Understanding the expansion of Italian metropolitan areas: A study based on entropy measures

Published Version:
Altieri Linda, Cocchi Daniela (2022). Understanding the expansion of Italian metropolitan areas: A study based on entropy measures. ENVIRONMENT AND PLANNING B-PLANNING & DESIGN, 49(2 (February)), 447-463 [10.1177/23998083211012699].

Availability:
This version is available at: https://hdl.handle.net/11585/905950 since: 2022-11-22

Published:
DOI: http://doi.org/10.1177/23998083211012699

Understanding the expansion of Italian metropolitan areas: a study based on entropy measures

Abstract

This work presents a study on the urban configuration of a number of Italian metropolitan areas and their development over time, with the aim of evaluating the size and shape of urban areas expansion. Raster data are used, produced by the European Environmental Agency within the CORINE Land Cover project. The study is based on a version of spatial entropy measures proposed and validated by a recent series of papers, aimed at the evaluation of spatial data heterogeneity; the methods assess the efficiency of the spatial configuration of urban areas. An innovative combination of two entropy measures is the tool for evaluating the urban development in Italy. Results allow both conclusive comments about each metropolitan area and comparisons across areas over space and time.

Keywords: Italian metropolitan areas; CORINE; spatial entropy; urban expansion; urban compactness

1. Introduction

A metropolitan area is an urban structure where a relevant amount of population lives (Tong, Plane 2014), formed by a main city and its influence area. Delimiting a metropolitan area is a difficult matter, due to the expansion of urban phenomena towards external areas (Pacione 2009). From the 1950s, Europe worked (Medina, Monclùs 2018) to control the urban development following industrialization and obtain an expansion with positive effects on its territory. The search of a rational city failed due to rapid and out of control urban development during the 60s and 70s, known as urban sprawl.

Inefficient urban expansion over rural or semi-rural areas is among the negative consequences of the metropolitan phenomenon (Bhatta, Saraswati, Bandyopadhyay 2010; Yamu, Frankhauser 2016). Two main social aspects cause such urbanisation: service and distribution networks take place near main streets and highways and people moving from the countryside to cities have different richness levels. This causes a dangerous tendency towards social division. Although a definition of urban sprawl is still debated (Jaeger et al. 2010; Torrens 2008; Banai, De Priest 2014), a general consensus is that a sprawled city is full of empty spaces denoting inefficient development (EEA 2011).

Urban expansion measurement is performed by different metrics (Dadashpoor, Azizi, Moghadasi 2019). The simplest indicator is the population density over an area, i.e. the number of residents per square kilometre. Further urban growth indicators (Ewing, Hamidi 2015; Jiang et al. 2007) can be compared in time and space (Altieri et al. 2014; Cheng, Masser 2004). A popular tool is Shannon's entropy: it comes from landscape ecology, but its transposition to urban expansion is direct. Ecological concepts of evenness and richness are related to heterogeneity, and entropy proved to be a rigorous and widely used technique, suitable for integration of remote sensing and GIS (Yeh, Li 2001), where information is limited but extensively available (Herold et al. 2003; Chong 2017). Land cover data are particularly suitable for measuring urban expansion: the territory is classified according to the prevailing land use, and split into urban or non-urban patches.

Shannon’s entropy does not consider space as a source of heterogeneity. It is only computed on the proportions of land use classes (estimates of land use probabilities), so territories with the same proportions but different compactness for the urban tissue share the same entropy value. Entropy measures need to incorporate spatial information (Heikkila, Hu 2006) and are the core of the present work.

In Italy, the metropolitan city is a well-defined entity: starting 2012, laws establish 10 metropolitan cities in the mainland (L. 135 in Gazzetta Ufficiale 189/2012; L. 56 in Gazzetta Ufficiale 81/2014). Regional laws established 4 metropolitan cities in Sicily and Sardinia. The 14 Italian metropolitan cities are centred on Bari, Bologna, Cagliari, Catania, Firenze, Genova, Messina, Milano, Napoli, Palermo, Reggio Calabria, Roma, Torino, Venezia. Metropolitan cities substitute the administrative units known as "provinces" (Balducci et al. 2017; Bettoni 2018). The main municipality is the core of each metropolitan city and an attraction pole for residents. It derives from the evolution of urban settlements historically organized in walled towns with recognizable borders and countryside, now rather being built up areas spread over the territory, including a city town, suburbs and other residential settlements. In this work, the administrative border of an Italian metropolitan city cannot be employed, since it regards a wide territory beyond the direct influence area of a main hub. Thus, we apply spatial entropy measures to the main municipality together with its commuting belt, i.e. the minor municipalities sharing a border with the main one and constituting its influence area (see Table S1). In the remainder of the paper, the terms "metropolitan area" and "city and commuting belt" indicate the main city plus its surrounding municipalities.

The main aim of this work is to present an extensive study on the configuration and level of urbanization of the 14 Italian metropolitan areas, using official European data and a rigorous statistical methodology. The use of administrative borders for all municipalities ensures comparability of the results to other areas. The methodological innovation lies in the combination of two entropy measures to describe the spatial development of metropolitan areas. Recently, Altieri, Cocchi, Roli (2019a) proposed the use of spatial entropy measures on urban/non-urban pixels. The urbanization intensity is summarized by Shannon's entropy of a transformation of the study variable, while the intensity level of urban dispersion is measured by the entropy-based measure named spatial mutual information (Altieri, Cocchi and Roli 2019b). The joint use of spatial mutual information and Shannon's entropy offers a new perspective on the analysis of urban data, via an evolution of commonly used entropy-based metrics for landscape studies.

Data come from the European CORINE land cover dataset. For years 1990 and 2012, the territory of the considered municipalities is divided into $250m$x$250m$ pixels, classified as urban/non-urban according to EEA, 2011. The availability of such dataset allows to employ a unified basis for assessing the spatial configuration and evolution of metropolitan areas.

The paper is organized as follows. Section 2 presents the European data for the study. Section 3 summarizes the methodology in the context of urban expansion, with a data example. Section 4 contains the study of Italian metropolitan areas, stressing the combination of urban intensity and compactness. Section 5 gives a final discussion.

2. European land cover data

European land cover data are the basis of the present work. Even if some cities have their own accurate system for monitoring land cover, a common dataset is fundamental for comparison. Two data collecting programmes about land use are run within the European Union: CORINE Land Cover programme by the European Environment Agency (EEA 1994), and LUCAS project under control of Eurostat (Eurostat 2010), used for improving CORINE outcome.

CORINE (COoRdination of INformation on the Environment) is a programme approved by the European Community Council in 1985 for gathering, coordinating and ensuring the consistency of information on the environment. CORINE Land Cover (CLC) is created for monitoring land characteristics: decision-makers need an overview of existing knowledge and information as complete and up-to-date as possible on certain features of the biosphere. In 1990, CORINE involved 12 countries and 2.3 million km^2 (EEA 1994); the 2012 update concerns 39 European countries and 5.8 million km^2. The Community area is divided into units classified in three levels according to CLC nomenclature. The first level includes five macro-categories: artificial surfaces, agricultural areas, forests/seminatural areas, wetlands, water bodies. The second level counts fifteen and the third one forty-four classes.

CORINE data come from remote sensing, i.e. acquisition of information by sensing devices not in physical contact with the object, for land use observation (Taubenböck et al. 2012). Georeferenced data from satellite images are analysed with Geographic Information Systems (GIS) and coded in vector or raster format. Vector data are formed by points, lines, polygons, coded based on their geographical coordinates. Raster data are visualised by a regular grid, or matrix, whose basic element is called cell or pixel. Each pixel contains information about a portion of land. The pixel size is specified in a mapping unit (e.g. metres, kilometres etc.) and is linked to data precision and reliability. The raster format is appropriate for studies involving entropy-type metrics. When a CLC vector dataset is rasterised, the nearest neighbour method is used: each cell is assigned to a land cover category based on the value of the polygon at its centre. Two datasets are produced: CLC European 250 metres grid, and CLC European 100 metres grid, given the precision of the vector dataset and considering they have to be small in order to reduce conversion errors.

EEA mapped urban areas across Europe; the selected areas are classified as "Urban Morphological Zones" (UMZs), made up of four Corine Land Cover core classes (EEA 2011): 'Continuous urban fabric', 'Discontinuous urban fabric', 'Industrial or commercial units', 'Green urban areas'. They are integrated with 'Port areas', 'Airports', 'Road and rail networks', 'Sport and leisure facilities', provided they are neighbours to the core classes. UMZs are useful to identify shapes and patterns of urban areas and detect the land cover lost because of urban development. They form a binary dataset that EEA recommends to use when studying urban development in Europe, especially at a regional scale (EEA 2011).

3. Entropy measures for land cover data

The present Section gives an introduction of the methodology with interpretation in the context of urban data (see Altieri, Cocchi, Roli 2019b for details). Steps are sided by a data example: the metropolitan area of Bologna in 2012 is chosen as a representative case in Italy, its level of urbanization intensity and compactness being close to the Italian mean values (see Figure 2 in Section 4). A simulation reorganizes Bologna's pixels according to a compact and a random spatial scenario (Figure 1), which have the same number of urban and non-urban pixels.

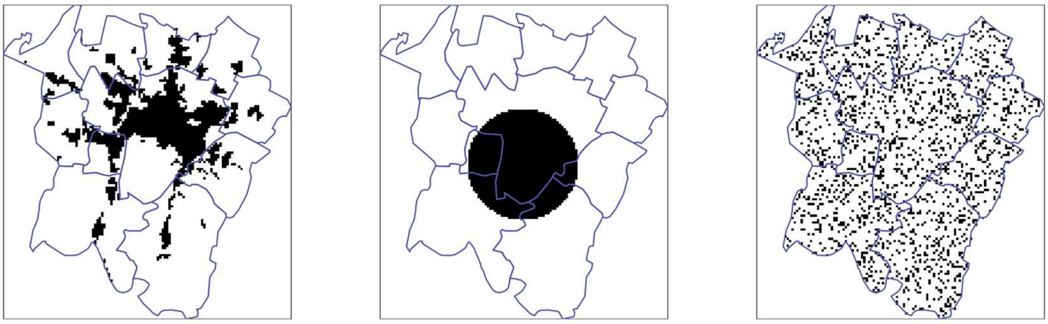

Figure 1: The metropolitan area of Bologna (left panel, 2012 data) and two simulated spatial arrangements of its urban tissue as compact (central panel) and scattered (right panel)

3.1. Non-spatial entropy for the urban tissue: *H*(*X*)

Let X be the categorical variable "land cover", with I land cover categories; let $p(x_i)$ be the probability of land cover type $i=1, \dots, I$. The entropy of land cover is the average information linked to finding one land cover category; it is the expected value of the variable "information function of X" (Cover and Thomas 2006):

$$H(X) = \sum_{i=1}^{I} p(x_i) \log\left(\frac{1}{p(x_i)}\right). \qquad (1)$$

When X has two categories x_1=*urban* and x_0=*non-urban*, with $p=p(x_1)$ being the probability of finding category “urban”, entropy reduces to

$$H(X) = p \log\left(\frac{1}{p}\right) + (1-p)\log\left(\frac{1}{1-p}\right) \qquad (2)$$

and ranges in [0, log2], where log2 is reached for $p=0.5$.

This setting suits to a binary raster dataset with N pixels: each pixel $u=1, \ldots, N$ can be either urban or non-urban. Let x_u be the realization of pixel u; the sum of all x_u is the number of urban pixels N_1. The only quantity that has to be known or estimated to compute entropy is p. Thus, two raster datasets with different urban configurations but a similar urban proportion p return similar entropies: the traditional entropy formula, though widely used in urban development studies, is not able to distinguish among urban patterns. In Bologna’s example, N=10000 pixels fall within the administrative borders, and N_1=1781; values are constant across the configurations of Figure 1. Thus, $p(x_1) = p$ is estimated by $\hat{p}= N_1 / N$ =0.178 for all patterns. Shannon’s entropy is

$$H(X)_{Bo} = 0.178 \log\left(\frac{1}{0.178}\right) + 0.822 \log\left(\frac{1}{0.822}\right) = 0.469$$

irrespective of the spatial arrangement.

Several approaches in the literature include some spatial information into entropy measures. The first milestones are authored by Batty (1974, 1976, 2010) and define a spatial entropy measure accounting for unequal space partition into sub-areas. Such entropy can only be computed for a binary variable, the local terms do not possess the properties of the global one, and results are heavily affected by the selected area partition. Another approach to spatial entropy starts from a transformation of the study variable (O’Neill et al. 1988) and accounts for distance between realizations. Recently, Leibovici and Claramunt (2019) extend the method to spatio-temporal studies with the consideration of the size of the land cover patch. The same approach is the basis for a series of papers by Altieri, Cocchi and Roli (2018, 2019a, b), which propose measures with desirable properties for investigating urban expansion and overcomes the limitations of the first approach.

3.2. Non-spatial entropy for urbanization intensity: *H*(*Z*)

Consider for simplicity a binary classification. When the study variable "land cover" has categories "urban" and "non-urban", three possible pairs can be built if order is irrelevant, i.e. z_1={*urban, urban*}, z_2={*urban, non-urban*} and z_3={*non-urban, non-urban*}. We use the word "order" in its standard mathematical meaning: the couple (*urban, non-urban*) is ordered and different from (*non-urban, urban*); conversely the set, or pair, {*urban, non-urban*} is unordered. Reasons for discarding the order within pairs are discussed in Altieri, Cocchi, Roli (2019b). In the remainder of the paper, the word "pair" refers to unordered sets. The combinations are categories of the new variable "pairs of land cover categories" *Z*.

The general formula for the entropy of *Z* is

$$H(Z) = \sum_{r=1}^{R} p(z_r) \log\left(\frac{1}{p(z_r)}\right) \qquad (3)$$

where *R* is the number of possible pairs. In the binary case $R=3$ and (3) reduces to

$$H(Z) = p(z_1)\log\left(\frac{1}{p(z_1)}\right) + p(z_2)\log\left(\frac{1}{p(z_2)}\right) + p(z_3)\log\left(\frac{1}{p(z_3)}\right) \qquad (4)$$

where the maximum value log3=1.0986 cannot be reached; the actual maximum is 1.0362, reached for $p=p(x_1)=0.5$ (Altieri, Cocchi, Roli 2018).

H(*Z*), as well as *H*(*X*), is the same irrespective of the data spatial configuration, being only based on the probability of occurrence of urban and non-urban pixels. In Bologna's example, the distribution of *Z* is estimated starting from N_1 and $N_0=N - N_1$, by relative frequencies:

$$\hat{p}(z_1) = \frac{n(z_1)}{n(z)}; \quad \hat{p}(z_2) = \frac{n(z_2)}{n(z)}; \quad \hat{p}(z_3) = \frac{n(z_3)}{n(z)},$$

where the computation of the number of pairs is the same for all configurations of Figure 1:

$$n(z)_{Bo} = \binom{N}{2} = \binom{10000}{2} = 49995000$$ total number of pairs in Bologna

$$n(z_1)_{Bo} = \binom{N_1}{2} = \binom{1781}{2} = 1585090$$ number of pairs *urban, urban*}

$$n(z_3)_{Bo} = \binom{N_0}{2} = \binom{8219}{2} = 33771871$$ number of pairs {*non-urban, non-urban*}

$$n(z_2)_{Bo} = \binom{N}{2} - \left(\binom{N_1}{2} + \binom{N_0}{2}\right) = 14638039$$ number of pairs {*urban, non-urban*}.

Then,

$$\hat{p}(z_1) = 0.032; \quad \hat{p}(z_2) = 0.293; \quad \hat{p}(z_3) = 0.675$$

are used to compute *H*(*Z*)

$$H(Z)_{Bo} = 0.032 \log\left(\frac{1}{0.032}\right) + 0.293 \log\left(\frac{1}{0.293}\right) + 0.675 \log\left(\frac{1}{0.675}\right) = 0.734.$$

This procedure allows efficient results in terms of computational time even for large datasets.

Entropy *H*(*Z*), as well as *H*(*X*), is suitable to quantify the intensity of urbanization over an area, intended as proportion of urban tissue, but not spatial compactness.

3.3. Toward a spatially-sensitive entropy measure: *H*(*Z*|*adjacency*)

Studying pairs of urban/non-urban pixels at different distances is relevant to address urban expansion: when most urban pixels are close to other urban pixels, i.e. many {*urban, urban*} pairs for adjacent pixels, the city configuration tends to be compact and efficient. Conversely, in a chaotic urban expansion, many adjacent pairs of pixels are {*urban, non-urban*}. The proportion of the three categories of Z for adjacent pixels can be used to evaluate urban expansion at a small scale. The consideration of adjacent pixels identifies a subset of the general variable Z, say (Z | *distance=adjacency*) which brings along spatial information and is the basis for O'Neill's spatial entropy. Entropy is analogous to $H(Z)$ except for conditioning, with the important consequence of returning different results according to the spatial pattern:

$$H(Z|adj) = p(z_1|adj)\log\left(\frac{1}{p(z_1|adj)}\right) + p(z_2|adj)\log\left(\frac{1}{p(z_2|adj)}\right) + p(z_3|\text{adj})\log\left(\frac{1}{p(z_3|\text{adj})}\right). \quad (5)$$

In Bologna's example, the number of adjacent pairs per category needs to be counted for each spatial settlement

$n(z_1\|adj)_{\text{obs}}=1412$;	$n(z_2\|adj)_{\text{obs}}=719$;	$n(z_3\|adj)_{\text{obs}}=7671$	"observed" Bologna
$n(z_1\|adj)_{\text{comp}}=1734$;	$n(z_2\|adj)_{\text{comp}}=94$;	$n(z_3\|adj)_{\text{comp}}=7974$	"compact" Bologna
$n(z_1\|adj)_{\text{rand}}=318$;	$n(z_2\|adj)_{\text{rand}}=2865$;	$n(z_3\|adj)_{\text{rand}}=6619$	"random" Bologna.

The total number of adjacent pairs $n(z|adj) = 9802$ (the sum of each line above) is the same across spatial configurations.

Consequently, entropy measures $H(Z$ | *adjacency*), estimated based on the relative frequencies of the categories of Z over adjacent pairs, give different results for each scenario:

$$H(Z|adj)_{obs} = 0.144\log\left(\frac{1}{0.144}\right) + 0.073\log\left(\frac{1}{0.073}\right) + 0.783\log\left(\frac{1}{0.783}\right) = 0.663$$

$$H(Z|adj)_{comp} = 0.177\log\left(\frac{1}{0.177}\right) + 0.010\log\left(\frac{1}{0.010}\right) + 0.813\log\left(\frac{1}{0.813}\right) = 0.519$$

$$H(Z|adj)_{rand} = 0.032\log\left(\frac{1}{0.032}\right) + 0.292\log\left(\frac{1}{0.292}\right) + 0.676\log\left(\frac{1}{0.676}\right) = 0.736.$$

The smallest value is found for the compact configuration, and the greatest for the random one. This measure is similar to O'Neill's entropy and is not satisfactory enough: our objective is a measure of compactness taking small values for a random scenario and large values for a compact one. Moreover, the extremes are absolute and unrealistic: value 0 is reached when the territory is entirely urban or entirely non-urban; the maximum is reached when the three types of pairs are present with the same proportion, impossible in real situations. We aim at a spatial measure calibrated on the urban context, that conditions the level of compactness on the intensity of urbanization.

3.4. A spatial entropy measure for urban compactness: $SPI(Z|w_k)$

The evaluation of urban compactness (Burton, 2002) can be enriched by looking in general at pairs lying at given distances; intuitively, the more compact a metropolitan area is, the more pairs of type {*urban, urban*} are present at any distance. A set of K distance ranges from w_1 to w_K can be fixed for investigating urban expansion in detail. They should be defined with more focus on small distances and less focus on very large distances, of scarce interest for the evaluation of urban development. This results in small ranges for the first classes and to wider ranges for the last classes. One example (discussed in Altieri, Cocchi and Roli 2019a, b) follows the standard "nearest neighbour system" (i.e. adjacent pixels) and "12 nearest neighbour system" (Anselin 1995): w_1 collects adjacent pairs of pixels and w_2 collects the immediately farther neighbours, i.e. pairs of pixels sharing a corner and pairs of pixels adjacent to a common one. This allows to understand whether blocks of buildings lie close by or close to rural/non-urban patches. Further classes w_3 to w_K are less relevant to urban studies and are consequently wider; the only requirement is that they cover all possible distances within the metropolitan area. In particular, the last class can be considered as residual and can be very large according to the study context.

By the construction of distance classes, K subsets of realizations of Z become available, formed by pairs of observations belonging to each distance range: $Z|w_k$, with $k=1, \ldots, K$. A measure of the compactness of the urban tissue at the kth distance range is

$$SPI(Z|w_k) = \sum_{r=1}^{R} p(z_r|w_k)\log\left(\frac{p(z_r|w_k)}{p(z_r)}\right), \qquad (6)$$

where $p(z_r|w_k)$ denotes the probability of the rth pair to fall within distance range w_k and $p(z_r)$ is the marginal probability of occurrence of the rth pair at any distance, i.e. a component of $H(Z)$. *SPI* stands for spatial partial information; the city is more and more compact when relevant *SPI* values are found for $k>1$. Note that $SPI(Z|w_k)$ is the Kullback-Leibler divergence between $p(z_r|w_k)$ relative to $p(Z)$ for a given w_k.

One particular case occurs for $w_k=adjacency$, and a binary X, i.e. three categories of Z:

$$SPI(Z|adj) = p(z_1|adj)\log\left(\frac{p(z_1|adj)}{p(z_1)}\right) + p(z_2|adj)\log\left(\frac{p(z_2|adj)}{p(z_2)}\right) + p(z_3|\mathrm{adj})\log\left(\frac{p(z_3|\mathrm{adj})}{p(z_3)}\right).$$

In Bologna's example:

$$SPI(Z|adj)_{obs} = 0.144\log\left(\frac{0.144}{0.032}\right) + 0.073\log\left(\frac{0.073}{0.293}\right) + 0.783\log\left(\frac{0.783}{0.676}\right) = 0.232$$

$$SPI(Z|adj)_{comp} = 0.177\log\left(\frac{0.177}{0.032}\right) + 0.010\log\left(\frac{0.010}{0.293}\right) + 0.813\log\left(\frac{0.813}{0.676}\right) = 0.423$$

$$SPI(Z|adj)_{rand} = 0.032\log\left(\frac{0.032}{0.032}\right) + 0.292\log\left(\frac{0.292}{0.293}\right) + 0.676\log\left(\frac{0.676}{0.676}\right) = 0.00001.$$

This example shows the key difference between entropy (5) and spatial partial information (6). Entropy (5) reaches its maximum when the distribution of $Z|adj$ is uniform, without considering the specific urban context. Conversely, $SPI(Z|adj)$ increases as the conditional distribution of $Z|adj$ steps away from the marginal distribution of Z, i.e. depends on the intensity of urbanization under study. The conditional spatial measure (6) is calibrated on the urban context: in *SPI,* each probability for adjacent pairs is compared to the marginal probability of Z, i.e. to a probability where space is not considered. In the random scenario the conditional and marginal probabilities are approximately the same, transforming *SPI* into a sum of zero values; the greater departure of each conditional probability from the marginal one, the greater value for *SPI*. Entropy (6) increases as the spatial configuration gets closer to a compact scenario, given the proportion of urban and non-urban pixels over the area: we have a measure of spatial compactness conditional on the urbanization intensity.

The weighted sum of *SPI* values results in a very well-known quantity in Information Theory, i.e. mutual information, in the present context renamed spatial mutual information:

$$SMI(Z,W) = \sum_{k=1}^{K} p(w_k) SPI(Z|w_k) \tag{7}$$

where the terms $p(w_k)$ weight each distance range. The second variable W summarizes the distribution of pairs of observations across a set of chosen distance ranges over the observation window. The construction of the distance classes w_k generalizes measures by Leibovici (2009); Leibovici et al. (2014); O'Neill et al. (1988), which do not enjoy the additivity property from local *SPI* terms to the overall *SMI*; the property holds for any choice of the distance classes. Moreover, O'Neill's and Leibovici's measures rely on the choice of a single distance for studying couples, without capturing the general behaviour of the variable of interest.

SMI can be added to the quantity known as residual entropy, $H(Z)_W$, in order to obtain $H(Z)$

$$\begin{aligned} H(Z) &= SMI(Z,W) + H(Z)_W \\ &= \sum_{k=1}^{K} p(w_k)[SPI(Z \mid w_k) + H(Z \mid w_k)] \end{aligned} \tag{8}$$

Differences in the overall level of compactness can be measured by the different contribution of *SMI*, and, in particular, of each *SPI* term, to the total $H(Z)$. An effective way to compare urban compactness across areas is to exploit the last line of (8): at each distance class w_k, the sum $SPI(Z|w_k)+H(Z|w_k)$ can be set to 1, so that the contribution of space given by each *SPI* may be appreciated in proportional terms and is comparable across datasets. For an example, see Altieri, Cocchi and Roli (2019b).

In conclusion, the properties of spatial entropy measures make them an appealing tool to evaluate urban expansion from a spatial perspective by an effective combination of $H(Z)$ for urban intensity and the *SPI* terms for urban compactness calibrated on the data urban proportion.

4. A comparative study on Italian metropolitan areas

The proposed measures are computed for the 14 Italian metropolitan areas for 1990 and 2012, with the aim of assessing the state and development of each main urban hub in terms of urbanization intensity and spatial compactness, by a combination of the information provided by the non-spatial entropy $H(Z)$ and the spatial partial information terms.

4.1. Design of the study

The 14 datasets are extracted from the database of CORINE Land Cover for 1990 and 2012, using the 250*m* pixel resolution. Every dataset contains the main municipality and its surrounding commuting belt towns. Each area is enclosed into a rectangle; all pixels outside the administrative borders are classified as “missing values” and do not affect computations. The datasets with 44 classes are then transformed into binary UMZ data following the CLC criteria.

Four distance ranges ($K=4$) are chosen; a discussion about the underlying motivation is in Section 5. Class $w_1=[0, 250]$ metres includes adjacent pairs of pixels. Class $w_2=]250, 500]$ covers the immediately farther neighbours. Class $w_3=]500, 1250]$ is a wider class including pairs at most 5 pixels apart, for checking whether any residual spatial information can be detected. The last class is w_4=]*1250,* d_{max}], d_{max} being the maximum distance, and collects the least interesting distances that will not be commented in detail.

The estimation procedure follows the steps illustrated in Section 3. The proportion of urban pixels for each city and time point is the estimated $p=p(x_1)$ used for $H(X)$; analogously, relative frequencies estimate the probabilities $p(z_1)$, $p(z_2)$ *and* $p(z_3)$, *with* z_1={both urban}, z_2={urban, non-urban}, z_3={both non-urban}. For each distance class, relative frequencies of pairs form an estimate of $p(Z|w_k)$. The total number of pairs for each distance range is counted, and its relative frequency is used as an estimate of $p(w_k)$, the probability distribution of distance ranges for each city.

All computations are run on the `R` software with the help of the packages `spatstat` (Baddeley, Rubak and Turner 2015) and `SpatEntropy` (Altieri, Cocchi and Roli 2018b). The computational time varies from a few minutes to a few hours for each metropolitan area.

Table 1 summarizes the information about each metropolitan area. The size in pixels of the enclosing rectangle gives an idea of the computational burden; in correspondence of each rectangle, we report the extension of the studied area both in square kilometres and in number of pixels within the administrative borders. They also show that the metropolitan areas of Palermo, Roma and Venezia are largely wider than the others. The urban proportion $\hat{p}$ at the two time points is "p(urb)": it increases from 1990 to 2012 for nearly all areas. The Table also includes the estimated marginal distribution of Z. The last columns of Table 1 show the population density and Shannon's entropy of Z for each area and time: this is the non-spatial information, as both density and entropy depend only on the amount of black and white pixels. Densities range from the low levels of Genova and Reggio Calabria to the crowded levels of Torino, Milano, and Napoli, and support the non-spatial nature of Shannon's entropy: the Table shows that the two measures are return a similar ranking of metropolitan areas. Values for $SPI(Z|w_1)$ - $SPI(Z|w_3)$ are in Table S2.

Main City	ER size	Ext	Pix	1990				2012				1990		2012	
				p(urb)	p(z_1)	p(z_2)	p(z_3)	p(urb)	p(z_1)	p(z_2)	p(z_3)	Den	H(Z)	Den	H(Z)
Bari	113x225	552	8837	0.169	0.029	0.281	0.690	0.192	0.037	0.310	0.653	1034	0.714	1015	0.763
Bologna	135x124	625	10000	0.155	0.024	0.262	0.714	0.178	0.032	0.293	0.676	913	0.682	909	0.734
Cagliari	110x201	484	7739	0.175	0.031	0.288	0.681	0.214	0.046	0.337	0.617	723	0.727	733	0.806
Catania	264x134	844	13497	0.143	0.021	0.246	0.734	0.160	0.026	0.269	0.705	648	0.651	639	0.694
Firenze	99x115	404	6466	0.190	0.036	0.307	0.657	0.23	0.054	0.357	0.589	1474	0.758	1388	0.837
Genova	136x224	731	11697	0.102	0.010	0.183	0.807	0.100	0.010	0.180	0.811	1024	0.531	892	0.524
Messina	135x112	366	5848	0.138	0,019	0.237	0.744	0.133	0.018	0.230	0.752	697	0.636	727	0.623
Milano	93x108	404	6466	0.604	0.365	0.478	0.157	0.650	0.423	0.455	0.122	4907	1.010	4622	0.979
Napoli	67x120	249	3980	0.534	0.285	0.498	0.217	0.596	0.354	0.482	0.164	6521	1.036	6148	1.016
Palermo	200x196	854	13658	0.124	0.015	0.217	0.768	0.131	0.017	0.228	0.755	921	0.599	913	0.620
Reggio C.	140x106	519	8301	0.056	0.003	0.105	0.892	0.095	0.009	0.171	0.820	420	0.356	419	0.507
Roma	284x273	2465	39441	0.200	0.040	0.321	0.639	0.230	0.053	0.354	0.593	1311	0.780	1355	0.833
Torino	111x113	414	6619	0.402	0.162	0.481	0.357	0.451	0.204	0.495	0.301	3254	1.014	3056	1.033
Venezia	242x214	762	12185	0.163	0.027	0.273	0.700	0.187	0.035	0.304	0.661	474	0.701	458	0.753

Table 1: Size of the enclosing rectangle in number of pixels (ER size), extension of the administrative metropolitan area within the rectangle in km^2 (Ext.), number of pixels within the metropolitan area (Pix.), urban proportion in 1990 and 2012, estimated distribution of Z in 1990 and 2012, population density (number of residents per km^2) and Shannon's entropy of Z for the metropolitan areas.

4.2. Intensity and compactness of Italian metropolitan areas

Conclusive comments can be drawn from Figure 2, which shows four scatterplots of the Italian metropolitan areas for SPI_1 and SPI_2 at both years, summarizing the overall situation and detecting the cold and hot spots.

The x axis presents $H(Z)$ for 1990 (left panels) and 2012 (right panels), which measures the urbanization intensity. The y axis shows the value of the spatial partial information at distance classes w_1 (upper panels) and w_2 (lower panels), synthesizing the compactness of urbanization; the dots corresponding to metropolitan areas can shift vertically from upper to lower panels. The scatterplot gives an exhaustive summary of the 14 Italian metropolitan areas as regards urban configuration and development.

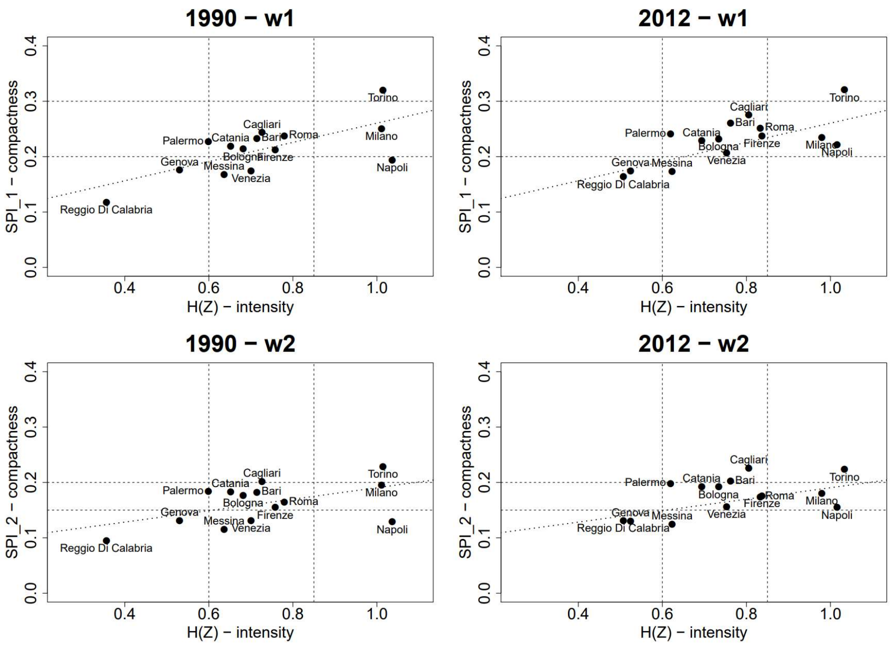

Figure 2: Scatterplot of Italian metropolitan areas in 1990 and 2012: Shannon's entropy of Z versus spatial partial information at distances w_1 and w_2

For the ease of interpretation, dashed vertical and horizontal lines mark three levels for each measure, chosen to evaluate the Italian scenario, not as absolute thresholds. For $H(Z)$, they divide each plot into low, intermediate and high urbanization intensity, with breaks 0.6 and 0.85; for the *SPI* terms, they divide each plot into low, intermediate and high urbanization compactness, with breaks 0.2 and 0.3 for SPI_1 and 0.15 and 0.2 for SPI_2.

Looking at the *x* axis, in 1990, $H(Z)<0.6$ is observed for Genova, Palermo and Reggio Calabria: the urban tissue covers a minority of the areas, with some belt municipalities so scarcely populated that no pixels are urban. Reggio Calabria witnesses a major increase in entropy from 1990 to 2012, but remains the least urbanized one, while Genova and Palermo have basically constant entropy values. The metropolitan areas with $H(Z)>0.85$ are Milano, Napoli and Torino. Entropies for Milano and Napoli decrease slightly from 1990 to 2012, while Torino's entropy increases. All other metropolitan areas present an intermediate level of urbanization in 1990, and their entropies increase in 2012 except Messina.

As regards the spatial information terms on the y axis, the least compact areas with a value under 0.2 for SPI_1 and under 0.15 for SPI_2 are Genova, Messina, Reggio Calabria, plus Napoli and Venezia in 1990. The most compact area, with a value above 0.3 for SPI_1 and 0.2 for SPI_2 in both years, is Torino; Cagliari increases its SPI_2 value in 2012, meaning that, though not extremely compact at a really small scale, its efficiency has increased. Focusing on time comparison, both SPI_1 and SPI_2 increase from 1990 to 2012 for most areas. This can be interpreted as a positive feature, witnessing the increase over time of the areas' compactness, since adjacent and neighbouring pixels tend to be more homogeneous. A slight decrease can be observed for Genova and Milano, as a sign of a less efficient (but nearly stable) urbanization.

For all Figure panels, a positive relationship is evident between $H(Z)$ and the *SPI*s. The diagonal dotted line summarizes the relationship according to a linear regression model for w_1 and w_2 respectively, and helps in assessing the relative behaviour of the 14 Italian areas. The regression line for w_2 is flatter than the one for w_1, since SPI_2 is not as strong as SPI_1, but leads to similar conclusions. Areas under the diagonal line present the worst behaviour, being more chaotic, while the ones above the diagonal line are less urbanized and more compact.

4.3. Discussion on the development of Italian metropolitan areas

When the goal is to evaluate urban expansion and prepare suitable urban plans and policy, usually a compact spatial development is regarded as positive while a scattered urban tissue is considered inefficient and negative. A general comment stresses that, as regards the 14 Italian metropolitan areas, more urbanization brings along more compactness, indicating positive urban development over space. In this study, dispersion is detected when urbanization is scarce, therefore areas with little urbanization but showing a compact development can be regarded as corresponding to a positive situation under the perspective of efficiency.

At this regard, Genova and Reggio Calabria are not only the least urbanized, but also among the least compact areas. They are the only areas in the bottom left quadrant of Figure 2, i.e. the ones with the slowest and most inefficient urban development. Palermo has a borderline low/medium level of urbanization intensity with a medium spatial efficiency. A consistent group of areas can be considered as the representative Italian average metropolitan situation, falling in the middle quadrant of Figure 2: Bari, Bologna, Cagliari, Catania, Firenze and Roma. The most urbanized and compact area in the upper right panel is Torino. Messina and Venezia are averagely urbanized according to the Italian levels, but they have very low SPI_1 and SPI_2 values and appear therefore inefficiently developed. Milano and Napoli are highly urbanized and have the highest population density, but their compactness is medium and low, respectively, even if Napoli improves its efficiency in 2012.

In conclusion, regarding the Italian metropolitan situation, the leading example as regards urban development is Torino. As per the contribution of class w_3 (Table S2), Torino has a SPI_3 value above 0.1, which confirms its strong homogeneity. Other metropolitan areas also present a value for $SPI_3>0.1$, however their SPI_1 and SPI_2 values are not as high, leading to the conclusion that the corresponding cities cannot be considered really compact.

Looking at the regression line of Figure 2, Genova and Reggio Calabria share another peculiarity: they are chaotic, but when considering their compactness and small level of urbanization jointly, their condition is not severe compared to the general Italian level. The most negatively performing areas, farther below the regression line, are Messina, Napoli and Venezia.

5. Discussion and conclusion

Our work focuses on the quantification of urban intensity and spatial compactness; we cannot however forget that several socio-economic factors contribute to the classification of Italian metropolitan cities in Figure 2, that may have ancient roots. For example, the municipality of Venezia has a peculiar condition, being characterized by canals and small islands, and partially under the sea level. Its geographical location certainly affected its expansion, while it also makes the city extremely appealing and precious under other points of view. In addition, all metropolitan areas with low compactness detected by SPI_1 are centred on coastal cities: Genova, Messina, Reggio Calabria, Napoli and Venezia. Their location encouraged them to be economic hubs in the past, but also impeded the development of a monocentric compact configuration that characterizes other areas such as Torino. In addition, the main municipalities of Genova and Reggio Calabria areas are very dense, while this does not occur for some of the municipalities at their borders, often situated in very steep hills. Other aspects may be considered, such as further insights on the historical origins of the cities or more recent local urban plans and laws, as well as delocalization of industrial activities, that contribute to a city configuration. The measures proposed in this work may be integrated with other indices from different areas of expertise to return a global evaluation of areas development.

The analogy between Shannon's entropy and population density shows that the most appropriate interpretation of entropy regards the intensity of the urban phenomenon for each metropolitan area and time point: the greater the entropy, the more difficult to predict if a pixel is urbanized or not. Cities with a high value for $H(Z)$ are more urbanized than others, not necessarily more chaotic or inefficiently developed. However, entropy is a better index of urbanization intensity than the population density, firstly because of its theoretical properties that allow a decomposition into spatial and non-spatial components, and secondly because density is based on the number of citizens, while entropy is based on the number of urban pixels, i.e. on the presence of buildings, consequently giving an idea of the amount of the metropolitan territory that is anthropized.

As regards spatial mutual information, the relevant quantities are the partial terms *SPI*, where the contribution of space can be investigated in detail, rather than the aggregated quantity $SMI(Z,W)$. Such weighted sum is meaningless for urban expansion, because it is affected by the residual class w_K, which is the least interesting but receives the largest weight due to its amplitude. We stress the meaningful minority of distances represented by classes w_1 and w_2: an evaluation of compactness and spatial efficiency of urban development can be drawn looking at pixels that lie close-by. In a chaotic configuration, pairs of neighbouring pixels do not exhibit any association, and the corresponding spatial mutual information term is low. In a compact configuration, pairs tend to be homogeneous, i.e. urban pixels lie close to other urban pixels, and non-urban pixels lie close to other non-urban pixels; this returns a high spatial mutual information term. High spatial partial information values for class w_3 would imply a remarkably compact metropolitan configuration, as in Torino. One of the crucial advantages of the present approach is its flexibility in choosing the most suitable classes for the study, without limitations to classes of equal size. The specific choice for the w_ks of the present paper is motivated by the tradition of spatial statistics and previous works. Indeed, w_1 and w_2 cover the standard neighbourhood systems in spatial studies, and, being the smallest distances, they allow to investigate small-scale spatial effects in detail. The choice of w_3 comes from Altieri, Cocchi, Roli (2019a): the work suggests that w_3 is the distance break where we can distinguish a polycentric (multicluster) urban configuration from a scattered one. Choosing regular spaced classes is not a satisfying option, as it aggregates small distances which are particularly informative, while disaggregating large, uninformative distances with a waste of computational time. Another option to avoid is the choice of a high number of classes, as it substantially increases the computational time. A large residual class w_K is computationally efficient because its elements are got as differences from the total number of pairs, once the numbers of pairs for all other distance classes are obtained. Conversely, wide intermediate classes imply identifying and counting a large number of pairs, which hugely increases the computational burden. The effect of aggregating two classes in *SPI* depends on the relative size of the two: if they are similar, the new aggregated class returns an intermediate *SPI* value; if one is remarkably larger than the other, its relative weight $p(w_k)$ dominates the resulting *SPI* value. In Altieri, Cocchi, Roli (2019b) an example is shown about the effect of aggregating/disaggregating classes.

A set of measures able to quantify both urbanization intensity and spatial compactness has been proposed. First, the entropy of the variable Z evaluates the urban intensity; then, compactness is measured as departure from the benchmarking distribution of Z at given distance ranges, so that the level of compactness is calibrated on the urbanization intensity over an area. When the attention is focused on a joint evaluation of urbanization intensity and spatial efficiency, our set of measures is complete, theoretically solid and flexible, both as regards properties of the measures and the ability to draw absolute and comparative conclusions. However, an interdisciplinary assessment of urban development needs integration of many disciplines and is beyond the scope of this work.

An extension to the present work involves consideration of a more detailed land cover information, with more than two categories. In Altieri, Cocchi, Roli (2019b) this is faced via simulation. With many categories, only a compact and a randomly scattered city scenario are distinguished: intermediate situations such as a polycentric/multicluster city become faded and similar to a random pattern. A desirable feature of the proposed method is that as the number of categories increases, a divergence between the two spatial configurations is detected: the *SPI*s and $H(Z)$ are able to quantify the level of compactness and intensity, respectively, and distinguish the two scenarios. In conclusion, the approach holds for any number of land cover categories; to compare results across datasets with a different number of categories, a normalization for $H(Z)$ and the *SPI* terms is needed. For further details, see Altieri, Cocchi, Roli (2019b) where an application with five land cover categories is considered.

References

Altieri, L., Cocchi, D., Pezzi, G., Scott, E. M. and Ventrucci, M. (2014). Urban sprawl scatterplots for Urban Morphological Zones data. *Ecological Indicators*, **36**, 315-323. doi:10.1016/j.ecolind.2013.07.011

Altieri, L., Cocchi, D. and Roli, G. (2018a). A new approach to spatial entropy measures. *Environmental and Ecological Statistics*, **25**(1), 95–110. doi: 10.1007/s10651-017-0383-1

Altieri, L., Cocchi, D. and Roli, G. (2018b). SpatEntropy: Spatial entropy measures [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=SpatEntropy (R package version 0.1.0)

Altieri, L., Cocchi, D. and Roli, G. (2019a). Measuring heterogeneity in urban expansion via spatial entropy. *Environmetrics*, **30**(2), e2548. doi: 10.1002/env.2548

Altieri, L., Cocchi, D. and Roli, G. (2019b). Advances in spatial entropy measures. *Stochastic Environmental Research and Risk Assessment*, **33**. doi: 10.1007/s00477-019-01686-y

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis*, **27**(2), 94-115.

Baddeley, A., Rubak, E. and Turner, R. (2015). *Spatial point patterns: Methodology and applications with R*. London: Chapman and Hall/CRC Press.

Balducci, A., Fedeli, V., Curci, F. (2017). *Oltre la metropoli. L'urbanizzazione regionale in Italia.* Guerini, Firenze-Milano.

Banai, R., DePriest, T. (2014). Urban Sprawl: Definitions, Data, Methods of Measurement, and Environmental Consequences. *Journal of Sustainability Education*, **7.**

Batty, M. (1974). Spatial entropy. *Geographical Analysis*, **6**, 1-31.

Batty, M. (1976). Entropy in spatial aggregation. *Geographical Analysis*, **8**, 1-21.

Batty, M. (2010). Space, scale, and scaling in entropy maximizing. *Geographical Analysis*, **42**, 395-421.

Bettoni, G. (2018). *L'émergence des cités metropolitaines italiennes,* in Bettoni, Le Bras, Navarre (eds.) *Métropoles en chantiers 2*, Paris: Berger Levrault, pp. 65-85

Bhatta, B., Saraswati, S. and Bandyopadhyay, D. (2010). Urban sprawl measurement from remote sensing data. *Applied Geography*, **30**, 731-740.

Burton, E. (2002). Measuring urban compactness in UK towns and cities. *Environment and planning B: Planning and Design*, **29**(2), 219-250.

Cheng, J. and Masser, I. (2004). Understanding Spatial and Temporal Processes of Urban Growth: Cellular Automata Modelling. *Environment and Planning B*, **31**(2), 167–194.

Chong, C. H.-S. (2017). *Comparison of spatial data types for urban sprawl analysis using Shannon's entropy*. University of Southern California: Dissertation.

Cover, T. M. and Thomas, J. A. (2006). *Elements of information theory. Second edition*. Hoboken, New Jersey: John Wiley and Sons, Inc.

Dadashpoor, H., Azizi, P. and Moghadasi, M. (2019). Land use change, urbanization, and change in landscape pattern in a metropolitan area. *Science of The Total Environment*, **655**, 707–719.

EEA. (1994). CLC methodology and nomenclature (Tech. Rep.). (Published by Commission of the European community. Available online at http://www.eea.europa.eu/publications)

EEA. (2011). Corine Land Cover 2000 raster data (Tech. Rep.). (Downloadable at http://www.eea.europa.eu/data-and-maps/ data/corine-land-cover-2000-raster-1)

Eurostat. (2010). Results on EU land cover and use published for the first time (Tech. Rep.). (Eurostat news release, Eurostat press office. Available online at http://epp.eurostat.ec.europa.eu/cache/)

Ewing, R. and Hamidi, S. (2015). Compactness versus sprawl. A review of recent evidence from the United States. *Journal of Planning Literature*, **30**, 413-432.

Heikkila, E. J. and Hu, L. (2006). Adjusting spatial-entropy measures for scale and resolution effects. *Environment and Planning B*, **33**(6), 845–861.

Herold, M., Goldstein, N. C., Clarke, K. C. (2003). The spatiotemporal form of urban growth: measurement, analysis and modeling. *Remote sensing of Environment*, **86**(3), 286-302.

Jaeger, J. A. G., Bertiller, R., Schwick, C. and Kienast, F. (2010). Suitability criteria for measures of urban sprawl. *Ecological Indicators*, **28**, 427-441.

Jiang, F., Liu, S., Yuan, H. and Zhang, Q. (2007). Measuring urban sprawl in Beijing with geospatial indices. *Journal of Geographical Sciences*, **17**, 469-478.

Leibovici, D. G. (2009). *Defining spatial entropy from multivariate distributions of co-occurrences*. Berlin, Springer: In K. S. Hornsby et al. (eds.): COSIT 2009, Lecture Notes in Computer Science 5756, 392-404.

Leibovici, D. G., Claramunt, C., LeGuyader, D. and Brosset, D. (2014). Local and global spatio temporal entropy indices based on distance ratios and co-occurrences distributions. *International Journal of Geographical Information Science*, **28**, 1061-1084.

Leibovici, D. G., Claramunt, C. (2019). On Integrating Size and Shape Distributions into a Spatio-Temporal Information Entropy Framework. *Entropy*, **21**(11), 1112; https://doi.org/10.3390/e21111112

Medina, C. D. and Monclùs, J. (2018). *Urban visions: From planning culture to landscape urbanism*. Springer.

O'Neill, R. V. et al. (1988). Indices of landscape pattern. *Landscape Ecology*, **1**, 153-162.

Pacione, M. (2009). *Urban geography: A global perspective*. Routledge.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A. and Dech, S. (2012). Monitoring urbanization in mega cities from space. *Remote sensing of Environment*, **117**, 162–176.

Tong, D. and Plane, D. (2014). A new spatial optimization perspective on the delineation of metropolitan and micropolitan statistical areas. *Geographical Analysis*, **46**, 230-249.

Torrens, P. (2008). A toolkit for measuring sprawl. *Applied Spatial Analysis and Policy*, **1**, 5-36.

Yamu, C. and Frankhauser, P. (2015). Spatial accessibility to amenities, natural areas and urban green spaces: using a multiscale, multifractal simulation model for managing urban sprawl. *Environment and Planning B,* **42**(6), 1054–1078.

Yeh, A. and Li, X. (2001). A constrained CA model for the simulation and planning of sustainable urban forms by using GIS. *Environment and Planning B*, **28**(5), 733-753.