



Upscaling ecosystem service maps to administrative levels: beyond scale mismatches



Michele Zen ^{a,*}, Sebastian Candiago ^b, Uta Schirpke ^{b,c}, Lukas Egarter Vigl ^b, Carlo Giupponi ^a

^a Ca' Foscari University of Venice, Department of Economics, S. Giobbe 873, 30121 Venezia, Italy

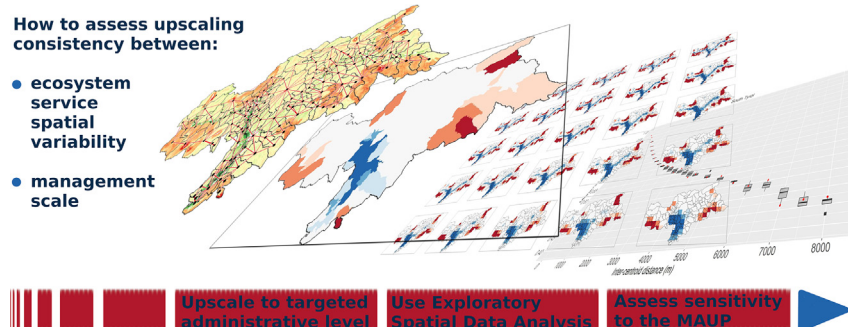
^b Institute for Alpine Environment, Eurac Research, Viale Druso 1, 39100 Bozen/Bolzano, Italy

^c Department of Ecology, University of Innsbruck, Sternwartestrasse 15, 6020 Innsbruck, Austria

HIGHLIGHTS

- We adopted a flexible methodology to understand how spatial relationships scale up.
- We used spatial autocorrelation to assess information loss through hypothesis testing.
- We explored three-grid displacements to estimate the modifiable areal unit problem.
- We detected clusters of the outdoor recreation ecosystem service at multiple scales.
- We found that only larger clusters persist up to the municipality level.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 30 July 2018

Received in revised form 31 December 2018

Accepted 8 January 2019

Available online 09 January 2019

Keywords:

Scale mismatches
Management levels
Outdoor recreation
Spatial autocorrelation
Upscaling
MAUP

ABSTRACT

As Ecosystem Services (ES) are the products of complex socio-ecological systems, their mapping requires a deep understanding of the spatial relationships and pattern that underpin ES provision. Upscaling ES maps is often carried out to avoid mismatches between the scale of ES assessment and that of their level of management. However, so far only a few efforts have been made to quantify how information loss occurs as data are aggregated to coarser scales. In the present study this was analyzed for three distinct case studies in the eastern Alps by comparing ES maps of outdoor recreation at the municipality level and at finer scales, i.e. high-resolution grids. Specifically, we adopt an innovative and flexible methodology based on Exploratory Spatial Data Analysis (ESDA), to disentangle the problem of the scale from the perspective of different levels of jurisdiction, by assessing in an iterative process how ES patterns change when upscaling high-resolution maps. Furthermore, we assess the sensitivity to the modifiable areal unit problem (MAUP) by calculating global statistics over three grid displacements. Our results demonstrate that spatial clusters tend to disappear when their extent becomes smaller than the features to which values are upscaled, leading to substantial information loss. Moreover, cross-comparison among grids and the municipality level highlights local anomalies that global spatial autocorrelation indicators fail to detect, revealing hidden clusters and inconsistencies among multiple scales. We conclude that, whenever ES maps are aggregated to a coarser scale, our methodology represents a suitable and flexible approach to explore clustering trends, shape and position of upscaling units, through graphs and maps showing spatial autocorrelation statistics. This can be crucial to finding the best compromise among scale mismatches, information loss and statistical bias that can directly affect the targeted ES mapping.

© 2019 Published by Elsevier B.V.

* Corresponding author.

E-mail address: michele.zen@unive.it (M. Zen).

1. Introduction

Ecosystem services (ES) are the products of complex interconnected socio-ecological systems that operate and interact at multiple scales (Scholes et al., 2013). The spatial visualization of ES in maps is an effective tool with great potential for the explanation of such complex phenomena (Burkhard et al., 2014), but spatial assessments are not often operationalized on solid or standardized frameworks (Primmer and Furman, 2012), neglecting, among others, important scale effects (Grêt-Regamey et al., 2014; Lü et al., 2013). A well-known scale-related issue is that policies can be effective only if their implementation matches the intrinsic scale of the problem under consideration (Wu and Li, 2006). Accordingly, several studies (Nahuelhual et al., 2015; Raudsepp-Hearne and Peterson, 2016) report that scale mismatches often arise when linking ES supply and related management objectives for societal needs, as the scale of environmental variation of an ES differs from the scale of the social organization that is responsible for its management (Cumming et al., 2006). Consequently, ES maps that are meant to drive ES management in policy-making need to prevent scale-related misinterpretation of their spatial pattern by selecting the scale, or the set of scales, that is both relevant for decision-makers and consistent with the mapped ES. The alignment of ES mapping outcomes to the related administrative unit (also referred to as “level of jurisdiction”) requires a solid knowledge of its management level and scaling rules over space (Grêt-Regamey et al., 2015). This process is often carried out by calculating the mean or sum of grid-based continuous values overlapping the shape of administrative units, either on primary data (Meacham et al., 2016) or on the output of ES calculation procedures (Nelson et al., 2009; Schirpke et al., 2018). In doing so, it is assumed that the fine-grain process is well described by its coarser aggregate, but this holds only if the process scales linearly (Cushman and Huettmann, 2010; Scholes et al., 2013).

To date, only a few studies have addressed the scaling behavior of ES spatial patterns, evaluating differences in information when aggregating ES maps to a coarser grain. These studies (Grêt-Regamey et al., 2014; Raudsepp-Hearne and Peterson, 2016; Rocés-Díaz et al., 2018) mainly assessed information loss through a family of techniques called Exploratory Spatial Data Analysis (ESDA). Thereby, the distribution and relationships among spatial features, i.e. spatial autocorrelation, can be analyzed using a series of indicators that highlight the level of spatial clustering of similar values (Anselin, 1996, 1995). These techniques were developed on the assumption that frequently spatial observations are clustered rather than independent, as explained by Tobler's first law of geography (Tobler, 1970). Like most spatial phenomena, ES tend to be heterogeneously distributed over space, forming clusters of different size and shape (Raudsepp-Hearne and Peterson, 2016). Therefore, to detect and compare variations of ES values at different scales, the above-mentioned studies used autocorrelation as proxy of the inherent scale-dependent information. More specifically, Raudsepp-Hearne and Peterson (2016) calculated Moran's I global spatial autocorrelation (Goodchild, 1986) on a 1 km and 3 km grid map and at municipality level. They found that most of the assessed ES show a high level of clustering and their pattern depends on the socio-ecological heterogeneity of the landscape. Both Grêt-Regamey et al. (2014) and Rocés-Díaz et al. (2018) calculated Moran's I incremental spatial autocorrelation, showing correlograms for a set of ES. They were able to show the spatial dependency per increasing distance class, respectively setting the bottom limit resolution at 25 m and at 1 km, where the upper lag sets the aggregation level. Despite the averaging effect that results in a loss of fine-grained information, both Raudsepp-Hearne and Peterson (2016) and Rocés-Díaz et al. (2018) proved that the municipality level still reproduces the underlying spatial pattern with reasonable accuracy. However, these studies do not estimate grid-based local spatial autocorrelation, i.e. the contribution of each individual cell with its neighbors to global indicators of spatial autocorrelation, such as Moran's I, which shows the overall degree of association of all features in a map

(Anselin, 1996). In fact, when calculating global statistics, spatial autocorrelation is estimated through a single score per each scale of analysis, which intuitively represents a global trend that might be in conflict with local associations of values or features. Moreover, despite the mentioned studies that described and assessed scale effects of different ES, practitioners in ES assessments still lack a common and standardized methodology to deal with the loss of information when upscaling maps from the scale of the ES process to the management level, i.e. aggregating information from finer scales (smaller grain size) to broader scales (larger grain size) (Wu and Li, 2006).

To enhance the understanding of scale in ES assessments, we draw on the outcomes of the Interreg Alpine Space project “AlpES - Alpine Ecosystem Services - mapping, maintenance and management”, which produced a collection of ES maps for the entire European Alps (Schirpke et al., 2019). Using the example of the ES “Outdoor recreation” we increase the spatial grain of the initial 100 m resolution map gradually to the municipality level, aiming to develop a novel, multi-step methodology based on existing ESDA techniques, in order to:

- evaluate whether statistically significant clusters of recreational values detected from a fine-grain representation (100 m grid) are consistent with their related coarser aggregates, up to the management level (in our example, the municipality level);
- understand how information loss occurs when upscaling grid-based data at local level i.e. if the aggregation process hides any important spatial pattern that might lead to a potential misinterpretation of spatial data when pursuing management objectives.

To evaluate the influence of site-specific characteristics, such as environmental variability, size and number of administrative units and socio-economic conditions, we test our approach in three distinct regions of the European Alps, namely Alto Bellunese, South Tyrol and Innsbruck. Related case-specific results are presented as a practical implementation of our methodology, which conceptually does not target specific objectives or ES. In fact it represents a novel, multi-step approach meant to deal with the loss of information when increasing the spatial resolution of any type of ES from the scale of the ES process to the management level.

2. Materials and methods

In this section, we first present the spatial autocorrelation indicators of ESDA to provide an overview of the algorithms grounded in existing literature and relevant to our study. Secondly, we describe our methodology, i.e. how we use existing spatial autocorrelation indicators in our novel multi-step approach to address the research objectives. Before assessing spatial autocorrelation, the methodology replicates the same aggregation process that scientists usually carry out while averaging ES values overlapping spatial features (Nelson et al., 2009; Schirpke et al., 2018). Consequently, the interpretation of the main results is rather simple, as the methodology highlights spatial inconsistencies among scales through graphs and maps, e.g. where a cluster of values disappears in the upscaling process. The modifiable areal unit problem (MAUP) is also assessed through an iterative approach. Lastly, we introduce the case study areas and the ES we selected to present our methodology. Since the applicability of the presented methodology does not depend on specific objectives and it does not target a particular ES, in the present study, the concept of management objective has to be intended as a general purpose fulfilled by an action or policy drawn on the outcomes of an ES assessment. In fact, management objectives targeting ES are often framed between a management level and one or multiple scales of observation that match the biophysical variability of the ES to be managed (Raudsepp-Hearne and Peterson, 2016). In this study, scale has to be intended as a hierarchical spatial

representation of an ES pattern. Among all components that qualify scale, the grain represents:

- the smallest unit of variability of a spatial pattern at a targeted scale, within which homogeneity is assumed;
- the size and number of cells included within the extent of a targeted area (Wu and Li, 2006).

Therefore, the grain is applied here as a convenient metric to match scale with pattern of ES. More specifically, in our analyses we refer to “scale” or “local level” when taking into consideration a spatial unit that is smaller, or at least comparable, with the size of a municipality. On the other hand, the term “municipality level” refers to municipalities as administrative units i.e. LAU2 level (Eurostat).

2.1. ESDA tools for spatial analyses

To explore spatial clustering over time, several indicators of spatial autocorrelation have been developed within the broader family of techniques called ESDA (Anselin, 1996): among these, Global Moran's I, Local Moran's I and Getis-Ord G_i^* .

Global Moran's I statistic is a measure of the spatial dependency and non-stationarity of spatial features and indicates spatial clustering of similar values (Anselin, 1995). It can be interpreted as the degree of linear association between observed values and the weighted average (spatial lag) of their neighbors (Anselin, 1996). Moran's I values range between 1 and -1 . Values close to 0 imply randomness; values close to 1 mean positive spatial autocorrelation, i.e. similarity among values; values close to -1 indicate negative spatial autocorrelation, i.e. dissimilarities among values. The spatial autocorrelation is an inferential statistic that is interpreted through hypothesis testing. The null hypothesis is the complete spatial randomness or, in other words, the independence of observed values from neighboring polygons or cells. The statistical significance of the presence of clusters is tested by comparing z-scores and p-values with the critical values of a normal distribution (Goodchild, 1986).

Anselin Moran's I local statistic differs from the Global Moran's I because it is calculated for each single deviation of observed values and not for their sum. Nevertheless, they share the same statistical interpretation. The Getis-Ord G_i^* statistic, also known as Hotspot Analysis, measures the degree of association resulting from the spatial concentration of the weighted values of each cell/polygon and its neighbors within a specified distance (Getis and Ord, 1992). The Getis-Ord G_i^* is already conceptualized as a Z-score (Getis and Ord, 1995) and implies positive spatial autocorrelation both for highly positive and negative Z-scores, i.e. high/high local relationships represent hotspots while low/low sets of values are coldspots. Both local statistics are calculated for every value x_i , resulting in a map where the clustering tendency can be visually explored to highlight local instabilities and multiple sources of spatial dependence (Anselin, 1996, 1995).

Further information and applied algorithms are reported in Supplementary material A.

2.2. The methodology

As reported by Cushman and Huettmann (2010), the upscaling of a spatial variable to a higher hierarchical level (i.e. municipality) can be carried out by calculating the mean of grid-based continuous values within the boundaries of selected administrative units. If we assume that grid-based cells are hypothetical sub-administrative units, even without a connection to any level of social organization, we can replicate the same aggregation process in an iterative and sequential procedure to monitor the loss of information across multiple scales. This is carried out with ESDA spatial autocorrelation statistics through global indicators and maps showing local statistics. With these tools we aim

at detecting clusters with statistically significant p-values in order to reject the null hypothesis, i.e. the assumption of complete spatial randomness.

To find a suitable mapping scale, consistent with its related fine-scale pattern and useful for management purposes, our multi-step procedure includes three steps (Fig. 1). First, ES values are aggregated to predefined grids with a sequentially increasing centroid distance (CD), covering a range of scales up to a grain size comparable with the mean of the areas of management units e.g. municipalities. Secondly, spatial indicators of autocorrelation are calculated over each scale of analysis as described in Section 2.1. The comparison of results among grids and with management units, and associated statistical inferences may reveal hidden patterns at coarser scales. At last, the MAUP can be assessed by displacing the original grids in alternative directions, to recalculate Global Moran's I and detect to which extent spatial indicators return different results. The methodology is flexible enough to be adjusted to the specificity of each individual study, e.g. selection of CD, targeted administrative level and ES to be assessed.

2.2.1. Aggregation of ES values

In our methodology, grid cells are assumed to be sub-administrative units with a regular shape, defined as an *a priori* level of jurisdiction without an existing connection with the real world. An empty grid is created with a targeted CD per each scale of analysis. The mean is calculated per each cell overlapping the raster map that displays the spatial variable to be upscaled. Control over borders is implemented, based on thresholds proposed by Raudsepp-Hearne and Peterson (2016): for CD equal or lower than 1 km, only the cells of the grid that entirely overlap the raster map are selected; for CD greater than 1 km, cells are kept if their area overlaps at least 50% of the raster map. The aggregation is carried out independently from the position of the grid. This is consistent with the fact that whenever the upscaling to a target organizational level is carried out, the shape and the position of administrative units cannot be chosen. As a general rule, the upper upscaling threshold is set to the CD that approximately equals the square of the mean of the administrative unit areas under investigation.

The aggregation procedures were coded in R, using the packages: sp, raster, rgdal and rgeos (R Core Team, 2018).

2.2.2. Calculation of spatial indicators of autocorrelation

The Global Moran's I spatial autocorrelation statistic is calculated per each scale of analysis. The approach is similar to that of correlograms, already explored by Grêt-Regamey et al. (2014) and Rocés-Díaz et al. (2018). However, instead of estimating Moran's I over each distance class (spatial lag, i.e. the mean of values of neighboring cells within a bounded upper distance) the statistics are calculated over upscaled maps whose i th cell is already the result of the aggregation process. Therefore, the statistic is calculated for a row-standardized spatial weight matrix based on first-order contiguity (Fig. 2), as in the case study proposed by Anselin (1995).

Local statistics aim at decomposing the contribution of each single cell to global autocorrelation outcomes. This lets us visualize the presence of statistically significant clusters at multiple scales. Local Moran's I spatial autocorrelation and Getis-Ord G_i^* are calculated per each scale of the analysis, respectively using a row-standardized and binary spatial weight matrix based on first-order contiguity (Anselin, 1995; Getis and Ord, 1995). Resulting maps report Z-score and p-value that can be visually explored, in order to detect clusters, hotspots and coldspots of the upscaled variable. Cluster detection is enhanced by applying the “False Discovery Rate” correction to formerly calculated p-values (de Castro and Singer, 2006). The comparative target is the chosen administrative level, acknowledged to be the management level under investigation, over which spatial autocorrelation statistics can be calculated and compared with grid results.

The spatial analyses were carried out in R, using the package: spdep (R Core Team, 2018).

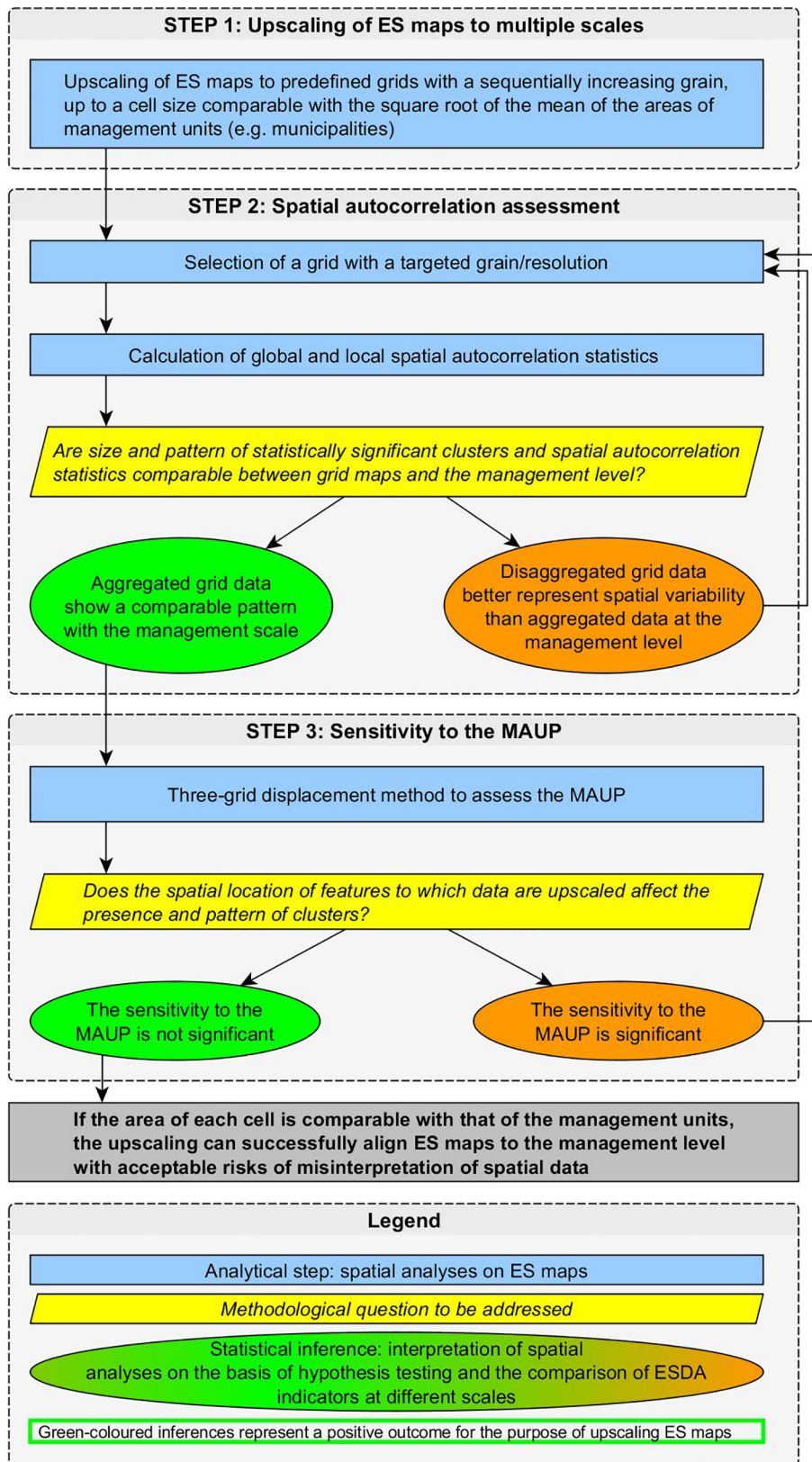


Fig. 1. Conceptual approach to identify the suitable mapping scale, consisting of three steps.

2.2.3. Assessment of the MAUP for randomly selected grids

The sequential, step-by-step upscaling to bigger cells is affected by a methodological bias, which is intrinsic to any kind of aggregation procedure at administrative level. The shape, the position and the number of

administrative units are always fixed. The consequence is that resulting aggregated values are affected by the MAUP, i.e. the outcomes of the upscaling process depend on shape and size of the aggregation units which might have strong implications on statistical analysis based on

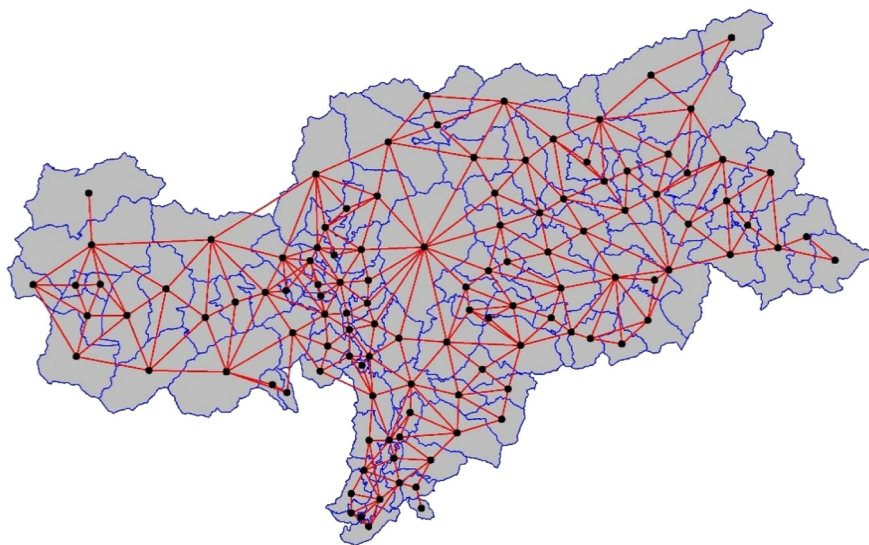


Fig. 2. Visual representation of the first-order contiguity: spatial relationships among units having a common boundary.

hypothesis testing (Dark and Bram, 2007; Spake et al., 2017). To entirely avoid the MAUP problem, according to Jelinski and Wu (1996), one should analyze each individual element linked to the process that generates the assessed spatial pattern, rather than aggregate them at some kind of “artificial” unit. This is not feasible while working with large spatial datasets. In our case, having assessed spatial autocorrelation on a single grid, it is worth investigating *ex-post* if grids with different positioning might reduce information loss. Therefore, we displace the extent of the original grid in three directions: 1/2 CD up, 1/2 CD right and together in both directions i.e. diagonally. By including the main grid configuration, it is possible to aggregate data and calculate Global Moran's I, on four different grids that overlap each other for 1/2 or 1/4 of their area. Moving the main grid in all Queen's case directions would result in redundant grid configurations. Results can be presented in a boxplot, showing the range of variability per each scale and the performance of the main grid. This range is a subset of the one that would be obtained from iterating all theoretically infinite grid positions. The method is meant to highlight the impact of the MAUP concerning the grid position in each scale assessed or, in other words, the effect of the size component, with a reasonable computational effort. This is consistent with the sensitivity analysis approach described by Jelinski and Wu (1996) since the MAUP affects the spatial relationships of cells/polygons under investigation and consequently the detection of clusters through autocorrelation indicators.

2.3. Outdoor recreation activities in the regions of Alto Bellunese, South Tyrol and Innsbruck

We selected the regions of Alto Bellunese (IT), South Tyrol (IT) and Innsbruck (AT) as our test sites. They are located in the central-eastern part of the European Alps (Fig. 3), straddling the border between Italy and Austria. The three areas represent contiguous territorial units subjected to different legislations and include municipalities with different size, topographic and socio-economic conditions. The overall region comprises 11,820 km² and spans an elevation gradient between 185 and 3851 m a. s. l. Its typical mountain geography considerably influences human activities, due to the high topographic variability and the harsh climate conditions. As a consequence, the land use composition is characterized by a high presence of natural and semi-natural ecosystems interspersed with urbanized areas and agricultural lands, usually located in the fertile valley floors.

The ES supply of outdoor recreation activities that was used to assess the scale effects within the selected test regions, was mapped in terms

of recreation opportunities (ES potential) provided by ecosystems and weighted by accessibility (Schirpke et al., 2018). Recreation opportunities were estimated on the basis of six spatial indicators: naturalness, protected areas, presence of water, landscape diversity, terrain ruggedness and density of mountain peaks. All spatial indicators were mapped as raster data with a spatial resolution of 100 m and overlaid after rescaling them to values between 0 and 100 (Fig. 3). The level of accessibility was calculated on the basis of the road and trail network by using travel time from residential areas and was rescaled to values between 0 and 1. Recreation supply was finally mapped by multiplying the recreation potential by the level of accessibility. Full details on methods and data sources can be found in Schirpke et al. (2018).

3. Results

3.1. Scaling relationships and spatial autocorrelation indicators

As described in Section 2.2.1, the square root of the mean of the municipality areas sets the upper upscaling threshold. Table 1 reports these thresholds and grids with target CD, starting from the finest scale which equals the resolution of the maps showing outdoor recreation activities in the three case study areas.

Global statistics of grid aggregates are presented in Fig. 4. Red points represent the Moran's I indices of the grids for which local statistics have been calculated. Each boxplot is created with four global results, the one of the grid in the original position and those whose cells are displaced in the Queen's case directions, as described in Section 2.2.3.

Global results at municipality level are shown in Table 2.

Fig. 5 summarizes the main results of the hotspot analysis, individually carried out over each selected case study area: it shows the fine-grain hotspots of recreational values at 100 m and its corresponding aggregate at the municipality level. All the other maps displaying local statistics are shown in Supplementary material A. These include both Local Moran's I and Getis-Ord G_i^* Z-score maps, along with their related p-values for both each scale of analysis and each case study, as reported in Table 1. Furthermore, Supplementary material A also reports spatial autocorrelation statistics of the ES supply of outdoor recreation activities over the whole Alpine arc. Despite being beyond the scope of this study, this analysis gives an overview of the hotspot pattern of recreational values at a higher hierarchical level, which share the same unit of variability as our test regions, i.e. the municipality level.

Since the presented methodology implies the upscaling of fine-grain data to a targeted administrative level, in the following sections we

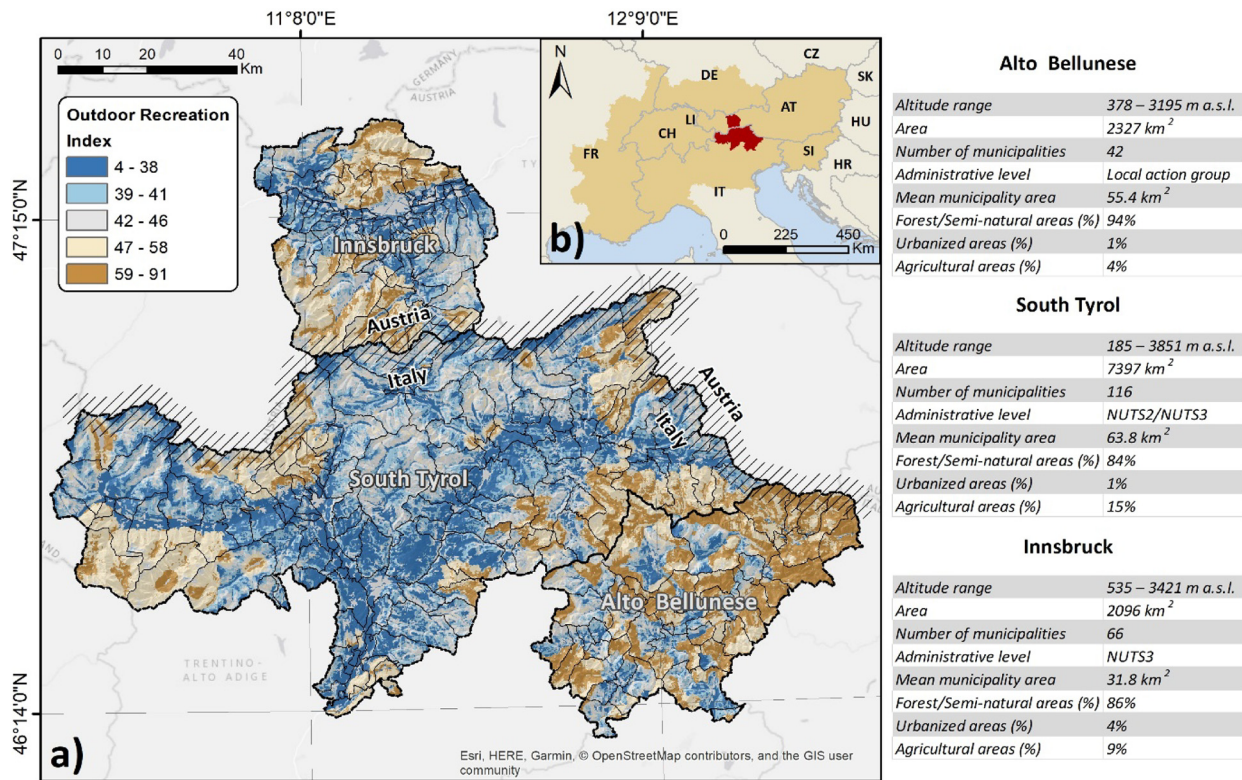


Fig. 3. Selected test regions, the supply ES outdoor recreation index (a) and their localization in the Alpine space cooperation area (b).

present our results from the finest to the coarsest scale, up to the comparative target.

3.2. Alto Bellunese

In the area of Alto Bellunese, local spatial autocorrelation analyses at 100 m CD revealed a substantial heterogeneity of recreational values with hotspots and coldspots closely associated along the whole region, except for the north-eastern part. This area corresponds to a hotspot that is statistically significant over all the scales of analysis, including the municipality level (Fig. 5). The aggregation implied a significant loss of information both in terms of quantity (the scale decreases together with the Z-score range) and allocation (the grain becomes unsuitable for providing spatial information concerning the hotspot location). Up to 1 km resolution, the smallest clusters were lost and the main hotspots and coldspots were still visible and increasingly easier to be identified. The information loss occurred almost linearly within 3000 m CD. After this threshold, it became non-linear and increasingly affected by the MAUP (Fig. 4). The number of clusters exponentially dropped up to the grid with 7500 m CD despite the different number, distribution and shape of administrative features, resulting in a comparable cluster distribution between the grid with the biggest CD and the municipality level. From 2000 to 4750 m, the grids returned an overall

deviation of around 0.05 Global Moran's I index units, but the MAUP became substantial only after 5000 m CD, resulting in a difference of up to 0.14 units. The MAUP was evident by comparing the hotspot analysis with global statistics at 6000, 6500, 7000 and 7500 m CD and at the municipality level (Fig. 5 and Supplementary material A). The grid with 7000 m CD showed the presence of three coldspots, two with confidence intervals of 90% and one with 95%. At 6000, 6500 m CD and at the municipality level only one coldspot was visible (p -value > 0.1), while at 7500 m no coldspots could be found. The Global Moran's I statistic (red points in Fig. 4 i.e. global results corresponding to grid maps in Fig. 5 and Supplementary material A) was above the median only at 7500 m CD and it was the highest among considered grids. This means that the related configuration better captures the autocorrelation signal of the recreational pattern under investigation. Since the median across the top four scales showed comparable values, the spatial autocorrelation turned out to be particularly sensitive to the displacements within each scale. Surprisingly, the Global Moran's I indicator at municipality level was even higher than that of the grid with 7500 m CD: it means that the shape and the location of the municipalities are more suitable for representing local clusters within Alto Bellunese. Practically, the local analysis returned that the fine-grain hotspot pattern in the south-western part was completely lost over 4000 m CD and in the municipality level.

Table 1

CD of grids up to the chosen upper threshold based on the square root of the mean of the municipality areas, per case study.

	Square root of the mean of the municipality areas (m)	Chosen upper CD threshold (m)	Grid cell CD
Alto Bellunese	7443.79	7500	<ul style="list-style-type: none"> • 100 m CD; • from 250 m to 4000 m CD with a step of 250 m; • from 4500 m to 7500 m CD with a step of 500 m.
South Tyrol	7985.61	8000	<ul style="list-style-type: none"> • 100 m CD; • from 250 m to 4000 m CD with a step of 250 m; • from 4500 m to 8000 m CD with a step of 500 m.
Innsbruck	5635.60	5750	<ul style="list-style-type: none"> • 100 m CD; • from 250 m to 5750 m CD with a step of 250 m;

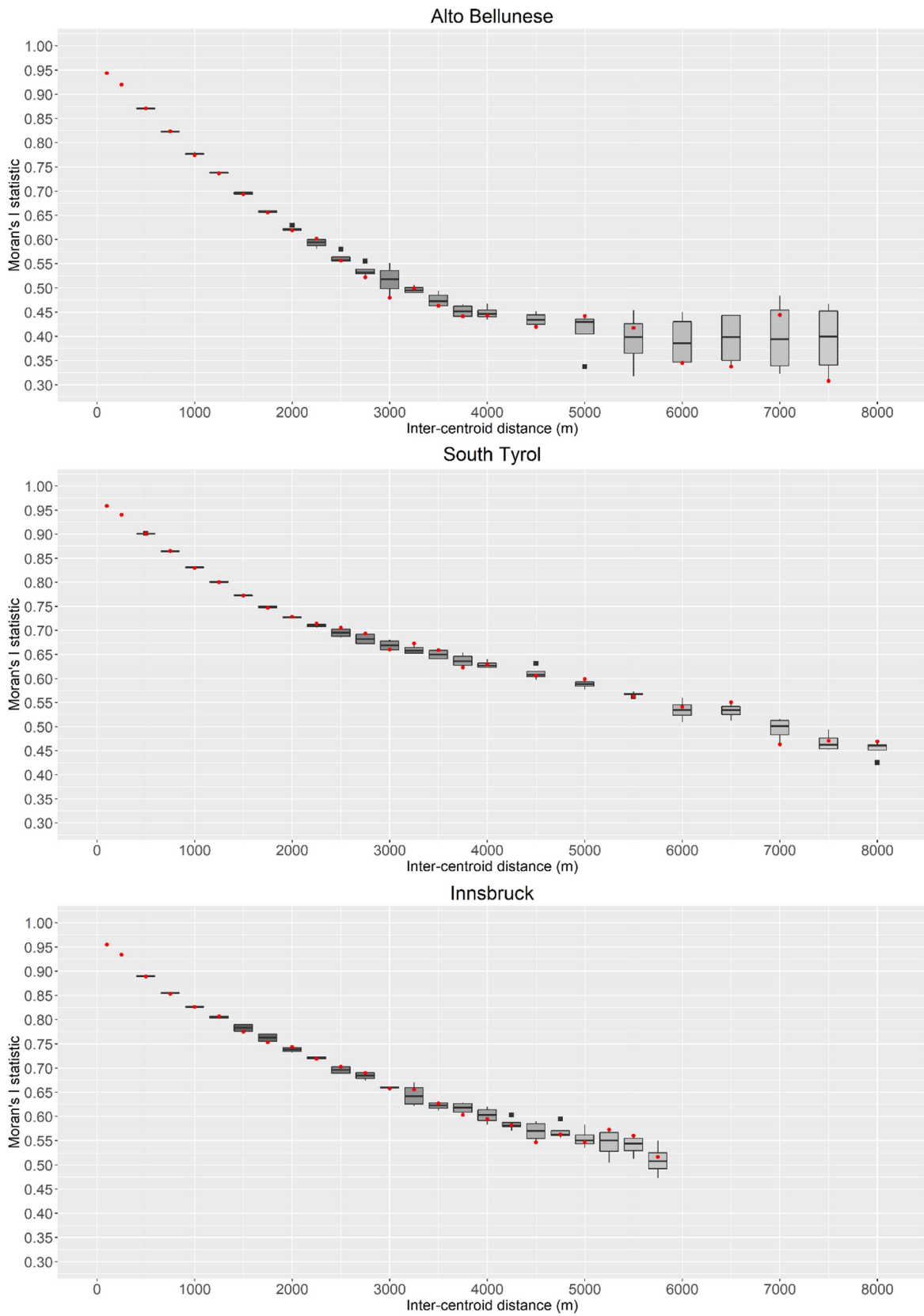


Fig. 4. Moran's I statistic calculated over sequentially aggregated grids; each boxplot is built with four grids with different position and same dimension, starting from 500 m CD; red points represent the grids without displacement; black squares are values that are statistically distant from other data within the same CD. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Global Moran's I at municipality level.

	Moran's I index	Z-score	p-value
Alto Bellunese	0.47	4.96	<0.01
South Tyrol	0.51	8.73	<0.01
Innsbruck	0.30	4.32	<0.01

3.3. South Tyrol

In South Tyrol, local spatial autocorrelation analyses showed a substantial agreement between the municipality level and the grids (Fig. 5). Local statistics returned highly significant clusters, homogeneously distributed over space, meaning that the aggregation of hotspots and coldspots do not influence each other greatly. Up to 1000 m CD, the smallest clusters disappeared or became part of the

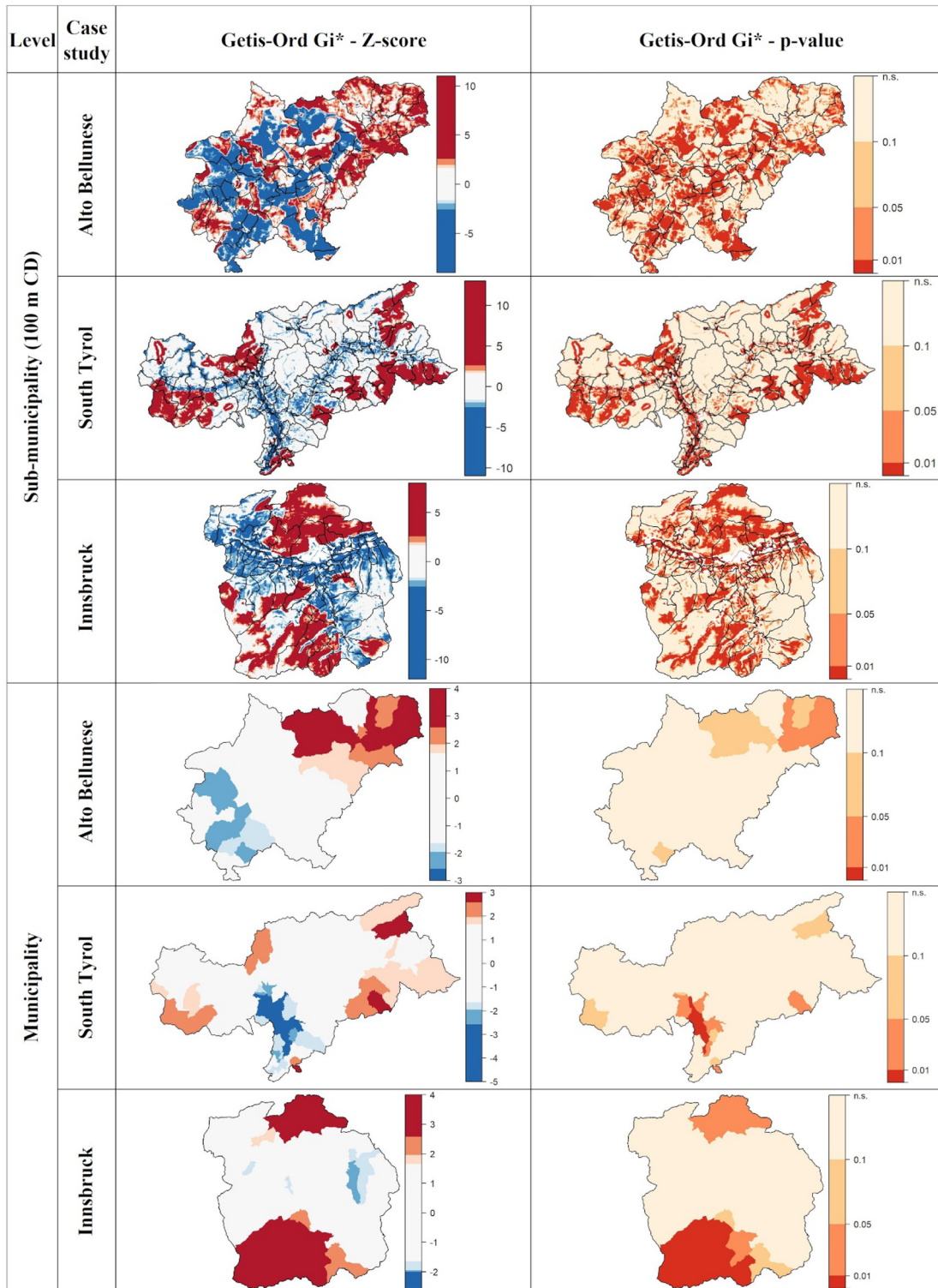


Fig. 5. Getis-Ord G_i^* Z-score maps and related p-values for each case study area: hotspot pattern at 100 m CD (top) and after the upscaling of recreational values to the municipality level (bottom).

main hotspots and coldspots, which were increasingly visible. Global Moran's I decreased linearly up to 2000 m CD, then once again almost linearly up to the threshold of 8000 m CD (Fig. 4). The aggregation clearly implied a loss of information but, with some exceptions, the main hotspots and coldspots were still visible in the grids with the highest CD as well as, by comparison, at the municipality level. This was also due to the reduced impact of the MAUP which accounts for about 0.05 Global Moran's I index units in the worst case (7000 m CD). The main difference between the fine-grain pattern and its related aggregate at coarser scales was the evident reduction of the extent of coldspots. At municipality level, they are limited to the central part of South Tyrol, between the municipalities of Merano and Bolzano.

3.4. Innsbruck

As for South Tyrol, local statistics in the Innsbruck region returned significant clusters homogeneously distributed over space, but with more closely associated hotspots and coldspots (Fig. 5). The Global Moran's I decreased almost linearly up to the threshold of 5750 m CD (Fig. 4), where local statistics returned four main clusters. The hotspots were clearly visible both in the grid and at the municipality level while no statistically significant coldspots were detected for the latter. Significant deviations were shown from 3250 m CD to 5750 m CD, reaching a maximum of 0.08 Global Moran's I index units. This means that the MAUP has some kind of impact in the aggregation, although smaller than in Alto Bellunese. Moreover, in the Innsbruck region, the grid-based aggregation performed better than the municipality level: even in the worst case for the grid with 5750 CD, the Global Moran's I index did not fall below 0.45 units while, at the municipality level, it reached only 0.3 units. It appeared that the polygon configuration of administrative units loses a significant amount of information.

Unlike in the previous test regions, in the area of Innsbruck it can be observed that (Fig. 4 and Table 2):

- the impact of the grid position is smaller than the difference between Global Moran's I values of the grids and the municipality level;
- the municipality level has a Global Moran's I smaller than the median of any other grids.

Since global statistics do not provide any information to explain these differences, local statistics were used to study global performances considering the local pattern. From Fig. 6, it is evident that statistically significant hotspots at the municipality level are surrounded by other municipalities with a relatively high G_i^* Z-scores. At the same time, coldspots show weak values, in some cases due to the aggregation over polygons with a longitudinal trend that covers areas with opposite Z-scores. By comparison, in grid maps it was observed that the coldspot pattern is reduced along with the upscaling of recreational values. On the other hand, the main hotspots are still visible in the top scale of the grid-based aggregation, which explains the higher Global Moran's I value. The most evident change in spatial pattern occurs at 3500 m CD, where the main coldspot that crosses the Innsbruck region from West to East splits into two parts, i.e. the last cell connecting them and showing a statistically significant p-value disappears (Supplementary material A).

4. Discussion and conclusion

ES maps are powerful tools that can drive the implementation of the ES framework in governance and decision-making processes. Specifically, to support management objectives and to be useful for policy-makers and stakeholders, maps must show the set of features/cells that represent the unit of variability over which a target organizational level might legislate, make decisions or act. In this

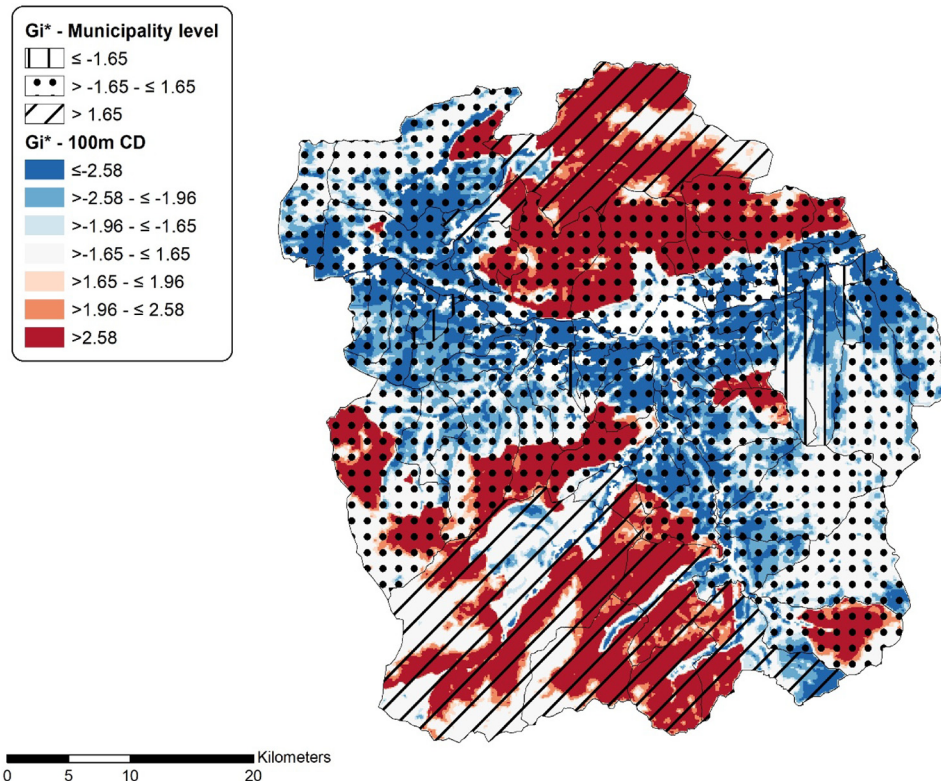


Fig. 6. Getis-Ord G_i^* comparison between the grid with 100 m CD and the municipality level in the Innsbruck NUTS3 region.

study, we reinterpreted existing ESDA tools in a novel multi-step methodology to assess upscaling performance when mapping ES at a targeted aggregated level. We focused only on the methodological issues without addressing specific targets, either management objectives or ES. As ESDA can be applied to any spatial data, consequently our methodology can be applied to any spatial variable or ES. This is important because scaling effects affect the assessment of ES depending on their unique relationship with terrain properties (Grêt-Regamey et al., 2014). For example, Grêt-Regamey et al. (2014) found minor differences when assessing carbon sequestration and a difference of up to 329% by quantifying timber production at alternative scales. If geomorphology plays a key role in the spatial quantification of ES, it is likely that sharp gradients would return a substantial spatial heterogeneity and consequently a loss of information when upscaling ES maps (Grêt-Regamey et al., 2014), as for outdoor recreation. Thus, if significant, such heterogeneity would be detected while applying our methodology to ES maps. Therefore, through global and local spatial autocorrelation analyses and the evaluation of the sensitivity to the MAUP, it was possible to decompose the problem of scale in ES mapping from the perspective of different levels of jurisdiction. Specifically, ESDA techniques were used to highlight spatial clusters among administrative features and grids with a sequentially increasing grain. Afterwards, the comparison between global indices among grids and the municipality level was used to analyze whether the latter was hiding some important patterns. Lastly, multiple grid displacements were applied to estimate the impact of the MAUP.

The methodology was tested on the case study of outdoor recreation in the Alps, where recreational values were aggregated prior to any spatial autocorrelation analysis, to address how spatial information of the selected ES behaves in the upscaling process. The first-order contiguity assumption helped us to simplify the interpretation of spatial autocorrelation statistics and enabled us to develop an iterative approach with reasonable computational effort that avoids sub-sampling procedures. The spatial pattern of the outdoor recreation ES at local level generally confirmed global outcomes, showing the step-by-step homogenization of the landscape of the assessed areas, from the finest to the coarsest scale. As the grain is expanded, the MAUP affects Global Moran's I results with an increasing magnitude. This was more evident in the area of Alto Bellunese where hotspots and coldspots were closely associated in a heterogeneous landscape with small and narrow valleys and sharp geomorphological gradients. Conversely, the South Tyrol region represents the example of an area where a linear aggregation shows a good performance due to relative homogeneity and clustering, with hotspots and coldspots that are not in opposition with each other. An intermediate situation was revealed in the Innsbruck region where, despite the fact that the main hotspots were detected, the municipality level was not able to show the presence of coldspots because of weak Z-scores and the MAUP. This seems to depend mainly on the shape of the municipalities, rather than on their location. In all three case studies, global spatial autocorrelation analyses returned an approximately linear loss of information below 2000 m CD (Fig. 4) and Moran's I indicators that converge to the same value among grids at equivalent scales. This CD accounts for roughly half of the range between the Global Moran's I indices of the coarsest and the finest scale considered, highlighting that at a finer scale the location of the grid is irrelevant to the result and that spatial processes are well described in terms of spatial allocation.

In our case study, and specifically in South Tyrol and in the region of Innsbruck, we can conclude that upscaled recreational values (and specifically, their statistically significant hotspots) are generally consistent with the related 100 m CD map. In Alto Bellunese, however, the heterogeneous pattern of relatively small clusters implies a substantial loss of information and either alternative aggregation strategies may be chosen to improve hotspot detection or the use of disaggregated data should be considered. A fine-grain spatial allocation of ES might be suitable to pursue management objectives. In this case, upscaling could be used to

highlight the main clusters, so as to spatially prioritize actions over sub-municipality areas and then switch to a more detailed cartography, e.g. for engineering purposes.

Although our methodology does not provide a unique solution to define the scale of detail for ES analysis, some generalizations can be inferred:

- clusters tend to disappear when their extent becomes smaller than the feature to which values are upscaled;
- the cross-comparison among grids and municipalities highlights local anomalies that grid displacements failed to detect through the Moran's I global spatial autocorrelation;
- in the grid-based approach, the median of the boxplots (Fig. 4) can be considered as an indicator of the overall level of clustered information preserved within each scale;
- the magnitude of the MAUP can be estimated with few iterations but cannot be prevented, unless a different aggregation strategy is chosen.

On the basis of these points, we can conclude that our methodology is effective to choose a suitable mapping scale in ES assessment. The comparative analysis between grids and the municipality level, through boxplots and local assessments, helped us to define the quality of the upscaling by using the spatial relationships as indicators of the change of the recreational pattern. This enabled us to estimate the loss of information through inferential statistics, using ESDA in a novel approach based on a multi-step procedure. Such a procedure is meant to provide a full set of data (maps and graphs) to study ES patterns in order to develop a global understanding of how spatial relationships scale up. This is crucial because avoiding scale mismatches to meet management objectives should not be the reason for information loss, which can lead to potential misinterpretation of mapping outcomes.

Acknowledgements

All authors are thankful to the editor and the two anonymous reviewers for their valuable comments. This study was co-financed by the European Regional Development Fund through the Interreg Alpine Space programme ('AlpES' project, CUP: D52I16000220007). The authors thank the Department of Innovation, Research and University of the Autonomous Province of Bozen/Bolzano for covering the Open Access publication costs. Furthermore, they thank Samuel Williams for the proofreading of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2019.01.087>.

References

- Anselin, L., 1995. Local indicators of spatial association — LISA. *Geogr. Anal.* 27, 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>.
- Anselin, L., 1996. The Moran Scatterplot as an ESDA tool to assess local instability in spatial association. *Spat. Anal. Perspect. GIS*.
- Burkhard, B., Kandziora, M., Hou, Y., Müller, F., 2014. Ecosystem service potentials, flows and demands—concepts for spatial localisation, indication and quantification. *Landsc. Online* 34, 1–32. <https://doi.org/10.3097/LO.201434>.
- de Castro, C.M., Singer, B.H., 2006. Controlling the false discovery rate: a new application to account for multiple and dependent test in local statistics of spatial association. *Geogr. Anal.* 38, 180–208.
- Cumming, G.S., Cumming, D.H.M., Redman, C.L., 2006. Scale mismatches in social-ecological systems: causes, consequences, and solutions. *Ecol. Soc.* 11.
- Cushman, S.A., Huettmann, F., 2010. Spatial complexity, informatics, and wildlife conservation. *Spatial Complexity, Informatics, and Wildlife Conservation* <https://doi.org/10.1007/978-4-431-87771-4>.
- Dark, S.J., Bram, D., 2007. The modifiable areal unit problem (MAUP) in physical geography. *Prog. Phys. Geogr.* 31, 471–479. <https://doi.org/10.1177/0309133307083294>.
- Eurostat, n.d. Nomenclature of Territorial Units for Statistics [WWW Document]. URL <https://ec.europa.eu/eurostat/web/nuts/local-administrative-units> (accessed 12.19.18).

- Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. *Geogr. Anal.* 24, 189–206. <https://doi.org/10.1111/j.1538-4632.1992.tb00261.x>.
- Getis, A., Ord, J.K., 1995. Local spatial autocorrelation statistics: distributional issues and an application. *Geogr. Anal.* 27, 286–306. <https://doi.org/10.1111/j.1538-4632.1995.tb00912.x>.
- Goodchild, M.F., 1986. *Spatial Autocorrelation. Concepts Tech. Mod. Geogr.* 47 pp. 1–56.
- Grêt-Regamey, A., Weibel, B., Bagstad, K.J., Ferrari, M., Geneletti, D., Klug, H., Schirpke, U., Tappeiner, U., 2014. On the effects of scale for ecosystem services mapping. *PLoS One* 9, 1–26. <https://doi.org/10.1371/journal.pone.0112601>.
- Grêt-Regamey, A., Weibel, B., Kienast, F., Rabe, S.E., Zulia, G., 2015. A tiered approach for mapping ecosystem services. *Ecosyst. Serv.* 13, 16–27. <https://doi.org/10.1016/j.ecoser.2014.10.008>.
- Jelinski, D.E., Wu, J., 1996. The modifiable areal unit problem and implications for landscape ecology. *Landsc. Ecol.* 11, 129–140. <https://doi.org/10.1007/BF02447512>.
- Lü, Y., Feng, X., Chen, L., Fu, B., 2013. Scaling effects of landscape metrics: a comparison of two methods. *Phys. Geogr.* 34, 40–49. <https://doi.org/10.1080/02723646.2013.778692>.
- Meacham, M., Queiroz, C., Norström, A.V., Peterson, G.D., 2016. Social-ecological drivers of multiple ecosystem services: what variables explain patterns of ecosystem services across the Norrström drainage basin? *Ecol. Soc.* 21. <https://doi.org/10.5751/ES-08077-210114>.
- Nahuelhual, L., Laterra, P., Villarino, S., Mastrángelo, M., Carmona, A., Jaramillo, A., Barral, P., Burgos, N., 2015. Mapping of ecosystem services: missing links between purposes and procedures. *Ecosyst. Serv.* 13, 162–172. <https://doi.org/10.1016/j.ecoser.2015.03.005>.
- Nelson, E., Mendoza, G., Regetz, J., Polasky, S., Tallis, H., Cameron, D.R., Chan, K.M.A., Daily, G.C., Goldstein, J., Kareiva, P.M., Lonsdorf, E., Naidoo, R., Ricketts, T.H., Shaw, M.R., 2009. Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Front. Ecol. Environ.* 7, 4–11. <https://doi.org/10.1890/080023>.
- Primmer, E., Furman, E., 2012. Operationalising ecosystem service approaches for governance: do measuring, mapping and valuing integrate sector-specific knowledge systems? *Ecosyst. Serv.* 1, 85–92. <https://doi.org/10.1016/j.ecoser.2012.07.008>.
- R Core Team, 2018. *R: A Language and Environment for Statistical Computing*.
- Raudsepp-Hearne, C., Peterson, G.D., 2016. Scale and ecosystem services: how do observation, management, and analysis shift with scale—lessons from Québec. *Ecol. Soc.* 21. <https://doi.org/10.5751/ES-08605-210316>.
- Roces-Díaz, J.V., Vayreda, J., Banqu e-Casanovas, M., D az-Varela, E., Bonet, J.A., Brotons, L., de-Miguel, S., Herrando, S., Mart nez-Vilalta, J., 2018. The spatial level of analysis affects the patterns of forest ecosystem services supply and their relationships. *Sci. Total Environ.* 626, 1270–1283. <https://doi.org/10.1016/j.scitotenv.2018.01.150>.
- Schirpke, U., Meisch, C., Marsoner, T., Tappeiner, U., 2018. Revealing spatial and temporal patterns of outdoor recreation in the European Alps and their surroundings. *Ecosyst. Serv.* 31C, 336–350. <https://doi.org/10.1016/j.ecoser.2017.11.017>.
- Schirpke, U., Candiago, S., Egarter Vigl, L., J ager, H., Labadini, A., Marsoner, T., Meisch, C., Tasser, E., Tappeiner, U., 2019. Integrating supply, flow and demand to enhance the understanding of interactions among multiple ecosystem services. *Sci. Total Environ.* 651, 928–941. <https://doi.org/10.1016/j.scitotenv.2018.09.235>.
- Scholes, R.J., Reyers, B., Biggs, R., Spierenburg, M.J., Duriappah, A., 2013. Multi-scale and cross-scale assessments of social-ecological systems and their ecosystem services. *Curr. Opin. Environ. Sustain.* 5, 16–25. <https://doi.org/10.1016/j.cosust.2013.01.004>.
- Spake, R., Lasseur, R., Crouzat, E., Bullock, J.M., Lavorel, S., Parks, K.E., Schaafsma, M., Bennett, E.M., Maes, J., Mulligan, M., Mouchet, M., Peterson, G.D., Schulp, C.J.E., Thuiller, W., Turner, M.G., Verburg, P.H., Eigenbrod, F., 2017. Unpacking ecosystem service bundles: towards predictive mapping of synergies and trade-offs between ecosystem services. *Glob. Environ. Chang.* 47, 37–50. <https://doi.org/10.1016/j.gloenvcha.2017.08.004>.
- Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit. *Econ. Geogr.* 46, 234–240. <https://doi.org/10.2307/143141>.
- Wu, J., Li, H., 2006. Concepts of scale and scaling. In: Wu, J., Jones, K.B., Li, H., Loucks, O.L. (Eds.), *Scaling and Uncertainty Analysis in Ecology: Methods and Applications*. Springer Netherlands, Dordrecht, pp. 3–15. https://doi.org/10.1007/1-4020-4663-4_1.