Journal of Official Statistics, Vol. 39, No. 3, 2023, pp. 285–324, http://dx.doi.org/10.2478/JOS-2023-0014

Looking for a New Approach to Measuring the Spatial Concentration of the Human Population

Federico Benassi[1], *Massimo Mucciardi*[2], *and Giovanni Pirrotta*[3]

In the article a new approach for measuring the spatial concentration of human population is presented and tested. The new procedure is based on the concept of concentration introduced by Gini and, at the same time, on its spatial extension (i.e., taking into account the concept of spatial autocorrelation, polarization). The proposed indicator, the Spatial Gini Index, is then computed by using two different kind of territorial partitioning methods: MaxMin (MM) and the Constant Step (CS) distance. In this framework an ad hoc extension of the Rey and Smith decomposition method is then introduced. We apply this new approach to the Italian and foreign population resident in almost 7,900 statistical units (Italian municipalities) in 2002, 2010 and 2018. All elaborations are based on a new ad hoc library developed and implemented in Python.

Key words: Spatial concentration; Gini index; constant step; maxmin distance; Italy.

1. Introduction

In population studies and, more generally, in the quantitative social sciences, concentration, and thus space, assumes a fundamental importance (Anselin 1999; Logan 2012; Howell et al. 2016). Indeed, people and firms tend to concentrate in space almost naturally, both to facilitate interactions, exchange ideas, goods and services, and to share the costs associated with survival itself. It is no coincidence that Aristotle, when defining man as a social animal ("zoon poolitkon"), refers to the concept of the arena, that is, space, and that he places this character of the individual at the birth of the polis (i.e., the city, concentrated in space, by definition). Population and space are therefore two closely interconnected and mutually dependent variables. Indeed, as Livi Bacci (1999) reminds us, not unlike other living species, humans need space to obtain the resources necessary for their survival, to maintain population growth and to organise themselves socially. On the other hand, space itself "depends" on human behaviour, not only in relation to its negative externalities, but also to the capacity to absorb human activity, the so-called "carrying capacity" (Verhulst 1838; Pearl and Reed 1920). From a more strictly statistical point of view, concentration is something intimately connected to the concept of variability, that is, the ability of a quantity to assume different values.

The statistical approach to measuring concentration is in fact essentially based on the concept of variability. But, as observed by Leti (1983), in reality the original concept of

[1] University of Naples Federico II, Deparment of Political Sciences, Via L. Rodinò 22, 80138, Naples, Italy. Email: federico.benassi@unina.it
[2] University of Messina, Department of Cognitive Science, Education and Cultural Studies, Messina, Italy. Email: massimo.mucciardi@unime.it
[3] University of Messina, IT Staff, Messina, Italy. Email: giovanni.pirrotta@unime.it

variability is, under certain conditions, entirely analogous to that of heterogeneity and homogeneity, as well as, to that of concentration.

In fact, for quantity distributions, any index of variability is also an index of concentration, and indeed indices of homogeneity, applied to quantitative characters, actually measure concentration (Novelli and Ocelli 1999). In this perspective, therefore, concentration would be nothing more than the variability associated with a transferable variable. On the basis of this approach to the study of concentration, which could be considered to be statistical, and which is aspatial, a vast literature was born and has developed. In this framework, one of the cornerstones is Gini's G index (Gini 1912, 1914, 1921, 2005), which in turn was based on the pre-existing Lorenz concentration curve (Lorenz 1905). This index has stimulated a growing school, especially among Italian statisticians and demographers, who have proposed many other indices of this type, that is, aspatial, for the study of concentration. One recalls among these, Bonferroni's index (Bonferroni 1938), and those of Zenga (1984, 1985) to which many others can be added as can be seen in Frosini (1996). But also in the international sphere, the work of Gini stimulated the definition of a number of indices, such as Wright's index (Wright 1937) and the well-known Hoover index (Hoover 1941).

However, the study of concentration according to an exclusively statistical approach has attracted increasing criticism over time. In particular, the major criticism levelled at these indices is that being aspatial indices, they deny the essentially spatial nature of the concentration process (Arbia 2001; Dawkins 2004, 2006). In essence, they do not consider the dimension of polarisation, that is, the geographical component of concentration that finds its analytical formulation in Tobler (1970) and the so called first law of geography. This lack assumes particular relevance in the measurement of all those processes, such as for instance residential segregation, where space is a foundational component (Reardon and O'Sullivan 2004). To make up for this shortcoming, spatial extensions of the concentration indices and the Gini index in particular have been proposed. Mention should be made here, among others, of those by Arbia and Piras (2009), Rey and Smith (2013), Crespo and Hernandez (2020), and Panzera and Postiglione (2020). Even more recently, Mucciardi and Benassi (2023), developing seminal ideas of Alleva (1987), Mucciardi and Bertuccelli (2007) and Mucciardi (2008b), have proposed an approach for measuring concentration by means of a spatial extension of the Lorenz curve that allows the definition of a spatial version of the Gini index.

In the present article a new approach for measuring spatial concentration of a human population is proposed and tested. The approach is partially based on that of Mucciardi (2008b) and Mucciardi and Bertuccelli (2007). Here, stress is placed on the effects of different neighbourhood structures on the computation of the Spatial Gini Index (SGI). In particular, two types of distances are proposed to determine the spatial structure of the territory: MaxMin (MM) and Constant Step (CS). Finally, an extension of the decomposition method of Rey and Smith (2013) is here proposed in the framework of SGI with MM and CS. The empirical approach is referred to Italian and foreign population resident in the Italian municipalities in 2002, 2010 and 2018. It is important to stress that given the huge number of statistical units (about 7,900) we have developed a specific computational procedure, writing an ad hoc library in Python.

This article is structured as follows. In the next section the Spatial Gini Index (SGI) is presented and discussed. Attention is paid to the way the territory can be partitioned. In this context two distance methods are presented: MM and CS. SGI is therefore presented as a measure to be built in this framework. Section 3 is devoted to the implementation of the decomposition method of Rey and Smith (2013) in the general framework of SGI and in the two decomposition methods here proposed. Section 4 presents the empirical results. The last section presents a discussion, conclusion and future developments. In the two appendices a description of the technical details of SGI with the two distance methods and the ad hoc library written in Python are provided.

2. The Spatial (Intrinsic) Dimension of Concentration and the Spatial Gini Index

Common to all spatial approaches is the recognition that, in the analysis of any phenomenon, consideration of the spatial dimension requires an adequate description of the spatial variability of the phenomenon itself (Matthews and Parker 2013). Stating that a phenomenon or a relation manifests spatial variability is equivalent to saying that that phenomenon or that relation is not spatially stationary. The presence of random fluctuations, the existence of differentiations in the perceptions and behaviour of individuals, the incompleteness or imprecise specification of the descriptive model assumed, are some of the possible causes of spatial non-stationarity (Fotheringham 1997). In consideration of this aspect, this new proposal of the Spatial Gini Index (SGI) is based on comparing how the contribution in terms of "connectivity" and "variability" varies as the geographical distance between spatial units increases. So, if the variable observed is not dependent on space, the variations between the connectivity and variability components should not differ much from each other. Before showing the construction of the SGI, in the next section we present the system of spatial weights used to determine the contribution of connectivity in spatial terms.

We believe that there are two major motivations for the implementation (first) and (after) the use of SGI by official statistics. The first major motivation is a general one, but, in our view, extremely relevant. We all know that space is a fundamental dimension to better grasp socio economic and demographic phenomena and processes. Referring only to population issues (for sake of simplicity) we know that modern demography is essentially a spatial social science (Voss 2007). The importance of space in the field of official statistics is also underpinning by Eurostat (see, for example Eurostat 2015). So, from that, second major motivation, it is quite surprisingly that, to the best of our knowledge, National Statistical Institutes (NSIs) don't have an "official" or at least a common measure of spatial concentration. In the measure of the concentration of the Gross Domestic Product for example, NSIs and other International Institutions like the United Nations or World Bank use the G index, but this is not true in terms of measuring the spatial dimension of concentration (that we proved in the article to be a fundamental dimension of concentration). So, in our view, there is a double trouble: (1) space is relevant for measuring process and phenomenon and especially concentration but (2) there aren't any "official" or at least commonly used measure (like in the case of G). This is why we proposed SGI. The methodology here proposed is simple and completely transparent (we didn't build a black box) and can therefore easily implemented and replicated by NSIs and other Institutions.

2.1. Partitioning the Territory: The "MaxMin" And "Constant Step" Distance Methods

According to the Lorenz curve approach (Lorenz 1905), we need a system that can quantify the contribution of connectivity in spatial terms. Our proposal is to consider buffer or threshold distances capable of progressively creating partitions of the territory (or territorial subsets). These partitions identify neighbouring and non-neighbouring units such that each partition is disjoint from the others and the sum of all the elements of all the partitions coincides with the number of all the possible pairs between the n spatial units.

2.1.1. Partitioning With MaxMin Distance Method

As we know, the Gini index is geometrically based on the Lorenz curve (Lorenz 1905). For a sample of n dimension, the curve takes the cumulative percentile of the n units (for example individuals) on the x-axis and cumulative percentile of the variable (for example income) on the y-axis. The idea of the Lorenz curve is very simple. Given a sample of n ordered units, it's a graph that compares the distribution of a variable with a hypothetical uniform distribution of that variable (the original contribution can be found in Lorenz 1905). Perfecting this graph would be a diagonal line at a 45° angle from the origin (meeting point of the x and y axis), indicating the population's perfect variable distribution (line of absolute equality).

To satisfy these conditions we use the MM method (Mucciardi 2008a) that we recall below.

Suppose we have n spatial units $u_{1}\ldots\ldots u_n$ in which we observe $x_{1}\ldots\ldots x_n$ data and let $\boldsymbol{E^0}$ be a $n\times n$ matrix of the Euclidean distances between these units such that $d_{ij}^0 = \|u_i - u_j\|_2$ (with $d_{ij}^0 \in E^0$, $i = 1\ldots n\, j = 1\ldots n$), where $\|\cdot\|_2$ is the Euclidean norm.

Then the MaxMin distance h_{MM} is defined by:

$$h_{MM} = max\big(d_1,\, d_2,\, \ldots d_j \ldots d_n\big), \tag{1}$$

where d_j denotes the minimum distance of the generic spatial unit i from the other units j (with $i \neq j$). As a consequence, the whole territory is connected and there are no isolated spatial units.

The h_{MM} represents the first distance; therefore, it will be called h_{MM}^1.

More formally, $h_{MM}^1 = max(d_1^1,\, d_2^1,\, \ldots d_j^1 \ldots d_n^1)$
with $d_j^1 = \min\limits_{i=1\ldots n}(\{d_{ij}^0\}\backslash\{0\})$ with $j = 1\ldots n$ and $d_{ij}^0 \in \boldsymbol{E^0}$.

Using h_{MM}^1, the generic element ω_{ij}^1 of the first order-spatial weight matrix Ω^1 is determined as follows:

$\omega_{ij}^1 = 1$ *if* $d_{ij}^0 \leq h_{MM}^1$ and $d_{ij}^0 \neq 0$;
$\omega_{ij}^1 = 0$ otherwise
$d_{ij}^0 \in \boldsymbol{E^0}$ and $h_{MM}^1 = max(d_1^1,\, d_2^1,\, \ldots d_i^1 \ldots d_n^1)$,
(with $\omega_{ij}^1 \in \Omega^1$).

This first distance h_{MM}^1 is the reference for the Euclidean distance matrix $\boldsymbol{E^1}$ where the generic element d_{ij}^1 is given by:

$d_{ij}^1 = d_{ij}^0 \in \boldsymbol{E^0}$ *if* $d_{ij}^0 > h_{MM}^1$;
$d_{ij}^1 = 0$ otherwise.

If we consider the k-distance $h_{MM}^k = max(d_1^k,\, d_2^k,\, \ldots d_j^k \ldots d_n^k)$

with $d_j^k = \min_{i=1...n}(\{d_{ij}^{k-1}\}\backslash\{0\})$ with $j = 1 \ldots n$ and $d_{ij}^{k-1} \in \boldsymbol{E^{k-1}}$,
for $k = 1 \ldots t$,
the generic element ω_{ij}^k of the k-order spatial weight matrix Ω^k is determined as follows:
$\omega_{ij}^k = 1$ *if* $d_{ij}^{k-1} \leq h_{MM}^k$ and $d_{ij}^{k-1} \neq 0$;
$\omega_{ij}^k = 0$ otherwise,
(with $\omega_{ij}^k \in \Omega^k$).

The k-distance h_{MM}^k is the reference for the Euclidean distance matrix $\boldsymbol{E^k}$ where the generic element d_{ij}^k is given by:
$d_{ij}^k = d_{ij}^{k-1} \in \boldsymbol{E^{k-1}}$ *if* $d_{ij}^{k-1} > h_{MM}^k$;
$d_{ij}^k = 0$ otherwise.

As a consequence: $E^0 \supset E^1 \supset \ldots E^{k-1} \supset E^k \supset \ldots E^{t-1} \supset E^t$.
So, by iterating this procedure it is possible to obtain the distances from h_{MM}^1 *to* h_{MM}^t.

With h_{MM}^t, all spatial units are linked with each other since the condition that $\forall d_{ij}^{k-1} \leq h_{MM}^t$ $(k = 1 \ldots t)$ holds true and the algorithm stops. This distance coincides with the maximum distance in the Euclidean distance matrix $\boldsymbol{E^0}$ $(h_{MM}^t = \max[d_{ij}^0])$. As a consequence, since the condition $d_{ij}^{k-1} > h_{MM}^t$ $(k = 1 \ldots t)$ cannot be verified, $\boldsymbol{E^t} = \emptyset$ *and* $\Omega^{t+1} = \emptyset$.

It is important to point out that this procedure generates a "threshold" or "buffer distance" h_{MM}^k (with $k = 1 \ldots t$) without imposing any constraint on the number of neighbours; thus, this is not arbitrary but based on the territorial pattern of the spatial units. So, the k-territorial partitions are disjoint from each other, that is, $\Omega^i \cap \Omega^{i+1} = \emptyset$ ($i = 1 \ldots t$ with $\Omega^{t+1} = \emptyset$).

It is important to underline that the algorithm sets the threshold to be the most distant nearest neighbour, consequently each unit has at least one neighbour included so there are no isolated units. For the sake of simplicity and to better illustrate the method, the "real" empirical application refers in fact to the municipality level, we apply this territorial partitioning to the 107 Italian provinces. The MM method applied to the territory creates 22 spatial lags (territorial partitions). The MM distances (h-distance), the relative joints (links) and the names of the provinces that originate the MM distances (province link name) are given in Table 1. Moreover, Figure 1 provides a graphical view of the method.

2.1.2. Partitioning with Constant Step Distance Method

The MM method creates threshold distances in relation to the natural shape of the territory. This criterion can sometimes produce spatial lags that present large gaps in terms of distance. For this reason, we introduce a variant of the MM method with constant increments. We call this procedure the CS method. Suppose we have n spatial units $u_{1......}u_n$ in which we observe $x_{1......}x_n$ data and let $\boldsymbol{E^0}$ be a $n \times n$ matrix of the Euclidean distances between these units such that $d_{ij}^0 = \|u_i - u_j\|_2$ (with $d_{ij}^0 \in E^0$, $i = 1 \ldots n\, j = 1 \ldots n$), where $\|\cdot\|_2$ is the Euclidean norm.

As in the previous procedure we determine the first distance of MM which in this method coincides with the first distance $\left(h_{cs}^1 \equiv h_{MM}^1\right)$. If we set the increment equal to the first distance, we can write the following relation for the generic spatial lag (k):

$$\mathrm{h}_{\mathrm{cs}}^{\mathrm{k}} = \mathrm{kh}_{\mathrm{cs}}^{1} \text{ (with } k = 1 \ldots t). \tag{2}$$

Table 1. Details of the MM method applied to the 107 Italian provinces[1].

Spatial Lag	h-distance (Km)	Links	Name of the provinces that determine the h-distance
1	77.96	516	('Palermo', 'Trapani')
2	116.74	554	('Palermo', 'Catania')
3	375.75	4140	('Grosseto', 'Oristano')
4	417.92	596	('Palermo', 'Cagliari')
5	454.58	510	('Chieti', 'Messina')
6	492.57	518	('Latina', 'Cremona')
7	528.30	452	('Pistoia', 'Sud Sardegna')
8	570.48	528	('Crotone', 'Perugia')
9	618.61	522	('Salerno', 'La Spezia')
10	666.72	490	('Forli'-Cesena', 'Lecce')
11	712.72	424	('Grosseto', 'Siracusa')
12	759.24	342	('Salerno', 'Asti')
13	804.87	302	('Pisa', 'Siracusa')
14	850.90	292	('Ravenna', 'Siracusa')
15	899.84	280	('Modena', 'Siracusa')
16	942.64	268	('Parma', 'Siracusa')
17	985.85	238	('Verona', 'Siracusa')
18	1036.66	222	('Brescia', 'Siracusa')
19	1087.97	102	('Lecco', 'Siracusa')
20	1111.11	34	('Bolzano', 'Siracusa')
21	1153.71	10	('Aosta', 'Siracusa')
22	1154.03	2	('Aosta', 'Ragusa')

[1] It is important to clearly explain that the MM distance method creates distances in relation to the spatial configuration of territorial units. This criterion can sometimes produce spatial lags that present large gaps in terms of distance. If we look, for example, to Table 1 a big jump is evident between lags 2 and lags 3 due to the link created between the Sardinia region (that is an island) and the rest of Italy. For a more technical explanation, please see Appendix (Section 6).

The procedure stops when $h^t_{cs} \geq \max\,[d^0_{ij}]$ where $d^0_{ij} \in \boldsymbol{E}^{\mathbf{0}}$

In the h^t_{cs} distance all the spatial units are linked with each other. For this method the details and a graphical illustration are also provided considering its application to the 107 provinces of Italy (Table 2 and Figure 2).

2.2. *From the Territorial Partitions to the Spatial Gini Index (SGI)*

The properties of the territorial partitions of the two distance methods discussed above makes the procedure compatible with the structure of the Gini index according to the definition of the ratio of the areas (Mucciardi and Bertuccelli 2007; Mucciardi 2008b). Indeed, the Gini index can then be thought of as the ratio of the area that lies between the line of equality and the Lorenz curve over the total area under the line of equality. Following the same approach, SGI can therefore be considered as the ratio between the area of spatial autocorrelation on the total area of the square of side 1.

We define $J_{(k)}$, the cumulated percentage of the total connectivity of the units in the generic distance h^k_{MM}, as

$$J_{(k)} = \frac{\Sigma^n_i \Sigma^n_j \omega^k_{ij}}{A} \downarrow \text{ with } k = 1 \ldots t \quad (J_{(0)} = 0 \text{ and } J_{(t)} = 1) \tag{3}$$

Fig. 1. Graphical view of the MM method applied to the 107 Italian provinces.

Table 2. Details of the CS distance method applied to the 107 provinces of Italy.

Spatial Lag	h-distance (Km)	Links
1	77.96	516
2	155.92	1170
3	233.89	1334
4	311.85	1250
5	389.81	1142
6	467.77	1066
7	545.74	1028
8	623.70	884
9	701.66	746
10	779.62	596
11	857.58	492
12	935.55	478
13	1013.51	412
14	1091.47	184
15	1169.43	44

where ω_{ij}^{k} denote the interconnection links (that is, the contiguous spatial units), in the generic k-territorial partition h_{MM}^{k} and $A = \Sigma_{i=1}^{t}\Omega^{i} = \Sigma_{k=1}^{t}\Sigma_{i=1}^{n}\Sigma_{j=1}^{n}\omega_{ij}^{k} = n(n-1)$ is the maximum number of links it is possible to obtain from a particular territorial configuration Please note that in the equations the arrows symbol means that the percentage are cumulated. Furthermore, we recall that the connections of the spatial units with themselves are excluded from the maximum number of links.

In the same way, we define $V_{(k)}$, the cumulated percentage of the variability of the phenomenon X "absorbed" by the linked elements in the distance h_{MM}^{k}, as

$$V_{(k)} = \frac{\Sigma_{i}^{n}\Sigma_{j}^{n}\left(x_i - x_j\right)^2\omega_{ij}^{k}}{D}\downarrow \text{ with } k = 1\ldots t \quad \left(V_{(0)} = 0 \text{ and } V_{(t)} = 1\right) \tag{4}$$

where ω_{ij}^{k} are as before and

$$D = \Sigma_{k=1}^{t}\Sigma_{i=1}^{n}\Sigma_{j=1}^{n}\left(x_i - x_j\right)^2\omega_{ij}^{k}.$$

Referring to a territorial system comprising n spatial units in which we observe $x_1 \ldots x_n$ data, the Spatial Gini Index (SGI) will be defined as

$$SGI = 1 - 0.5\sum_{k=1}^{t}\left(V_{(k)} + V_{(k-1)}\right)\left(J_{(k)} - J_{(k-1)}\right)$$

with

$$V_{(k)} = \frac{\Sigma_{i}^{n}\Sigma_{j}^{n}\left(x_i - x_j\right)^2\omega_{ij}^{k}}{D}\downarrow \quad \text{and} \quad J_{(k)} = \frac{\Sigma_{i}^{n}\Sigma_{j}^{n}\omega_{ij}^{k}}{A} \quad k = 1\ldots t. \tag{5}$$

Its construction is based on the computation of the area of spatial autocorrelation, which has been proposed, to the best of our knowledge, (see for more details, Alleva 1987; Mucciardi and Bertuccelli 2007; Mucciardi 2008b).

Following the Lorenz curve, which is the basis of the Gini index, in a condition of no spatial autocorrelation, the cumulated percentage of variability $V_{(k)}$ should not differ from the cumulated percentage of connectivity $J_{(k)}$. SGI can assume the minimum value of 0

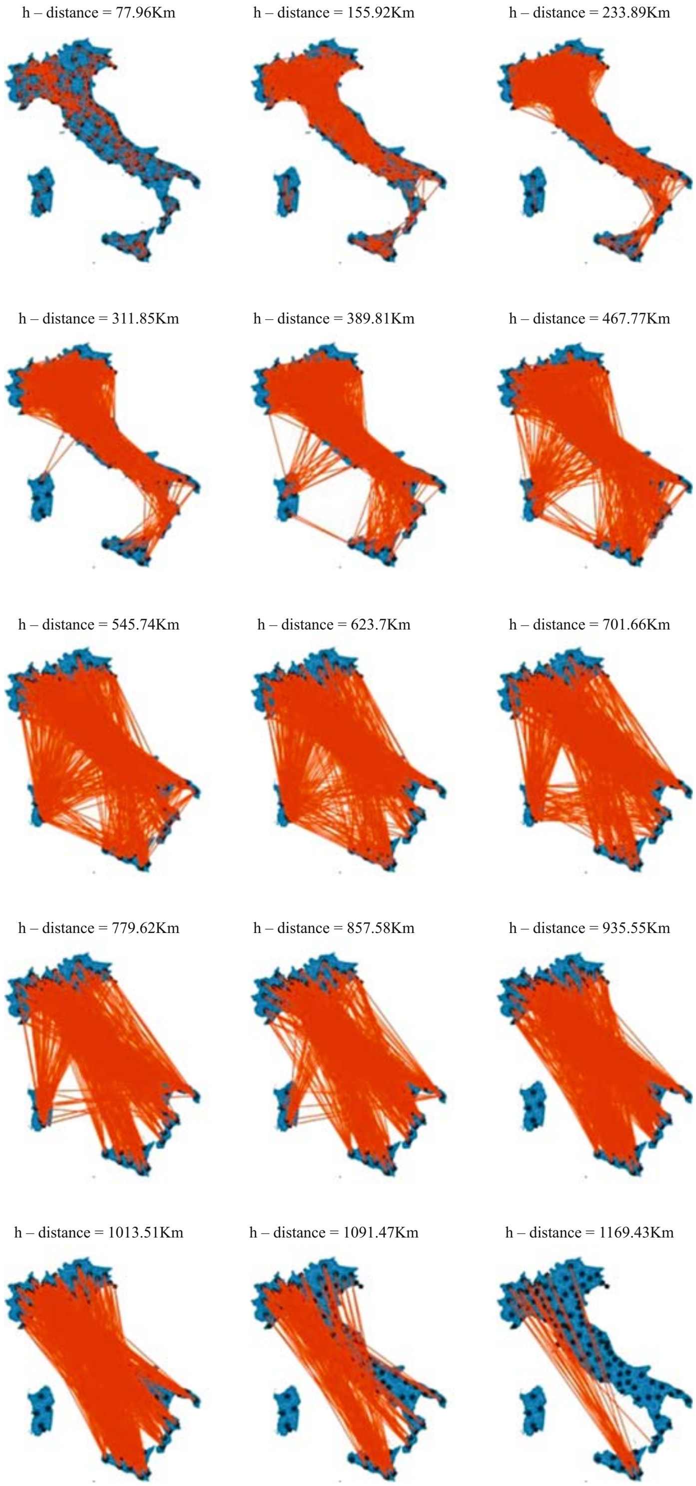


Fig. 2. Graphical illustration of the CS distance method applied to the 107 provinces of Italy.

and the maximum value of 1, but in terms of spatial autocorrelation we distinguish three cases:

1. Case of "negative spatial autocorrelation": if the relative contributions in terms of variability are larger than the contribution in terms of connectivity in the generic distance h^k_{MM}, we will graphically obtain a "convex curve" with respect to the ordinate axis and $0 \leq \text{SGI} < 0.5$,
2. Case of "no spatial autocorrelation": if the relative contributions in terms of variability and connectivity increase proportionally to the variation of distance h^k_{MM}, we will graphically obtain a straight line at a perfect angle of 45° and SGI = 0.5 (area of the curve exactly equal to 0.5), and
3. Case of "positive spatial autocorrelation": if the relative contributions in terms of variability are smaller than the contribution in terms of connectivity in the generic distance h^k_{MM} we will graphically obtain a "concave curve" with respect to the ordinate axis and $0.5 < \text{SGI} \leq 1$.

To better understand the relation between the concept of a Lorenz curve and that of spatial correlation, Figure 3 shows three scenarios of spatial autocorrelation with the related range of values for SGI.

From a geometric point of view, these three forms of spatial autocorrelation may be assessed, as the h^k_{MM} (h-distance) varies, by considering the tangent of the angle formed by the straight line with the x-axis:

$$\tan^k(\alpha) = \frac{V_{(k)}}{J_{(k)}} \; k = 1 \ldots t. \tag{6}$$

So, we can have:

1. $\tan^k(\alpha) < 1$ (angle $< 45°$) indicating positive spatial autocorrelation,
2. $\tan^k(\alpha) = 1$ (angle $= 45°$) indicating no spatial autocorrelation, and
3. $\tan^k(\alpha) > 1$ (angle $> 45°$) indicating negative spatial autocorrelation.

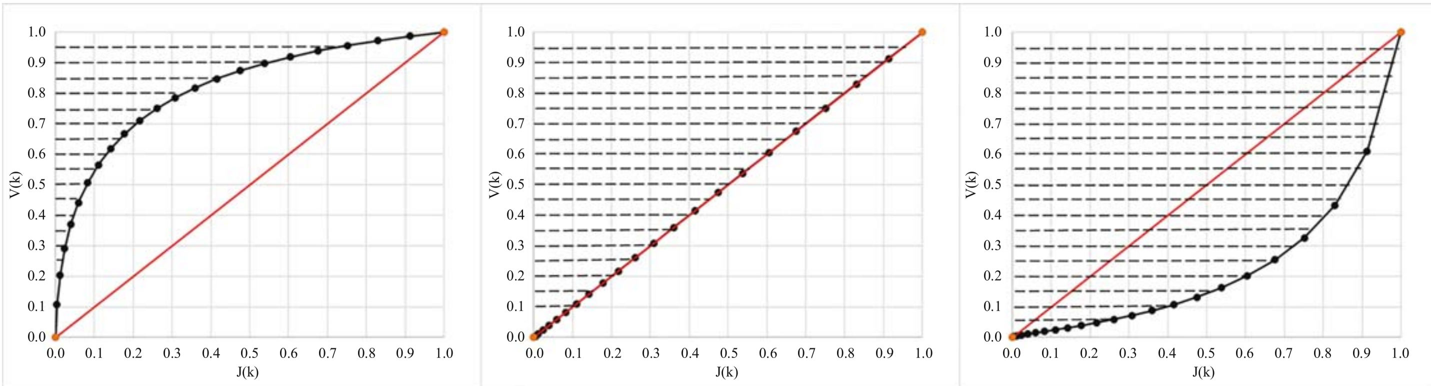
For the calculation of the angle, we use the arctan function. We recall that the arctan function is the inverse of the tangent function. It returns the angle whose tangent is a given number. Figure 4 shows the expected trend of the arctan function (degrees) as the h^k_{MM} vary for the three scenarios of spatial autocorrelation.

By using the CS distance method instead of MM in all these relations, we obtain the SGI with this territory partition method. In the context of this research, we will distinguish the calculation of the SGI using this notation:

1. SGI_{MM} when the k-partitioning method is that of the MM distance, and
2. SGI_{CS} when the k-partitioning method is that of the CS distance

3. The Decomposition Method of Rey and Smith in the Framework of SGI

A spatial decomposition of the Gini coefficient has recently been proposed by Rey and Smith (2013). In this work the authors suggest an alternative approach towards considering the joint effects of inequality and spatial autocorrelation that relies on a decomposition of the classic Gini coefficient. This decomposition involves the splitting of



(A) Simulated data for 22 spatial lags with three assumed scenarios: SGI = 0.2 (case of negative spatial autocorrelation); SGI = 0.5 (case of no spatial autocorrelation); SGI = 0.8 (case of positive spatial autocorrelation)

Fig. 3. Curve of spatial autocorrelation (dashed line): 1) case of negative spatial autocorrelation – 0 < SGI < 0.5 (left); 2) case of no spatial autocorrelation – SGI = 0.5 (centre); 3) case of positive spatial autocorrelation 0.5 < SGI < 1 (right) (A).

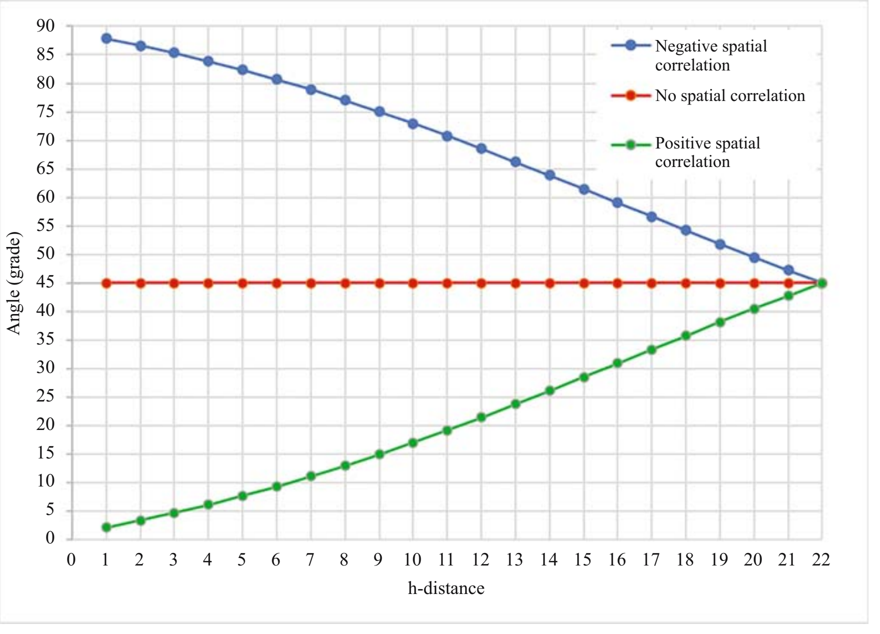


Fig. 4. Expected trend of the arctan function (degrees) as the h-distances vary for the three scenarios of spatial autocorrelation – (simulated data).

the Gini index into two mutually exclusive components: into contiguous units (w_{ij}) and the non-contiguous units ($1 - w_{ij}$).

Equation 7 shows this decomposition.

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right|}{2n^2\mu} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\left|x_i - x_j\right|}{2n^2\mu} + \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(1 - w_{ij})\left|x_i - x_j\right|}{2n^2\mu}. \quad (7)$$

According to SGI we extend this decomposition applying to the two distance methods (MM and CS) shown above. Therefore, it is possible to make this decomposition of each "*k-territorial partition*" imposed by the h^k-distances.

Now, if we denote by G_C^k the Gini index calculated inside the contiguous units in the h^k -distances and by G_{NC}^k the Gini index calculated in the non-contiguous units (or outside the contiguous units) in the h^k-distances, we can rewrite the Gini index with the Rey and Smith decomposition (G_T^k):

$$G_T^k = G_C^k + G_{NC}^k \quad (k = 1 \ldots t). \quad (8)$$

As a consequence,

$$G_T^k = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i - x_j|}{2n^2\mu} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\omega_{ij}^k|x_i - x_j|}{2n^2\mu} + \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(1 - \omega_{ij}^k)|x_i - x_j|}{2n^2\mu} \quad (9)$$

$$(k = 1 \ldots t)$$

with

$$G_C^k = \frac{\sum_{i=1}^n \sum_{j=1}^n \omega_{ij}^k |x_i - x_j|}{2n^2\mu}; \; G_{NC}^k = \frac{\sum_{i=1}^n \sum_{j=1}^n (1 - \omega_{ij}^k)|x_i - x_j|}{2n^2\mu}$$

and $G_T^k \equiv G \; (\forall k = 1 \ldots t)$.

To distinguish the decomposition based on the partition used (MM and CS), we will differentiate the calculation of the G_T^k index using this notation:

1. $G_{T(MM)}^k = G_{C(MM)}^k + G_{NC(MM)}^k$ when the k-partitioning method is that of the MM distance
2. $G_{T(CS)}^k = G_{C(CS)}^k + G_{NC(CS)}^k$ when the k-partitioning method is that of the CS distance.

Our task here is both to propose an extension of the decomposition method of Rey and Smith in the framework of the two distance approaches (MM and CS) here discussed and to evaluate similarities and differences in the achieved results. Our general idea is that all these measures can be used in a complementary way in measuring the spatial concentration of the population. This is somewhat similar to what happens with measuring residential segregation, where the use of different measures and approaches is highly recommended (Brown and Chung 2006)

4. Empirical Application

In this section we show the results of the empirical application. The application regards three points in time, 2002, 2010 and 2018, and it is realized at the municipality level, the finest territorial scale possible, that is to say almost 7,900 spatial units. The application was done for two groups of population selected on the basis of country of citizenship: Italians and foreigners. As is known, these two populations typically have different geographical patterns of spatial distribution and therefore they are particularly useful for our tasks (Massey and Denton 1988). The data are provided by the Italian National Institute of Statistics (Istat) and disseminated by the institutional website. In more detail, we use data on the resident population broken down by the country of citizenship (Italian/Foreign) for the years 2002, 2010, and 2018.

4.1. The Spatial Gini Index

4.1.1. The MM Approach

In Figure 5 we can appreciate the evolution over time of the level of the classical Gini index and the spatial Lorenz curve of Italian and foreign population in the observed period. In the first year, 2002, the foreign population in Italy was about 1.3 million and the level of its spatial concentration was lower than the Italian population (SGI = 0.446 versus SGI = 0.485). In the following years the Italian population remained quite stable and so did its level of spatial concentration: SGI = 0.485 in 2002 and SGI = 0.483 in 2018. In contrast, the foreign population grew significantly, reaching more than 5 million in 2018. Its level of spatial concentration grew too: SGI = 0.446 in 2002 and SGI = 0.450 in 2018. The dynamics of the arctan functions (Figure 6) inform us about the evolution of the level of spatial autocorrelation for both populations in the selected years for each h-distance. As can

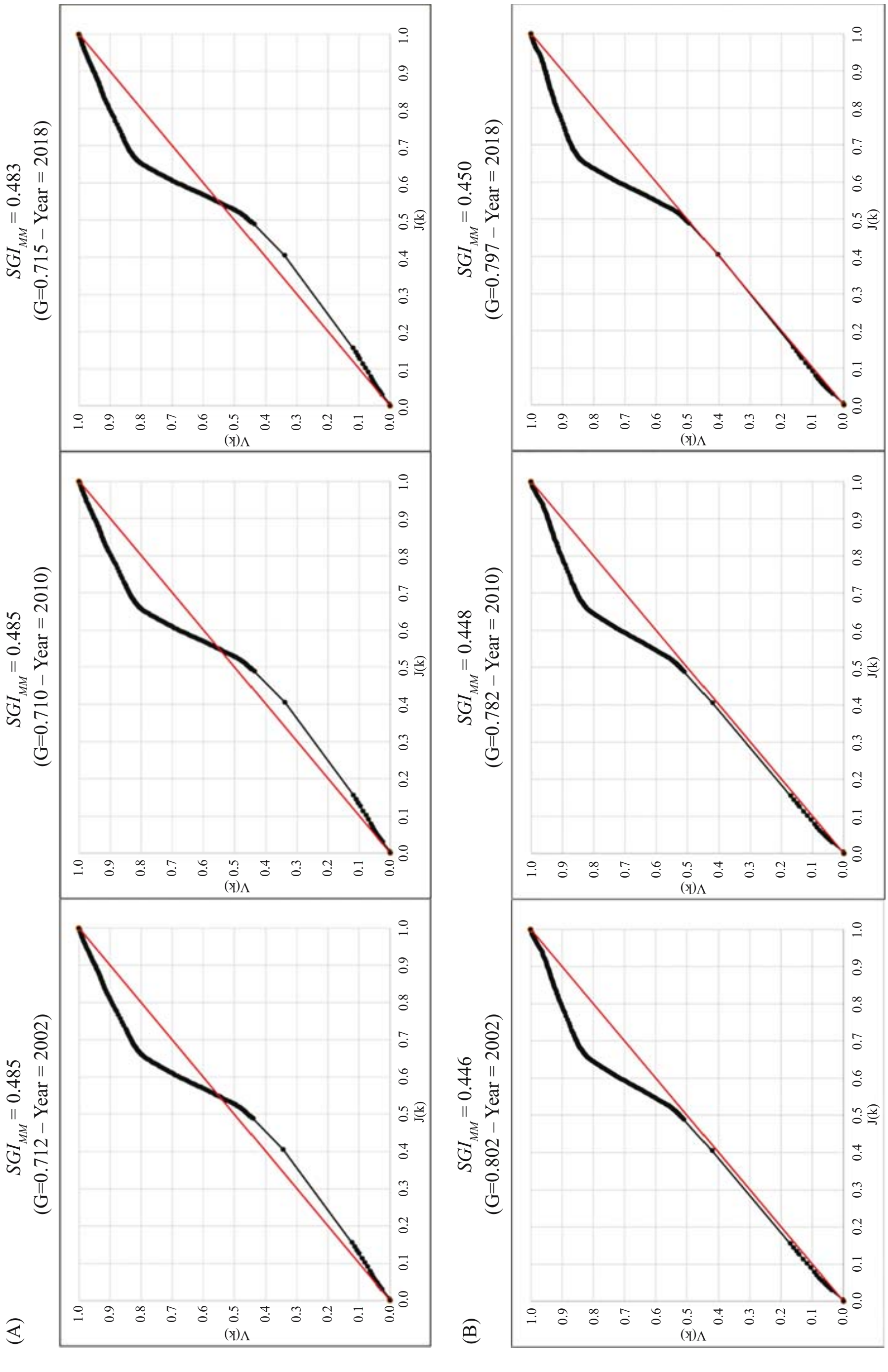


Fig. 5. Spatial Lorenz Curve, SGI_{MM} and aspatial Gini's G (in brackets). Italian population (A) and foreign population (B).

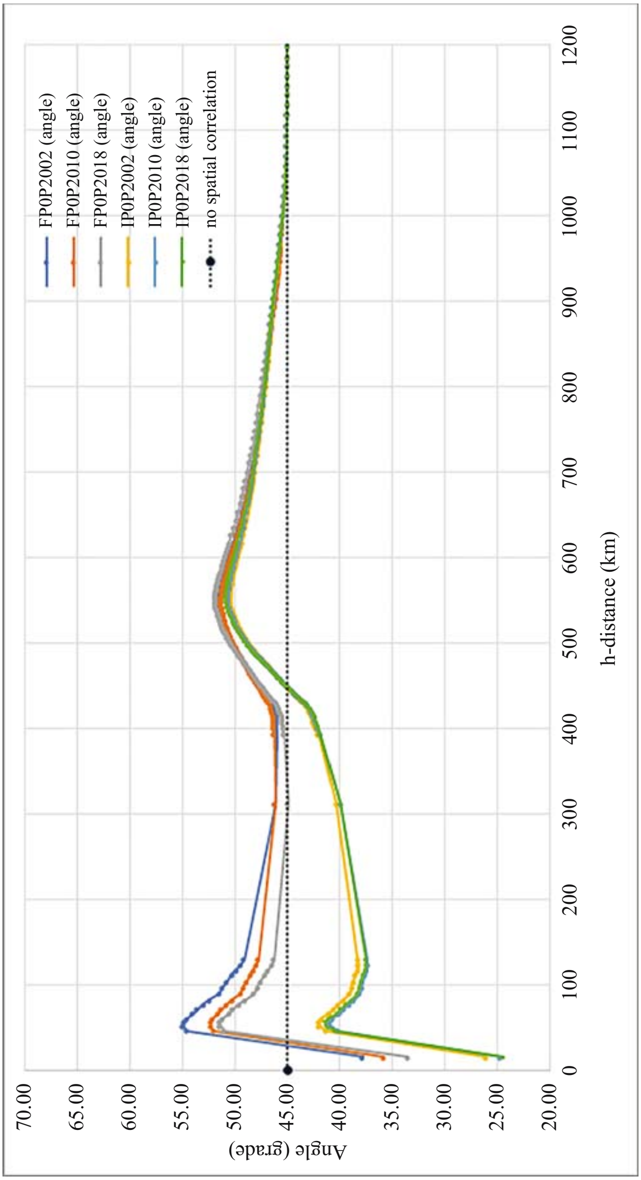


Fig. 6. Value of arctan (degrees). Italian (IPOP) and foreign (FPOP) population (MM Distance Method).

be clearly seen, the two populations start (16 km) from a different condition of positive spatial autocorrelation: greater positive spatial autocorrelation for the Italian population and less positive spatial autocorrelation for the foreign population (see the angles less than 45 degrees in Figure 6). From the second h-distance (47 km) the situation changes: the positive spatial autocorrelation effect is attenuated for the Italian population while the foreign population shows spatial instability with negative values of arctan (see in Figure 6 the angles greater than 45 degrees for foreign populations only). By increasing the h-distance, a process of convergence towards a level of no spatial autocorrelation comes up. This condition is obtained starting from an h-distance equal to 400 km and obtained when the h-distance is equal to 1,000 km. From the point of view of the classic Gini index (G), we note an inverse behaviour: lower values of SGI are followed by higher values of G. Therefore, from these results it emerges that the foreign population, although concentrated at an "aspatial level" (G), instead shows a moderate tendency to have a negative autocorrelation with the territory. The Italian population shows a lower level of the aspatial concentration (G). This behaviour manifests itself in the SGI with a tendency towards no spatial autocorrelation (SGI tends to 0.5). However, we verify the spatial behaviour of these two populations by calculating the Moran's I index for the same h-distances (Figure 7). As we can see at 16 km (first h-distance) the different manifestation of positive spatial autocorrelation is confirmed, albeit in a more attenuated manner (Moran's I indices, although statistically significant, are very small). The spatial autocorrelation level then converges to the expected values exactly from a distance of 400 km.

4.1.2. Constant Step Approach

From a territorial partitioning point of view, the main difference between the two methods (MM and CS) is in the number of h-distances produced: 162 for MM and 76 for CS (see Appendix, Subsection 6.1, for more details). However, despite this difference in the number of h-distances, the results of the CS procedure are quite similar to the ones obtained by MM (see Subsection 4.1.1) with slight substantial differences. Therefore, the explanation of the spatial concentration process of the Italian and foreign populations in the years considered (2002, 2010 and 2018) remains almost unchanged with respect to the SGI index calculated with the MM method (see Figure 8–10).

In general, and based on some empirical evidence known in the literature (Miller 2004), positive autocorrelations on a small spatial scale and negative autocorrelations on a larger spatial scale are all in all expected results. The first law of geography states that "everything is related to everything else, but near thing is more related than distant things" (Tobler 1970, 235). As known, scale can play important effect on results, nevertheless our results seem to be coherent with Tobler's law. Indeed, we know that the spatial distribution of human population, is a process particularly affected by such kind of "spatial" effects (distance and scale). This is particularly true for the foreign population that tend to have specific settlement models to maximize opportunity and minimize cost. As know the scientific debate about that is wide and rich but, at least for Italy, there are clear evidence about the attractive role played by urban areas on foreign and immigrant population (Strozza et al. 2016) and on the spatial heterogeneity and dependence that characterized the settlement models of foreign communities at different spatial scales (Benassi et al. 2019, 2020).

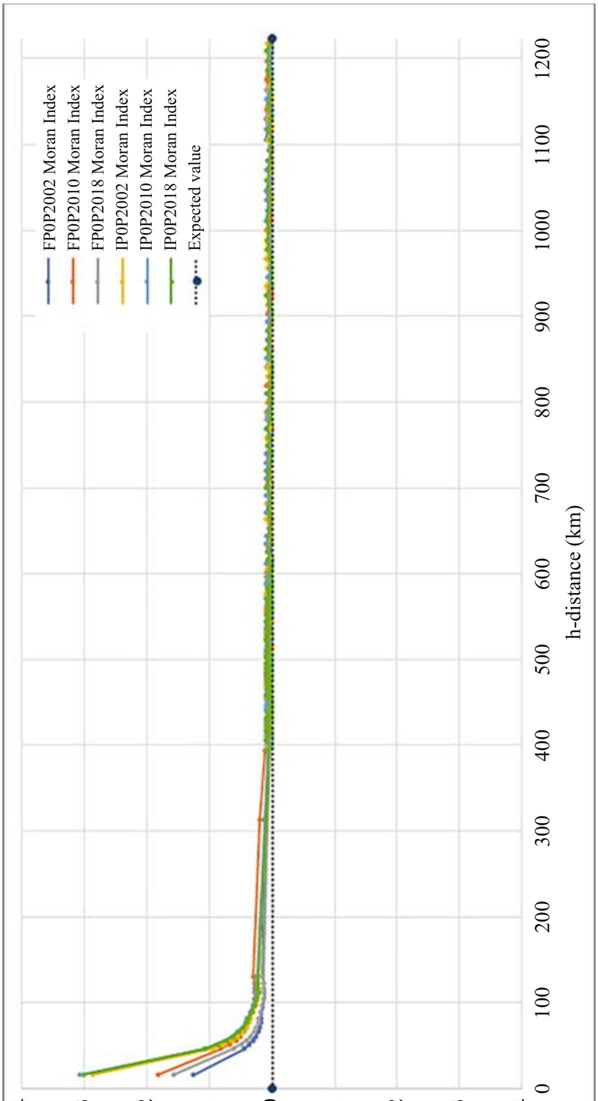


Fig. 7. Moran's I index. Italian (IPOP) and Foreign (FPOP) populations (a) (MM Distance Method).
(a) Moran's I indices are almost all statistically significant ($p < 0.05$). All the Moran's I indices are processed with Arcgis10.8 software for the same MM distances used for SGI. Data are available upon request.

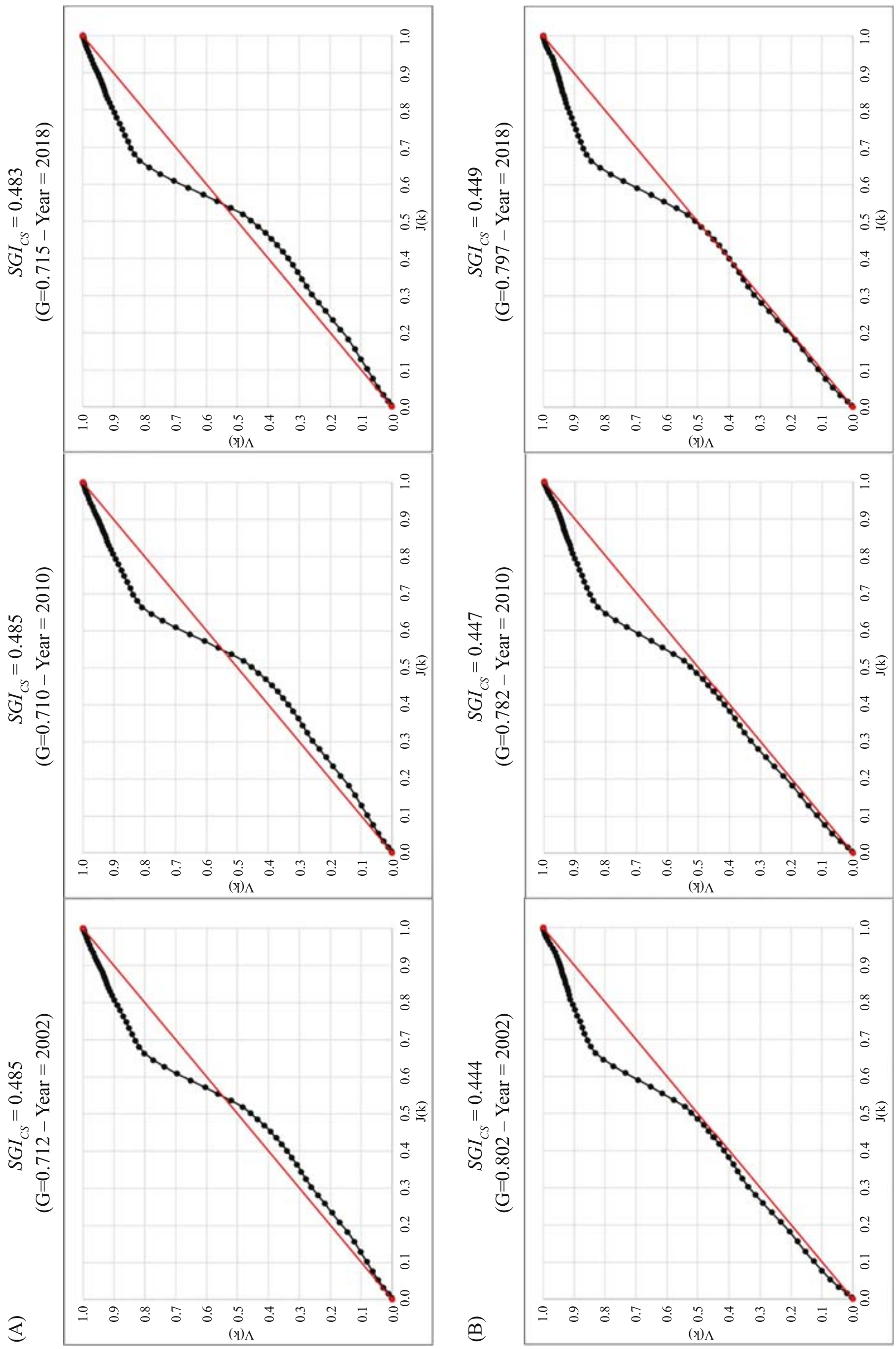


Fig. 8. Spatial Lorenz Curve, SGI_{CS} and aspatial Gini's G index (in brackets). Italian population (A) and foreign population (B).

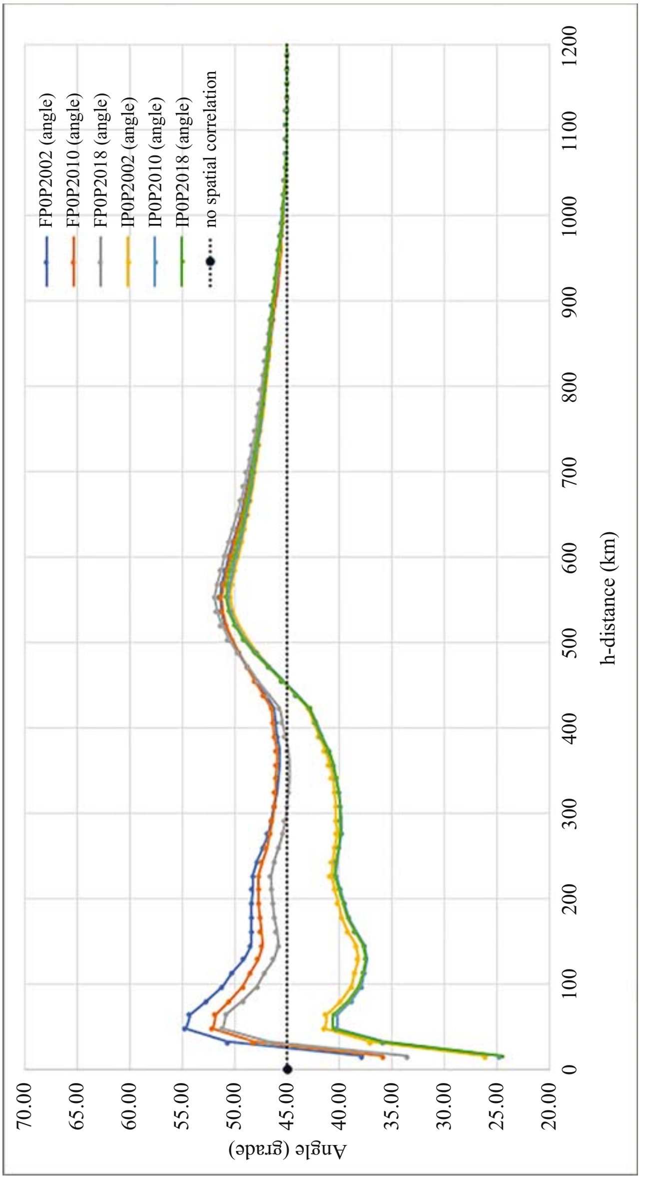


Fig. 9. Values of arctan (degrees). Italian (IPOP) and foreign (FPOP) population (CS Distance Method).

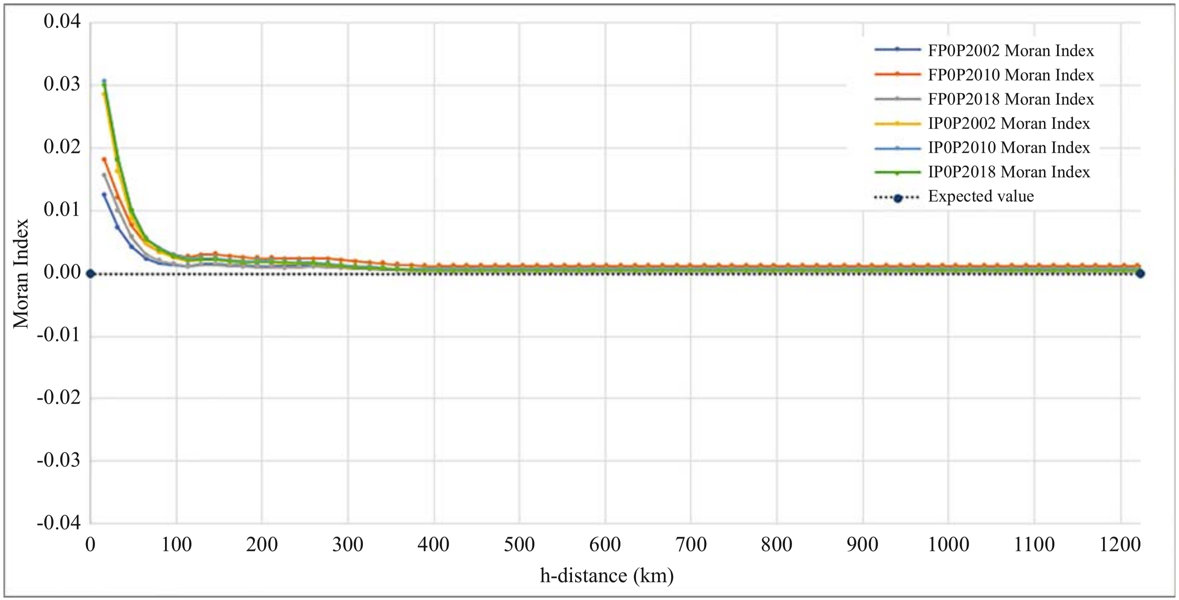


Fig. 10. Moran's I index. Italian (IPOP) and Foreign (FPOP) population (a) (CS Distance Method).
(a) Moran's I indices are almost all statistically significant ($p < 0.05$). All the Moran's I indices were processed with Arcgis10.8 software for the same CS distances used for SGI. Data are available upon request.

4.2. Rey and Smith Decomposition Method With MM and CS Approaches

As mentioned above, let's extend the Rey and Smith decomposition (G_T^k) applying the two distance methods: $G_{T(MM)}^k$ with MM approach and $G_{T(CS)}^k$ with CS approach. As can be seen from Figures 11 and 12, considering the same h-distances with which we build SGI, the G_T^k in the spatial version is less sensitive to the quantification of inequality in the various spatial lags. In fact, the components of the Gini index (in the contiguous G_C^k and non-contiguous G_{NC}^k units) remain almost unchanged (the lines are almost parallel). Only considering the $G_{T(MM)}^k$, around 300 km, we have a peak of the Gini index in the component calculated in the contiguous units (G_C^k). In our opinion this occurs because the two indices measure spatial inequality differently. While SGI is a global measure of spatial inequality, that is, it is (completely) evaluated considering all the spatial lags (k), the G_T^k (with $G_T^k = G_C^k + G_{NC}^k$) is evaluated in each k-spatial lag. So, from this point of view, SGI and the G_T^k can be seen as complementary measures (or even tools) to better grasp and detect the level of (spatial) concentration and its dynamics for a given population.

5. Discussion and Conclusions

Measuring the concentration of population is an old and traditional activity of applied statistics. Many developments have been proposed from the seminal contributions of Lorenz and Gini. In relative recent times, also due to the GIS revolution, more attention has been paid to the spatial aspects of concentration, especially when it refers to a human population (Arbia 2001). The present article lies in this stream of the literature, proposing a new approach for measuring the spatial concentration of a human population, based on a spatial version of Gini's G index, which we call the Spatial Gini Index (SGI). In any kind of spatial approach, the definition of spatial neighbourhood and spatial weight matrixes are crucial and pivotal issues. This article proposes two new approaches to partitioning the territory, and therefore obtains two different kinds of spatial connectivity: one is based on a MM approach and allows computing the $\mathrm{SGI_{Mm}}$ version of the SGI. The other is based on a CS approach, and it is the basis for the computation of $\mathrm{SGI_{CS}}$. From the results obtained in the previous sections, SGI results are different from the ones produced by classic (aspatial) Gini's G index, indicating: (1) the importance of the spatial dimension in detecting the concentration of a population in space; (2) higher values of the level of concentration when measured using a traditional aspatial approach; (3) that this difference (G vs SGI) is higher for the foreign population. The higher values reached by the aspatial version of the concentration index are due to the fact that this approach is essentially based on the statistical concept of variability separated from the influence of the territory. So, this kind of measuring of (aspatial) concentration assumes that the space is independent of the distribution of the variable (and vice versa). From this point of view, the SGI index can be viewed as, at least, a complementary tool for better measuring and detecting the spatial concentration of a population. In particular, given the peculiarity of the spatial distribution of the foreign population in a destination country like for example Italy (Strozza et al. 2016) and the relevance of the spatial concentration for such a population (Reardon and O'Sullivan 2004), the proposed approach seems to add new methodological perspectives on measuring spatial concentration. In evaluating SGI, we also used well known spatial statistical measures of global spatial autocorrelation (Moran's I). The index has been

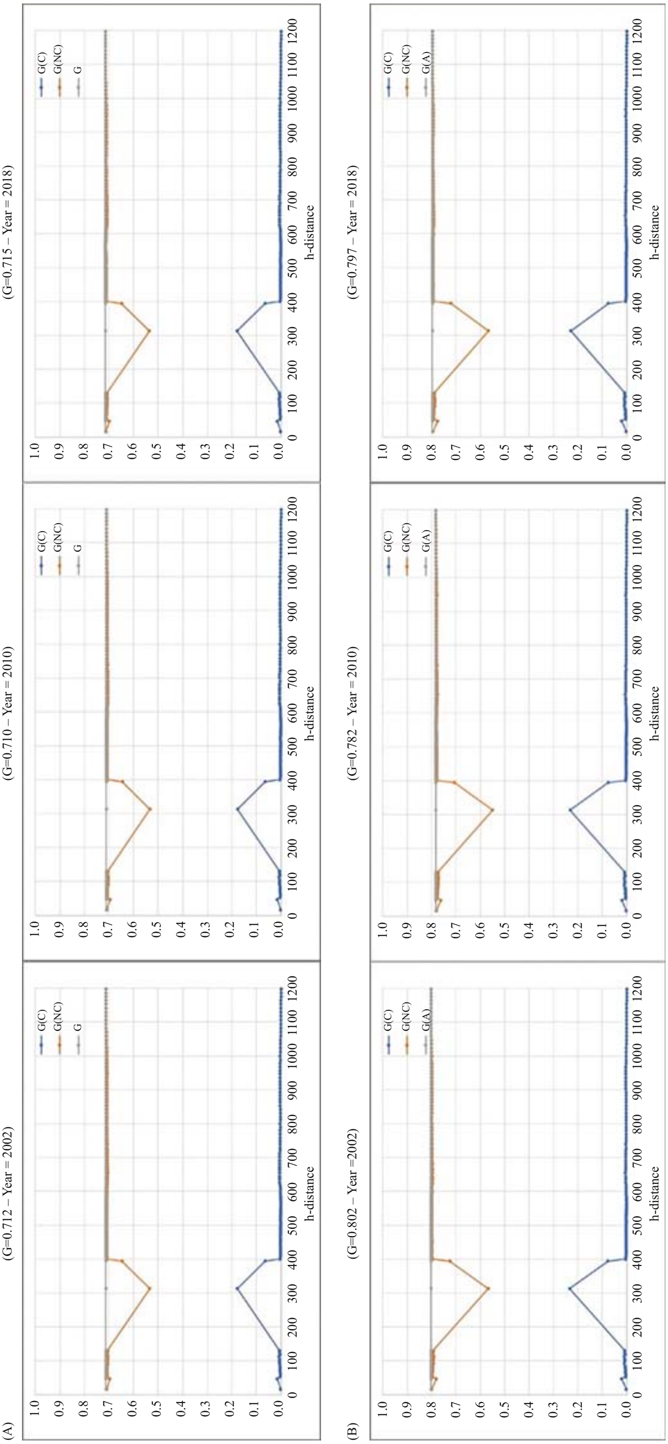


Fig. 11. $G^k_{T(MM)}$ *and aspatial Gini's G index (in brackets). Blue line –* $G^k_{C(MM)}$*; Orange line* $G^k_{NC(MM)}$*; Grey line – G (Italian population (A) and foreign population (B)).*

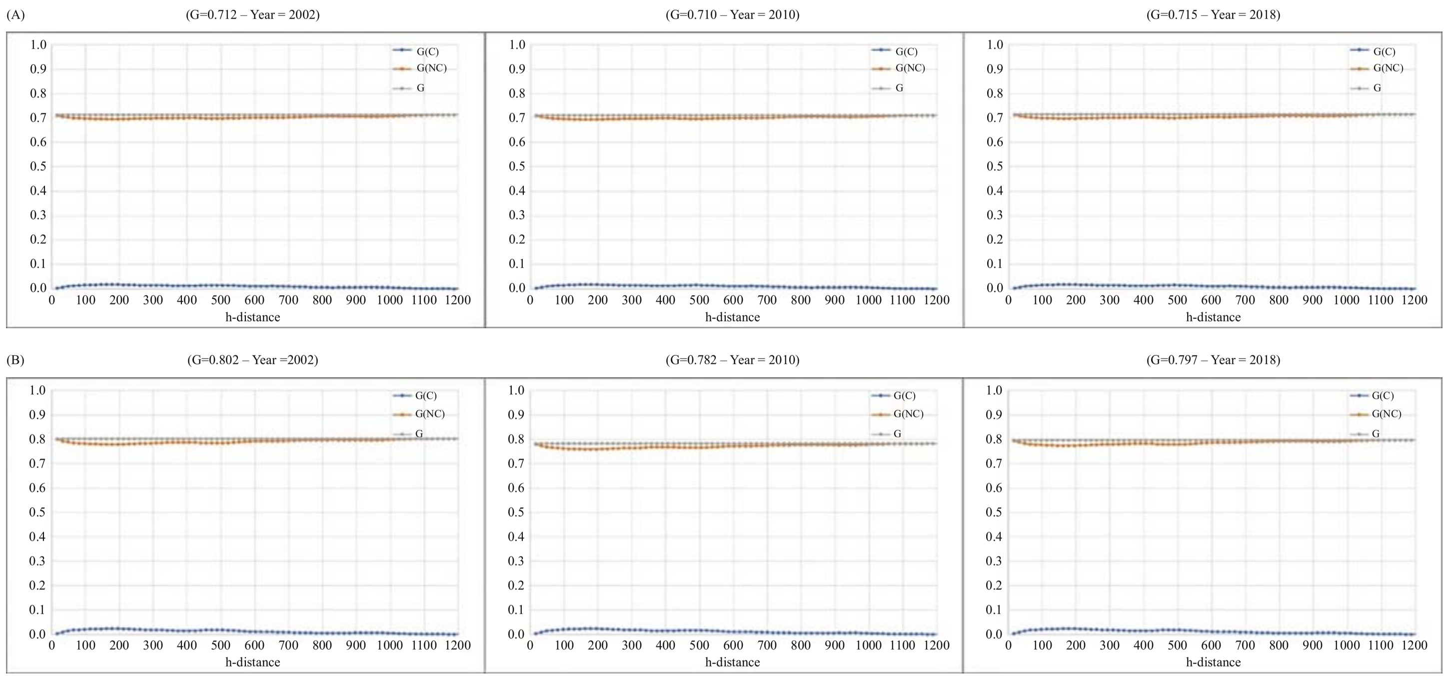


Fig. 12. $G^k_{T(CS)}$ *and aspatial Gini's G index (in brackets). Blue line –* $G^k_{C(CS)}$*; Orange line* $G^k_{NC(CS)}$*; Grey line – G (Italian (A) and foreign population (B)).*

computed on the resident population (Italian and foreigners) of Italian municipalities in 2002, 2010 and 2018. The results indicate that, at least using population as a variable, the level of global spatial autocorrelation is quite low and obviously tends to decrease with increasing h-distance. Another point faced in this article is the extension of the decomposition method of the Gini's G index recently proposed by Rey and Smith (2013) in the context of the territorial partitioning. Here the main objective was to propose an extension of this recent method in the framework of the MM and CS distances procedure and to evaluate similarities and differences between them both. Future developments should address the extension/optimization of the Monte Carlo test (so far only tested for 100 statistical units) and solving problems of computational requirements. There is also the possibility of using "classical binary weight" based on portions of shared boundaries of spatial unity (e.g., Queen and Rook method) and "non binary weights" in the spatial matrix (e.g., a kernel matrix with a distance decay function, etc.) and do simulation experiments in order to grasp the behaviour of SGI in terms of statistical distribution. In this direction it will be possible to make comparisons between the different methods of partitioning of the territory. Further developments can regard the local decomposition of the SGI and proposing other functionalities like semivariogram and similar. From an interpretative point of view, we have to underline that high levels of spatial concentration of foreign population can lead to different processes and behaviours that can act as detrimental to social cohesion. In conclusion, researching new approaches to measuring the spatial concentration of the human population is still an open challenge and improvements are currently in progress by the authors.

6. Appendix

6.1. Details of the SGI Procedure

6.1.1. MM Distance

Number of municipalities = 7,890 – Total links = 62,244,210

Spatial Lag	h-distance (km)	Links	% Links	J(k)	Name of the Italian municipalities that determine the h-distance
1	16.3	269852	0.004	0.004	(Campagnatico, Grosseto)
2	47.0	1609960	0.026	0.030	(Piraino, Malfa)
3	52.7	416318	0.007	0.037	(Follonica, Marciana)
4	57.0	337466	0.005	0.042	(Bibbona, Porto Azzurro)
5	61.7	373534	0.006	0.048	(San Severo, Vieste)
6	67.3	470188	0.008	0.056	(Campo nellElba, Monteverdi Marittimo)
7	72.0	417136	0.007	0.063	(Statte, Nard)
8	77.0	447572	0.007	0.070	(Martina Franca, Surbo)
9	83.5	600656	0.010	0.079	(Massafra, Carmiano)
10	91.4	766296	0.012	0.092	(Alberobello, Lecce)
11	98.3	675754	0.011	0.103	(Cavallino, Massafra)
12	105.6	727020	0.012	0.114	(Alberobello, Vernole)

Continued

Spatial Lag	h-distance (km)	Links	% Links	J(k)	Name of the Italian municipalities that determine the h-distance
13	113.9	852748	0.014	0.128	(Alberobello, Melendugno)
14	119.0	521240	0.008	0.136	(Massafra, Cannole)
15	124.5	566798	0.009	0.145	(Massafra, Otranto)
16	131.2	682664	0.011	0.156	(Ortueri, Palau)
17	312.8	15407126	0.248	0.404	("Campo nellElba", Ghilarza)
18	394.9	5325726	0.086	0.489	(Erice, Buggerru)
19	400.9	375570	0.006	0.496	(Castellammare del Golfo, Iglesias)
20	405.3	278344	0.004	0.500	(Mazara del Vallo, Portoscuso)
21	407.7	152894	0.002	0.502	(Santa Ninfa, San Giovanni Suergiu)
22	411.7	252966	0.004	0.507	(Vita, Calasetta)
23	413.8	135790	0.002	0.509	(Giungano, Portopalo di Capo Passero)
24	416.8	195508	0.003	0.512	(Santa Ninfa, Calasetta)
25	419.5	171870	0.003	0.515	(Cagli, Vignone)
26	421.9	160844	0.003	0.517	(Partanna, Calasetta)
27	424.2	146760	0.002	0.520	(Torrita di Siena, Perito)
28	427.4	209442	0.003	0.523	(Umbertide, San Bernardino Verbano)
29	429.8	160176	0.003	0.525	("Campo nellElba", Masainas)
30	431.9	132634	0.002	0.528	(Camerano, Taceno)
31	434.1	146030	0.002	0.530	(Orbetello, Sala Consilina)
32	436.5	160766	0.003	0.533	(Borgo Pace, "SantAngelo a Fasanella")
33	438.9	155548	0.002	0.535	(Guanzate, Montecassiano)
34	440.8	129086	0.002	0.537	(Borgo Pace, Corleto Monforte)
35	443.1	154146	0.002	0.540	(Teora, Ragusa)
36	446.0	191916	0.003	0.543	(Città di Castello, Futani)
37	448.9	197168	0.003	0.546	(Castiglione del Lago, Domodossola)
38	451.4	165464	0.003	0.549	(Bibbona, Castelfranci)
39	453.5	143382	0.002	0.551	(Castiglione del Lago, Bognanco)
40	455.5	136260	0.002	0.553	(Senigallia, Nemoli)
41	457.3	116818	0.002	0.555	(Varese, Viterbo)
42	459.9	183474	0.003	0.558	(Cagli, Trasquera)
43	462.1	144276	0.002	0.560	(Città di Castello, San Giovanni a Piro)
44	464.0	136254	0.002	0.562	(Peglio, Grassano)
45	466.0	137450	0.002	0.565	(Rancio Valcuvia, Sirolo)
46	467.8	122780	0.002	0.567	(Cassano Valcuvia, Sirolo)
47	469.8	140424	0.002	0.569	(Villongo, Montefino)
48	471.5	115082	0.002	0.571	(Treviglio, Mandela)
49	473.9	167052	0.003	0.573	(Colli al Metauro, Castelsaraceno)
50	476.2	160610	0.003	0.576	(Pieve Santo Stefano, Irsina)
51	478.3	148072	0.002	0.578	(Tavullia, Casaletto Spartano)

Continued

Spatial Lag	h-distance (km)	Links	% Links	J(k)	Name of the Italian municipalities that determine the h-distance
52	480.3	141082	0.002	0.581	(San Godenzo, Ricigliano)
53	482.6	164560	0.003	0.583	(Orbetello, Latronico)
54	485.1	172450	0.003	0.586	(San Giovanni Valdarno, Genzano di Lucania)
55	487.8	197618	0.003	0.589	("Colle di Val dElsa", "SantAngelo Le Fratte")
56	489.7	131192	0.002	0.591	(Castiglione della Pescaia, Anzi)
57	491.6	135422	0.002	0.593	(Costa Serina, Montefino)
58	493.4	126432	0.002	0.595	(Cappelle sul Tavo, Librizzi)
59	495.3	134642	0.002	0.598	(Monterotondo Marittimo, Brienza)
60	497.8	171850	0.003	0.600	(Cesena, Abriola)
61	499.9	148382	0.002	0.603	(Bionaz, Mondavio)
62	502.6	191420	0.003	0.606	(Etroubles, Cantiano)
63	504.7	144616	0.002	0.608	(Statte, Fano)
64	507.2	179144	0.003	0.611	(Peglio, Pisticci)
65	509.4	152160	0.002	0.613	(Cassiglio, Montefino)
66	511.4	139900	0.002	0.616	(Borgo Pace, Praia a Mare)
67	513.3	134538	0.002	0.618	(Bibbiena, Cirigliano)
68	515.8	173082	0.003	0.621	(Galeata, Marsicovetere)
69	517.7	139306	0.002	0.623	(Greve in Chianti, Castelmezzano)
70	520.2	173528	0.003	0.626	(Lizzano, Mondolfo)
71	522.4	149706	0.002	0.628	(Ischia di Castro, Montemesola)
72	525.0	182560	0.003	0.631	(Torricella, Mondolfo)
73	527.0	135842	0.002	0.633	(Gradara, San Paolo Albanese)
74	529.2	155050	0.002	0.636	(San Giovanni Valdarno, Miglionico)
75	532.0	193266	0.003	0.639	(Cellere, Lizzano)
76	534.8	193452	0.003	0.642	(Greve in Chianti, Cirigliano)
77	537.6	189660	0.003	0.645	(Monteiasi, Tavullia)
78	539.5	133562	0.002	0.647	(Morgex, Fabriano)
79	542.4	196648	0.003	0.650	(Cellere, Maruggio)
80	544.6	156582	0.003	0.653	(Torrita di Siena, Mongrassano)
81	547.1	165624	0.003	0.655	(Colli al Metauro, Fuscaldo)
82	549.2	146582	0.002	0.658	(Poggibonsi, Aliano)
83	551.2	135322	0.002	0.660	(Colli al Metauro, San Benedetto Ullano)
84	553.5	156334	0.003	0.662	(Piombino, Grumo Appula)
85	555.8	158972	0.003	0.665	(Sarteano, Torricella)
86	558.3	167494	0.003	0.668	(Tuoro sul Trasimeno, Zumpano)
87	560.3	132346	0.002	0.670	(Bettona, Pianopoli)
88	562.6	159984	0.003	0.672	(Cesena, Noepoli)
89	564.8	144134	0.002	0.675	(Terricciola, Corleto Perticara)
90	567.1	157634	0.003	0.677	(Forlimpopoli, Tursi)
91	569.3	144650	0.002	0.680	(Roccalbegna, Longobucco)
92	571.8	167254	0.003	0.682	(Bagno di Romagna, Statte)

Continued

Spatial Lag	h-distance (km)	Links	% Links	J(k)	Name of the Italian municipalities that determine the h-distance
93	573.7	127608	0.002	0.684	(Siena, Amendolara)
94	576.3	168020	0.003	0.687	(Crespina Lorenzana, Corleto Perticara)
95	578.4	137270	0.002	0.689	(Sava, Pennabilli)
96	581.3	187026	0.003	0.692	(Forlimpopoli, Nova Siri)
97	583.2	129670	0.002	0.694	(Bagno di Romagna, Taranto)
98	587.4	266656	0.004	0.699	(Nepi, Lentini)
99	591.7	282090	0.005	0.703	(Collevecchio, Lentini)
100	597.6	376858	0.006	0.709	(Ronciglione, Lentini)
101	602.3	298352	0.005	0.714	(Caprarola, Lentini)
102	607.4	320232	0.005	0.719	(Vasanello, Lentini)
103	612.7	339328	0.005	0.725	(Bassano in Teverina, Lentini)
104	618.1	337658	0.005	0.730	(Viterbo, Lentini)
105	626.8	548464	0.009	0.739	(Ladispoli, Pozzallo)
106	635.8	555474	0.009	0.748	(Scandriglia, Pachino)
107	644.8	558236	0.009	0.757	(Anguillara Sabazia, Pachino)
108	654.7	605934	0.010	0.766	(Monterosi, Pachino)
109	664.2	565212	0.009	0.776	(Civitavecchia, Pachino)
110	672.8	493060	0.008	0.783	(Villa San Giovanni in Tuscia, Pachino)
111	682.5	532642	0.009	0.792	(Soriano nel Cimino, Pachino)
112	691.6	473420	0.008	0.800	(San Gemini, Pachino)
113	701.5	487816	0.008	0.807	("Civitella dAgliano", Pachino)
114	712.1	495336	0.008	0.815	(Capalbio, Pachino)
115	721.5	412406	0.007	0.822	(Onano, Pachino)
116	729.5	337244	0.005	0.827	(Allerona, Pachino)
117	740.2	421080	0.007	0.834	(Scansano, Pachino)
118	749.5	347336	0.006	0.840	(Arcidosso, Pachino)
119	758.3	310714	0.005	0.845	(Castiglione del Lago, Pachino)
120	768.4	341190	0.005	0.850	(Castiglione della Pescaia, Pachino)
121	780.0	374746	0.006	0.856	(Portoferraio, Pachino)
122	790.1	313472	0.005	0.861	(Massa Marittima, Pachino)
123	800.9	326350	0.005	0.867	(Monterotondo Marittimo, Pachino)
124	810.8	296858	0.005	0.871	(Predoi, Lotzorai)
125	820.9	305822	0.005	0.876	(Predoi, Ortueri)
126	831.3	311214	0.005	0.881	(Cecina, Pachino)
127	842.0	318078	0.005	0.886	(Predoi, Villaurbana)
128	852.5	313336	0.005	0.891	(Predoi, Setzu)
129	863.3	329662	0.005	0.897	(Santa Maria a Monte, Pachino)
130	873.9	338908	0.005	0.902	(Predoi, Sanluri)
131	884.6	353692	0.006	0.908	(Predoi, Gonnosfanadiga)
132	894.5	340330	0.005	0.913	(San Benedetto Val di Sambro, Pachino)
133	905.2	386932	0.006	0.919	(Fabbriche di Vergemoli, Pachino)

Continued

Spatial Lag	h-distance (km)	Links	% Links	J(k)	Name of the Italian municipalities that determine the h-distance
134	915.5	392128	0.006	0.926	(Pieve Fosciana, Pachino)
135	926.4	425120	0.007	0.933	(Predoi, Santadi)
136	936.8	414170	0.007	0.939	(Predoi, Calasetta)
137	947.6	430800	0.007	0.946	(Pignone, Pachino)
138	957.2	376818	0.006	0.952	(Bomporto, Pachino)
139	968.1	408870	0.007	0.959	(Maissana, Pachino)
140	979.4	407604	0.007	0.965	(Leivi, Pachino)
141	990.4	370348	0.006	0.971	("Santo Stefano dAveto", Pachino)
142	1002.2	360116	0.006	0.977	(Pompeiana, Pachino)
143	1014.6	337812	0.005	0.983	(Bajardo, Pachino)
144	1026.2	270758	0.004	0.987	(Garessio, Pachino)
145	1037.9	225794	0.004	0.991	(Malvicino, Pachino)
146	1049.9	180950	0.003	0.993	(Niella Belbo, Pachino)
147	1061.7	133604	0.002	0.996	(Montelupo Albese, Pachino)
148	1073.6	94374	0.002	0.997	(Caraglio, Pachino)
149	1085.5	62782	0.001	0.998	(Lagnasco, Pachino)
150	1097.5	44050	0.001	0.999	(Rifreddo, Pachino)
151	1109.7	30902	0.000	0.999	(Bellino, Pachino)
152	1121.7	19482	0.000	1.000	(Frossasco, Pachino)
153	1133.3	11220	0.000	1.000	(San Francesco al Campo, Pachino)
154	1145.0	6466	0.000	1.000	(Germagnano, Pachino)
155	1157.2	3808	0.000	1.000	(Salbertrand, Pachino)
156	1168.2	1896	0.000	1.000	(Valprato Soana, Pachino)
157	1179.4	932	0.000	1.000	(Cogne, Pachino)
158	1191.0	492	0.000	1.000	(Gressan, Pachino)
159	1201.7	196	0.000	1.000	(Gignod, Pachino)
160	1213.6	80	0.000	1.000	(Prè Saint-Didier, Pachino)
161	1219.7	12	0.000	1.000	(Courmayeur, Pachino)
162	1224.0	2	0.000	1.000	(Courmayeur, Portopalo di Capo Passero)

6.1.2. CS Distance

Number of municipalities = 7,890 – Total links = 62,244,210

Spatial lag	h-distance (km)	Links	% Links	J(k)
1	16.3	269852	0.004	0.004
2	32.6	707810	0.011	0.016
3	48.9	1037740	0.017	0.032
4	65.2	1285378	0.021	0.053
5	81.5	1456148	0.023	0.076
6	97.8	1576398	0.025	0.102
7	114.1	1645254	0.026	0.128

Continued

Spatial lag	h-distance (km)	Links	% Links	J(k)
8	130.4	1674670	0.027	0.155
9	146.7	1670984	0.027	0.182
10	163.0	1643270	0.026	0.208
11	179.3	1585280	0.025	0.234
12	195.6	1509674	0.024	0.258
13	211.9	1439704	0.023	0.281
14	228.2	1363314	0.022	0.303
15	244.5	1298370	0.021	0.324
16	260.8	1244432	0.020	0.344
17	277.1	1200846	0.019	0.363
18	293.4	1168032	0.019	0.382
19	309.6	1144248	0.018	0.400
20	325.9	1115892	0.018	0.418
21	342.2	1089716	0.018	0.436
22	358.5	1059130	0.017	0.453
23	374.8	1029234	0.017	0.469
24	391.1	1015950	0.016	0.486
25	407.4	1026014	0.016	0.502
26	423.7	1051418	0.017	0.519
27	440.0	1071476	0.017	0.536
28	456.3	1093704	0.018	0.554
29	472.6	1123552	0.018	0.572
30	488.9	1146420	0.018	0.590
31	505.2	1147878	0.018	0.609
32	521.5	1143024	0.018	0.627
33	537.8	1126598	0.018	0.645
34	554.1	1115188	0.018	0.663
35	570.4	1097028	0.018	0.681
36	586.7	1066738	0.017	0.698
37	603.0	1043360	0.017	0.715
38	619.3	1029782	0.017	0.731
39	635.6	1017356	0.016	0.748
40	651.9	1004440	0.016	0.764
41	668.2	965410	0.016	0.779
42	684.5	900172	0.014	0.794
43	700.8	820590	0.013	0.807
44	717.1	753354	0.012	0.819
45	733.4	683142	0.011	0.830
46	749.7	620716	0.010	0.840
47	766.0	564658	0.009	0.849
48	782.3	524888	0.008	0.857
49	798.6	499496	0.008	0.865
50	814.9	491084	0.008	0.873
51	831.2	489564	0.008	0.881
52	847.5	483856	0.008	0.889
53	863.8	495646	0.008	0.897
54	880.1	526450	0.008	0.905

Continued

Spatial lag	h-distance (km)	Links	% Links	J(k)
55	896.4	557884	0.009	0.914
56	912.6	602946	0.010	0.924
57	928.9	637266	0.010	0.934
58	945.2	649362	0.010	0.945
59	961.5	634632	0.010	0.955
60	977.8	597054	0.010	0.964
61	994.1	541932	0.009	0.973
62	1010.4	472068	0.008	0.981
63	1026.7	390872	0.006	0.987
64	1043.0	297590	0.005	0.992
65	1059.3	208124	0.003	0.995
66	1075.6	130388	0.002	0.997
67	1091.9	75550	0.001	0.998
68	1108.2	46484	0.001	0.999
69	1124.5	25958	0.000	1.000
70	1140.8	12398	0.000	1.000
71	1157.1	5718	0.000	1.000
72	1173.4	2452	0.000	1.000
73	1189.7	878	0.000	1.000
74	1206.0	264	0.000	1.000
75	1222.3	60	0.000	1.000
76	1238.6	2	0.000	1.000

6.2. SGI Python Routine

To compute the Spatial Gini Index (SGI), an ad-hoc library was developed and implemented in Python to allow the immediate use of the new indicator. The ease of use and operational flexibility of this language are the main features of the routine. The library is particularly fast because it uses the well-known “NumPy” (Harris et al. 2020) and “SciPy” (Virtanen et al. 2020) libraries which, thanks to the vectorization of the functions, guarantee a high execution speed and represent the standard in matrix numerical calculation. One of the most famous and high performing libraries for Exploratory Spatial Data Analysis is certainly the PySAL library (Python Spatial Analysis Library) (Rey and Anselin 2010). Nevertheless, for efficiency reasons required by our use case it was necessary to create some “in-house” modules in order to better implement the SGI algorithm. The *k*-partitions obtained are also used for the spatial decomposition of the Gini index proposed by Rey and Smith (2013).

The principle of operation of the library is as follows:

The library can be downloaded as a Python package (Pirrotta 2022).

1. The SGI class accepts as input the geographic points (optionally the labels), the target variable and the type of partitioning (MM approach is considered here) – the geographic points are the centroids of the territorial units considered,
2. The Euclidean distance matrix is calculated,
3. The variability matrix is calculated,

4. The algorithm finds all the n * (n-1) connections between the n territorial units. This process takes place inside a loop,
5. According to the MM method, for each territorial unit the minimum distance with all the other units is taken into consideration. Among these n minimum values, the maximum is taken. This value represents the h-distance,
6. For each iteration, only connections inside the range distance between minimum threshold (previous h-distance) and maximum threshold (current h-distance) are selected,
7. For each iteration the variability matrix is multiplied element-wise by the spatial weight matrix generated in point 6. By adding the values obtained, the total variability for each spatial-lag is obtained,
8. In order to calculate the successive h-distances according to the SGI approach, at each iteration it is necessary to filter from the Euclidian distance matrix the distances between units less than or equal to the previous h-distance. and
9. When all the units are connected, the process ends with the calculation of the index. The result obtained is the Spatial Gini Index (SGI).

The flow chart of the procedure is presented in the figure below.

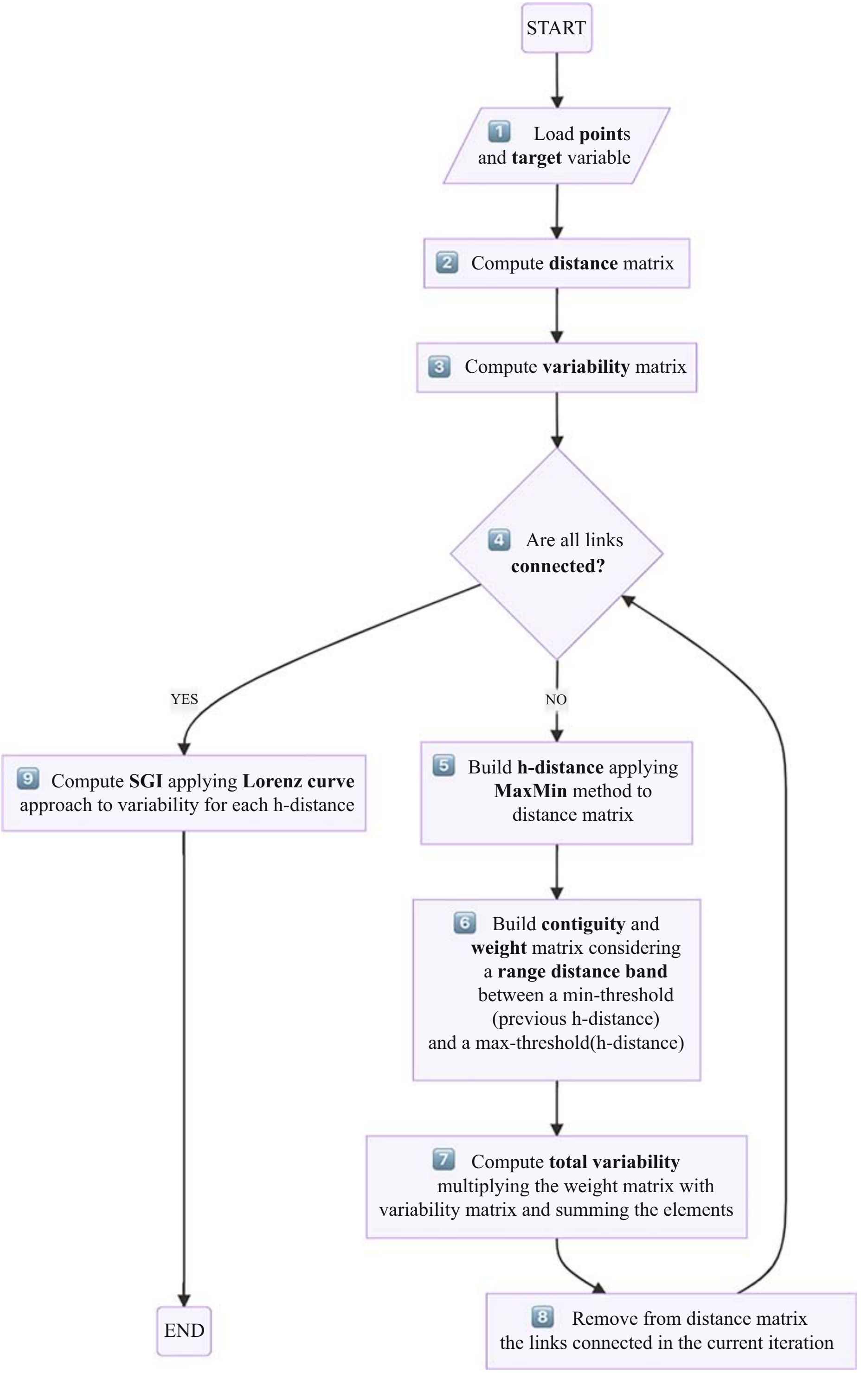


Fig. 13. Flow chart of the procedure for the computation of SGI.

6.3. Algorithm and Pseudo Code

SGI Max Min Algorithm

INPUT
$U := \{u_1, u_2 ... u_n\}$ with $u_i \in R^2$, n geographic points
$X := \{x_1, x_2 ... x_n\}$ with $x_i \in R$, n target variable values

OUTPUT
$sgi \leftarrow$ the Spatial Gini Index

START
$t = 0$
$E^0 = (d_{ij})_{n\times n}$ where $d_{ij} = ||u_i - u_j||_2$ with $u \in U, i = 1 \ldots n, j = 1 \ldots n$

while $E^t \neq 0_{n\times n}$ **do**
    $t = t + 1$
    $h^t_{MM} = \max(d_1, d_2 ... d_n)$ where $d_j = \min\limits_{i=1 \ldots n}(\{d_{ij}\} \setminus \{0\}),\ j = 1 \ldots n, d_{ij} \in E^{t-1}$
    $\Omega^t = (\omega_{ij})_{n\times n}$ where $\omega_{ij} = 1$ if $d_{ij} \leq h^t_{MM}$ and $d_{ij} \neq 0$ otherwise $\omega_{ij} = 0$ where $d_{ij} \in E^{t-1}$
    $E^t = (d^t_{ij})_{n\times n}$ where $d^t_{ij} = d^{t-1}_{ij} \in E^{t-1}$ if $d^{t-1}_{ij} > h^t_{MM}$ otherwise 0
end

$S = (s_{ij})_{n\times n}$ where $s_{ij} = (x_i - x_j)^2$ with $x \in X, i = 1 \ldots n, j = 1 \ldots n$
$D = \sum_{k=1}^{t} \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij} w^k_{ij}$ with $s \in S$ and $w^k \in \Omega^k$
$A = \sum_{k=1}^{t} \sum_{i=1}^{n} \sum_{j=1}^{n} w^k_{ij}$ with $w^k \in \Omega^k$
$J_{(k)} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w^k_{ij}}{A}$ with $k = 1 \ldots t$
$V_{(k)} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} w^k_{ij}}{D}$ with $k = 1 \ldots t$ 2

SGI $= 1 - 0.5 \sum_{k=1}^{t} (V_{(k)} + V_{(k-1)})(J_{(k)} - J_{(k-1)})$

SGI Constant Step Algorithm

INPUT
$U := \{u_1, u_2 ... u_n\}$ with $u_i \in R^2$, n geographic points
$X := \{x_1, x_2 ... x_n\}$ with $x_i \in R$, n target variable values

OUTPUT
$sgi \leftarrow$ the Spatial Gini Index

START
$t = 0$
$E^0 = (d_{ij})_{n \times n}$ where $d_{ij} = ||u_i - u_j||_2$ with $u \in U, i = 1 \ldots n, j = 1 \ldots n$
$h_{CS} = \max(d_1, d_2 ... d_n)$ where $d_j = \min\limits_{i=1...n}(\{d_{ij}\} \setminus \{0\}),\ j = 1 \ldots n, d_{ij} \in E^0$

while $E^t \neq 0_{n \times n}$ **do**
    $t = t + 1$
    $h_{CS} = h_{CS} \times t$
    $\Omega^t = (\omega_{ij})_{n \times n}$ where $\omega_{ij} = 1$ if $d_{ij} \leq h_{CS}$ and $d_{ij} \neq 0$ otherwise $\omega_{ij} = 0$ where $d_{ij} \in E^{t-1}$
    $E^t = (d^t_{ij})_{n \times n}$ where $d^t_{ij} = d^{t-1}_{ij} \in E^{t-1}$ if $d^{t-1}_{ij} > h_{CS}$ otherwise 0
end

$S = (s_{ij})_{n \times n}$ where $s_{ij} = (x_i - x_j)^2$ with $x \in X, i = 1 \ldots n, j = 1 \ldots n$
$D = \sum_{k=1}^{t} \sum_{i=1}^{n} \sum_{j=1}^{n} s_{ij} w^k_{ij}$ with $s \in S$ and $w^k \in \Omega^k$
$A = \sum_{k=1}^{t} \sum_{i=1}^{n} \sum_{j=1}^{n} w^k_{ij}$ with $w^k \in \Omega^k$
$J_{(k)} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w^k_{ij}}{A}$ with $k = 1 \ldots t$
2
$V_{(k)} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij} w^k_{ij}}{D}$ with $k = 1 \ldots t$

SGI $= 1 - 0.5 \sum_{k=1}^{t} (V_{(k)} + V_{(k-1)})(J_{(k)} - J_{(k-1)})$

SGI MaxMin Pseudocode

```
1  INPUT
2  points ← geographic points (centroids) of the territorial units
3  target ← phenomenon variable to analyze
4
5  OUTPUT
6  sgi ← the Spatial Gini Index
7
8  INITIZIALIZATION
9  progressive_num_links ← 0
10 min_threshold ← 0
11 links ← [ ]
12 contiguity_variabilities ← [ ]
13
14 START
15 distance_matrix ← euclidean distance matrix build from points
16 target_matrix ← (s_ij)_{n×n} where s_ij = (x_i − x_j)^2 with x ∈ target, i = 1...n, j =
     1...n
17 total_links ← ntu × (ntu − 1) // ntu = number territorial units
18
19 while progressive_num_links < total_links do
20   |  max_threshold ← max of minimums of distance_matrix
21   |  num_links ← number of elements of distance_matrix ≤ max_threshold
22   |  variability ←
     |    spatial_lag_total(target_matrix, distance_matrix, min_threshold, max_threshold)
23   |  progressive_num_links ← progressive_num_links + num_links
24   |  min_threshold ← max_threshold
25   |  distance_matrix ← update_distance_matrix(distance_matrix, max_threshold)
26   |  links[] ← num_links
27   |  contiguity_variabilities[] ← variability
28 end
29
30 area ← trapezoidify(contiguity_variabilities, links)
31 sgi ← 1 − area
32
33 return sgi
```

6.2.1. Pseudocode Explanation

The library was developed using programming techniques and strategies to ensure speed, robustness, scalability, high performance, and efficiency. Below the explanation of the algorithm.

2–3: The algorithm accepts as input the geographic **points** (centroids of the territorial units considered) and the **target** variable

15: the Euclidean **distance matrix** is calculated

16: the **target matrix** is calculated

17: the algorithm finds all the $n \times (n - 1)$ connections between the n territorial units

19: the loop ends when **all links** between territorial units are **connected**

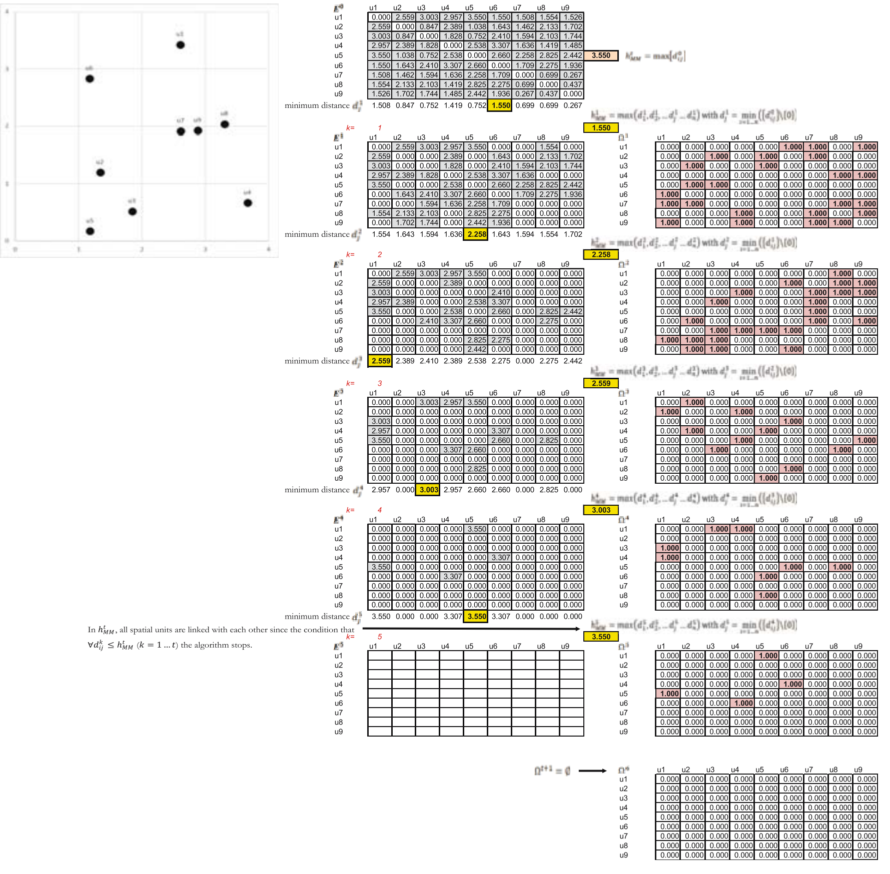
20: Starting from the **distance matrix** (dim. $ntu \times ntu$), for each territorial unit the **minimum** distance with all the other units is taken into consideration. Among these n

minimum values (array), the **maximum** (scalar) is taken. This value represents both the max **threshold** and the current **h distance**
21: for each iteration the number of links corresponds to the number of values in the distance matrix less or equal to the maximum threshold (current h distance);
22: In the **spatial_lag_total** procedure a spatial boolean weights matrix is created (RangeDistance-Band object) selecting connections between minimum and maximum threshold. The target matrix is so multiplied **element-wise** by the spatial weight matrix. By adding the values obtained, the **total variability** for each spatial-lag is obtained;
25: for each iteration we update the distance matrix removing (reset to zero) values less or equal to the maximum threshold;
30: for the estimation of **Gini coefficients** we calculate the **Lorenz curve** area using the **trapezium rule**
31: the **Spatial Gini Index** is obtained subtracting area from 1

6.3. Summary of the Results

Years	Italian population				Foreign population			
	MM		CS		MM		CS	
	G	SGI	G	SGI	G	SGI	G	SGI
2002	0.712	0.485	0.712	0.485	0.802	0.446	0.802	0.444
2010	0.710	0.485	0.710	0.485	0.782	0.448	0.782	0.447
2018	0.715	0.483	0.715	0.483	0.797	0.450	0.797	0.449

6.4. Toy Example MaxMin Distance Method Procedure (Simulated Nine Spatial Units)



7. References

Alleva, G. 1987. “Autocorrelazione spaziale nel caso di contiguità di ordine superiore al primo.” In Proceedings of the Italian Statistical Society Conference “Informazione ed analisi statistica per aree regionali e sub regionali”, Galeno Editrice, Perugia, Italia: 1–10.

Anselin, L. 1999. “The Future of Spatial Analysis in the Social Sciences.” *Geographic Information Sciences* 5(2): 67–76. DOI: https://doi.org/10.1080/10824009909480516.

Arbia, G. 2001. “The role of spatial effects in the empirical analysis of regional concentration.” *Journal of Geographical Systems* 3(3): 271–281. DOI: https://doi.org/10.1007/PL00011480.

Arbia, G., R. Benedetti, and G. Espa. 1996. “Effect of MAUP on image classification.” *Geographical System* 3: 123–141.

Arbia, G., and G. Piras. 2009. “A new class of spatial concentration measures.” *Computational Statistics and Data Analysis* 53(21): 4471–4481. DOI: https://doi.org/10.1016/j.csda.2009.07.003.

Benassi, F., F. Lipizzi, and S. Strozza. 2019. "Detecting foreigners' spatial residential patterns in urban contexts: two tales from Italy." *Applied Spatial Analysis and Policy* 12(2): 301–319. DOI: https://doi.org/10.1007/s12061-017-9243-5.

Benassi, F., R. Iglesias-Pascual, and L. Salvati. 2020. "Residential segregation and social diversification: Exploring spatial settlement patterns of foreign population in Southern European cities." *Habitat International* 101:102200. DOI: https://doi.org/10.1016/j.habitatint.2020.102200.

Bonferroni, C.E. 1938. *Elementi di Statistica Generale*. Litografai Feilice Gili: Torino.

Brown, L.A., and S.Y. Chung. 2006. "Spatial segregation, segregation indices and the geographical perspectives." *Population, Space and Place* 12(2): 125–143. DOI: https://doi.org/10.1002/psp.403.

Crespo, R., and I. Hernandez. 2020. "On the spatially explicit Gini coefficient: The case study of Chile – a high-income developing country." *Letters in Spatial and Resource Sciences* 13(1): 37–47. DOI: https://doi.org/10.1007/s12076-020-00243-4.

Dawkins, C.J. 2004. "Measuring the spatial pattern of residential segregation." *Urban Studies* 41(4): 833–851. DOI: https://doi.org/10.1080/0042098042000194.

Dawkins, C.J. 2006. "The spatial pattern of black-white segregation in US metropolitan areas: An exploratory analysis." *Urban Studies* 43(11): 1943–1969. DOI: https://doi.org/10.1080/00420980600897792.

Eurostat. 2015. *Geospatial analysis at Eurostat*. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Geospatial_analysis_at_Eurostat#Spatial_analysis_for_statistical_purposes.

Fotheringham A.S. 1997. "Trends in Quantitative Methods I: Stressing the Local." *Progress in Human Geography* 21(1): 88–96. DOI: https://doi.org/10.1191/030913297676693207.

Frosini, B.V. 1996. "Recenti contributi Italiani alla statistica descrittiva." *Statistica Applicata* 8(2): 361–396.

Gini, C. 1912. "Variabilità e mutabilità. Contributo allo studio delle distribuzioni e delle relazioni statistiche." *Studi economico-giuridici, Anno III, parte II. Facoltà di giurisprudenza della Regia Università di Cagliari*, Cuppini, Bologna.

Gini, C. 1914. "Sulla misura della concentrazione della variabilità dei caratteri." *Atti del reale istituto Veneto di scienze, lettere e arti* 73(2): 1203–1248.

Gini, C. 1921. "Measurement of inequality of incomes." *Economic Journal* 31: 124–126. DOI: https://doi.org/10.2307/2223319.

Gini, C. 2005. "On the measurement of concentration and variability of characters." *Metron* 63(1): 3–38. Available at: https://www.dss.uniroma1.it/RePec/mtn/articoli/2005-1-1.pdf.

Harris, C.R., K.J. Millman, S.J. van der Walt. R. Gommers, P. Virtanen, David Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus, S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T.E. Oliphan. 2020. "Array programming with NumPy." *Nature* 585: 357–362. DOI: https://doi.org/10.1038/s41586-020-2649-2.

Hoover, E. 1941. "Interstate redistribution of population, 1850–1940." *Journal of Economic History* 1: 199–205. DOI: https://doi.org/10.1017/S0022050700052980.

Howell, F.M., J.R. Porter, and S.A. Matthews. 2016. *Recapturing space: New middle-range theory in spatial demography*. Berlin: Springer. DOI: https://doi.org/10.1007/978-3-319-22810-5.

Livi Bacci, M. 1999. *The Population of Europe: A History*. Oxford: Blackwell

Leti G. 1983 *Statistica descrittiva*, Bologna: Il Mulino.

Logan J.R. 2012. "Making a Place for Space: Spatial Thinking in Social Science." *Annual Review of Sociology* 38: 507–524. DOI: https://doi.org/10.1146/annurev-soc-071811-145531.

Lorenz, M.O. 1905. "Methods of measuring the concentration of wealth." *Publications of the American Statistical Association* 9(70): 209–219. DOI: https://doi.org/10.2307/2276207.

Matthews, S.A., and D.M. Parker. 2013. "Progress in spatial demography." *Demographic Research* 28: 271–312. DOI: https://doi.org/10.4054/DemRes.2013.28.10.

Massey, D.S., and N.A. Denton. 1988. "The dimensions of residential segregation." *Social forces* 67(2): 281–315. DOI: https://doi.org/10.1093/sf/67.2.281.

Miller, H.J. 2004. "Tobler's first law and spatial analysis." *Annals of the association of American geographers* 94(2): 284–289. DOI: https://doi.org/10.1111/j.1467-8306.2004.09402005.x.

Mucciardi, M. (2008a). "Use of a flexible weight matrix in a local spatial statistic." In: First joint meeting of the Société Francophone de Classification and the Classification and Data Analysis Group of the Italian Society of Statistics. 5: 385–388. Naples: Edizioni Scientifiche Italiane. DOI: https://doi.org/10.1007/978-3-642-13312-1.

Mucciardi, M. (2008b). "Geographic information and global index of spatial variability." Proceedings of the VI International Conference on Stochastics Geometry, Convex Bodies, Empirical Measures. In *Rendiconti del Circolo dei Matematici di Palermo 80*: 241–250.

Mucciardi, M., and F. Benassi. 2023. "Measuring the spatial concentration of population. A new approach based on the graphical representation of the Gini index." *Quality and Quantity*. DOI: https://doi.org/10.1007/s11135-022-01607-2.

Mucciardi, M., and P. Bertuccelli. 2007. "S-Joint: a new software for the analysis of spatial data". In Proceedings of the Italian Statistical Society Annual Conference, Venice, 6–8 June: 647–648. Available at: http://sis2007.unive.it/.

Novelli, E. and S. Ocelli. 1999. "Describing spatial patterns' profile: diversification measure." *European Journal of Geography*. DOI: https://doi.org/10.4000/cybergeo.4930.

Panzera, D., and P. Postiglione. 2020. "Measuring the spatial dimension of regional inequality: An approach based on the Gini correlation measure." *Social Indicators Research* 148(2): 379–394. DOI: https://doi.org/10.1007/s11205-019-02208-7.

Pearl, R., and L.J. Reed. 1920. "On the rate of growth of the population of the USA since 1970 and its mathematical representation." *Proceedings of the National Academy of Science* 6: 275-288. DOI: https://doi.org/10.1073/pnas.6.6.275.

Pirrotta, G. 2022. *The Spatial Gini Index, a Python library to compute a new indicator to measure the spatial concentration*. Available at: https://github.com/gpirrotta/spatial-gini-index.

Reardon, S.F., and D. O'Sullivan. 2004. "Measures of spatial segregation." *Sociological Methodology* 34(1): 121–162. DOI: https://doi.org/10.1111/j.0081-1750.2004.00150.x.

Rey, S.J., and L. Anselin. 2010. "PySAL: A Python Library of Spatial Analytical Methods." In *Handbook of Applied Spatial Analysis,* edited by *M*. Fischer, and A. Getis. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-642-03647-7_11.

Rey, S.J., and R.J. Smith. 2013. "A spatial decomposition of the Gini coefficient." *Letters in Spatial and Resource Sciences* 6(2): 55–70. DOI: https://doi.org/10.1007/s12076-012-0086-z.

Strozza, S., F. Benassi, R. Ferrara, and G. Gallo. 2016. "Recent demographic trends in the major Italian urban agglomerations: The role of foreigners." *Spatial Demography* 4(1): 39–70. DOI: https://doi.org/10.1007/s40980-015-0012-2.

Tobler, W.R. 1970. "A computer movie simulating urban growth in the Detroit region." *Economic Geography* 46: 234–240. DOI: https://doi.org/10.2307/143141.

Verhulst, P.F. 1838. "Notice sur la loi que la population poursuit dans accroissement." *Correspondance Mathématique et Physique* 10: 113-121.

Voss, P.R. 2007. "Demography as a spatial social science." *Population research and policy review* 26(5): 457–476. DOI: https://doi.org/10.1007/s11113-007-9047-4

Wright, J.K. 1937. "Some measures of distributions." *Annals of the Association of American Geographers* 27(4): 177–211.

Virtanen, P., R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, İ. Polat, Y. Feng, E.W. Moore, and J. VanderPlas. 2020. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." *Nature Methods*, 17: 261–272. DOI: https://doi.org/10.1038/s41592-019-0686-2.

Zenga, M. 1984. "Proposta per un indice di concentrazione basato sui rapporti fra quantili di popolazione e quantili di reddito." *Giornale degli Economisti e Annali di Economia* 43(5/6): 301–326. Available at: http://www.jstor.org/stable/23246074.

Zenga, M. 1985. "Un secondo indice di concentrazione basato sui rapporti fra quantili di popolazione e quantili di reddito." *Rivista di Statistica Applicata* 18: 143–154.

Received June 2022
Revised October 2022
Accepted January 2023