European Journal of Geography

The publication of the European Journal of Geography (EJG) is based on the European Association of Geographers' goal to make European Geography a worldwide reference and standard. Thus, the scope of the EJG is to publish original and innovative papers that will substantially improve, in a theoretical, conceptual, or empirical way the quality of research, learning, teaching, and applying geography, as well as in promoting the significance of geography as a discipline. Submissions are encouraged to have a European dimension. The European Journal of Geography is a peer-reviewed open access journal and is published quarterly.

Received: 21/02/2025
Revised: 07/07/2025
Accepted: 19/07/2025
Published: 24/07/2025

Academic Editor:
Dr. Alexandros Bartzokas-Tsiompras

DOI: 10.48088/ejg.n.tou.16.2.268.285

ISSN: 1792-1341

Research Article

Statistical Reliability of the Modified Areal Weighted by Control Zones Method to Spatially Downscale Individual Social Data

Najla Touati [1], Wilda Jean Baptiste [2] & Julia Hidalgo [2]✉

[1] LISST, UMR 5193, Toulouse Jean Jaures University, 5 allée A. Machado, Toulouse, 31058, France
[2] LISST, UMR 5193, CNRS, 5 allée A. Machado, Toulouse, 31058, France

✉ Correspondence: julia.hidalgo@univ-tlse2.fr

Abstract: This study evaluates the modified areal weighting by control zones method (MAW-CZ) often involved in downscaling social data from a large spatial mesh, to a smaller mesh. This method has been extensively used in literature but the impossibility, until recently, of accessing individual data makes it so that it has not been evaluated. In this study it is applied to two case studies, Toulouse and Grenoble-Alpes Metropoles, using the census INSEE data at the IRIS scale and the building islet or topographical reference units (RSU) scale. The study found that 27.2% of RSUs in the Toulouse metropolis and 21.9% in the Grenoble-Alpes metropolis are inhabited, with mean populations of 122 and 116 residents, and maximum populations of 2,429 and 6,451 residents, respectively in 2018. The chosen downscaling approach introduces small errors for small and medium-sizedRSUs. For example, 94%, 78%, and 72% of RSUs of <100, 101–255, and 256–500 inhabitants, respectively, are correctly classified by the modified areal weighting by control zones method in the Toulouse Metropole. However, there are significant differences for the most populated RSUs (the performance decreases to 60% for RSUs with more than 500 inhabitants), with this category having a representativeness of 8.4% and 7.2% of the total number of inhabited RSUs in the Toulouse and Grenoble-Alpes metropoles, respectively. The spatial distribution of the biased RSUs are nevertheless homogeneous throughout the two territories. These discrepancies are due to both the upscaling/downscaling methods used and the nature of the data (points in the upscaling and polygons in the downscaling).

Keywords: Downscaling social data; upscaling social data; aggregation; disaggregation; mesh; spatial analysis; topographical reference unit (RSU)

Highlights:

- Population is downscaled from the census IRIS scale, to the RSU scale.
- Modified areal weighted by control zones approach is evaluated.
- Downscaling performs for small and medium-sized RSUs, <500 inhabitants with errors between 2 and 28%.
- RSUs with >500 inhabitants — where errors reach 40% — represent <10% of inhabited RSUs in both case studies.

1. Introduction

Today, public policymakers, urban planners, and researchers are challenged with integrating social and environmental data that have diverse spatial resolutions. One example is evaluations of human vulnerability to the various impacts of climate change at finer scales, such as at the neighborhood or islet level in urbanized areas. In recent years, environmental data have benefited from increasingly fine resolutions, thanks to advances in spatial imaging and numerical modeling. These data also enjoy a high level of open access, exemplified by initiatives such as the European Union's Copernicus program and various national research projects. However, because of the confidentiality status of social data, publicly available variables required for vulnerability studies are often aggregated into larger statistical spatial units. These units tend to be too coarse to be effectively combined with and analyzed alongside environmental spatial data.

To address this issue, social data are often spatially downscaled. Downscaling looks to transfer social data to a new framework while maintaining the quality of the information. This process requires the use of auxiliary data to ensure an effective transfer. Consequently, the choices of the alternative control dataset and the methodological framework are essential to estimation of the disaggregation quality and the relevance of the destination scale.

The process of the spatial transfer of coarser social data, to finer spatial scales, such as geographic data, administrative boundaries, or statistical units, often encounters issues of harmonization across different spatial supports (Fotheringham & Sachdeva, 2022). The geometric dimension of geographic data can lead to statistical biases when social data are transferred to this type of framework. This phenomenon is referred to as the change of support problem (COSP) (Arbia, 1989).

Two types of COSPs have been identified. The first, ecological bias inference, is a concept that comes from environmental sciences (King et al., 2004) and refers to errors related to downscaling methodologies. The second is the modifiable areal unit problem (MAUP). The effect of this COSP depends on how the data are upscaled to the chosen spatial unit. The choice of aggregation level can alter interpretations of spatial analyses and the relationships between variables (Fotheringham & Wong, 1991; Openshaw et al., 1979; Openshaw, 1981; Pivano et al., 2015; Robinson, 1950; Wong, 2004). This issue is prevalent across numerous disciplines that utilize data aggregation (e.g., geosciences, environmental sciences, geography, and ecology) and resembles the ecological inference problem, where individual patterns are inferred from group data. The type of spatial data input also has an impact on the effectiveness of the methods used for the spatial modeling of social data (Louvet, 2015; Monteiro, 2018; Plumejeaud, 2010).

It is important to consider these effects when evaluating the upscaling and downscaling quality. According to Louvet, three potential approaches to reduce the effects of COSPs are (1) utilizing individual data with access to confidential sources, (2) adapting statistical methods to account for MAUPs, and (3) assessing the sensitivity to MAUPs based on the spatial partitioning of the results.

In the context of this paper, the first proposition is used to evaluate the MAUP effect and the statistical sensitivity of downscaling French social data. In France, as in many countries in Europe and abroad, social data (e.g., fiscal or health data and census data) are not freely accessible at the individual level. Nevertheless, the French National Institute for Statistics and Economic Studies (INSEE) provides multiple types of open social data on its website; these include census data (https://www.insee.fr/en/accueil) aggregated into larger statistical units, called Ilots de Regroupement de l'Information Statistique (IRIS). This is the smallest sub-municipal unit used for public statistical analyses in France. The IRIS framework divides communes, particularly those with at least 10,000 residents, and includes many communes with populations between 5,000 and 10,000 (INSEE, 2016). Despite their utility, IRIS units do not always correspond to real-world territorial or urban planning needs, including difficult informed urban development decisions, such as addressing heat islands or establishing local services. Moreover, the IRIS boundaries often do not align with other spatial scales used in environmental studies, such as urban overheating, where the spatial unit of the physical phenomena is finer. Such studies frequently rely on finer topographical reference units (RSUs) based on urban and architectural morphology (Bocher et al., 2018; Masson, 2015), which better suit the realities of urban and environmental planning.

The use of individual social data, as proposed by Louvet (2015), is now possible in France because social and physical data at individual and local levels have become more easily accessible to researchers via the Secure Data Access Center (CASD), a new controlled means of access under conditions of confidentiality (https://www.casd.eu/). INSEE provides some of these sensitive data to researchers and data scientists, including official population data, ensuring statistical confidentiality in line with the Data Protection Act and under the supervision of the French Data Protection Authority (CNIL).

This paper uses these individual social data as a set of control data to test and reproduce the modified areal weighted by control zones method (Goodchild, 1993; Langford, 1992; Plumejeaud, 2010). This approach has been used in the literature (Goodchild, 1993; Langford, 1992; Nordhaus, 2002; Plumejeaud, 2010) without being evaluated for statistical and spatial discrepancies and biases caused by transferring data from one grid (IRIS) to another (in our case, the RSU).

This paper is organized as follows. In Section 2, the case studies and the utilized datasets are described. In Section 3, the theoretical frameworks of the transfer methods (upscaling and downscaling) are presented and briefly reviewed. For the downscaling, the modified areal weighted by control zones method is explained in detail. Then, in Section 4, the application of the transfer method to the inhabitant number is described. In Section 5, to estimate the sensitivityand the reliability of the modified areal weighted by control zones method, the results of the disaggregation (from the IRIS scale to the RSU scale) are compared with the individual scale data aggregated on the RSU mesh.

2. Case studies and datasets

2.1. Case studies: The Toulouse and Grenoble-Alpes metropoles

To evaluate the specificities and generality of the results, two case studies of the Toulouse and Grenoble-Alpes metropoles are analyzed (Figure 1). Toulouse, the capital of the district of Occitanie with a population of 520,896 in 2024, is located in southwestern Occitanie in the south of France. It is the fourth most populous city in France (after Paris, Marseille, and Lyon) and is the largest municipality of the Toulouse Metropole. With 37 municipalities and close to 0.8 million inhabitants in an area of approximately 460 km2, this intermunicipality contains more than 57% of the population in the Haute-Garonne district.

Grenoble, located in the eastern Rhône-Alpes region of France, is smaller than Toulouse, with 156,064 inhabitants in 2024. The Grenoble-Alpes Metropole includes 49 municipalities in the Isère district and has a population of 450,000 and a surface area of 546 km2.

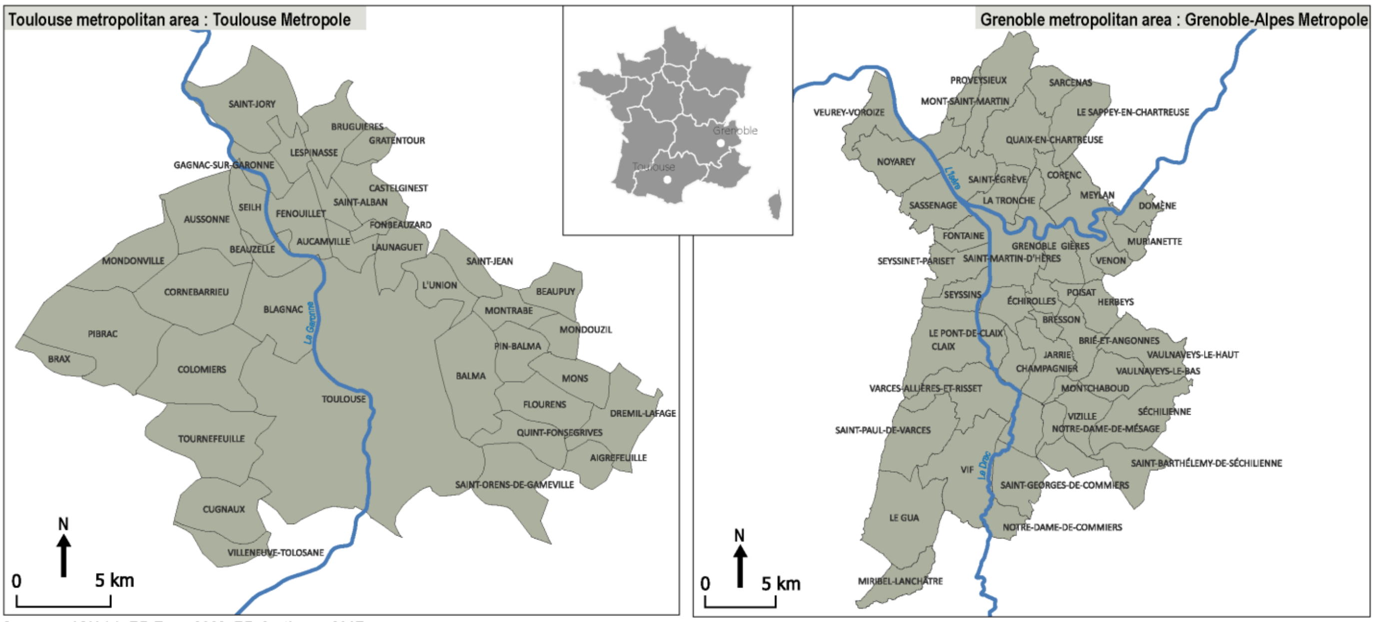
Despite differences in their population sizes, these two metropoles have been the subject of a number of studies in urban climatology and urban planning (Rome, 2021; Hidalgo, 2023). The available studies and data, if at the appropriate spatial resolution, should make it possible to observe and compare statistical trends in the social vulnerability of these areas with respect to climate change.

2.2. Datasets for territorial analyses of the two intermunicipalities

The datasets used in this study enable comparison of population estimates derived from spatial disaggregation methods with georeferenced reference data at the RSU level. This study uses individual social data as a benchmark to evaluate the modified areal weighting by control zones method (Goodchild, 1993; Langford, 1992; Plumejeaud, 2010), a technique applied in French research projects (e.g., MApUCE, PAENDORA) but rarely assessed for statistical and spatial biases caused by data transfers between grids (from IRIS to RSU).

- To conduct this evaluation, we utilize high-resolution geographic, urban morphology, and socio-demographic data for the Toulouse and Grenoble urban agglomeration. Spatial data are obtained from open-access platforms, while individual socio-demographic data are accessed through secure data portals (CASD). Table 1 and the following list summarize the main data sources: Geographic files of communes in 2020 from the Institut Géographique National (IGN). The 2018 INSEE data correspond to the geography on January 1, 2020. These data consist of polygon shapefiles.

- Geographical indicators produced with GeoClimate tool for the RSU mesh and buildings. GeoClimate is an opensource geospatial processing tool for environmental and climate studies (https://github.com/orbisgis/geoclimate/wiki). Using vector-based inputs, the workflow uses the Open Street Map database or the French BD Topo database. These data consist of FlatGeobuf polygons.



Sources : IGN (c), BD Topo 2023, BD Carthage, 2017

Figure 1. Maps of the Toulouse and Grenoble-Alpes metropoles.

- FIDELI data for 2018 (demographic files on dwellings and individuals). This database is special because it contains data at the individual level, including information from tax authorities on taxes and built-up properties. This provides a better understanding of the housing stock and the demographics of residents. The FIDELI files describe the characteristics of dwellings (e.g., the surface area, location, and lifts) and the households living in these dwellings (e.g., marital status and income) and identifies individuals by their location (https://www.insee.fr/fr/metadonnees/source/serie/s1019). These data consist of point shapefiles.

Table 1. Number of features for each data source for the two metropoles.

Grenoble-Alpes	Toulouse	
201	254	Number of IRIS units (INSEE/IGN data) [1]
19,048	24,827	RSU (GeoClimate) [2]
130,946	355,517	Buildings (GeoClimate)
445,376	774,971	Inhabitants (FIDELI)

[1] Here, IRIS indicates the smallest sub-municipal statistical unit available for public statistical analyses in France.
[2] RSU indicates the topographical reference unit.

Together, these datasets provide a basis for evaluating population estimates derived from spatial disaggregation methods against reference data at the RSU level. The following section details the transfer methods and data processing techniques used in this study.

3. Transfer method and data processing

3.1. Downscaling and upscaling operations

Two main operations are used in this study: downscaling and upscaling. Downscaling, or disaggregation, enables data to be redistributed onto smaller spatial units (points or polygons), entirely contained within the source and target units, by creating a smaller common spatial denominator between the two supports. Upscaling, or aggregation, combines data from source units into target units based on a geometric inclusion rule. This operation is generally applied after disaggregation to estimate a variable on a misaligned grid or when moving from a smaller to larger scale (Plumejeaud et al., 2010; Vignes et al., 2013).

3.2. Different disaggregation methods

Various approaches have been developed to address spatial disaggregation (Patil et al., 2024; Do, 2015), primarily focused on proxy data-based methods, machine learning techniques, and model-based geostatistical methods (Louvet, 2015; Monteiro, 2018; Zhang, 2022). Proxy methods rely on ancillary variables such as land use, road density, night-time lights, while machine learning and geostatistical models capture spatial patterns and correlations through predictive modeling and spatial interpolation. These methods are often used to disaggregate demographic, socio-economic, and environmental data.

Among these approaches, simple areal weighting is the most basic technique. It assumes a uniform distribution within the source zones and reallocates values proportionally based on the area of overlap with the target geometries, which can vary in size and shape. This method is easy to implement using GIS software such as QGIS and requires minimal data, but it often results in imprecise spatial distributions.

A variation of this method, the modified areal weighting using control zones (MAW-CZ), introduces constraints based on predefined zonal structures to improve spatial accuracy. This method extends on the principle of simple area weighting by incorporating an auxiliary variable to better control the data transfer process (Do, 2015; Flowerdew, 1992). The auxiliary variable must satisfy three criteria: (1) it should be spatially correlated with the variable to be transferred; (2) its distribution or spatial dispersion should be known in both the source and target zones; and (3) it should be spatially similar to the source mesh.

While this method significantly improves the quality of the transferred information, finding an auxiliary variable that is both spatially correlated and shares the same distribution as the variable to be transferred can be challenging. To overcome some of these limitations, the MAW-CZ method proposes a refined proxy-based approach that leverages an intermediate control zone to guide the disaggregation process more effectively. Despite its conceptual simplicity and potential advantages—particularly in contexts with limited ancillary data—the MAW-CZ method has received relatively little systematic evaluation in the spatial disaggregation literature. Few empirical studies have thoroughly assessed its performance, which highlights the need for further investigation, such as that undertaken in this study (Patil et al., 2024; Monteiro et al., 2018).

In this study, the MAW-CZ method is applied as an alternative to simple areal weighting. Known information about the variable (support A) is first transferred to an intermediate zone (support C) or to geographical objects, such as buildings, before estimating the value of the variable in the target zone (support B) (Goodchild et al., 1993; Langford et al., 1992). This method involves creating two zones, the source and target zones, along with an intermediate zone or object called the control zone. By downscaling through the control zone followed by upscaling, it allows for better estimations of inhabitants or other variables. Figure 2 illustrates this data transfer process, as proposed by Plumejeaud et al. (2010).

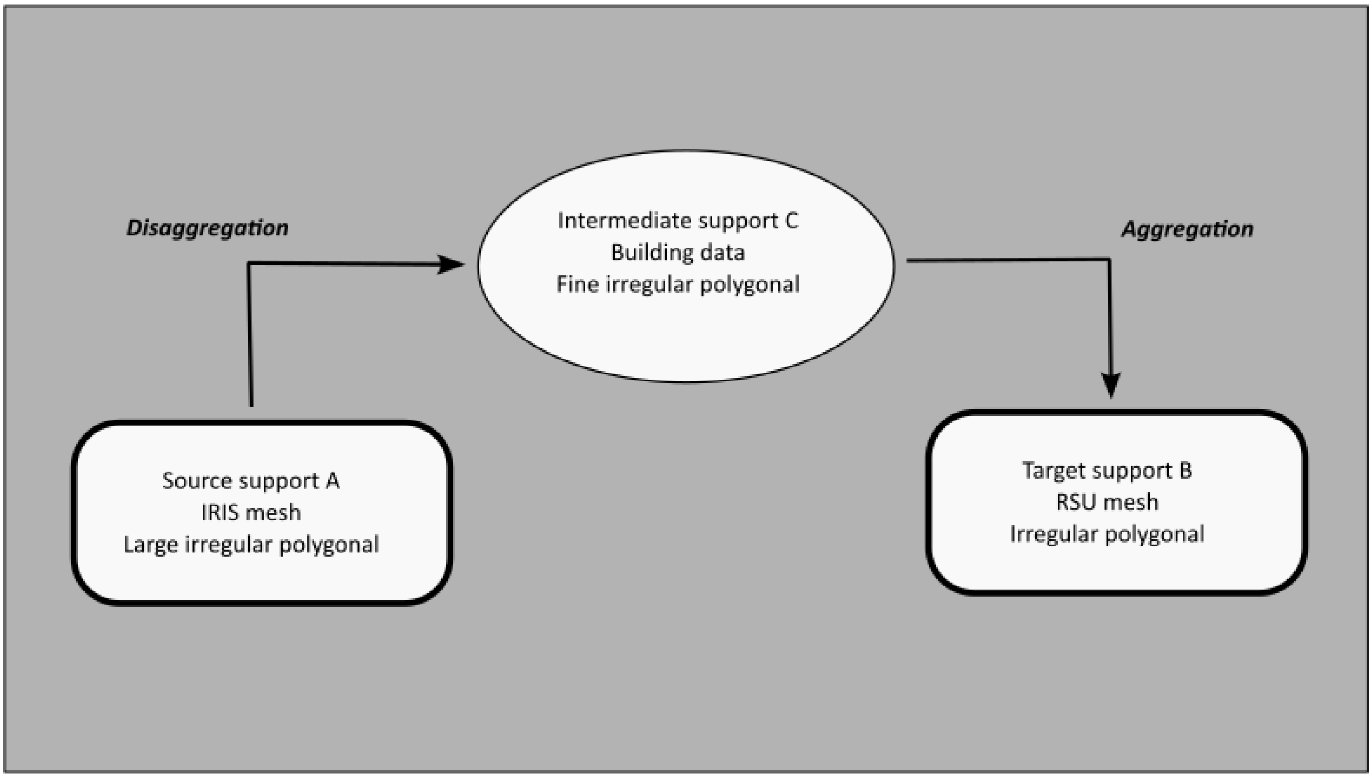


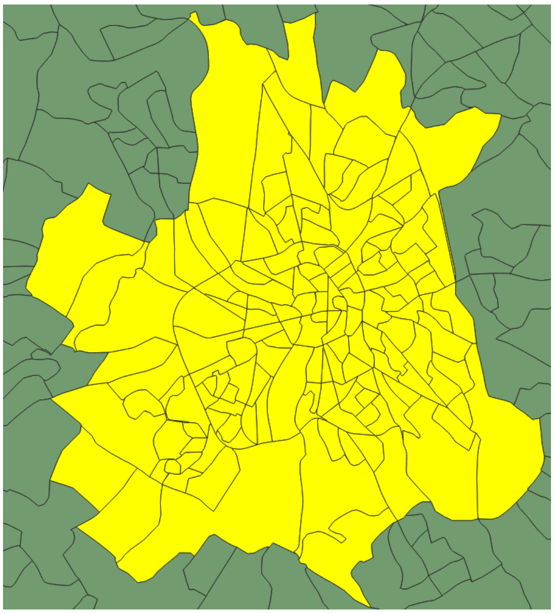
Figure 2. Support transformation data transfer (as proposed by Plumejeaud et al., 2010). IRIS indicates the larger statistical unit (the smallest sub-municipal unit available for public statistical analyses in France), and RSU indicates the smaller topographical reference unit.

3.3. Choice of support mesh

In this study, the information to be transferred is represented by a social variable: the inhabitant count. The transfer of the population or inhabitant count occurs from a source zone where this information is known (the IRIS unit; Figure 3a) to a target zone where this information is unknown (the RSU; Figure 3b). The RSU is a partitioning of the urban territory based on the Delaunay triangulation (Bocher et al., 2018). It offers a more precise mesh than the IRIS statistical division and is often used as a reference spatial division in French environmental studies, in particular, in urban climate studies. The RSU geometry was computed for each case study using the GeoClimate open source software. IRIS and RSU correspond to two non-aligned grids, which means that they do not overlap (see the yellow surfaces in Figure 3b) and that there are no common rules for their construction.

https://doi.org/10.48088/ejg.n.tou.16.2.268.285

The choice of the RSU target source (support B) is motivated by its good adaptation to architectural and urban morphology compared with a regular grid mesh. As explained in Figure 2, IRIS units, the source support (support A), are large statistical units, comparable to large neighborhoods. RSUs, conversely, are small units representing a city islet (a building block surrounded by streets). Because of this difference in size, intermediate regular polygon building entities (support C), which are included in the target support, are needed.



(a)

(b)

Figure 3. (a) Larger-scale IRIS units and **(b)** finer-scale RSUs for the Toulouse Metropole.

3.4. Data processing scheme

This study evaluates the accuracy of the modified areal weighting by control zones method when disaggregating the inhabitant count from the IRIS scale to the RSU scale using the methodology proposed in Figure 2. Accordingly, using basic statistical approaches, the results of this disaggregation are compared with the aggregation of individual-level data from the FIDELI database to the RSUs (Figure 4); these data are considered as being the most accurate data available.

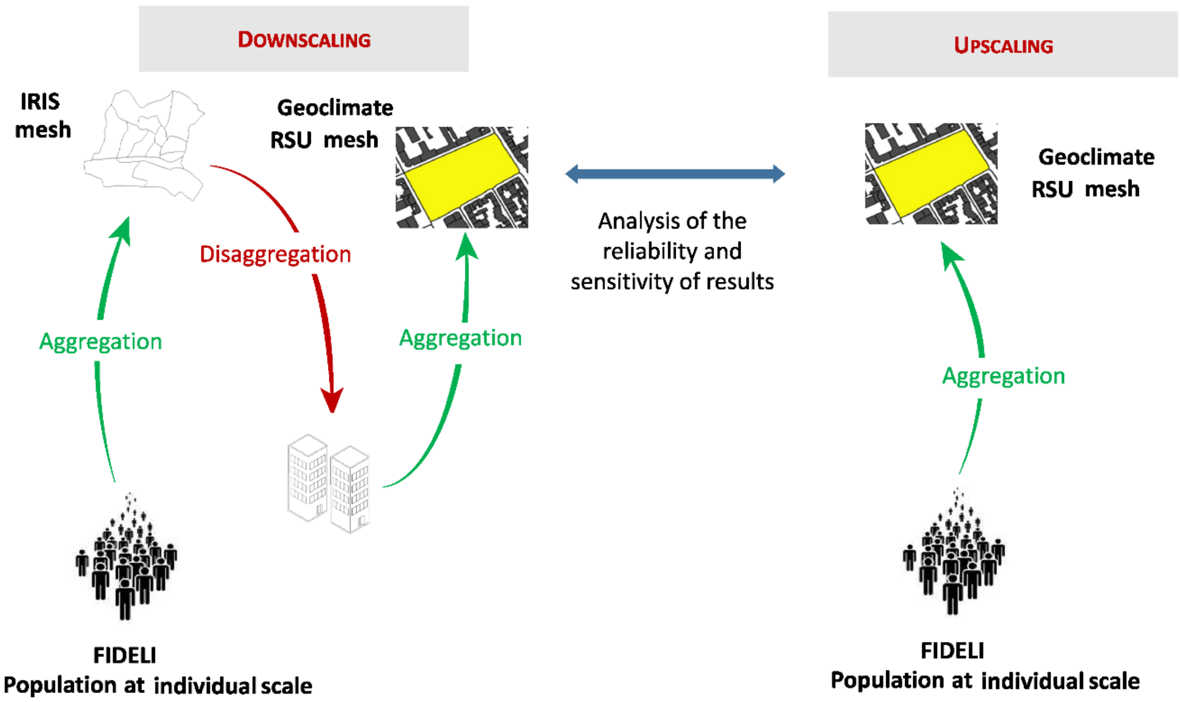


Figure 4. Description of the data processing performed in this study to evaluate the downscaling of social data from a coarse mesh to a finer mesh.

https://doi.org/10.48088/ejg.n.tou.16.2.268.285

To analyze the sensitivity of the results, the individual data are used in two steps. First, the FIDELI database populations are aggregated onto the RSU mesh, the target support for the analysis. Next, the process for transferring the population count presented in Figure 2 is started by first aggregating the FIDELI database inhabitants onto the IRIS mesh and then disaggregating them via the modified areal weighting by control zones method using the building scale as the intermediate support before aggregating them again at the RSU scale for comparison with the cited upscaled results.

3.5. Upscaling the FIDELI database population to the RSU and IRIS meshes

Using the population variable from the FIDELI database, the population can be summed within each RSU mesh:

$$pop_R = \sum(pop_{F \subset R}), \tag{1}$$

where $pop_F$is the FIDELI population and $pop_R$is the population by RSU.

The same calculation is applied to the aggregation of the FIDELI data to the IRIS mesh:

$$pop_I = \sum(pop_{F \subset I}), \tag{2}$$

where $pop_I$is the population by IRIS unit.

The aggregation of the population estimates by building for each RSU is calculated as:

$$pop_B pop_R = \sum(pop_{B \subset R}), \tag{3}$$

where pop_B is aggregated for each RSU.

3.6. Downscaling the IRIS population to the RSU mesh

Following Plumejeaud (2010), a population potential was created using building footprints as control zones. Two variables extracted from the GeoClimate toolbox (building area and number of floors) were used to estimate the building developed surface ($Sdev_B$), defined as the footprint area multiplied by the number of floors, representing the total constructed surface.

The number of floors per building(nbf_B) is estimated using building elevation and average floor height, depending on the building's construction period. This developed surface is then used to weight the disaggregation of IRIS-level population data (pop_I)onto building objects, in order to estimate the number of inhabitants per building (pop_B). Finally, the building-level population can be aggregated at the block level, allowing an estimation of population per RSU (pop_R). $Sdev_B$

The formula used to calculate the building developed surface is:

$$Sdev_B = area_B \times nbf_B, \tag{4}$$

where $area_B$ indicates the area of building B and nbf_Bindicates the number of floors in building B.

The population by building (pop_B)can then be estimated using the population in each IRIS unit (pop_I)and the building developed surface $Sdev_B$:

$$(pop_B)(pop_I)Sdev_B pop_B = pop_I \times Sdev_B / \sum_I Sdev_B, \tag{5}$$

Finally, as detailed in Section 3.5, the building-level population is aggregated to the RSU mesh to estimate the population per RSU (pop_R):

$$pop_B pop_R = \sum(pop_{B \subset R}), \tag{6}$$

where the sum is taken over all buildings B contained within each RSU R.

4. Results

In this section, a descriptive statistical analysis is conducted to identify the central tendencies, variabilities, and any potential imbalances or anomalies within the data sample. This step is important to ensure a correct interpretation of the subsequent results and to contextualize more complex analyses. Next, to ensure the robustness and generalizability of the modified areal weighting by control zones method, the results obtained from the upscaling and downscaling approaches are compared in Figure 4. This allows to verify the method's ability to maintain its performance across different subsets of the data while minimizing the risk of overfitting. Then, a bias analysis can enable the differences between the model performances for the different classes of inhabitant counts to be understood.

Statistical computations, including spatial statistical analyses, were carried out using R 4.3.1 (CRAN), a free software environment for statistical computing. Spatial data exploration and map production were performed using QGIS 3.28, an open-source software developed by the Open Source Geospatial Foundation (OSGeo).

4.1. Descriptive statistics of the two case studies

The downscaling method to the RSU level was applied to both the Toulouse and Grenoble-Alpes metropoles. According to the upscaling method, 27.2% and 21.9% of the RSUs in the Toulouse and Grenoble-Alpes metropoles, respectively, are inhabited (Table 2).

https://doi.org/10.48088/ejg.n.tou.16.2.268.285

Table 2. Summary of the RSUs by metropole.

Grenoble-Alpes		**Toulouse**		
Downscaling	Upscaling	Downscaling	Upscaling	
19,048		24,827		Total RSUs
4197 (22%)	4164 (21.9%)	6302 (25.4%)	6746 (27.2%)	Inhabited RSUs
14,851 (78%)	14,884 (78.1%)	18,525 (74.6%)	18,081 (72.8%)	Uninhabited RSUs

The analysis is applied to a sample of RSUs that have at least one inhabitant according to both the upscaling and downscaling methods to illustrate, if an RSU has two inhabitants according to the upscaling method but no inhabitant according to the downscaling method, this RSU cannot be included in the sample. The final sample contains 5,993 RSUs in the Toulouse Metropole and 3,661 RSUs in the Grenoble-Alpes Metropole.

(a)

(b)

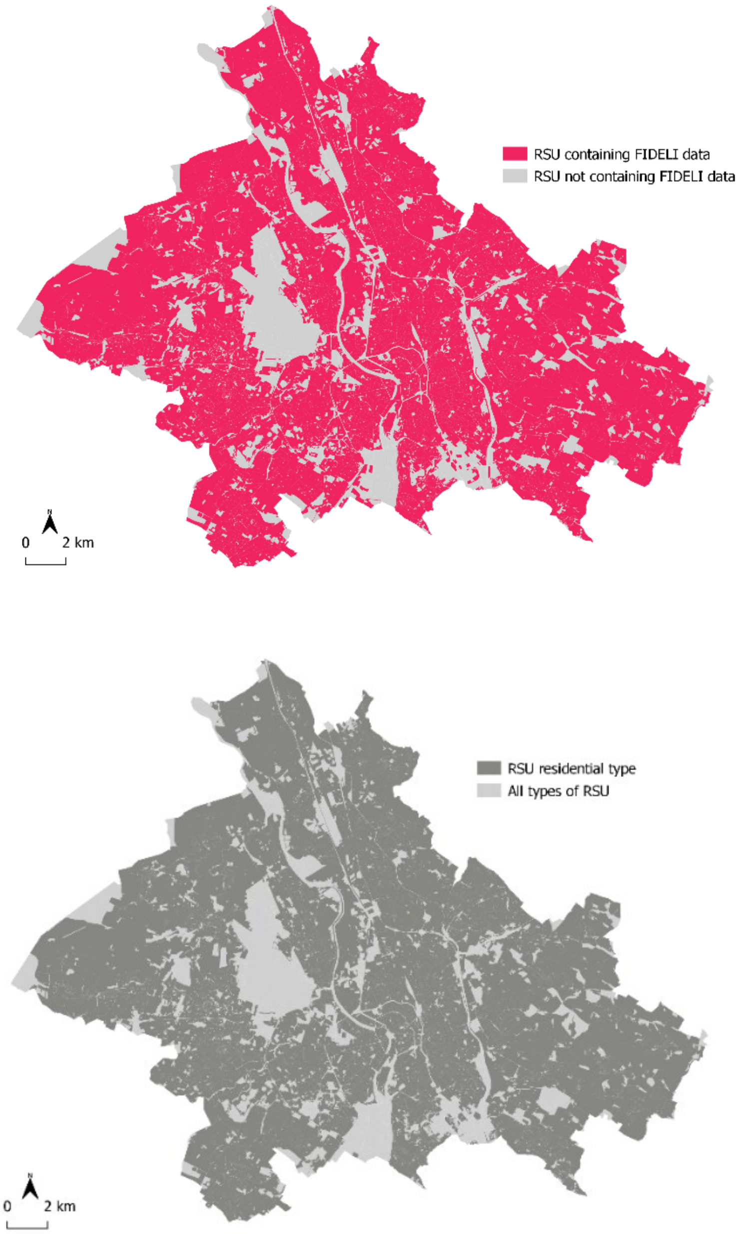


Figure 5. Spatial distributions of **(a)** inhabited and uninhabited RSUs and **(b)** residential RSUs in the Toulouse Metropole. (For the Grenoble-Alpes Metropole, see Figure A.1. in appendix)

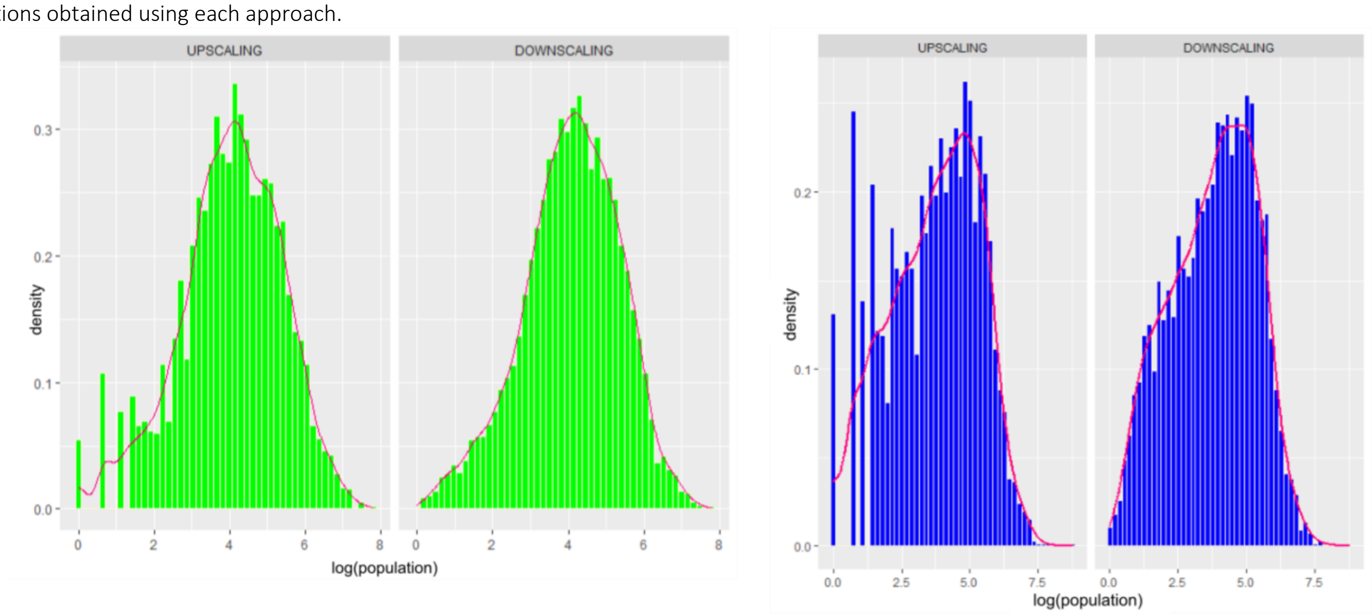
https://doi.org/10.48088/ejg.n.tou.16.2.268.285

Figures 5 (and Figure A.1 in appendix) illustrate the spatial distributions of the inhabited and uninhabited RSUs for the two case studies. Figure 5a shows the contrast between RSUs with and without inhabitants, while Figure 5b shows the contrast between RSUs with and without residential buildings.

4.2. Statistical Distribution

A significant positive correlation was observed between population size and total developed surface area at the IRIS level in both metropolitan areas. Specifically, the Toulouse metropolitan area exhibited a correlation coefficient of $r = 0.85$ ($t(252) = 26.17$, $p < 2.2 \times 10^{-16}$; 95% CI: [0.82–0.88]), while the Grenoble metropolitan area showed $r = 0.80$ ($t(195) = 18.64$, $p < 2.2 \times 10^{-16}$; 95% CI: [0.74–0.85]), indicating a strong and robust linear association in both cases.

To compare the statistical distributions resulting from the upscaling and downscaling methods, log-transformed density curves and histograms were produced for both metropolitan areas (Figure 6). These visualizations reveal a high degree of overall similarity between the distributions obtained using each approach.



(a) **(b)**

Figure 6. Density curves and histograms of the inhabitants in the RSUs of the **(a)** Toulouse and **(b)** Grenoble-Alpes metropoles using the downscaling (left) and upscaling (right) methods.

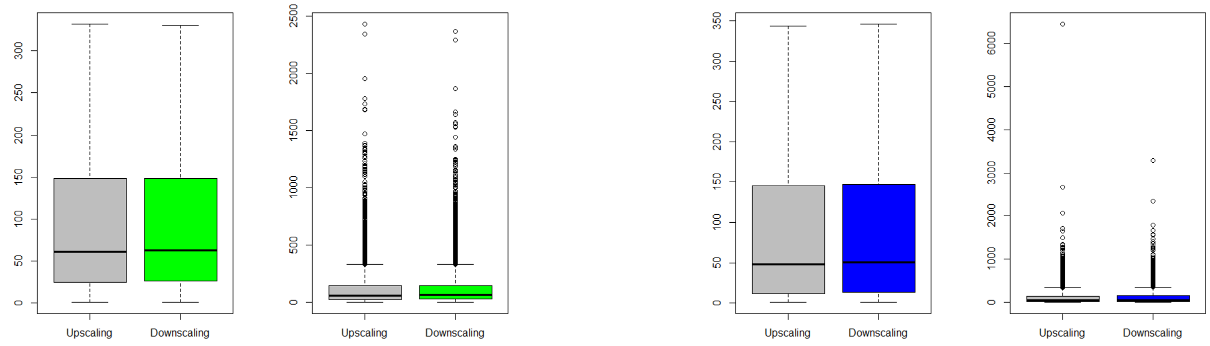
Statistical tests further supported these observations. In Toulouse, the Student's *t*-test comparing the mean population per RSU between aggregated and disaggregated datasets showed no significant difference ($t(5992) = 1.47$, p = 0.14; mean difference = 1.23; 95% CI: [-0.41, 2.87]). However, the Kolmogorov–Smirnov (KS) test indicated a slight but statistically difference in distribution shapes (D = 0.025; p = 0.042). A similar pattern was observed in Grenoble, where the t-test again revealed no significant difference in means ($t(3660) = 0.51$, p = 0.61; mean difference = 0.71; 95% CI: [-2.01, 3.43]), whereas the KS test showed a more pronounced significant difference in distribution (D = 0.045; p = 0.0012).

Table 3 presents statistical indicators for both methods and metropolitan areas.

Average population per RSU is nearly identical between the two approaches: 122 vs. 121 in Toulouse, and 116 vs. 115 in Grenoble. However, extreme values, particularly in the Grenoble metropolitan area, exhibit greater variability depending on the method used (Figure 7).

Table 3. Statistical indicators of the upscaling and downscaling methods for the Toulouse and Grenoble-Alpes metropoles.

Statistical Indicators	**Toulouse**		**Grenoble-Alpes**	
	Upscaling	Downscaling	Upscaling	Downscaling
Minimum	1	1.1	1	1
1st quartile	25	26.5	12	13.1
Median	61	62.9	48	50.4
Mean	122	120.8	116	115.2
3rd quartile	148	148.3	145	146.7
Maximum	2429	2363.8	6451	3291.7
Standard deviation	178	170	207	185
Variance	31,862	28,958	43,046	34,258



(b) **(a)**

Figure 7. Boxplots without (left) and with (right) outliers for the upscaling (aggregation-AG; grey boxplots) and downscaling (disaggregation-DG; blue or green boxplots) methods for the (a) Toulouse (TM) and (b) Grenoble-Alpes (GAM) metropoles

Furthermore, Figure 7 and Table 4 highlight the presence of statistical units (RSUs) considered outliers in terms of population, representing approximately 7% of RSUs in each metropolitan area, regardless of the method employed. (see Table A.1 in appendix).

Table 4. Number of RSUs with the number of inhabitants considered as outliers[1].

Grenoble-Alpes	Toulouse	Method
7.2% (265/3661)	8.4% (504/5993)	Upscaling
7.1% (258/3661)	7.9% (474/5993)	Downscaling

[1] In this case, the outliers are the RSU with a large population compared to the average.

In summary, although mean population values per RSU do not differ significantly between the upscaling and downscaling methods, the Kolmogorov–Smirnov test reveals subtle but significant differences in distributional form, especially in Grenoble. These results underscore the importance of a comprehensive evaluation that considers both central tendency and distributional characteristics when comparing spatial disaggregation methods. Further analyses of bias, absolute error, and spatial structure are presented in the following sections.

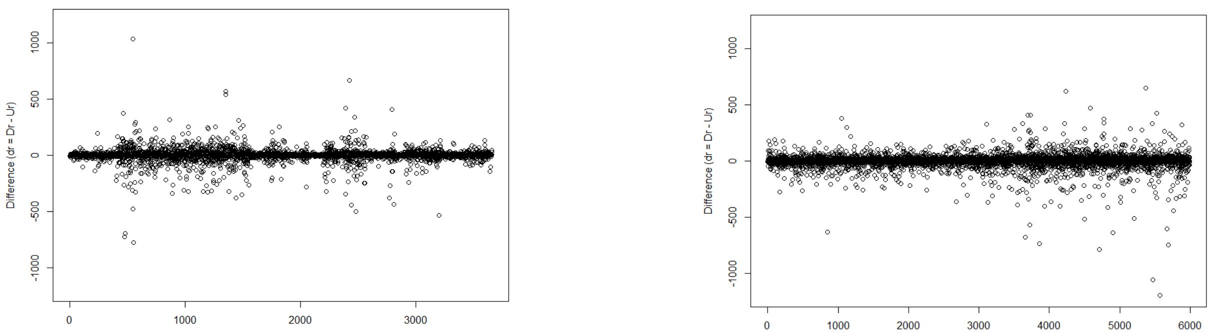
4.3. Error and bias analysis of the MAW-CZ method

To assess the accuracy of the modified areal weighting by control zones (MAW-CZ) method, we compared the estimated number of population obtained through downscaling to the reference values produced via upscaling. Two discrepancy indicators were calculated at the RSU level: the absolute difference and the relative deviation (percentage), using the following formulas:

$$d_r = D_r - U_r, \text{ and } rd_r = \frac{|D_r - U_r|}{U_r} \times 100 \quad (7)$$

where D_r is the population estimate from downscaling, and U_r the reference value from upscaling of inhabitants calculated by the upscaling, for D_r the r^{th} RSU. For example, D_{64}= 132 and U_{64}= 120, then d_{64}= 12 and rd_{64}= 10%.

Figure 8 presents the distribution of absolute differences across RSUs in Toulouse and Grenoble-Alpes. Most RSUs show relatively small errors, but some, especially those with high population densities, display larger discrepancies.



(b) **(a)**

Figure 8. Absolute error per RSU (Dr−UrD_r - U_rDr−Ur) in the Toulouse (a) and Grenoble-Alpes (b) metropoles.

Descriptive statistics of the relative deviation are presented in Table 5. Average errors are relatively high: 58% in Toulouse and 63% in Grenoble-Alpes. However, median errors are substantially lower (23% and 29%, respectively), indicating a positively skewed distribution. These error measures offer a global assessment of model accuracy across all RSUs. However, to provide a more interpretable view, we grouped RSUs by error thresholds (Table 6). This allows identification of zones with high, moderate, or low accuracy. For instance, only 24% of RSUs have a deviation below 10%, and over 40% exceed 30%.

Table 5. Statistical indicators of the relative deviation (%).

Statistical Indicators	Toulouse	Grenoble-Alpes
Minimum	0	0
1st quartile	11	13
Median	23	29
Mean	58	63
3rd quartile	46	57
Maximum	>100%	>100%
Standard deviation	555	268
Variance	3081	720

To better describe the heterogeneity in model performance, RSUs were grouped by relative error thresholds (Table 6). This classification is purely descriptive and does not correspond to spatially contiguous zones. It provides an overview of how accuracy varies across the urban areas: for instance, only 24% of RSUs have a relative deviation below 10%, while over 40% exceed a 30% error rate. A non-negligible share of RSUs (10.1% in Toulouse and 16.0% in Grenoble-Alpes) present deviations greater than 80%. Outliers (RSUs with errors well above the mean) represent 6.8% of the sample in Toulouse and 8.9% in Grenoble-Alpes (see Table A.1 in Appendix).

Table 6. Percentage of RSUs under different thresholds.

Threshold	Toulouse	Grenoble-Alpes
Less than 5%	12.0%	10.2%
Between 5% and 10%	11.8%	10.2%
Between 10% and 15%	11.4%	9.9%
Between 15% and 30%	25.2%	21.9%
Between 30% and 50%	17.5%	18.7%
Between 50% and 80%	12.0%	13.1%
More than 80%	10.1%	16.0%

These figures reflect a spatially heterogeneous pattern of errors, which is further illustrated in Figure 9. Panel (a) displays the downscaled population estimates in Toulouse, while panel (b) maps the relative error rates using natural breaks (Jenks classification). Most RSUs display low errors, especially in central areas, but larger deviations are observed in peripheral zones or areas with unusual urban morphology. Equivalent maps for the Grenoble-Alpes Metropole are provided in Appendix (Figures A.3 and A.4).

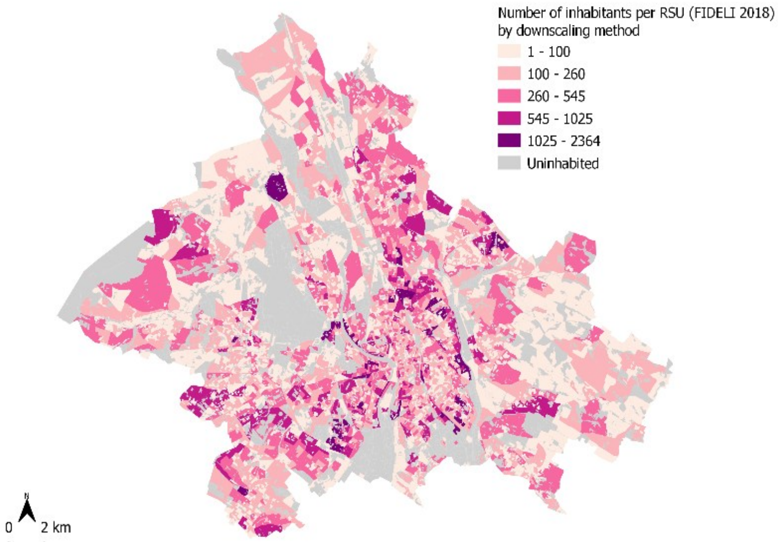
These results confirm that the downscaling method broadly preserves the spatial structure of population distribution. However, discrepancies at the local level remain, particularly in areas with complex urban forms or atypical population densities. This spatial heterogeneity in the error distribution highlights the need for further spatial diagnostics. In the next section, global and local spatial autocorrelation metrics are used, Moran's I and LISA indicators) to assess whether the spatial patterns derived from the disaggregation process are consistent with those obtained through aggregation approach.

4.4. Spatial autocorrelation of population and errors patterns

The Moran's I index and Local Indicators of Spatial Association (LISA) were used to determine whether the disaggregation process preserves the spatial structure of the aggregated population data at the reference statistical unit (RSU) level (Moran, 1950; Anselin, 1995). Moran's I assesses global spatial autocorrelation, while LISA detects local clusters and outliers. These measures provide spatial diagnostics that complement the previously seen error metrics, helping to determine whether the disaggregated data reproduces the spatial organization of the aggregated reference data. Two types of spatial weight matrices were used: the Queen contiguity matrix, which defines neighbors based on shared edges or corners, and the k-nearest neighbors (k-NN) matrix, which links each RSU to its nearest neighbors based on distance. These two configurations capture distinct spatial relationships, morphological (Queen) and metric (k-NN), allowing evaluation of whether disaggregation reproduces the spatial population patterns observed in the aggregated data.

Specifically, Moran's I provides a global measure of spatial autocorrelation, reflecting the degree to which population values are spatially clustered across all RSUs. LISA complements this analysis by identifying local spatial clusters and outliers, offering a finer understanding of how well the local spatial structure is preserved. Together, these indices provide a spatially explicit comparison that goes beyond descriptive statistics and helps determine whether the disaggregated data retains the spatial coherence of the reference dataset.

(a)



(b)

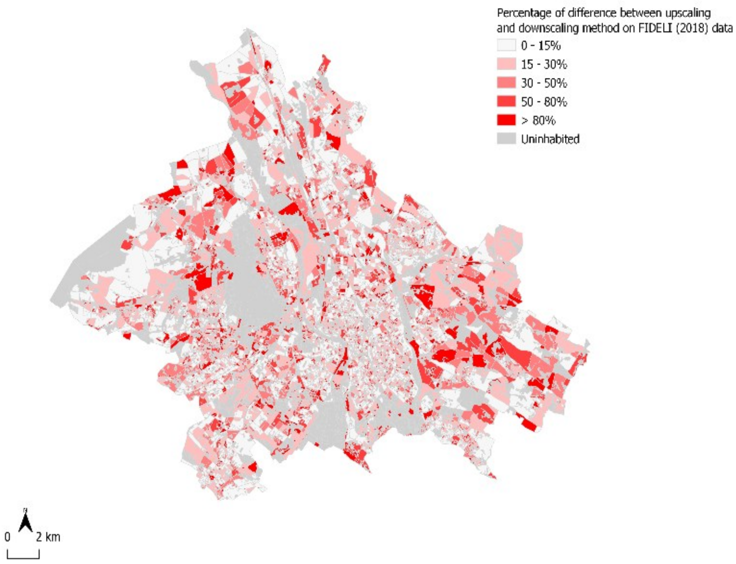


Figure 9. Spatial distribution of the number of inhabitants in the Toulouse Metropole by RSU when downscaling with the modified areal weighting by control zones method **(a)** and map of the error rate (natural thresholds) compared with the reference dataset obtained by upscaling the FIDELI data **(b)**. (For the Grenoble-Alpes Metropole, see Figure A.2. in appendix)

Table 7 summarizes Moran's I values for both cities and methods. For Toulouse, Moran's I using the Queen matrix is 0.04 for aggregation and slightly lower at 0.03 for disaggregation, while using the k-NN matrix, the values are 0.12. These results indicate a strong similarity in global spatial autocorrelation between the reference and disaggregated data, suggesting that the MAW-CZ method effectively preserves the spatial structure in Toulouse. In Grenoble, Moran's I values are relatively higher overall, reflecting a stronger spatial autocorrelation in the population distribution. Using the Queen matrix, Moran's I is almost similar, being 0.16 in the aggregated data and 0.14 after disaggregation. With the k-NN matrix, Moran's I shows a slight increase from 0.20 to 0.26, indicating that the disaggregation process may slightly enhance spatial autocorrelation when considering a fixed number of nearest neighbors. These results suggest that the global spatial structure of the population remains comparable between the disaggregated and reference datasets, particularly in Toulouse. The slightly higher spatial autocorrelation in Grenoble reflects its more compact and topographically constrained urban form. Therefore, the MAW-CZ method, when used for disaggregation, seems to preserve the global spatial organization of the population distribution.

https://doi.org/10.48088/ejg.n.tou.16.2.268.285

Table 7. Moran's I on population values (RSU level)

City	Method	Moran's I (Queen)	Moran's I (k-nearest neighbors)
Toulouse	Upscaling	0.04	0.12
	Downscaling	0.03	0.12
Grenoble	Upscaling	0.16	0.20
	Downscaling	0.14	0.26

To evaluate whether the disaggregation process introduces spatially structured biases, we computed Moran's I not only on the population values but also on the error distributions between the disaggregated and reference datasets (table 8). We used Moran's I on the absolute error to detect whether areas of higher inaccuracy are spatially concentrated. A low Moran's I on errors would indicate a random distribution of inaccuracies, suggesting no spatial bias introduced by the method. In contrast, a significant positive autocorrelation would imply that the method performs poorly in specific spatial contexts, revealing spatial weaknesses in the disaggregation model.

Table 8. Moran's I on absolute errors values (RSU level)

City	Moran's I (abs. errors)	p-value
Toulouse	0.01	0.36
Grenoble	0.14	<0.0001

In Toulouse, errors appear to be spatially random, suggesting that the disaggregation does not introduce spatial bias. However, in Grenoble, the significant Moran's I value indicates that errors tend to cluster slightly spatially, particularly in morphologically complex or heterogeneous areas. This suggests that the method may under- or overestimate population in specific urban contexts, where the proxy used (e.g., developed building surface) may not fully capture local demographic variability.

While Moran's I captures global structure, LISA (Anselin, 1995) identifies local spatial clusters such as high-population RSUs surrounded by high values (High-High) or outlier patterns (High-Low, Low-High). Table 9 compares LISA classifications between the aggregated and disaggregated datasets to assess the method's ability to replicate localized population structures (For the maps of LISA, see Figure A.5 in appendix).

Table 9. LISA cluster classification (% of inhabited RSUs)

City	Method	High-High(%)	High-Low (%)	Low-High(%)	Low-Low(%)	Non-significant(%)
Toulouse	Upscaling	0.59	15.85	17.64	1.45	64.67
	Downscaling	0.76	16.80	17.10	1.10	63.64
Grenoble	Upscaling	0.72	11.17	14.84	1.03	72.24
	Downscaling	0.41	10.20	13.63	0.95	74.82

Although the overall distribution of cluster types is similar between the two datasets, a substantial proportion of RSUs fall into spatial outlier categories (High-Low and Low-High), indicating localized mismatches in population clustering. This finding corroborates previous research (e.g., Mennis, 2015; Eichhorn, 2020), which demonstrates that while dasymetric methods are effective in capturing broad spatial patterns, they may face challenges in accurately representing finer local spatial variations, particularly in areas where proxy data provide limited information.

4.5. Classifying RSU categories

Finally, an additional comparative approach was employed by classifying RSUs into six categories based on population thresholds to evaluate the effectiveness of the downscaling method. These classes were initially generated using the K-means clustering algorithm on the number of inhabitants and then manually adjusted to enhance the interpretability and visual coherence of the resulting maps.

Tables 10 and 11 are confusion matrices of the population according to the upscaling and downscaling methods. The diagonals of the confusion matrices, in bold, show the percentage of RSUs correctly classified by the downscaling method.

For the cluster with 1–92 inhabitants, 94% of the RSUs were classified in the same cluster according to the downscaling method in Grenoble-Alpes metropole. The downscaling method appears to classify the RSUs correctly based on the population categories. The percentages on the diagonals are higher than 50% for all categories in both metropoles.

These confusion matrices allow to calculate an indicator of accuracy of the downscaling method, defined as the percentage of RSUs correctly classified into their respective population categories. Specifically, accuracy (Acc) is calculated as the ratio of correctly classified RSUs to the total number of RSUs, multiplied by 100:

$$Acc = \frac{number\ of\ RSU\ correctly\ classified}{total\ number\ of\ RSU} \times 100, \tag{8}$$

The calculated accuracies are:

$$Acc_{TM} = \frac{5229}{5993} \times 100 = 87.25\%, \tag{9}$$

which corresponds to 87.25% accuracy in Toulouse Metropole;

$$Acc_{GAM} = \frac{3198}{3661} \times 100 = 87.35\%, \tag{10}$$

which corresponds to 87.35% accuracy in Grenoble-Alpes Metropole. These high accuracy rates confirm the method's ability to reliably classify RSUs according to population categories, reinforcing the notion that the fine-scale spatial distribution is broadly well reproduced. This finding complements the previous results obtained with Moran's I and LISA indices, which showed good preservation of both global and local spatial structure in the disaggregated data. Indeed, the consistency between accurate RSU classification and preserved spatial autocorrelation suggests that the method maintains not only population quantities but also their spatial distribution — a critical aspect for reliable and detailed geographic analyses.

Table 10. Confusion matrix for the population in the Toulouse Metropole. The bold elements show the percentage of RSUs correctly classified by the downscaling method.

Upscaling/Downscaling	1–100	100–260	260–545	545–1,025	1,025–2,364	2,364–2,500	Number of RSUs	Number of inhabitants
1–100 [1]	**94%** [2]	6.2%	0.2%	<0.1%	0%	0%	3,888	148,041
100–260	13%	**78%**	8.6%	0.1%	0%	0%	1,389	227,336
260–545	2.3%	21%	**73%**	3.1%	0%	0%	521	189,780
545–1,025	1.3%	4.5%	26%	**65%**	3.8%	0%	156	113,089
1,025–2,364	0%	5.3%	0%	21%	**74%**	0%	38	50,203
2,364–2,500	0%	0%	0%	0%	0%	**100%**	1	2,429
Number of RSUs	3,835	1,444	551	128	35	0	5,993	
Number of inhabitants	151,328.7	233,697.6	197,399.8	94,308.21	46,771.5	0		
Overall Accuracy (%)								**87.25**

[2]1-100 : RSUs category with 1 to 100 inhabitants

[3] 94% of the RSU from the category 1 to 100 inhabitants are correctly classified by the modified areal weighting by control zones method in Toulouse metropole.

Table 11*.* Confusion matrix for the population in the Grenoble-Alpes Metropole. The bold elements show the percentage of RSUs correctly classified by the downscaling method.

Upscaling/Downscaling	1–92	92–265	265–620	620–1,792	1,792–3,292	3,292–6,500	Number of RSUs	Number of inhabitants
1–92	**94%**	5.7%	<0.1%	0%	0%	0%	2,340	64,037
92–265	14%	**76%**	9.5%	0.3%	0%	0%	884	142,581
265–620	3.4%	20%	**73%**	3.7%	0%	0%	350	130,842
620–1,792	1.2%	1.2%	19%	**79%**	0%	0%	84	75,850
1,792–3,292	0%	0%	0%	0%	**100%**	0%	2	4,743
3,292–6,500	0%	0%	0%	0%	0%	**100%**	1	6,451
Number of RSUs	2,343	876	357	82	3	0	3,661	
Number of inhabitants	68,886.4	139,653.5	132,436.4	73,959.3	7,982.3	0		
Overall Accuracy (%)								**87.35**

5. Discussion

The methodological and spatial implications of downscaling census data from larger spatial units (IRIS) to finer scales (residential spatial units, or RSUs) were analyzed using the Modified Areal Weighting by Control Zones (MAW-CZ) method. The objective was to assess the method's ability to downscale population distributions at high spatial resolution while preserving key statistical and spatial properties of the original source.

Three complementary analytical steps were undertaken to evaluate the performance of the MAW-CZ method in the context of comparing aggregated and downscaled data derived from the same method. First, a descriptive statistical analysis was conducted to examine central tendencies and dispersion of population values at the finer RSU scale. This preliminary assessment confirmed the sparsity of inhabited units—27.2% of RSUs in the Toulouse and 21.9% in Grenoble—and comparable mean populations of 122 and 116 inhabitants per RSU, respectively, providing a reference baseline for evaluating the method's ability to reproduce plausible local population distributions.

Second, analysis of global spatial autocorrelation, measured by Moran's I, revealed an almost complete absence of spatial autocorrelation with the MAW-CZ method, except for a slight signal detected in Grenoble (0.143 with Queen contiguity). A Moran's I value close to zero indicates an almost random spatial distribution of values, suggesting that the method does not generate artificial clustering or marked dispersion effects in the downscaled data. The analysis of Local Indicators of Spatial Association (LISA) revealed a predominance of High-Low and Low-High clusters, indicating substantial local spatial heterogeneity characterized by sharp contrasts between adjacent units. This pattern suggests that the MAW-CZ disaggregation method successfully preserves the complexity of the underlying spatial structure, maintaining localized discontinuities rather than producing artificially smoothed or homogenized distributions.

Finally, the classification bias analysis by population classes provides important nuance to the previous results. It shows that while the method generally reproduces central and spatial trends well, its accuracy decreases for the most densely populated units. For example, in Toulouse, 94% of RSUs with fewer than 100 inhabitants are correctly classified, 78% for RSUs between 101 and 255 inhabitants, 72% for those between 256 and 500 inhabitants, while accuracy drops to 60% for RSUs with more than 500 inhabitants. However, this limitation affects only a small proportion of RSUs—8.4% in Toulouse and 7.2% in Grenoble—and does not exhibit spatially localized bias, indicating overall stability of the method across the territory. This step thus refines understanding of the optimal application conditions of the MAW-CZ method, while highlighting its limitations.

6. Conclusions

This study proposes a statistical assessment of the *Modified Areal Weighting by Control Zones* (MAW-CZ) method, applied to the disaggregation of population data aggregated at the IRIS mesh to spatial units close to the block (RSU) of finer resolution. By using individual-level data from the FIDELI dataset, aggregated at the IRIS mesh, this work extends previous research by Plumejeaud (2010) and Mennis (2015), testing the robustness of the method when applied to existing aggregated census data such as those published by INSEE (Plumejeaud, 2010; Louvet, 2015).

The results indicate that MAW-CZ produces reliable estimates in sparsely to moderately populated RSUs. However, its accuracy tends to decline in highly dense units, likely due to increased spatial heterogeneity and the effects of the Modifiable Areal Unit Problem (MAUP) (Monteiro et al., 2018; Fotheringham & Sachdeva, 2022). Nonetheless, the generally homogeneous distribution of classification errors suggests that no major localized spatial bias is introduced. These findings also highlight the limitations of the method in morphologically complex or heterogeneous urban areas, where the proxies used (e.g., built-up area) may not accurately reflect local demographic variability.

Despite these limitations, the results confirm the relevance of the MAW-CZ method for disaggregating census data aggregated at coarser spatial units, while emphasizing the importance of accounting for MAUP effects in spatial analysis (Openshaw, 1981; Fotheringham & Sachdeva, 2022). The method offers a more cost effectively approach to producing a more accurate representation of population distribution in urban environments, particularly when combined with detailed morphological and architectural urban units.

The successful application of this method to French social data suggests its potential transferability to other national contexts with statistical units comparable to IRIS. This opens up possibilities for international comparison and collaborative research in urban planning and environmental studies. Furthermore, the integration of tools such as Geoclimate, based on globally available OpenStreetMap data (Bocher, 2021), enhances the ability to link urban indicators, social data, and spatial models.

This work thus contributes to ongoing efforts to more finely integrate social and environmental data. It provides methodological and empirical insights useful to researchers, urban planners, and public decision-makers engaged in addressing the challenges of urban and climate dynamics at the local scale.

Funding: The authors thank LISST and the Transversal Scientific Initiatives (Urbanized Areas) at the Midi-Pyrénées Observatory for funding for the VUSOTEM research project and CASD (Centre d'accès sécurisé aux données) for access to the FIDELI confidential data through a secure environment offered by contract Ref. 10.34724/CASD.

Data Availability Statement: As noted in section 2.2, three datasets were used : IGN ; INSEE; Urban and Geographical indicators produced with GeoClimate tool and FIDELI data from CASD.

Acknowledgments: We thank Martha Evonuk, PhD, from Evonuk Scientific Editing (http://evonukscientificediting.com) for editing a draft of this manuscript.

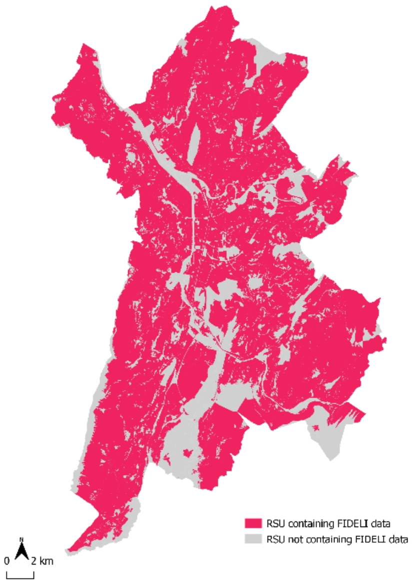
Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A.1. Number of outliers for the difference between the methods.

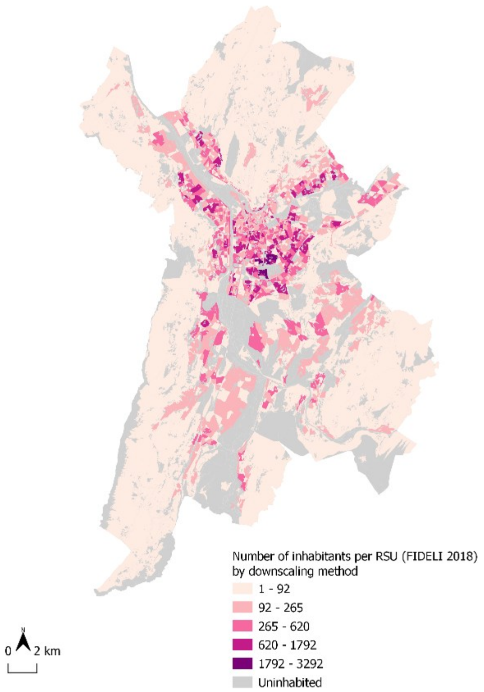
Method	Toulouse	Grenoble-Alpes
Outliers	410/5993*	329/3661

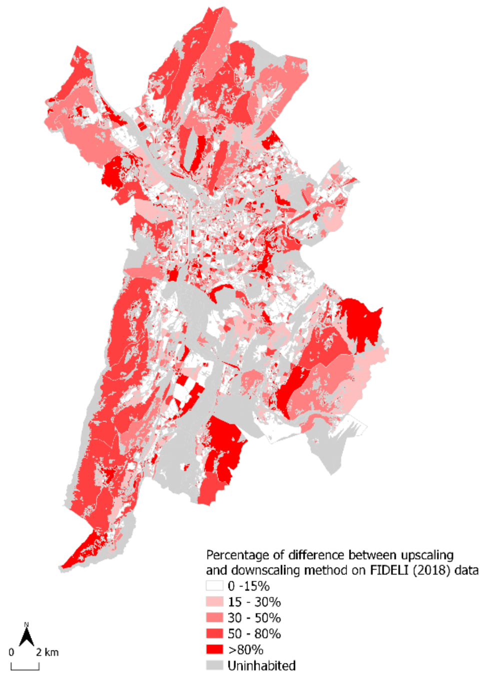
*410/5993 (6.8%) of RSUs have a value of the number of inhabitants that is considered to be a outlier far from the average.



(a) **(b)**

Figure A.1. Spatial distributions of **(a)** inhabited and uninhabited RSUs and **(b)** residential RSUs in the Grenoble-Alpes.





(a) **(b)**

Figure A.2. Spatial distribution of the number of inhabitants in the Toulouse Metropole by RSU when downscaling with the modified areal weighting by control zones method **(a)** and map of the error rate (natural thresholds) compared with the reference dataset obtained by upscaling the FIDELI data **(b)**.

https://doi.org/10.48088/ejg.n.tou.16.2.268.285

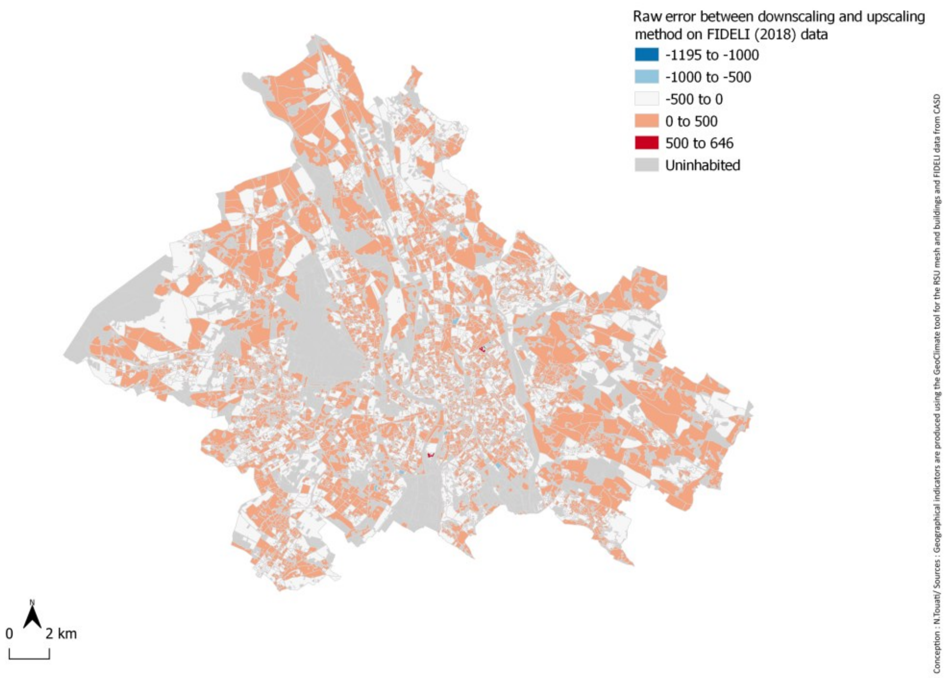


Figure A.3. Map of the raw error between the two methods in the Toulouse Metropole.

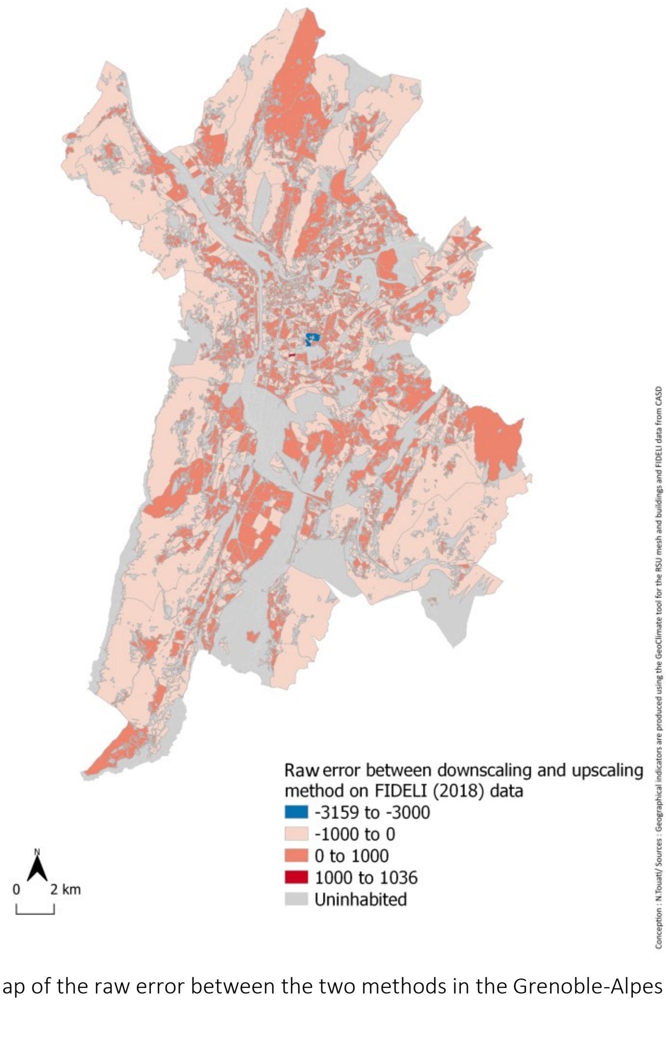
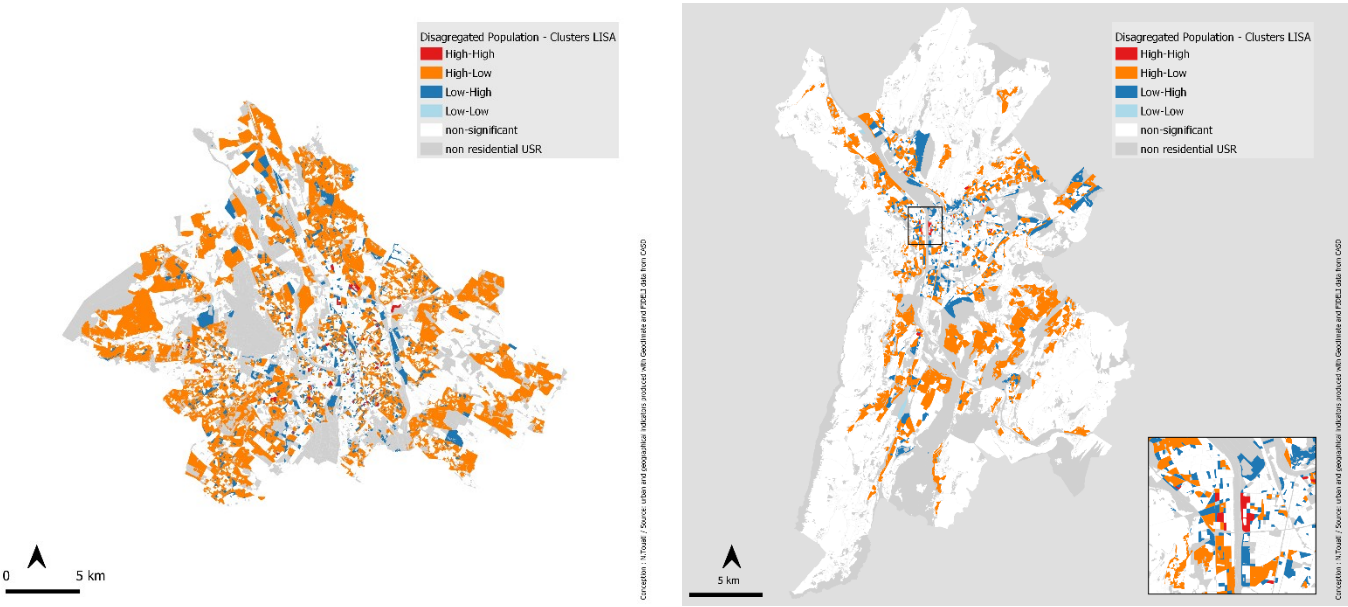


Figure A.4. Map of the raw error between the two methods in the Grenoble-Alpes Metropole.



(a) **(b)**

Figure A.5. LISA Cluster Typologies for Downscaled Population Data in the Toulouse Metropole **(a)** and the Grenoble-Alpes Metropole **(b)**

References

Anselin, L. (1995). Local Indicators of Spatial Association — LISA. *Geographical Analysis*, 27(2), 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Arbia, G. & others. (1989). Statistical effect of spatial data transformations: A proposed general framework. In Accuracy of spatial databases (p. 249-249). Taylor and Francis New York.

Batsaris, M., Zafeirelli, S., Vaitis, M., & Kavroudakis, D. (2023). PoD: A Web Tool for Population Downscaling Using Areal Interpolation and Volunteered Geographic Information. *European Journal of Geography*, *14*(4), 22–36. https://doi.org/10.48088/ejg.m.bat.14.4.022.036

Bocher, E., Petit, G., Bernard, J., & Palominos, S. (2018). A geoprocessing framework to compute urban indicators: The MApUCE tools chain. *Urban Climate*, 24, 153-174. https://doi.org/10.1016/j.uclim.2018.01.008

Bocher, E., Bernard, J., Wiederhold, E. L. S., Leconte, F., Petit, G., Palominos, S., & Noûs, C. (2021). GeoClimate: A Geospatial processing toolbox for environmental and climate studies. *Journal of Open Source Software*, 6(65), 3541. https://doi.org/10.21105/joss.03541

Do, V. H., Thomas-Agnan, C., & Vanhems, A. (2015). Spatial reallocation of areal data – another look at basic methods. *Revue d'Économie Régionale & Urbaine*, 1, 27-58. https://doi.org/10.3917/reru.151.0027

Fotheringham, A. S., & Wong, D. W. S. (1991). The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. *Environment and Planning A: Economy and Space*, 23(7), 1025-1044. https://doi.org/10.1068/a231025

Fotheringham, A. S., & Sachdeva, M. (2022). Scale and local modeling: New perspectives on the modifiable areal unit problem and Simpson's paradox. *Journal of Geographical Systems*, 24(3), 475-499.

Flowerdew, R., & Green, M. (1992). Developments in areal interpolation methods and GIS. *The Annals of Regional Science*, 26(1), 67-78. https://doi.org/10.1007/BF01581481

Goodchild, M., Anselin, L., & Deichmann, U. (1993). A Framework for the Areal Interpolation of Socioeconomic Data. *Environment and Planning A*, 25, 383-397. https://doi.org/10.1068/a250383

Hidalgo, J., Touati, N., Haouès-Jouve, S., Jégou, L., Bretagne, G., Bocher, E., Masson, V., Mayis, A., Jougla, R., Petit, G., & Schoetter, R. (2023). Mapping the urban climate to address summer comfort management in French urban planning documents. *Cybergeo : Revue européenne de géographie / European journal of geography*. https://doi.org/10.4000/cybergeo.40319

Insee & Ministère des Finances (DGFiP) [Producteur], Fichier DÉmographique sur les Logements et les Individus - 2018 [Fichiers de données], Centre d'Accès Sécurisé aux Données (CASD) [Diffuseur], http://doi.org/10.34724/CASD.295.3257.V1

King, G., Rosen, O., & Tanner, M. A. (Éds.). (2004). Ecological inference: New methodological strategies. Cambridge University Press.

Langford, M., & Unwin, D. (1994). Generating and Mapping Population Density Surfaces within a GIS. *The Cartographic Journal*, 31, 21-26. https://doi.org/10.1179/000870494787073718

Louvet, R., Josselin, D., Genre-Grandpierre, C., & Aryal, J. (2015). Impact du changement d'échelle sur l'étude des causes des feux de forêts du sud-est de la France. SAGEO1. https://ceur-ws.org/Vol-1535/paper-18.pdf

Masson, V., Hidalgo, J., Amossé, A., Belaid, F., Bocher, E., Bonhomme, M., Bourgeois, A., Bretagne, G., Caillerez, S., Cordeau, E., Demazeux, C., Faraut, S., Gallato, C., Haouès-Jouve, S., Lambert, M.-L., Lemonsu, A., Lestringant, R., Levy, J.-P., Long, N., … Vye, D. D. (2015, juillet). Urban Climate, Human behavior & Energy consumption: From LCZ mapping to simulation and urban planning (the MapUCE project). 9th International Conference on Urban Climate. https://hal.science/hal-01252761

Monteiro, J., Martins, B., & Pires, J. M. (2018). A hybrid approach for the spatial disaggregation of socio-economic indicators. *International Journal of Data Science and Analytics*, 5(2), 189-211. https://doi.org/10.1007/s41060-017-0080-z
Moran, P. (1950) A Test for the Serial Independence of Residuals. *Biometrika*, 37, 178-181. http://dx.doi.org/10.1093/biomet/37.1-2.178
Nordhaus, W. D. (2002). Alternative Approaches to Spatial Rescaling. (Yale University, New Haven, CT) https://gecon.yale.edu/sites/default/files/files/gecon_data_20051206_3.pdf
Openshaw, S., Taylor P.J. (1979). A million or so correlation coefficients: three experiments on the modifiable areal unit problem, *Statistical applications in the spatial sciences*, 1979, vol. 21, p. 127 -144.
Openshaw, S. (1981). Le problème de l'agrégation spatiale en géographie. *L'Espace géographique*, 10(1), 15-24. https://doi.org/10.3406/spgeo.1981.3599
Pivano, C., Bonin, O., & Hubert, J.-P. (2015). Transfert de support de données de déplacements : Quelles données de contrôle, et jusqu'à quelles échelles? *Journées de méthodologie statistiques*, 26p. https://hal.science/hal-01217297
Plumejeaud, C., Prud'homme, J., Davoine, P.-A., & Gensel, J. (2010). Transferring Indicators into Different Partitions of Geographic Space. In D. Taniar, O. Gervasi, B. Murgante, E. Pardede, & B. O. Apduhan (Éds.), *Computational Science and Its Applications – ICCSA 2010* (Vol. 6016, p. 445-460). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-12156-2_34
Robinson, W. S. (1950). Ecological Correlations and the Behavior of Individuals. *American Sociological Review*, 15(3), 351-357. https://doi.org/10.2307/2087176
Rome, S., Bigot, S., Foissard, X., Madelin, M., Duche, S., & Fouvet, A.-C. (2020). Les deux épisodes caniculaires de l'été 2019 à Grenoble : Constat et perspective pour une gestion des extrêmes thermiques futurs. *Climatologie*, 17. https://doi.org/10.1051/climat/202017012
Simpson, E. H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 13(2), 238-241.
Vignes, C., & Rimbourg, S. (2013). Méthodes statistiques d'allocation spatiale : Interpolation de données surfaciques. 79.
Wong, D. W. S. (2004). The Modifiable Areal Unit Problem (MAUP). In D. G. Janelle, B. Warf, & K. Hansen (Éds.), *WorldMinds: Geographical Perspectives on 100 Problems: Commemorating the 100th Anniversary of the Association of American Geographers 1904–2004* (p. 571-575). Springer Netherlands. https://doi.org/10.1007/978-1-4020-2352-1_93

European Journal of Geography