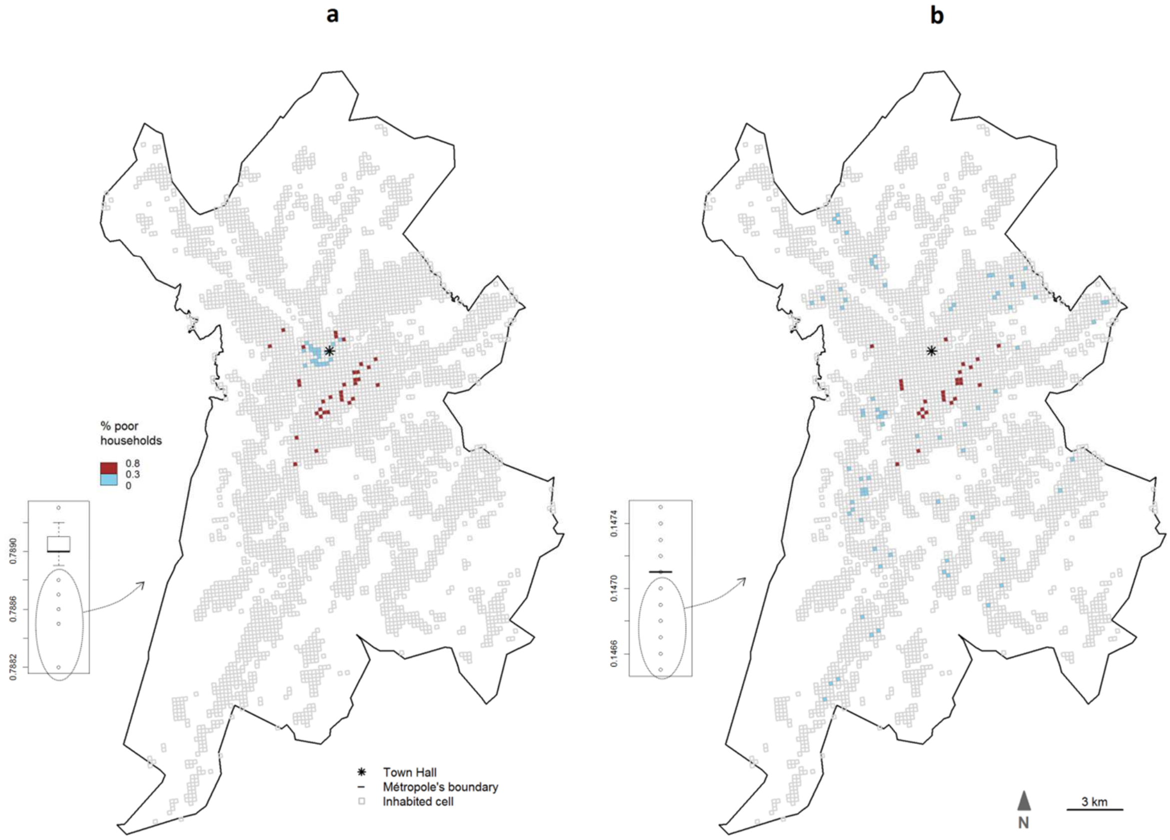

Fig. 6. Local dissimilarity analysis for the vegetation cover: box plots and hotspots

Notes: Subfigure *a* shows the box plot of Jackknife simulations of ED_0 for the poor, and the map of those hotspots whose removal is associated with a lower environmental segregation (i.e., low outliers); subfigure *b* shows the same box plot for ΔED_0 and the map of those hotspots whose removal is associated with a lower environmental inequality (i.e., low outliers). The value of 0.3 retained to highlight cells with high shares of poor households corresponds to the threshold of high outliers (mean+2*SD) in the distribution of this variable in the study zone (see Table 1).

Table 4

Global centralization analysis for industrial hazards: observed and simulated indices

	Distance bound (d)	Empirical EC_d	Simulated EC_d (mean)	Rank [a]	Pseudo P-value
*A*ll dangerous industrial sites	Maximum distance [b]	0.036	0	500	0.002
	15 km	0.036	0	500	0.002
	12 km	0.035	0	500	0.002
	9 km	0.033	0	500	0.002
	6 km	0.01	0	497	0.008
	3 km	-0.066	0	1	0.002
	1 km	0.07	0	500	0.002
High-risk industrial sites only	Maximum distance [b]	0.099	0	500	0.002
	15 km	0.099	0	500	0.002
	12 km	0.098	0	500	0.002
	9 km	0.096	0	500	0.002
	6 km	0.059	0	500	0.002
	3 km	-0.046	0	1	0.002
	1 km	0.097	-0.001	500	0.002

Notes: [a] The rank corresponds to the position of the observed index value in the distribution of simulated values; [b] The maximum distance between a grid cell and a dangerous site in the study zone equals 17.9 km; the corresponding EC_d index covers all spatial units and population comprised within the study zone.

To confirm this conclusion, we now consider spatially constrained EC_d indices, with the spatial scope of the analysis – d – varying from 1 to 15 km by steps of 200m. For each d, the analysis is restricted to spatial units located at a distance less than d from the closest hazardous site. In other words, we make a focus on 'intra-buffers' environmental inequalities. The results obtained (subfigures *c* and *d* on Figure 7 and Table 4, which offers significance tests) show as expected that either one group or the other can be significantly more centralized for small value of d. But as d becomes sufficiently large ($d \geq 9$ km), the constrained indices converge toward the unconstrained ones, and the poor appear significantly more centralized.

The sensitivity for small values of d reflects complex social segregation patterns around and moving away from dangerous sites: spatial units where poor people are overrepresented are followed by spatial units where they are underrepresented and so on; as the spatial scope of the analysis is extended, new rings of cells are included gradually in the calculation of the index, and the relative centralization of poor households decreases (resp. increases) when these new and most distant rings have high (resp. low) shares of poor households. However, non-poor households are unambiguously overrepresented in the spatial units that are the farthest from the dangerous sites, what explains the convergence toward unrestricted values.

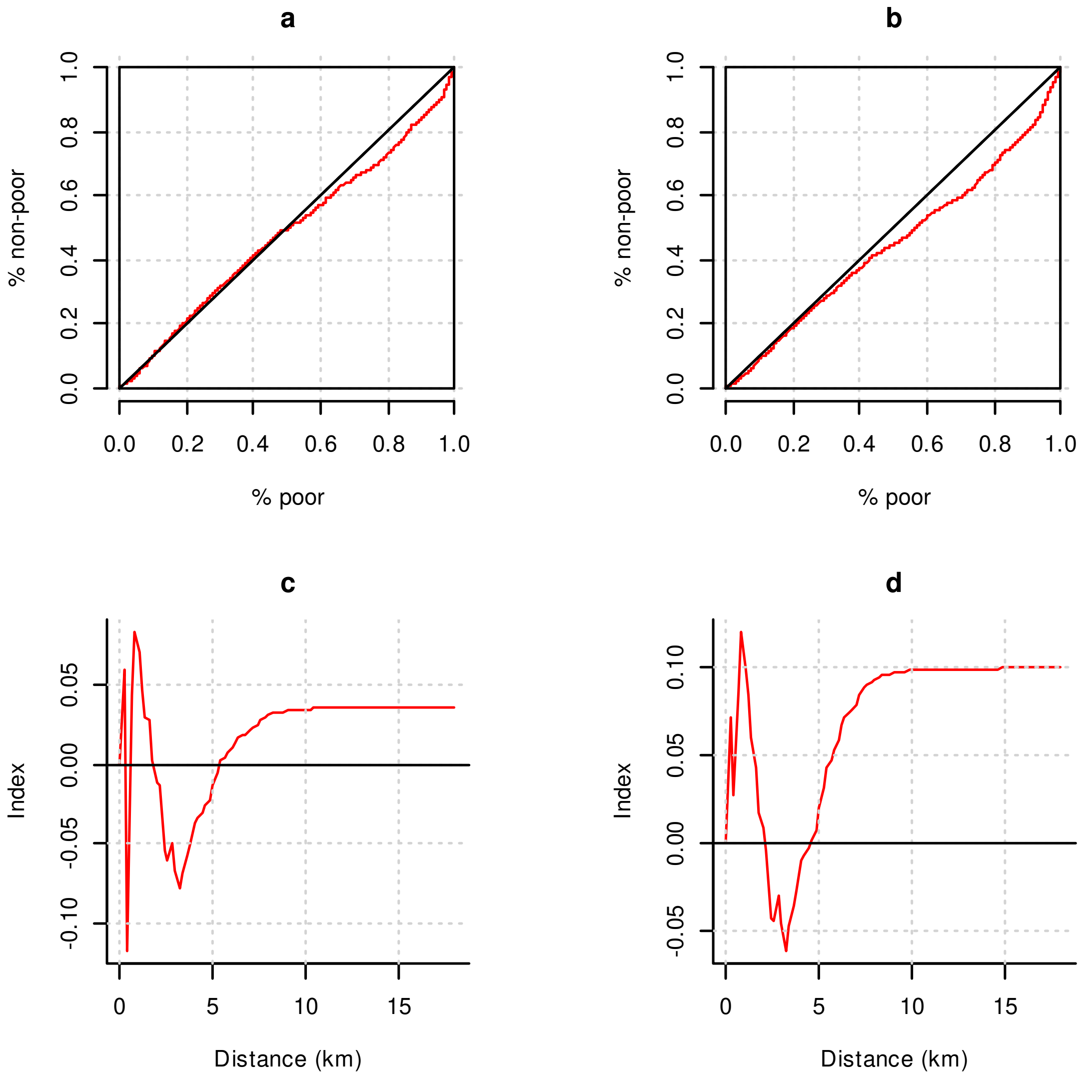

Fig. 7. Global centralization analysis for industrial hazards: graphical approaches
Notes: The left panel refers to all hazardous sites and the right one to very dangerous sites; Subfigure *a* and *b* are Lorenz-like curve: the curve on *b* shows that the spatial units which are the closest from hazards and that house, e.g., 80% of the poor, only house 70% of the non-poor (subfigure *a* reads in the same way); $EC_{d=dmax}$ is the area between the curve and the diagonal; Subfigures *c* and *d* provide the values of EC_d indices calculated by varying *d* in steps of 200m.

4.2.4. Local Centralization Analysis for Industrial Hazards

Since the *EC* index measures a *relative* centralization, it is tailored to inform a policy aiming at reducing the inequality between the poor and the non-poor. The local analysis allows identifying the hotspots that have the largest influences on this inequality. In this example, we focus specifically on the uneven exposure to *very* dangerous industrial sites, without assuming any spatial constraint for the analysis.

Figure 8 shows the box plot of the distribution of simulated *EC* values, and maps the spatial units associated with *low* (map a) and *high* (map b) outlier values of this distribution (see section 3.2.2). The former spatial units have the largest *positive* impacts on the environmental centralization of the poor (i.e., their removal makes *EC* increase), whereas the latter have the largest *negative* impacts (i.e., it makes *EC* decrease).

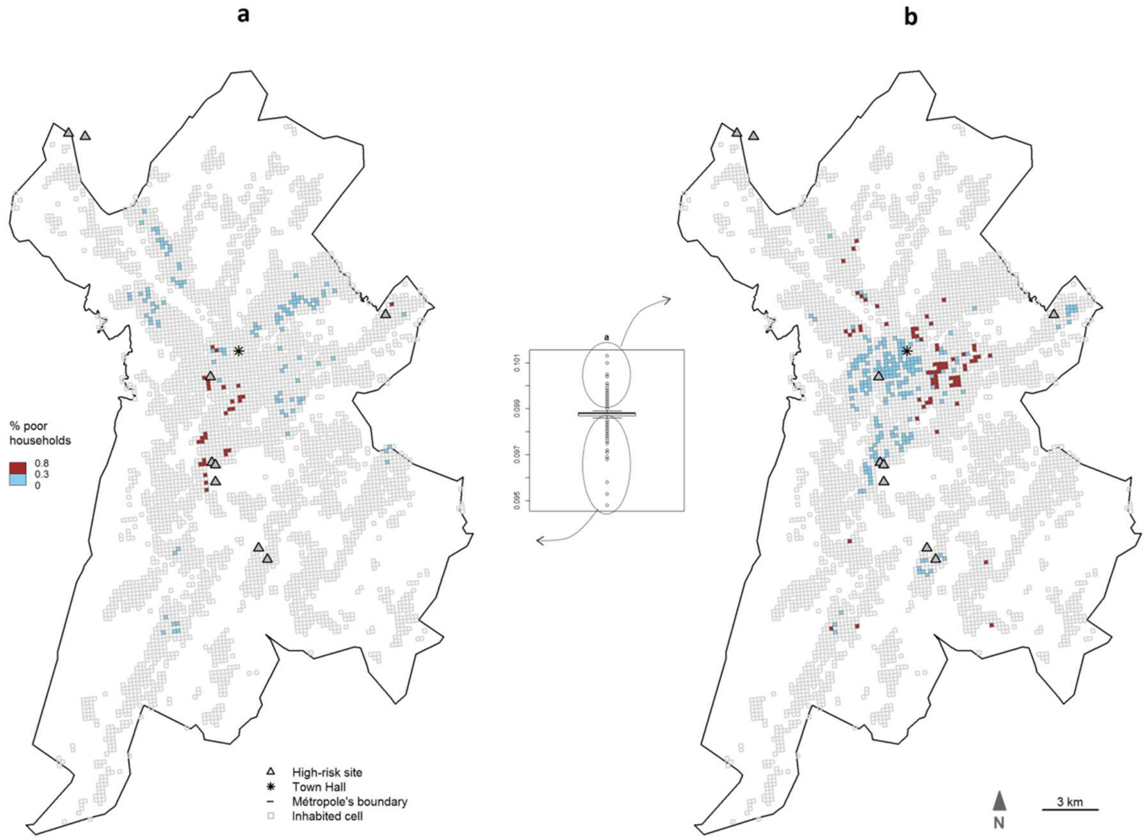

Fig. 8. Local centralization analysis for industrial hazards: box plot and hotspots

*Notes: S*ubfigure *a* shows the map of hotspots whose removal is associated with a lower environmental inequality (i.e. low outliers on the box plot of the Jackknife simulations of *EC* values); subfigure *b* presents the hotspots associated with a higher environmental inequality (i.e. high outliers). The value of 0.3 retained to highlight cells with high shares of poor households corresponds to the threshold of high outliers (mean+2*SD) in the distribution of this variable in the study zone (see Table 1).

On the map *a*, we can distinguish spatial units with higher shares of low-income households (say ‘C’ cells, coloured in red), and others with lower shares (say ‘D’ cells, coloured in blue). As expected, C cells are close to dangerous sites, whereas D cells are quite distant from all sites. This map highlights well the hotspots which make the poor more centralized than the non-poor. It may help policy-makers target their policies: for instance, if a new social housing programme were to be implemented, or if new hazardous facilities were to be installed, then the political goal of reducing environmental inequality with respect to hazards would recommend avoiding C cells and wherever possible choosing locations in D cells.

Conversely, the map *b* shows cells that 'play against' the environmental inequality. Although the poor are globally more centralized than the non-poor with respect to dangerous sites, there are cells nearby high-risk sites where non-poor households are numerous and more distant cells with high shares of poor households. This map illustrates the complexity of the environmental segregation patterns already emphasized in the previous section. It also shows the 'egalitarian' hotspots in the current spatial distribution of households that should not be compromised by inappropriate actions.

5. Conclusion

Robust methods for assessing environmental inequalities are needed to help define and evaluate environment justice policies. We have developed such methods to measure between-group environmental inequalities related to residential segregation patterns. All will soon be available in a dedicated R package called *SegEnvIneq*. This approach is particularly suited to environmental inequality assessments in urban context.

Inspired by the residential segregation literature, we propose the Environmental Dissimilarity index (ΔED_K) and the Environmental Centralization index (EC_d) to analyse 'segregation-based' environmental inequalities. These indices allow working with, respectively, areal-level (e.g., vegetation cover) or multiple-points (e.g., dangerous sites) environmental data. Both indices are genuinely spatial: ΔED_K can incorporate local interactions between people and environment across spatial units' borders, and EC_d can be spatially constrained to better understand segregation patterns around and moving away from environmental centralities. To ensure the robustness of this analysis, we also suggest a distribution-free significance test based on Monte Carlo experiments, relevant for ΔED_K and EC_d indices.

Beyond these useful tools for carrying global (city-wide) analyses, we propose a consistent method to identify local (neighbourhood scale) hotspots. The procedure based on Jackknife simulations highlights the spatial units most responsible for environmental inequalities, as measured by ΔED_K or EC_d.

These methods are applied in Grenoble-Alpes Métropole, France, and reveal inequalities between poor and non-poor households in terms of access to green spaces and exposure to industrial risks. This illustration uses data available at very fine spatial scales (i.e. a 200 m x 200 m grid for population data, and even finer scales for environmental data). To explore the sensitivity of our methods, an extension left for future research would be to compare these results with additional results obtained at coarser and less regular spatial scales (e.g. using population data from French census tracks).

We believe that designing global and local segregation-based approaches to measuring environmental inequalities is a useful step forward in the environmental justice literature. Past research has shown that taking the environment into account help understanding residential segregation dynamics (Wu, 2006; Schaeffer *et al.*, 2016). Further researches are still required to deepen our comprehension of theoretical and empirical relationships between residential segregation and segregation-based environmental inequalities.

Acknowledgments

This work is part of the EFFIJIE project, funded jointly by the French National Research Agency (grant ANR-13-SENV-0001-01) and Irstea. We would like to thank Laurent Martinez and Frédéric Bray for research assistance and the Grenoble Urban Planning Agency for stimulating talks at the beginning of the project and for help with the data. We also thank an anonymous reviewer for relevant suggestions.

References

1. Apparicio, P., Pham, T., Séguin, A., & Dubé, J. (2016). Spatial distribution of vegetation in and around city blocks on the Island of Montreal: A double environmental inequity? *Applied Geography*, *76*, 128-136. doi:10.1016/j.apgeog.2016.09.023
2. Bowen, W. (2002). An Analytical Review of Environmental Justice Research: What Do We Really Know? *Environmental Management*, *29*(1), 3-15. doi:10.1007/s00267-001-0037-8
3. Boyce, J. K., Zwickl, K., & Ash, M. (2016). Measuring environmental inequality. *Ecological Economics*, *124*, 114-123. doi:10.1016/j.ecolecon.2016.01.014
4. Brown, L. A., & Chung, S. (2006). Spatial segregation, segregation indices and the geographical perspective. *Population, Space and Place*, *12*(2), 125-143. doi:10.1002/psp.403
5. Carrier, M., Apparicio, P., Séguin, A., & Crouse, D. (2014). The application of three methods to measure the statistical association between different social groups and the concentration of air pollutants in Montreal: A case of environmental equity. *Transportation Research Part D: Transport and Environment*, *30*, 38-52. doi:10.1016/j.trd.2014.05.001
6. Chakraborty, J., & Armstrong, M. P. (2001). Assessing the Impact of Airborne Toxic Releases on Populations with Special Needs. *The Professional Geographer*, *53*(1), 119-131. doi:10.1111/0033-0124.00274
7. Chakraborty, J., Maantay, J. A., & Brender, J. D. (2011). Disproportionate Proximity to Environmental Health Hazards: Methods, Models, and Measurement. *American Journal of Public Health*, *101*(S1), S27-S36. doi:10.2105/ajph.2010.300109
8. Crowder, K. (2000). The Racial Context of White Mobility: An Individual-Level Assessment of the White Flight Hypothesis. *Social Science Research*, *29*(2), 223-257. doi:10.1006/ssre.1999.0668
9. Downey, L., Dubois, S., Hawkins, B., & Walker, M. (2008). Environmental Inequality in Metropolitan America. *Organization & Environment*, *21*(3), 270-294. doi:10.1177/1086026608321327
10. Duncan, O. D., & Duncan, B. (1955a). A Methodological Analysis of Segregation Indexes. *American Sociological Review*, *20*(2), 210. doi:10.2307/2088328
11. Duncan, O. D., & Duncan, B. (1955b). Residential Distribution and Occupational Stratification. *American Journal of Sociology*, *60*(5), 493-503. doi:10.1086/221609
12. Duncan, O.D., & Duncan, B. (1961). Statistical Geography: Problems in Analyzing Area Data. Free Press, Glencoe, Illinois.
13. Folch, D. C., & Rey, S. J. (2014). The centralization index: A measure of local spatial segregation. *Papers in Regional Science*, *95*(3), 555-576. doi:10.1111/pirs.12145
14. Frey, N. (2016). Equity in the distribution of urban environmental amenities: the case of Washington, D.C. *Urban Geography*, *38*(10), 1534-1549. doi:10.1080/02723638.2016.1238686
15. Hajat, A., Hsia, C., & O'Neill, M. S. (2015). Socioeconomic Disparities and Air Pollution Exposure: a Global Review. *Current Environmental Health Reports*, *2*(4), 440-450. doi:10.1007/s40572-015-0069-5
16. Harner, J., Warner, K., Pierce, J., & Huber, T. (2002). Urban Environmental Justice Indices. *The Professional Geographer*, *54*(3), 318-331. doi:10.1111/0033-0124.00333
17. Hwang, J., & Lin, J. (2016). What Have We Learned About the Causes of Recent Gentrification? *Cityscape*, *18*(3), 9-26. http://www.jstor.org/stable/26328271

18. Jacobson, J. O., Hengartner, N. W., & Louis, T. A. (2005). Inequity Measures for Evaluations of Environmental Justice: A Case Study of Close Proximity to Highways in New York City. *Environment and Planning A*, *37*(1), 21-43. doi:10.1068/a36225
19. Jennings, V., Johnson Gaither, C., & Gragg, R. S. (2012). Promoting Environmental Justice Through Urban Green Space Access: A Synopsis. *Environmental Justice*, *5*(1), 1-7. doi:10.1089/env.2011.0007
20. Landry, S. M., & Chakraborty, J. (2009). Street Trees and Equity: Evaluating the Spatial Distribution of an Urban Amenity. *Environment and Planning A*, *41*(11), 2651-2670. doi:10.1068/a41236
21. Laurent, É. (2011). Issues in environmental justice within the European Union. *Ecological Economics*, *70*(11), 1846-1853. doi:10.1016/j.ecolecon.2011.06.025
22. Lopez, R. (2002). Segregation and Black/White Differences in Exposure to Air Toxics in 1990. *Environmental Health Perspectives*, *110*(s2), 289-295. doi:10.1289/ehp.02110s2289
23. Maantay, J. (2002). Mapping Environmental Injustices: Pitfalls and Potential of Geographic Information Systems in Assessing Environmental Health and Equity. *Environmental Health Perspectives*, *110*(s2), 161-171. doi:10.1289/ehp.02110s2161
24. Maantay, J. (2007). Asthma and air pollution in the Bronx: Methodological and data considerations in using GIS for environmental justice and health research. *Health & Place*, *13*(1), 32-56. doi:10.1016/j.healthplace.2005.09.009
25. Massey, D. S., & Denton, N. A. (1988). The Dimensions of Residential Segregation. *Social Forces*, *67*(2), 281-315. doi:10.1093/sf/67.2.281
26. Mitchell, G., & Walker, G. (2007). Methodological issues in the assessment of environmental equity and environmental justice. *Sustainable Urban Development: The Environmental Assessment Methods*, 2, 447-472.
27. Mohai, P., Pellow, D., & Roberts, J. T. (2009). Environmental Justice. *Annual Review of Environment and Resources*, *34*(1), 405-430. doi:10.1146/annurev-environ-082508-094348
28. Morrill, R. (1991). On the Measure of Geographic Segregation. *Geography Research Forum*, 11, 25-36. http://raphael.geography.ad.bgu.ac.il/ojs/index.php/GRF/article/view/91
29. Morello-Frosch, R., & Lopez, R. (2006). The riskscape and the color line: Examining the role of segregation in environmental health disparities. *Environmental Research*, *102*(2), 181-196. doi:10.1016/j.envres.2006.05.007
30. Park, Y. M., & Kwan, M. P. (2017). Multi-Contextual Segregation and Environmental Justice Research: Toward Fine-Scale Spatiotemporal Approaches. *International Journal of Environmental Research and Public Health*, *14*(10), 1205. doi:10.3390/ijerph14101205
31. Pham, T., Apparicio, P., Séguin, A., Landry, S., & Gagnon, M. (2012). Spatial distribution of vegetation in Montreal: An uneven distribution or environmental inequity? *Landscape and Urban Planning*, *107*(3), 214-224. doi:10.1016/j.landurbplan.2012.06.002
32. Reardon, S. F., & O'Sullivan, D. (2004). Measures of Spatial Segregation. *Sociological Methodology*, *34*(1), 121-162. doi:10.1111/j.0081-1750.2004.00150.x
33. Saporito, S., & Casey, D. (2015). Are There Relationships Among Racial Segregation, Economic Isolation, and Proximity to Green Space? *Human Ecology Review*, *21*(2). doi:10.22459/her.21.02.2015.06

34. Schaeffer, Y., & Dissart, J. (2018). Natural and Environmental Amenities: A Review of Definitions, Measures and Issues. *Ecological Economics*, *146*, 475-496. doi:10.1016/j.ecolecon.2017.12.001
35. Schaeffer, Y., Cremer-Schulte, D., Tartiu, C., & Tivadar, M. (2016). Natural amenity-driven segregation: evidence from location choices in French metropolitan areas. *Ecological Economics*, *130*, 37-52. doi:10.1016/j.ecolecon.2016.05.018
36. Schlosberg, D. (2013). Theorising environmental justice: the expanding sphere of a discourse. *Environmental Politics*, *22*(1), 37-55. doi:10.1080/09644016.2013.755387
37. Schlosberg, D., & Collins, L. B. (2014). From environmental to climate justice: climate change and the discourse of environmental justice. *Wiley Interdisciplinary Reviews: Climate Change*, *5*(3), 359-374. doi:10.1002/wcc.275
38. Schwarz, K., Fragkias, M., Boone, C. G., Zhou, W., McHale, M., Grove, J. M., … Cadenasso, M. L. (2015). Trees Grow on Money: Urban Tree Canopy Cover and Environmental Justice. *PLOS ONE*, *10*(4), e0122051. doi:10.1371/journal.pone.0122051
39. Shanahan, D., Lin, B., Gaston, K., Bush, R., & Fuller, R. (2014). Socio-economic inequalities in access to nature on public and private lands: A case study from Brisbane, Australia. *Landscape and Urban Planning*, *130*, 14-23. doi:10.1016/j.landurbplan.2014.06.005
40. Sheppard, E. R., Leitner, H. E., McMaster, R. B., & Tian, H. O. (1999). GIS-based measures of environmental equity: Exploring their sensitivity and significance. *Journal of Exposure Analysis and Environmental Epidemiology*, *9*(1), 18-28. doi:10.1038/sj.jea.7500023
41. Timmons Roberts, J., Pellow, D., & Mohai, P. (2018). Environmental Justice. *Environment and Society*, 233-255. doi:10.1007/978-3-319-76415-3_11
42. Tivadar, M. OasisR: an R Package to Bring Some Order to the Word of Segregation Measurement. Forthcoming in *Journal of Statistical Software*.
43. Wen, M., Zhang, X., Harris, C. D., Holt, J. B., & Croft, J. B. (2013). Spatial Disparities in the Distribution of Parks and Green Spaces in the USA. *Annals of Behavioral Medicine*, *45*(S1), 18-27. doi:10.1007/s12160-012-9426-x
44. White, M. J. (1983). The Measurement of Spatial Segregation. *American Journal of Sociology*, *88*(5), 1008-1018. doi:10.1086/227768
45. Wong, D. W. (1993). Spatial Indices of Segregation. *Urban Studies*, *30*(3), 559-572. doi:10.1080/00420989320080551
46. Wong, D. W. (2008). A Local Multidimensional Approach to Evaluate Changes in Segregation. *Urban Geography*, *29*(5), 455-472. doi:10.2747/0272-3638.29.5.455
47. Wu, J. (2006). Environmental amenities, urban sprawl, and community characteristics. *Journal of Environmental Economics and Management*, *52*(2), 527-547. doi:10.1016/j.jeem.2006.03.003
48. Zhang, X., Lu, H., & Holt, J. B. (2011). Modeling spatial accessibility to parks: a national study. *International Journal of Health Geographics*, *10*(1), 31. doi:10.1186/1476-072x-10-31
49. Zhou, X., & Kim, J. (2013). Social disparities in tree canopy and park accessibility: A case study of six cities in Illinois using GIS and remote sensing. *Urban Forestry & Urban Greening*, *12*(1), 88-97. doi:10.1016/j.ufug.2012.11.004
50. Zwickl, K., Ash, M., & Boyce, J. K. (2014). Regional variation in environmental inequality: Industrial air toxics exposure in U.S. cities. *Ecological Economics*, *107*, 494-509. doi:10.1016/j.ecolecon.2014.09.01