GEOMORPHOMETRY 2020 PERUGIA, ITALY JUNE 22 - 26 2020

# Comparative study of delineation of urban areas using imperviousness products and open data

Massimiliano Alvioli
Istituto di Ricerca per la Protezione Idrogeologica
Consiglio Nazionale delle Ricerche
via Madonna Alta 126, I-06128 Perugia, Italy
massimiliano.alvioli@irpi.cnr.it

*Abstract*—**City boundaries are not self-manifest, and typically do not coincide with administrative boundaries. A sound delineation of cities, more generally of urban areas, is a non-trivial task. A delineation method should comply with a well-defined metric, in order to reduce subjectivity, to favour reproducibility, and to allow assimilation with other methods. In fact, many existing city delineation methods rely heavily on numerical parameters such as population density thresholds. Here, we present city delineation for the whole of Italy performed with two different methods. On the one hand, we consider delineation based on terrain imperviousness, as a proxy for the existence of continued human presence, which is an inherently parametric method. On the other hand, we adopt a strictly data-driven method known as "natural cities", based on head-tail breaks of areas extracted from road junctions. We compare results from the two methods by considering numerical figures from the two delineated set of cities. We further propose an additional metric for assessing the results, namely a scaling relation between area and population of individual cities in the two sets. We show that the two results are similar in terms of number and total areal extent of cities, while area-population relations highlights substantial differences which can be ascribed to the parametric character of delineation from imperviousness.**

## I. INTRODUCTION

The size, shape and geographical location of cities is relevant to many studies including demographic and social issues, labor trends, natural hazards, urban planning, to mention some. To date, little consensus exists on where city boundaries are located, and how the criteria to delineate them should be formalized. A recent review [1] compiled a survey of 32,231 studies of urban agglomerations, with a wide range of variability regarding the definition of city itself. United Nations [2] acknowledges, "no standardized international criteria exist for determining the boundaries of a city and often multiple boundary definitions are available for any given city". Urban areas exist in three nested levels, in the definition of United Nations: City Proper, Urban Agglomeration and Metropolitan Area, in order of increasing sizes with respect to both planimetric area and population, but these definitions are not standard.

Delineation of cities has been performed in the relevant literature using a number of different methods, relying on very heterogeneous data sources. Batty [3] distinguished city delineation methods by three different criteria: (a) population and/or urbanization density, (b) interactions, described by different kinds of networks, either physical or non-physical, and (c) geographical proximity and/or contiguity.

In this work, we applied two specific methods, falling in the categories (a) and (c) above, respectively. The first method makes use of artificial surfaces obtained from satellite data, inferring that sealed (or impervious) surfaces are a proxy for continued human presence, *i.e.* urbanization. The second method [4] uses street nodes to build geographically contiguous areas, which we consider as part of cities based on a head-tail break rule applied to their planimetric size. The first method is intrinsically parametric, while the second is parameter-free.

We investigated the outcome of the two methods within the framework of area-population scaling relations. Scaling relations among a number of different urban indicators exist [5, 6] and many authors used them and critically analyzed them [7, 8]. A scaling relation for planimetric area, $A$, as a function population, $P$, of a set of cities has the following form:

$$A = \beta \, e^{\alpha}. \qquad (1)$$

It implies that a city twice as large of another city, in terms of population $P$, is expected to cover a planimetric area $A_2 = 2\,A_1$, with $A_1$ and $A_2$ the area of the smaller and larger city, respectively. Different urban indicators exhibit scaling with respect to city size, taken as the population, with different values of the exponent. If a scaling relations as in Eq. (1) is in effect, the value scaling exponent $\alpha$ being larger, smaller or consistent with unity has different implications. A value $\alpha < 1$ signals a diseconomy of scale, while $\alpha > 1$ signals an economy of scale. A value of $\alpha$

*Massimiliano Alvioli (2020) Comparative study of delineation of urban areas using imperviousness products and open data:*

*in Massimiliano Alvioli, Ivan Marchesini, Laura Melelli & Peter Guth, eds., Proceedings of the Geomorphometry 2020 Conference, doi:10.30437/GEOMORPHOMETRY2020_1.*

consistent with unity would correspond to constant returns to scale.

Figure 1 shows, for illustrative purposes, a log-log plot of area-population ($A$-$P$) relations for world's cities. In the figure, we used data from http://www.demographia.com, which lists a data set of 1,750 cities compiled a few years ago. We used this data source to show $A$-$P$ relations with data collected in a homogeneous way. We show data separately for five European countries, along with the whole data set. The straight lines in Fig. 1 are linear regressions of the log $A$ versus the log $P$. In fact, taking the log of Eq. (1) one obtains a linear relation between the logarithms $A$ and $P$ as follows:

$$\log A = \beta' + \alpha \log P, \qquad (2)$$

with $\beta' = \log \beta$. Figure 1 shows that different countries may have substantially different scaling exponents, as far as $A$-$P$ scaling relations are concerned. In this work, we focused on two independent ways of delineating cities and, thus, of obtaining area and population, in the case of Italy.

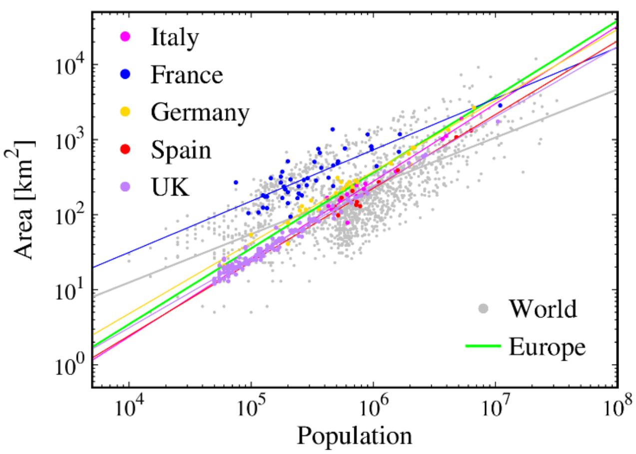


**Figure 1**. Area-population relations of a few European countries and corresponding linear fits as in Eq. (2); "Europe" refers to the five countries shown in this Figure; "World" refers to the whole set of 1,750 cities. The table below lists the coefficient of the fits, Eq. (2) (*after Ref. [4]*).

| **Country** | **$\alpha$** | **$R^2$** | **# cities** |
|---|---|---|---|
| Italy | 1.03 | 0.91 | 14 |
| France | 0.68 | 0.66 | 47 |
| Germany | 0.95 | 0.95 | 25 |
| Spain | 0.98 | 0.94 | 12 |
| United Kingdom | 0.94 | 0.97 | 138 |
| Europe | 1.01 | 0.78 | 236 |
| World | 0.64 | 0.50 | 1,750 |
| World – Ref. [5] | 0.56-1.04 | - | - |

## II. Methods

In the following, we describe in detail the two methods we used in this work to delineate cities in Italy, in two separate paragraphs. The outcome of both methods is a set of polygons, representing urban areas in Italy. We will compare them in the framework of area-population scaling relations, Eqs. (1) and (2).

### A. *City delineation from imperviouness*

Imperviouness is a measure of the degree (percentage) of soil sealing. Impervious surfaces are both built-up and non-built-up, and include a variety of objects that we identify with human locations or activities [9]. Artificial surfaces can be detected using remote sensing [10], and assuming that any impervious (or sealed) surface is part of an urban system [11]. Separate clusters of impervious terrain can be identified with individual cities.

The Copernicus programme makes available imperviousness data as raster layers [9] with a resolution of 20 m. Each grid cell in the raster is a percent value, which introduces the need for a parameter: one can introduce a percent threshold over which one can flag a grid cell as an urban area. In this work, we considered any non-zero value as indicator of an urban area. Next, we need to cluster grid cells, in order to obtain individual cities. This step introduces the additional difficulty of delineating boundaries between areas who might actually have relations, either spatial or regarding human activities, which highlights that cities are difficult to delineate, and also difficult to study in isolation [12]. We overcome the difficulty by introducing an additional parameter. We selected disjoint clusters resulting from generating a buffer with negative radius (GIS raster reduction operation) around the original cells with non-zero imperviousness, and then a positive buffer (GIS grow operation). The radius of the two operations was arbitrarily set to five grid cells, which is the additional parameter.

### B. *Delineation of "natural cities"*

The original algorithm of "natural cities" [14] implements city delineation starting from a collection of point-like populated sites, and performs an iterative clustering of sites within a given radius. The requirement for a radius was dropped by an improvement in Ref. [15], who applied the algorithm selecting streets nodes as starting points and using the head-tail rule to select polygons corresponding to urban areas. The polygons occurring in the generalized algorithm were subsequently singled out by using either city blocks as clustering domains [15], triangulated irregular network (TIN) [16], or Thiessen polygons [17].

In this work, and in Ref. [4], we used street nodes obtained from the OpenStreetMap vector layer as a starting point, and generated a TIN network separately for the peninsular Italy and the major islands, Sicily and Sardinia. Application of the head-tail break rule consisted in considering the planimetric area distribution of the triangles. A head-tail break rule applies because an unbalanced ratio exists within the three sub regions of Italy for the number of triangles with planimetric area above and below the average value. Numerical figures for the number above/number below average were as follows: 0.11 for peninsular Italy, 0.14 for Sicily, and 0.21 for Sardinia. Jiang and Liu found for comparable

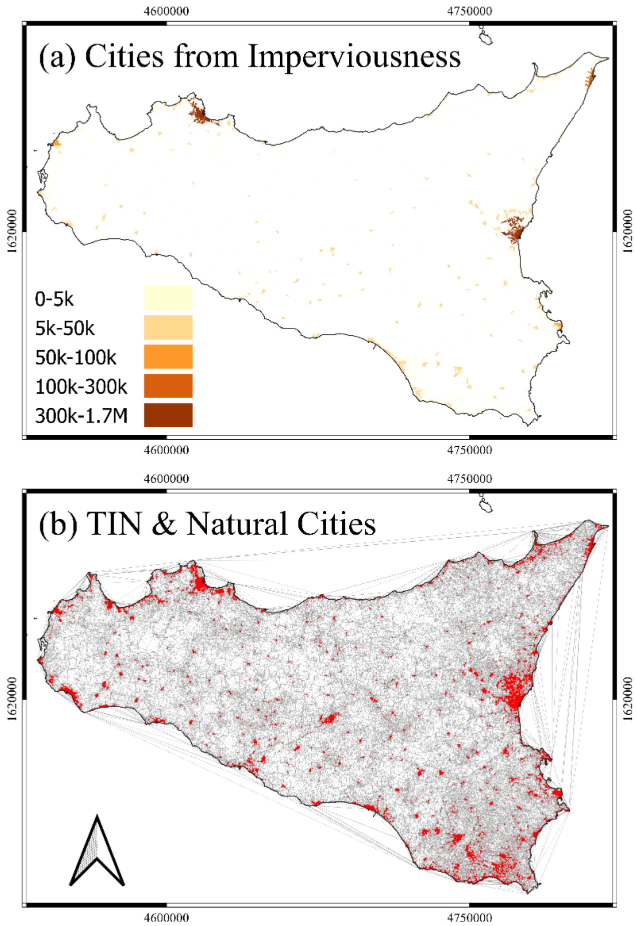


**Figure 2**. Sample of city delineation in Italy: we show a subset of the results, in Sicily, obtained from the two methods used in this work. (a) Results from the method based on imperviousness; the scale indicates number of citizens. (b) Results from the method of natural cities; grey lines show the TIN network generating natural cities (*after Ref. [4]*). Maps are in LAEA projection, EPSG:3035

quantities in France, Germany and UK the values of 0.05, 0.14 and 0.09 for the ratio, respectively [15].

We discarded all of the TIN polygons with area above average, along with all of the polygons with area below average, which were adjacent to a polygon with large area, as in Ref. [15]. This last step is a replacement for utilizing a clustering radius to single out cities in this work, at variance with Ref. [14].

## III. Results and conclusions

Figure 2 shows results of city delineation, limited to the subset in Sicily (one of the 20 Italian administrative Regions).

Note that Fig. 2 (a) also reports the population of each delineated city. Population data at municipality level is available from the Italian Institute for statistics (ISTAT, http://www.istat.it). We calculated population at city level by distributing the population, known at municipality level, on a grid aligned with the imperviousness layer.

Delineation of cities based on the imperviousness layer and on the natural cities method produces different numbers of agglomerations, listed in the following Table.

| **Imperviousness** | **Number of cities** | **Max. area [km$^2$]** | **Mean area [km$^2$]** |
|---|---|---|---|
| Peninsular | 54,379 | 443,602 | 0.232 |
| Sicily | 9,808 | 53,002 | 0.080 |
| Sardinia | 2,467 | 45,023 | 0.204 |
| Italy | 66,654 | 443,602 | 0.201 |
| **Natural cities** (*after Ref. [4]*) | **Number of cities** | **Max. area [km$^2$]** | **Mean area [km$^2$]** |
| Peninsular | 77,103 | 1,311,601 | 0.244 |
| Sicily | 6,190 | 112,829 | 0.189 |
| Sardinia | 5,979 | 86,762 | 0.182 |
| Italy | 89,272 | 1,311,601 | 0.224 |

Figure 3 shows *A-P* relations calculated from the two methods, in a log-log scale as in Fig. 1. In both Fig. 3 (a), corresponding to the imperviousness method, and Fig. 3 (b), for natural cities, we show separately the ($P$, $A$) data points for peninsular Italy, Sicily and Sardinia. The straight lines, instead, are linear fit of the merged data sets. Both boxes also contain the linear fit for European cities (green curve), also shown in Fig. 1.

One can immediately appreciate that the two methods provide substantially different results, as far as the distribution of data points is concerned. In the case of Fig. 3 (a), corresponding to cities in Fig. 2 (a), data not only have a lower limit for the area, 400 m$^2$, dictated by the resolution of the imperviousness layer, they also show peculiar patterns, which are specific of the three distinct sub sets and can hardly be explained with simple arguments. On the other hand, in the case of Fig. 3 (b), corresponding to cities in Fig. 2 (b), data nicely distribute on the (P, A) plane, hinting to a linear correlation [4]. In this case, some patterns seem to emerge as well; however, they mostly emerge as collinear structures to the overall linear fit. Moreover, they seem

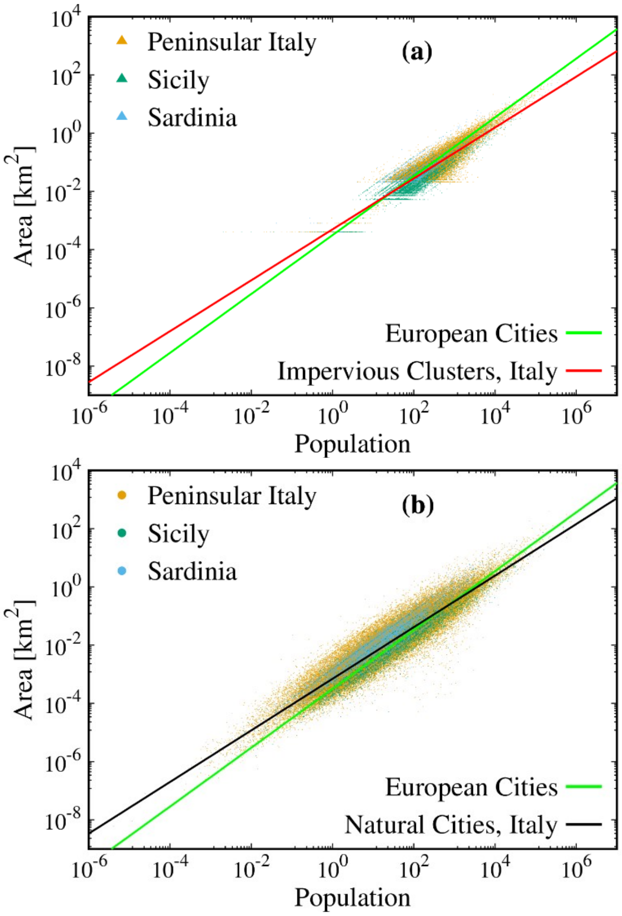


**Figure 3**. *A-P* relations from the two different approximations to delineate cities considered in this work, compared to results for municipalities (black and red curves). (a): cities delineated as disjoint clusters in the imperviousness layer; (b) natural cities (*after Ref. [4]*). A comparison of the geographical distributions of the two results, for Sicily, is in Fig. 2. Data points consist of three subsets, but the linear fits correspond to the aggregate data set. Green curve: fit to ($P$, $A$) data of European cities, from an independent data source, as in Fig. 1.

to exist only in the Sicily and Sardinia sub sets, suggesting some bias might occur for delineation of natural cities in smaller areas.

In conclusion, we gave proof that a data-driven method for delineating cities such as natural cities, adopted in this work and in Ref. [4], represents an objective method for city delineation. The method provides advantages with respect to the delineation of cities by means of imperviousness products, as far as area-population scaling relations are concerned.

## References


[1] Fang, C., D. Yu, 2017. "Urban agglomeration: An evolving concept of an emerging phenomenon". Landscape and Urban Planning 162, 126–136. DOI: 10.1016/j.landurbplan.2017.02.014

[2] United Nations, 2018. "The world's cities in 2018". Department of Economic and Social Affairs, Population Division. Data booklet, ST/ESA/ SER.A/417

[3] Batty, M., 2018. "Inventing future cities". The MIT Press, Cambridge, MA, USA. 304 pp. ISBN: 9780262038959

[4] Alvioli, M., 2020. "Administrative boundaries and urban areas in Italy: a perspective from scaling laws". Landscape and Urban Planning (*in press*)

[5] Bettencourt, L., 2013. "The origins of scaling in cities". Science 340, 1438-1441. DOI: 10.1126/science.1235823

[6] Barthelemy, M. et al., 2013. "Self-organization versus top-down planning in the evolution of a city". Scientific Reports 3, 2153. DOI:10.1038/srep02153

[7] Arcaute, E. et al., 2015. "Constructing cities, deconstructing scaling laws". Journal of the Royal Society Interface. DOI: 10.1098/rsif.2014.0745

[8] Cottineau, C. et al., 2017. "Diverse cities or the systematic paradox of urban scaling laws". Computers, Environment and Urban Systems 63, 80-94. DOI:10.1016/j.compenvurbsys.2016.04.006

[9] Steidl, M. et al., 2016. "Copernicus Land Monitoring Service - High Resolution Layer Imperviousness". Technical Report version 1 of 985 2018-12-21. European Environment Agency. https://land.copernicus.eu

[10] Lu, D. et al., 2014. "Methods to extract impervious surface areas from satellite images". International Journal of Digital Earth 7, 93-112. DOI: 10.1080/17538947.2013.866173

[11] Ma, D.et al., 2018. "Why topology matters in predicting human activities". Environment and Planning B: Urban Analytics and City Science 0, 1-17. DOI: 10.1177/2399808318792268

[12] Thomas, I. et al., 2018. City delineation in European applications of LUTI models: review and tests. Transport Reviews 38, 6-32. DOI: 10.1080/01441647.2017.1295112

[13] Bettencourt, L., G. West, 2010. "A unified theory of urban living". Nature 467, 897 912-913. DOI: 10.1038/467912a

[14] Rozenfeld, H.D. et al., 2009. "The Area and Population of Cities: New Insights from a Different Perspective on Cities". Working Paper 15409. National Bureau of Economic Research. DOI: 10.3386/w15409

[15] Jiang, B., X. Liu, 2012. "Scaling of geographic space from the perspective of city and field blocks and using volunteered geographic information". International Journal of Geographical Information Science 26, 215-229. DOI: 10.1080/13658816.2011.575074

[16] Jiang, B., Y. Miao, 2015. "The evolution of natural cities from the perspective of location-based social media". The Professional Geographer 67, 295-306. DOI: 10.1080/00330124.2014.968886

[17] Jiang, B., 2018. "A topological representation for taking cities as a coherent whole". Geographical Analysis 50, 298-313. DOI: 10.1111/gean.12145