Landscape and Urban Planning 253 (2025) 105205

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Research Paper

# Urban Sprawl and Routing: A Comparative Study on 156 European Cities

Roberto Maria Rosati *

*DPIA, University of Udine, via delle Scienze 206, 33100, Udine, Italy*
*Research Institute for Supply Chain Management, WU Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna, Austria*

## HIGHLIGHTS

- We propose a novel methodology that integrates operations research into urban studies.
- We find a significant correlation between urban sprawl and routing times.
- Addressing urban sprawl can reduce transportation's economic, social, and environmental costs.
- Shannon's entropy emerges as the best predictor of the route length.
- We gain insights about the geographical distribution of sprawl in Europe.

## ARTICLE INFO

*Keywords:*
Urban sprawl
Routing
Impacts of urban sprawl
Geography of urban sprawl
City logistics
Traveling salesman problem

## ABSTRACT

To address the growing challenges urban sprawl poses, it is essential to understand its influence on urban transportation, a primary source of economic, social, and environmental impact. This study fills this gap by quantifying the consequences of sprawl on transportation efficiency, proposing an interdisciplinary methodology that integrates knowledge from operations research.

Specifically, adopting a broad European perspective, we investigate how urban sprawl correlates with travel distances and optimal routes in 156 spatially heterogeneous cities across 28 European countries.

We discover a significant correlation between five sprawl indicators (Land usage, Gini coefficient, Shannon entropy, Moran I index, and Bribiesca index) and both travel distances and routes by car and bicycle: transportation is inherently less efficient in cities with higher levels of sprawl. Among the considered indicators, Shannon entropy emerges as the best predictor of route efficiency.

We offer insights into the geography of sprawl in Europe, finding that many Spanish cities stand out for their compactness and route efficiency, while hotspots of sprawl are present in many Western and Central European countries.

Our results underline the underestimated importance of addressing urban sprawl to reduce transportation's economic, social, and environmental costs and encourage policymakers and urban planners to prioritize compact city development to foster sustainable urban growth.

## 1. Introduction

Urban sprawl consists of the expansion of urban settlements into low-density residential areas, either in a deregulated or regulated fashion. Sprawled cities are affected by weak public transportation, a lack of walkable and bicycle-friendly routes, long distances from services and workplaces, and dependency on private vehicles. Many studies have identified it as a significant issue of modern urban development, and urban planners increasingly view it negatively. Nevertheless, sprawl is on the rise (Behnisch, Krüger, & Jaeger, 2022).

This paper investigates how urban sprawl influences transportation efficiency. The mobility of people and goods is integral to daily routines in every urban context. It is a source of economic and social cost, contributing negatively to the energy and $CO_2$ balance. Companies are working to optimize urban transportation and city logistics, developing algorithms for routing (Crainic, Ricciardi, & Storchi, 2009; Cattaruzza, Absi, Feillet, & González-Feliu, 2017) with the goals of reducing both costs and environmental impacts (Turan, Hemmelmayr, Larsen, & Puchinger, 2024). However, the concrete risk is that uninterrupted urban sprawl undermines these optimization efforts by making distances

* Address: DPIA, University of Udine, via delle Scienze 206, 33100, Udine, Italy.
*E-mail addresses:* robertomaria.rosati@uniud.it, robertomaria.rosati@wu.ac.at.

https://doi.org/10.1016/j.landurbplan.2024.105205
Received 30 December 2023; Received in revised form 4 September 2024; Accepted 5 September 2024
Available online 15 October 2024

inherently longer.

Hence, we aim to answer the following research questions:

1. Under what methodological framework can we measure the correlation between sprawl and the efficiency/length of vehicle routes?
2. What is the nature and statistical significance of such a correlation, and what does it imply?
3. Which indicator of sprawl emerges as the best predictor of route efficiency in a city?
4. What specific considerations and geographical insights can we derive for the context of Europe?

In this study, we adopt a broad European perspective comparing sprawl and routes increasingly 156 urban areas[1] located across the majority of European countries. For every city, we compute five distinct sprawl indicators: Land Usage ($L$), Gini coefficient ($G$), Adapted Shannon Entropy ($\eta'$), Moran I index ($I$), and Bribiesca index ($B$). Routing time is estimated on road graphs based on real population density and distances. On them, we solve the Asymmetric Traveling Salesman Problem (ATSP) and use the average value of the optimal solutions as a benchmark for route efficiency. We do this for different graph sizes (10 and 50) and different transportation means (*car* and *bicycle*). A Pearson correlation test measures the correlation between sprawl and routes.

The results indicate that sprawl and routes correlate positively with statistical significance. Our study also highlights the spatial differences across European countries. A rich set of graphics, tables, and a discussion complete our analysis. The main implication for policymakers is that the contrast to urban sprawl is a lever to improve transportation efficiency, reducing times, costs, and externalities.

The remainder of this paper is organized as follows. Section 2 presents a literature review. Section 3 describes the methodology. Computational results are described in Section 4 and discussed in Section 5. The conclusions are drawn in Section 6. The paper is completed by Appendix A, which contains the data for all the cities in the sample.

## 2. Related work

The acknowledgment of an urban sprawl phenomenon dates back to the 1950s, with the work of Burckhardt, Frisch, and Kutter (1955) and Whyte (1958). Nowadays, there is a large consensus that urban sprawl causes more damage than benefits (Hamidi, Ewing, Preuss, & Dodds, 2015).

Urban sprawl has traditionally been studied within the North American context, where extensive evidence has been gathered on its detrimental effects. For instance, Pendall (2000) and Power (2001) identified a connection between urban sprawl in the United States (US) and issues such as social exclusion and racial segregation. Additionally, Putnam (2000) shed light on the link between sprawl and marginalization in suburban areas.

Many scholars have associated urban sprawl with increased travel distances and longer commuting times. Nilles (1991), for example, observed that while residential development in the US has expanded outward, employment opportunities have remained concentrated in city centers, resulting in extended commuting times. Consequently, urban sprawl leads to increased time spent sitting in transportation (Zolnik, 2011) and is associated with higher obesity rates in US cities (Ewing, Schmid, Killingsworth, Zlot, & Raudenbush, 2003; Zhao & Kaestner, 2010). The forced car ownership imposed by the limited accessibility of sprawling areas places a significant financial burden on low-income households (Tiznado-Aitken, Lucas, Munoz, & Hurtubia, 2022). Lee (2020) examined the connection between urban form and travel behavior in the US, showing that urban form significantly influences modal split, particularly in large metropolitan statistical areas. Furthermore, Lee (2020) found that urban form factors have a statistically significant association with commuting trips and emissions from road traffic, with higher sprawl leading to higher emissions. The reliance on private vehicles in sprawled urban areas has also been linked to increased pollution and $CO_2$ emissions in the US (Schweitzer & Zhou, 2010), with similar findings reported in Japan (Makido, Dhakal, & Yamagata, 2012). Moreover, urban sprawl in the US is correlated with higher rates of road and pedestrian fatalities (Frumkin, 2002; Ewing & Hamidi, 2015; Ewing, Hamidi, & Grace, 2016), and, as reported by Trowbridge, Gurka, and O'connor (2009), sprawled cities face the additional challenge of longer ambulance response times.

The existence of an urban sprawl phenomenon in Europe was highlighted in the 2000s. The European Environment Agency raised concern about the lack of awareness by policymakers and defined it as "the ignored challenge" (European Environment Agency, 2006; European Environment Agency, 2016). Analogously, scholars raised concerns regarding the applicability to Europe of findings from the United States (see, e.g. Schwanen, 2002; Patacchini, Zenou, Henderson, & Epple, 2009; Hennig et al., 2015) and pushed for context-specific research. However, most studies examining the relationship between travel distance, commuting, and urban sprawl in Europe are limited to the national scale. For example, Travisi, Camagni, and Nijkamp (2010) highlighted that empirical data from Italy support the expectation that intensive travel movements are associated with urban sprawl. Similarly, Marique, Dujardin, Teller, and Reiter (2013) reported similar findings for Belgium, where De Vos and Witlox (2013) also observed a decrease in cycling and walking in highly sprawled urban areas. In Spain, Hortas-Rico and Solé-Ollé (2010) highlighted the increased cost of public services. Moreover, in Germany, Van Ommeren and Gutiérrez-i Puigarnau (2011) found that reducing commuting time to a negligible level could decrease absenteeism at work. Despite growing evidence, housing policies in Europe, which are determined at the national level, remain generally lax concerning urban sprawl, as shown in studies on France, Germany, and the Netherlands (Bas Waterhout & Sykes, 2013), Italy (Salvati, 2015), Greece (Colantoni, Grigoriadis, Sateriano, Venanzoni, & Salvati, 2016), and the United Kingdom (Ferm, Clifford, Canelas, & Livingstone, 2021).

Some scholars have begun to investigate urban sprawl from a European perspective, producing studies based on data from various cities across the continent (Oueslati, Alvanides, & Garrod, 2015; Pourtaherian & Jaeger, 2022). Moreover, Hennig et al. (2015) have proposed a European *de-sprawling* strategy. However, to the best of our knowledge, no research has yet focused on the impact of urban sprawl on routing efficiency in Europe.

## 3. Methodology

We answer the first research question by proposing a novel methodological framework that integrates knowledge from operations research into urban studies, which building blocks are: (i) a sample of cities, (ii) one or more measures of sprawl, (iii) routing times, and (iv) a measure of correlation.

To address (i), we use a sample of 156 European cities and data from the JRC-GEOSTAT population grid[2]. The most recent validated version contains the population data updated to 2018, with a resolution of 1 km$^2$, covering the whole European Union plus a few neighboring countries. This dataset is produced by the European Commission Joint Research Center in collaboration with Eurostat. The JRC-GEOSTAT 2018 was created by integrating data from national institutes of statistics regarding the population figures for 2018 with the population

[1] Throughout this paper, we use interchangeably the words "urban area" and "city".

[2] Available as a shapefile with coordinate system EPSG:3035 at: https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography/geostat

density per 1 km$^2$ cells in 2011 (contained in the GEOSTAT 2011) and with data on the density of built-up areas in 2012 and 2018 obtained from Copernicus earth observation data. Data quality is discussed in the Appendix A, and further details about the construction of the dataset are available in the technical report by Batista e Silva, Dijkstra, and Poelman (2021). Section 3.1 addresses the details behind data selection. For the complete list of cities, their details, and their density heatmaps, we redirect the reader to Appendix A.

Knowing the population distribution of a city, we can compute a measure of urban sprawl to address (ii). However, the literature has not agreed on a single best indicator of sprawl. For this reason, we study five different indicators of sprawl: Land Usage ($L$), Gini coefficient ($G$), Adapted Shannon Entropy ($\eta'$), Moran I index ($I$), and Bribiesca index ($B$). The first three measures rely solely on population data, while the last two consider population, distances, and geometry. The indicators are discussed in Section 3.2.

The routes are the third component (iii). By exploiting the available data on the population distribution, we generate realistic road graphs such that the nodes are more concentrated in denser areas of the city. The underlying hypothesis is that the demand for products and services (like good/food delivery, waste collection, home healthcare…) is more likely to arise where the population is denser. On them, we compute the optimal route that passes through all points, a solution to the Asymmetric Traveling Salesman Problem (ATSP). Given that no solution to the ATSP is better than the optimal one, we use the cost of the optimal route to measure routing efficiency. A similar generation procedure has been recently employed by Rosati and Schaerf (2024) and Ceschia, Di Gaspero, Rosati, and Schaerf (2024) to generate realistic graphs for the Capacitated Dispersion Problem and for the Home Healthcare Routing and Scheduling Problem. Section 3.1 focuses on graph generation, while Section 3.3 explains the mathematical background.

Finally, we use the Pearson correlation (iv) between the five indicators and the average optimal routes, and we fit a linear regression model for route prediction. These points are discussed in Section 3.4.

This section introduces several concepts and uses many terms taken from the information theory and operations research literature. Table 1 summarizes the notation.

### 3.1. City selection and road graph generation

We define a city as a subset of 1 km$^2$-cells with coordinates $(x_i, y_i)$ of the GEOSTAT dataset such that $(x_i - x)^2 + (y_i - y)^2 \leqslant r^2$, where $(x, y)$ are the center's coordinate and $r$ is the desired radius. The resulting shape is a circle centered in $(x, y)$ and with radius $r$. We suitably set the $r$ to a value of 30 km. All urban areas, therefore, include 2809 cells and cover a surface of 2809 km$^2$. However, if a city lies on the sea, cells classified as water are excluded, and the radius is suitably enlarged to include 2809 land cells. Therefore, all cities are identical in size but not shape, which is a not circle in coastal cities. Furthermore, we enforce geographical disjointness by ensuring no cell appears in two areas to avoid biases and overfitting. Fig. 1 shows four cities, all having the same surface but different shapes due to the presence of the sea.

**Table 1**
Notation used.

| Symbol/ Term | Definition |
|---|---|
| *cell* | The smallest unit of the grid, with squared shape, side 1 km and size 1 km$^2$. For every cell, its coordinates and its population are known. |
| *city* | The subset of cells from the GEOSTAT dataset centered in the coordinates $(x, y)$ that lies within a distance $r$. |
| *urban area* | Used as a synonym of *city*. |
| $(x, y)$ | Coordinates of a cell in the GEOSTAT dataset |
| $n$ | Number of cells in the city |
| $p_i$ | Population of the $i$-th cell |
| $l(p_i)$ | 1 if $p_i > 0$, 0 otherwise |
| $n_+$ | Number of cells with positive population |
| $w_{ij}$ | 1 if cells $i$ and $j$ are neighbors, 0 otherwise |
| $\eta$ | Shannon entropy |
| $T$ | Number of sides of the cells |
| $P_C$ | Contact perimeter |
| $P$ | Perimeter of the shape of the city |
| $(V, E)$ | A (road) graph |
| $V$ | Set of vertices of the graph |
| $E$ | Set of edges of the graph |
| $d_{ij}$ | Distance between nodes $v_i$ and $v_j$ |

Graphs composed of geolocated nodes are generated through Algorithm 1. The procedure takes as input the urban area $A$, the transportation mean $T$ (car or bicycle), the number of nodes $c$, and returns a graph $(V, E)$. First, we initialize the empty graph (line 1) and we assign every cell with a probability of $p^*[i] = p_i / \sum_{j=1}^{n} p_j$, where $p_i$ is the population of the $i$-th cell (line 2). Based on the probabilities, we sample, through a biased random selection with replacement, $c$ cells from the urban area (lines 3—9). For each sampled cell $a_k$, we determine a point at random by drawing uniformly random coordinates $(x'_k, y'_k)$ in the cell boundaries $[x_{min}, x_{max}]$ and $[y_{min}, y_{max}]$. The actual coordinates of the nodes $(x_k, y_k) \in V$ are obtained by matching them to the nearest point with road access. Finally, we populate $E$ computing the distance matrix (lines 10—15) as the shortest paths between pairs of nodes in the road network. We employ the Open Source Routing Machine, an open-source C++ routing engine. In this work, we use it through the R package `osrmr`. The matrix is asymmetric due to one-way streets, traffic lights, steepness, and other real-world constraints.

For every city, we generate 10 graphs by each transportation mode $T \in \{car, bicycle\}$ and each size $c \in \{10, 50\}$, obtaining a total of 6420 graphs.

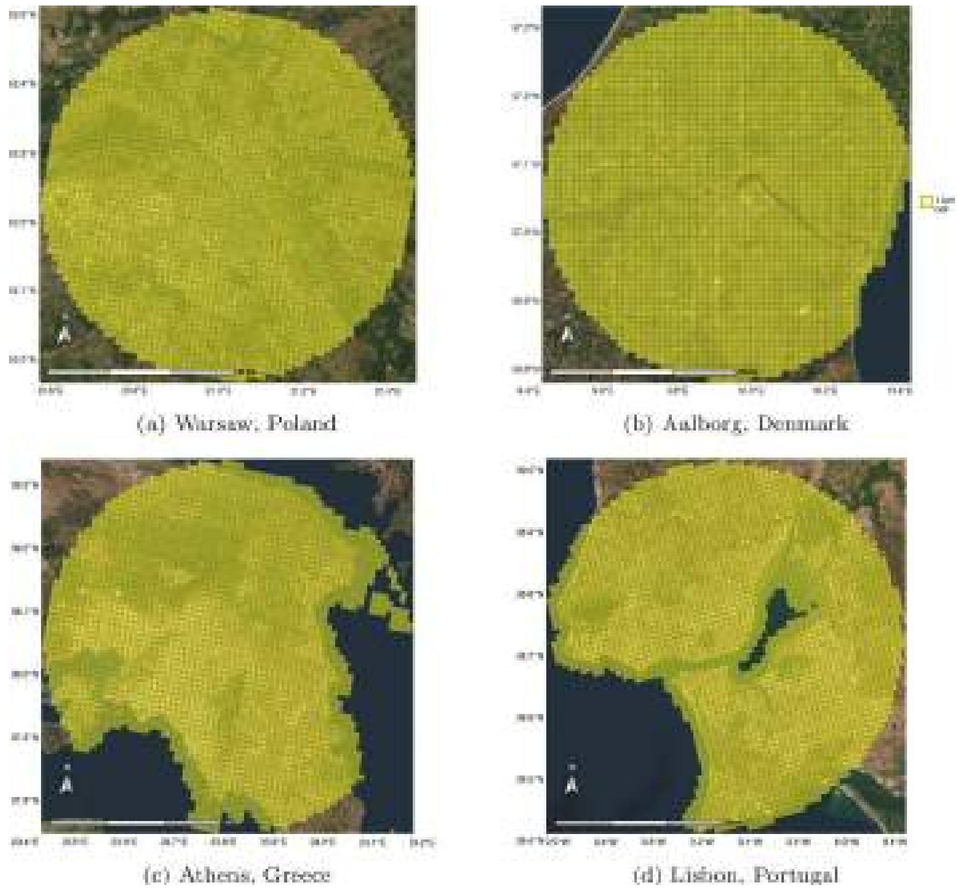
**Algorithm 1.** Realistic road graph generator

**input:** Area $A$ (with $n$ cells), transportation mean $T$, number of nodes $c$
**output:** $(V, E)$ directed graph with $|V| = c$ nodes
1: $V \leftarrow \emptyset, E \leftarrow \emptyset$
2: $p^* \leftarrow \left[\frac{p_i}{\sum_{j=1}^{n} p_j} \text{ for } i \leftarrow 1, \dots, n\right]$ //Compute the probabilities
3: **for** $k \leftarrow 1, \dots, c$ **do**
4: $\quad a_k \in A \leftarrow \text{Random}(A, p^*)$ //Random coordinates within the urban area
5: $\quad (x'_k, y'_k) \leftarrow \text{Random}(a_k, \cdot)$ //Random coordinates within the cell
6: $\quad (x_k, y_k) \leftarrow \text{Nearest}(x'_k, y'_k)$ //Nearest street address
7: $\quad v_k \leftarrow (x_k, y_k)$
8: $\quad V \leftarrow V \cup \{v_k\}$
9: **end for**
10: **for** $i \leftarrow 1, \dots, c$ **do**
11: $\quad$ **for** $j \leftarrow 1, \dots, c$ **do**
12: $\quad\quad d_{ij} \leftarrow \text{Distance}(v_i, v_j, T)$ //Real road distance
13: $\quad\quad E \leftarrow E \cup \{d_{ij}\}$
14: $\quad$ **end for**
15: **end for**
16: **return** $(V, E)$

### 3.2. Measures of urban sprawl

Many definitions and measures of sprawl have been proposed. Some of these are adapted from information theory, spatial statistics, and econometrics, such as the Shannon Entropy, the Gini coefficient, the Moran I Index, and the Bribiesca contact perimeter (Yeh & Li, 2001; Tsai, 2005; Steurer & Bayr, 2020). Several articles have treated and compared indicators (Davis & Schaub, 2005; Frenkel & Ashkenazi, 2008; Kasanko et al., 2006; Schneider & Woodcock, 2008; Torrens, 2008). However, most of these measures are subject to limitations and none of them has been universally accepted as a measure of sprawl (Wolman et al., 2005; Nazarnia, Harding, & Jaeger, 2019).

Jaeger, Bertiller, Schwick, and Kienast (2010b) undertook a significant effort to address the lack of a unified definition and reliable



(a) Warsaw, Poland (b) Aalborg, Denmark (c) Athens, Greece (d) Lisbon, Portugal

**Fig. 1.** Examples of selection of cities. Satellite map images courtesy of `Esri`.

measures by proposing a classification of existing urban sprawl measures based on 13 suitability criteria. Subsequently, Jaeger, Bertiller, Schwick, Cavens, and Kienast (2010a) proposed four new metrics that try to capture the multiple dimensions of the phenomenon: degree of urban dispersion (DIS), total sprawl (TS), degree of urban permeation of the landscape (UP) and sprawl per capita (SPC). In subsequent work, Jaeger and Schwick (2014) summarized the four metrics into a new measure called weighted urban proliferation (WUP).

Hereby we adopt five indicators compatible with the GEOSTAT dataset, resumed in Table 2. All indicators take values between 0 and 1. For all indicators but the Gini coefficient, a higher value indicates higher sprawl.

### 3.2.1. Land usage

Land usage is the most elementary measure among the ones considered. We first define a function $l(p_i)$ equal to 1 if the $i$-th cell counts at least one inhabitant and 0 otherwise. In other terms, we differentiate between inhabited and uninhabited cells without considering the actual population of the cell.

**Table 2**
The five indicators of sprawl, and the measured sprawl dimensions.

| Indicator | Name | Measured dimension |
|---|---|---|
| $L$ | Land usage | Estimation of residential land use |
| $G$ | Gini coefficient | Inequality of the population distribution |
| $\eta'$ | Adapted Shannon entropy | Uncertainty in the population distribution |
| $I$ | Moran I index | Degree of spatial clustering |
| $B$ | Bribiesca Index | Compactness of the city shape |

$$l\left(p_i\right) = \begin{cases} 1, & \text{if } p_i > 0 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

We use a variable $n_+$ to count the number of cells with positive population.

$$n_+ = \sum_{i=1}^{n} l\left(p_i\right) \tag{2}$$

Finally, we denote by the letter $L$ the land usage:

$$L = \frac{n_+}{n} \tag{3}$$

For example, consider an urban area of 100 km$^2$, divided into 100 cells of 1 km$^2$ each. If half of the cells are inhabited, $L = 50/100 = 0.5$. If the same city only occupies 20 cells, $L = 20/100 = 0.2$. Therefore, a higher value of $L$ indicates a higher level of sprawl because the population spreads over a greater percentage of land. While this measure can be quite imprecise as it does not take into account the actual population of a lived cell, cells with no inhabitants are often encountered in the case of compact cities surrounded by unused or arid land, as it is often the case in Spain, Greece, Scandinavia or in cities that lie in Alpine valleys.

### 3.2.2. Gini coefficient

The Gini coefficient is a notorious index defined by the Italian statistician Corrado Gini in Gini (1912), which is often used in socioeconomic science to capture the inequality in the distribution of income or wealth. It has been employed to measure urban sprawl (see, e.g., Tsai, 2005), as it captures the inequality in the distribution of the population over the urban area. A higher level of inequality in the population distribution indicates a lower level of sprawl because the population within the urban area is concentrated in a few cells. The Gini coefficient is computed as:

$$G = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|p_i - p_j\right|}{2\sum_{i=1}^{n}\sum_{j=1}^{n}p_j} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left|p_i - p_j\right|}{2n^2\overline{p}} \tag{4}$$

In this formula, $\left|p_i - p_j\right|$ represents the absolute difference between the population of two cells, $i$ and $j$, $\sum_{i=1}^{n}\sum_{j=1}^{n}\left|p_i - p_j\right|$ represents the sum of the absolute differences between the population of all pairs of cells, $\sum_{j=1}^{n}p_j$ represents the sum of all populations in the cells, that is, the total population of the city, and $\overline{p}$ is the average cell population. The resulting measure $G$ is comprised in the interval $[0,1)$, where 0 indicates a perfect equality in the distribution of the population, that is, all cells have the same population, while values near to 1 indicate a perfect inequality, that is, all the population is concentrated in a single cell.

### 3.2.3. Adapted Shannon entropy

In a broad sense, entropy is a physical measure associated with the *disorder* of a system. Shannon entropy or information entropy (Shannon, 1948), used in information theory to measure the uncertainty of a random variable, has also been employed as a measure of urban sprawl (Yeh & Li, 2001; Tsai, 2005; Steurer & Bayr, 2020). When applied to the spatial distribution of the population in a city, a higher value of entropy indicates higher sprawl. One drawback, as noted by Yeh and Li (2001), is that entropy is affected by the size of the area being measured. Additionally, Nazarnia et al. (2019) are particularly critical of using entropy as a measure of urban sprawl, noting that it is strongly influenced by the choice of zones with regard to their center, shape, and location. However, in this context, the cities have the same number of cells, making entropy suitable as a measure of sprawl.

To compute Shannon entropy, we first divide the population of the cells by the maximum cell population so that all values are normalized in the interval $[0,1]$:

$$p_i^{'} = \frac{p_i}{\max_{j=1}^{n}p_j} \tag{5}$$

Then, to compute the urban sprawl, we consider the following relation.

$$\eta := -\sum_{i=1}^{n}p_i^{'}\log p_i^{'} = \mathbb{E}\left[-\log p_i^{'}\right] \tag{6}$$

Finally, we divide $\eta$ by the maximum entropy that can be theoretically obtained, that is, the entropy obtained when all cells have the same population ($p_i^{'} = 1/n$ for all $i$).

$$\eta^{'} := \frac{-\sum_{i=1}^{n}p_i^{'}\log p_i^{'}}{-\sum_{i=1}^{n}\frac{1}{n}\log\frac{1}{n}} = \frac{-\sum_{i=1}^{n}p_i^{'}\log p_i^{'}}{\log n} = \frac{\eta}{\log n} \tag{7}$$

The numerator $-\sum_{i=1}^{n}p_i^{'}\log p_i^{'}$ calculates the entropy of the probability distribution $p_i^{'}$. The denominator $-\sum_{i=1}^{n}\frac{1}{n}\log\frac{1}{n}$ calculates the maximum possible entropy for a distribution with $n$ events. It assumes an equal population for all cells ($1/n$). The numerator is divided by the denominator to normalize the entropy.

The resulting measure $\eta^{'}$ is in the interval $[0,1]$. We call it *adapted Shannon entropy*. A higher value of $\eta^{'}$ indicates a higher level of sprawl, as it is related to the uncertainty of the distribution of the population.

### 3.2.4. Moran I index

Moran I index (Moran, 1950) measures spatial autocorrelation, a quantification of cell similarity. It was first associated with urban sprawl by Tsai (2005), who stated that the Moran I index can distinguish compactness from sprawl in the metropolitan form. More recently, however, Steurer and Bayr (2020) criticized its hard interpretability. The index requires a weighting function that takes higher values for cells that are closer in space. Analogously to Tsai (2005) and Steurer and Bayr (2020), we define a function $w_{ij}$ that takes value 1 if two cells are neighbors, 0 otherwise.

$$w_{ij} = \begin{cases} 1, & \text{cells } i \text{ and } j \text{ are neighbors} \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

The Moran I index is defined as:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}\left(p_i - \overline{p}\right)\left(p_j - \overline{p}\right)}{\sum_{i=1}^{n}\left(p_i - \overline{p}\right)^2} \tag{9}$$

where $n$ is the number of cells, $\overline{p}$ is the mean population of the cells in the city, and $p_i$ is the population of the $i$-th cell. A value of $I$ close to 1 indicates a high level of sprawl, as it is related to clusters of cells with high or low populations.

### 3.2.5. Bribiesca index

The Bribiesca index was proposed by Bribiesca (1997) to distinguish the compactness of discrete shapes composed of identical cells (or *pixels*) by quantitatively analyzing their perimeters. It is often used in the domain of image processing. Steurer and Bayr (2020) recently proposed it as a sprawl indicator. In the following, we adopt their procedure and their terminology. The Bribiesca index is based on the concept of contact perimeter, which is the sum of the lengths of the cell sides that are in contact with other cells. The intuition is that compact cities have a higher ratio of contact perimeter to total perimeter, while sprawled cities have a larger external perimeter. In general, the most compact shapes are squares and circles, while a shape with completely disconnected cells that do not border each other will have a contact perimeter equal to zero. We consider only cells with $p_i > 0$. If $n_+$ is the number of cells such that $p_i > 0$, and $T$ is the number of sides of the cells, it holds the relation:

$$2P_C + P = Tn_+ \tag{10}$$

where $P_C$ is the contact perimeter and $P$ is the perimeter. Given that the minimum contact perimeter for a continue shape is $P_{Cmin} = n_+ - 1$ and the maximum contact perimeter as $P_{Cmax} = \frac{Tn_+ - 4\sqrt{n_+}}{2}$, the normalized Bribiesca index is defined as follows:

$$B = \frac{P_C - P_{Cmin}}{P_{Cmax} - P_{Cmin}} \tag{11}$$

However, the formula mentioned above applies to connected figures. In our case, we have to consider the possibility that cells are not connected, thus we set $P_{Cmin} = 0$, and the modified normalized Bribiesca index becomes:

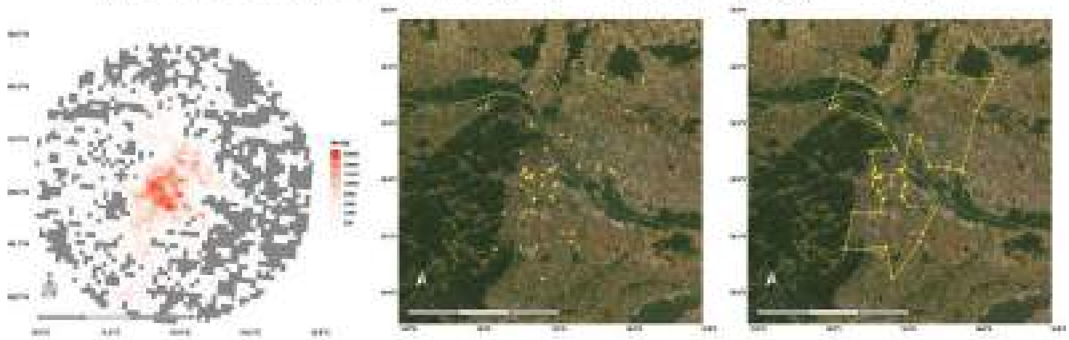
$$B = \frac{P_C}{P_{Cmax}} \tag{12}$$

## 3.3. Route length estimation

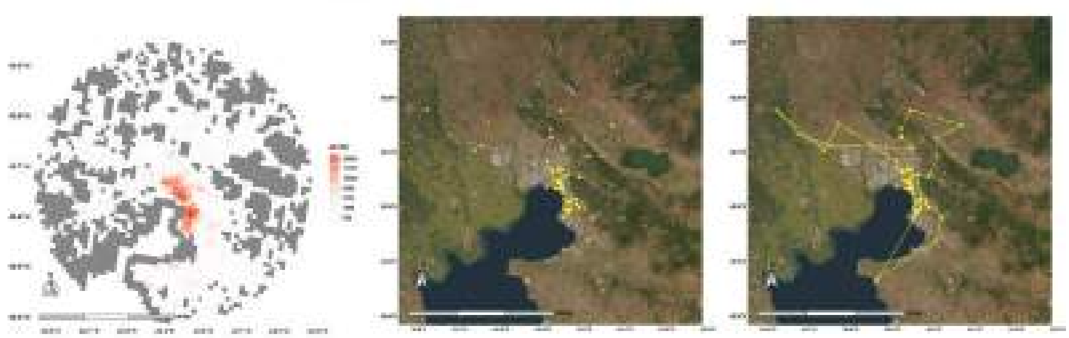
While many real-world routing problems exist, to compare route efficiency in cities, we use as a benchmark the most fundamental one, the Traveling Salesman Problem (TSP), which deals with finding the shortest Hamiltonian cycle that covers all nodes in a graph. A Hamiltonian cycle (also known as Hamiltonian circuit) is a path that visits every vertex exactly once, starting and ending on the same node. Therefore, the optimal solution to the TSP, a solution of minimum cost, is the Hamiltonian cycle of minimum length on a graph.

The TSP was proven to be NP-hard by Karp (1972). Given that road graphs are asymmetric, we refer to the asymmetric version of the problem, the Asymmetric Traveling Salesman Problem (ATSP). We describe the integer linear programming (ILP) formulation proposed by Dantzig, Fulkerson, and Johnson (1954). Given a directed graph $(V, E)$, where $V$ is the set of nodes and $E \subseteq V \times V$ is the set of edges, we define a binary variable $x_{uv}$ that takes value 1 if the directed edge $(u, v)$ belongs to the solution, 0 otherwise.

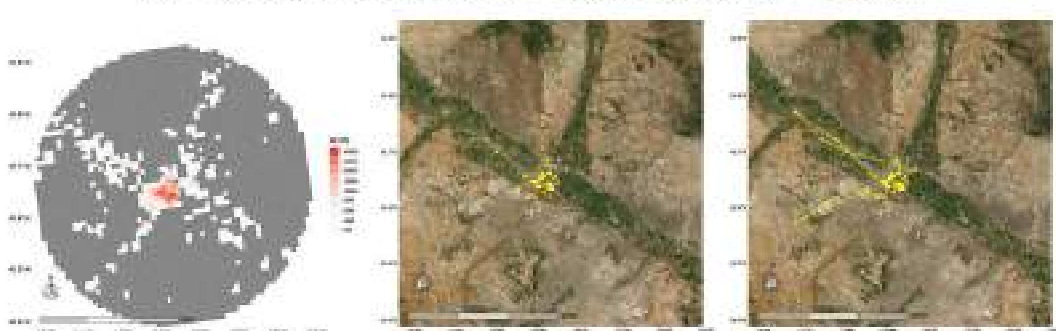
(a) Venice-Padua-Treviso, Italy, $\eta' = 0.900$, $\bar{C}(\mathcal{G}_{ATc}) = 33172\ s$



(b) Vienna, Austria, $\eta' = 0.748$, $\bar{C}(\mathcal{G}_{ATc}) = 26820\ s$



(c) Thessaloniki, Greece, $\eta' = 0.654$, $\bar{C}(\mathcal{G}_{ATc}) = 19679\ s$



(d) Zaragoza, Spain, $\eta' = 0.528$, $\bar{C}(\mathcal{G}_{ATc}) = 15648\ s$

**Fig. 2.** Optimal routes in four cities. From top to bottom: Venice-Padua-Treviso, Vienna, Thessaloniki, and Zaragoza. From left to right: density heatmap, red cells are denser, grey cells hare not inhabited; a random graph with $|V| = 50$, the yellow dots indicate the position of the nodes; optimal solution to the ATSP (for sake of simplicity, the points are connected linearly). Satellite image is courtesy of `Esri`.

$$\min \sum_{u \in V} \sum_{v \in V, v \neq u} d_{uv} x_{uv} \tag{13a}$$

$$\sum_{u \in V, u \neq v} x_{uv} = 1 \qquad v \in V; \tag{13b}$$

$$\sum_{v \in V, v \neq u} x_{uv} = 1 \qquad u \in V; \tag{13c}$$

$$\sum_{u \in Q} \sum_{v \neq u, v \in Q} x_{uv} \leqslant |Q| - 1 \qquad \forall Q \subsetneq V, |Q| \geqslant 2 \tag{13d}$$

The ILP model for the ATSP is shown in Eq. (13). Constraints 13b and 13c impose that the route should go in and go out every node, and Constraints 13d are the sub-tour elimination constraints, where $Q$ is any possible proper subset of $V$ with cardinality at least 2. The Objective 13a imposes the minimization of the length of the route.

Finally, given that we have a family of graphs $\mathscr{G}_{ATc} = \{(V,E)_{ATc_1}, \dots, (V,E)_{ATc_{10}}\}$ for every combination of urban area $A$, transportation mode $T$ and size $c$, we call $\overline{C}(\mathscr{G}_{ATc})$ the average cost of the optimal ATSP routes on the graphs in $\mathscr{G}_{ATc}$, and we define it as:

$$\overline{C}\left(\mathscr{G}_{ATc}\right) := \frac{1}{|\mathscr{G}_{ATc}|} \sum_{(V,E) \in \mathscr{G}_{ATc}} \left(\min \sum_{u \in V} \sum_{v \in V, v \neq u} d_{uv} x_{uv}\right) \tag{14}$$

### 3.4. Inference analysis

We adopt Pearson correlation to correlate the five measures of sprawl with $\overline{C}(\mathscr{G}_{ATc})$. It takes values between -1 and 1, with values near -1 and 1 indicating a strong negative and positive correlation, respectively, and values near 0 indicating no correlation. We also evaluate the correlation between the indicators and the average distance between nodes, denoted as $\overline{d}$. Finally, we perform linear regression for the prediction of $\overline{d}$ and $\overline{C}(\mathscr{G}_{ATc})$. The resulting linear models allow an estimation of the route efficiency in a generic city without solving the corresponding mathematical model.

### 3.5. Visual example

We discuss a visual example of the described methodology on four cities from the sample with very different levels of sprawl. The most sprawled area is a part of the Veneto region in Italy, which comprises the cities of Venice, Padua, and Treviso. It is characterized by a net of small- and medium-sized cities and towns and the absence of a predominant city core. The other cities are Vienna, in Austria, Thessaloniki, in Greece, and Zaragoza, in Spain. The latter is a very compact city that lies in the Ebro Valley, with a very dense urban core and a sharp transition between the urban core and the surrounding land, with a limited presence of suburbs.

Fig. 2 shows the four urban areas. For each city, from left to right, the first graphic is the density heatmap, where the color spans from white to red, being darker for denser cells and grey if a cell is not inhabited. The second graphic shows a random graph, with $|V| = 50$, generated according to the methodology described in Section 3. Finally, the third graphic is the optimal solution to the ATSP on the graph. With a simplification, the picture's points are connected linearly. For every city, we report in the caption $\eta'$ and $\overline{C}(\mathscr{G}_{ATc})$. The four urban areas are ordered from higher to lower sprawl. In Venice-Padua-Treviso, we observe that practically all cells are inhabited, without a predominant city core. The sampled points are very sparse, and routes are inevitably long. The other extreme, Zaragoza, is a very compact city. Therefore, the points are very close, and the routes are short. Vienna and Thessaloniki are in an intermediate situation.

## 4. Computational results

Our code is written in R and runs on Ubuntu 22.04. To handle the geographical database, we use the `sf` package, while to solve the ATSP, we use the solver Concorde (Applegate, Bixby, Chvatal, & Cook, 2006), written in C, which guarantees the optimality of the solution[3]. The handling and elaboration of the extensive data was possible thanks to a machine equipped with over 116 GB of RAM.

All the generated graphs (in TSPLIB format) and the tables with the computational results are publicly available at https://github.com/robertomrosati/urban-sprawl-public.

### 4.1. General results on sprawl and routes

The values of the sprawl indicators $L, G, \eta', I$, and $B$, the average distances $\overline{d}$ and the mean ATSP routes $\overline{C}(\mathscr{G}_{ATc})$ for all the 156 cities are shown in Tables A.9 to A.11 in Appendix A.

Sprawl and routes are compared in Fig. 3. The adapted Shannon entropy expresses the sprawl, and the route duration is converted into hours for better readability. We show the plot for all cities in the sample (Fig. 3a) and and for a subset of cities taken from Germany, Italy, and Spain, the three most represented countries (Fig. 3b).

Fig. 4 displays the geographical location of the 156 urban areas. The color indicates the average routing time $\overline{C}(\mathscr{G}_{ATc})$, with green associated with shorter routes and red associated with longer ones. The diameter of the point is proportional to the population of the urban area. The colors in the figure help us identify geographical differences in the distribution of sprawl.

Additionally, Fig. 5 shows the distribution of the route lengths if $c = 50$ and $T = car$ for countries with at least five cities in the sample (at least 50 graphs).

We observe sharp differences between cities in Europe. In order to visit the same number of customers, a company in a sprawled city needs twice or more the time needed in a compact city. Interestingly, we observe that routing times by bicycle in the most compact cities are comparable to those obtained by car in sprawled cities. We discuss further insights in Section 5.

### 4.2. Correlation between sprawl and routes

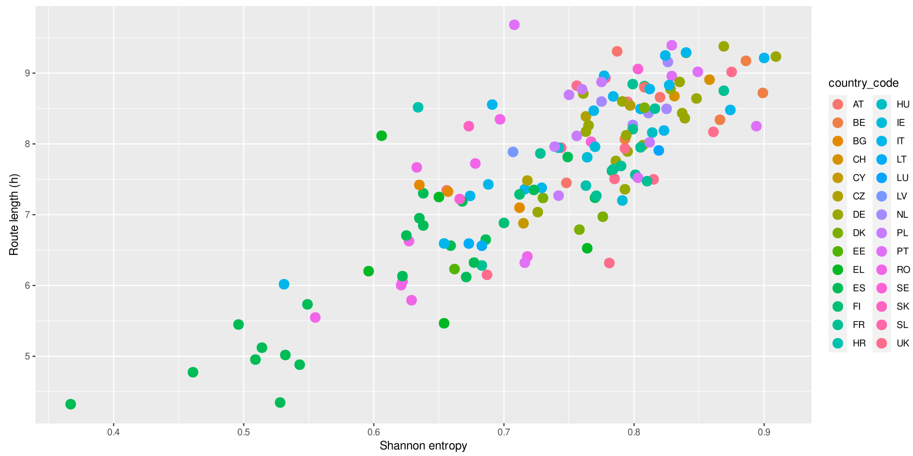
The results of the correlation test between the indicators and $\overline{C}(\mathscr{G}_{ATc})$ for the graphs with $c = 50$ are shown in Table 3. The Table also reports the p-values and the boundaries of the 95%-confidence intervals. We mark in bold the highest correlation, given by Shannon entropy, and all p-values below the significance threshold of 5%. They are always very near to zero, highlighting the statistical significance of the correlation, being slightly higher only for the Moran I index (but still well below the significance threshold). All indicators yield a significant correlation: higher sprawl bears longer routes, which answers the second research question.

In Table 4, we repeat the correlation test using only specific subgroups of cities. The first subset, labeled *no sea*, includes 107 cities with a radius of exactly 30.0 km. What we observe is a slight increase in the correlation. We also show the correlation test results for $c = 10$, which shows a lower correlation than $c = 50$ for all indicators but $I$. The p-values in this test are near zero, confirming the statistical significance of the correlation test.

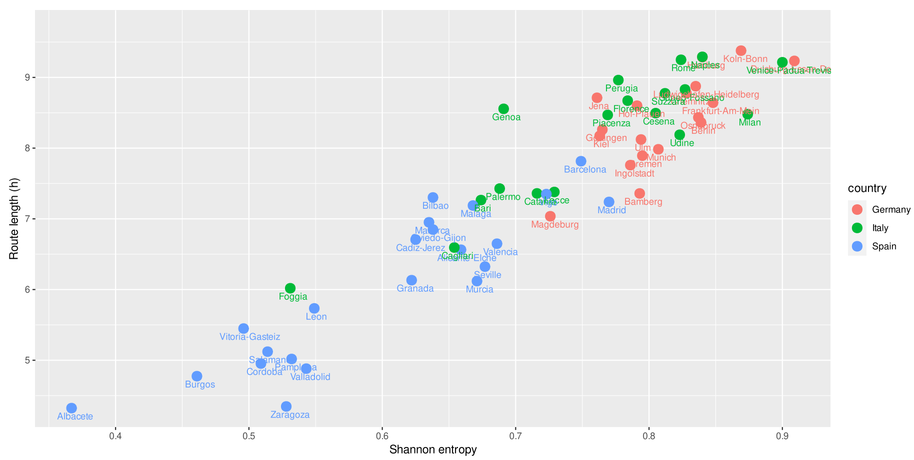
### 4.3. Correlation between sprawl and distances

In Table 5, we present the results of the Pearson correlation test

[3] Concorde is available online at http://www.math.uwaterloo.ca/tsp/concorde.html




(a) All 156 cities from the sample.




(b) Cities in the three most represented countries (Germany, Italy and Spain).

**Fig. 3.** Relationship between the adapted Shannon entropy $\eta'$ on the $x$-axis and the average routing time $\overline{C}(\mathscr{G}_{ATc})$ is on the $y$-axis for the cities in the sample. Legend for country codes is available in Table 8.

between the sprawl indices and the average shortest paths between pairs of points in the city. We observe that, while all indicators but $I$ have a higher correlation with $\overline{C}(\mathscr{G}_{ATc})$, $I$ is more correlated with $\overline{d}$. The p-values are well below the significance threshold for all the indicators, highlighting that the correlation between sprawl and distances is statistically significant.

### 4.4. Linear regression

The resume of the linear models obtained fitting $\overline{d}$ on the sprawl indicators are presented in Table 6. The table shows the parameter $R^2$ and the adjusted $R^2$ values, the $F$-statistic, and the coefficients $\beta_0$ and $\beta_1$ of the linear model $\overline{d} = \beta_0 + \beta_1 \cdot \sigma$, where $\sigma$ is either $L, G, \eta', I$ or $B$. The table also shows the following information about the residuals: the minimum, the first quartile, the median, the third quartile, and the maximum. The model's predictive accuracy is higher for the Shannon entropy.

We repeat the procedure for optimal ATSP routes $\overline{C}(\mathscr{G}_{ATc})$, in the case of $c = 50$, for both the car and the bicycle. The results are shown in Table 7, which has the same columns as Table 6. The (adjusted) $R^2$ are generally higher than for $\overline{d}$, indicating higher accuracy. Among the models compared in Table 7, $R^2$ is higher for Shannon entropy. In response to the third research question, Shannon entropy is the best predictor of routing efficiency.

Considering, in particular, Shannon entropy, the linear model tells that for every increase of 0.1 in the value of $\eta'$, the predicted distance between any pair of points by car increases by 232 seconds, which is almost four additional minutes more that are needed to reach any other point. Regarding routes, we obtain that for every increase of 0.1 in the value of $\eta'$, the route length increases by 3518 seconds, almost one hour.

Fig. 6 plots the distribution of the values of, respectively, $\overline{d}$ (Fig. 6a) and and $\overline{C}(\mathscr{G}_{ATc})$ (Fig. 6b) against Shannon entropy $\eta'$. In both cases, $T = car$ and $c = 50$. The red line is the prediction made by the linear model. The diagnostic plots for the model (the residuals vs. the fitted values, the normal Q-Q plot, the scale-location plot, and the residuals vs. the leverage) are on the right.

## 5. Discussion

The main discovery of this study is that the population distribution in a city plays a crucial role in routing efficiency: compact cities are efficient in transportation. Corporations and citizens operating and living in cities affected by higher sprawl are paying a *sprawl toll*, of which they are probably unaware.

A major implication for policymakers is that reducing urban sprawl can significantly enhance transportation efficiency and mitigate its externalities. Conversely, urban sprawl presents substantial challenges to the effectiveness of transportation optimization, as solutions are conditioned by the geography introduced by sprawling urban landscapes. Addressing urban sprawl is also crucial within the broader global

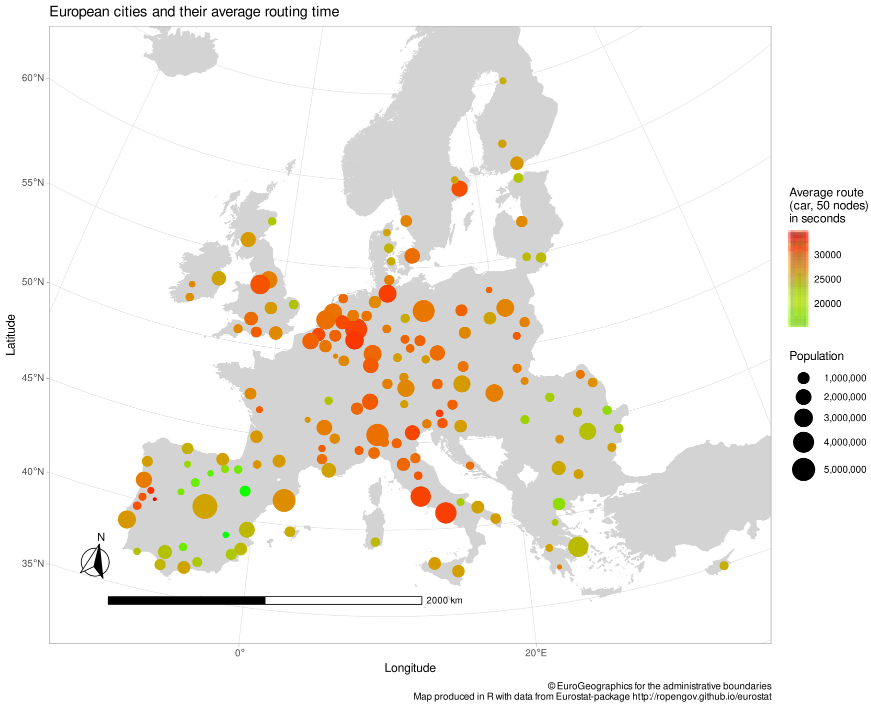



**Fig. 4.** All the 156 cities used in the analysis. The color represents the average routing time, while the diameter of the circle is proportional to the population.

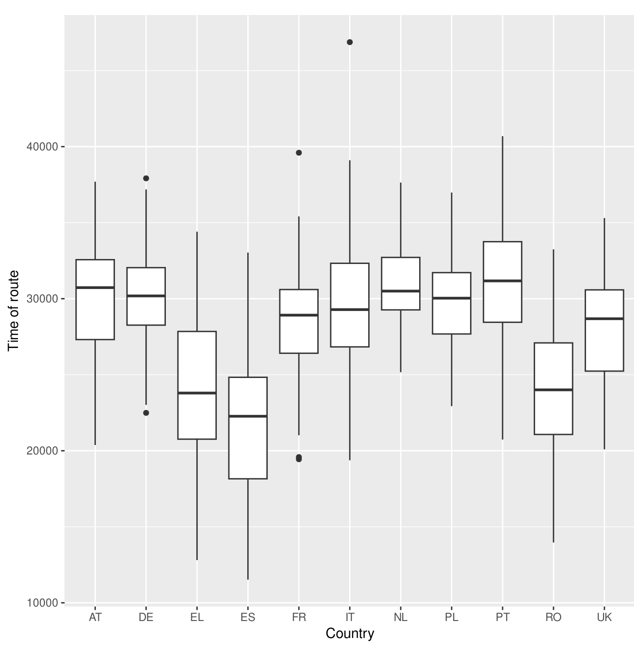



**Fig. 5.** Distribution of route length in countries represented by at least five cities. Country codes are explained in Table 8.

strategy to limit the rise in global temperatures to 1.5°C above pre-industrial levels, a critical threshold identified by scientists to prevent severe consequences of climate change (Intergovernmental Panel on Climate Change, 2022). Proactive measures to curb sprawl today could avert future environmental and logistical challenges in cities.

In what follows, we discuss the implications of the results and the geographical insights we can derive, answering the fourth research question.

### 5.1. Geographical insights

Europe is not a monolithic block, but it is made up of diverse socioeconomic situations and urban structures (the list of countries used in this study can be visualized in Table 8). Therefore, for the heterogeneity of the urban forms, choosing a dataset that covers most of Europe strengthens our results. At the same time, the fact that data comes from a unique and trustworthy source ensures that the differences in routing times are not due to discrepancies in data.

From Tables A.9 to A.11 and Figs. 4,5, we observe that many Spanish cities score better than other European cities in all the sprawl indicators and route lengths. We also see that sprawl increases from Southern and Eastern Europe to Central and North-Western Europe. Portugal, Belgium, the Netherlands, and Germany are particularly affected by sprawl. We observe a North-South divide in Italy, with some Southern cities being relatively compact and many Northern Italian cities being among the most sprawled in Europe. Romania, Greece, and Spain have the fastest routes among the countries represented by at least five cities.

**Table 3**
Pearson correlation between sprawl indices and $\overline{C}(\mathcal{G}_{ATc})$, graphs with $c = 50$.

| | bicycle | | | | car | | | |
|---|---|---|---|---|---|---|---|---|
| M. | corr | p.value | int.min | int.max | corr | p.value | int.min | int.max |
| $L$ | 0.514 | $<$ **0.001** | 0.389 | 0.621 | 0.710 | $<$ **0.001** | 0.622 | 0.780 |
| $G$ | −0.607 | $<$ **0.001** | −0.698 | −0.497 | −0.749 | $<$ **0.001** | −0.811 | −0.671 |
| $\eta'$ | **0.689** | $<$ **0.001** | 0.597 | 0.764 | **0.823** | $<$ **0.001** | 0.764 | 0.868 |
| $I$ | 0.336 | $<$ **0.001** | 0.189 | 0.468 | 0.233 | **0.003** | 0.078 | 0.376 |
| $B$ | 0.483 | $<$ **0.001** | 0.352 | 0.595 | 0.655 | $<$ **0.001** | 0.555 | 0.736 |

**Table 4**
Pearson correlation between sprawl indices and the average optimal ATSP route duration, for subsets of cities

| | cities | M. | bicycle corr | bicycle p.value | bicycle int.min | bicycle int.max | car corr | car p.value | car int.min | car int.max |
|---|---|---|---|---|---|---|---|---|---|---|
| 50 | no sea | $L$ | 0.577 | < **0.001** | 0.435 | 0.691 | 0.772 | < **0.001** | 0.682 | 0.839 |
| | | $G$ | −0.662 | < **0.001** | −0.757 | −0.540 | −0.793 | < **0.001** | −0.854 | −0.710 |
| | | $\eta'$ | **0.754** | < **0.001** | 0.658 | 0.825 | **0.872** | < **0.001** | 0.817 | 0.911 |
| | | $I$ | 0.433 | < **0.001** | 0.265 | 0.575 | 0.294 | **0.002** | 0.110 | 0.458 |
| | | $B$ | 0.558 | < **0.001** | 0.412 | 0.677 | 0.743 | < **0.001** | 0.644 | 0.818 |
| 10 | all | $L$ | 0.404 | < **0.001** | 0.263 | 0.527 | 0.533 | < **0.001** | 0.410 | 0.637 |
| | | $G$ | −0.504 | < **0.001** | −0.613 | −0.377 | −0.560 | < **0.001** | −0.659 | −0.442 |
| | | $\eta'$ | **0.562** | < **0.001** | 0.445 | 0.661 | **0.647** | < **0.001** | 0.545 | 0.730 |
| | | $I$ | 0.469 | < **0.001** | 0.336 | 0.583 | 0.421 | < **0.001** | 0.283 | 0.542 |
| | | $B$ | 0.366 | < **0.001** | 0.222 | 0.495 | 0.526 | < **0.001** | 0.402 | 0.631 |

**Table 5**
Pearson correlation between sprawl indices and $\overline{d}$.

| M. | bicycle corr | bicycle p.value | bicycle int.min | bicycle int.max | car corr | car p.value | car int.min | car int.max |
|---|---|---|---|---|---|---|---|---|
| $L$ | 0.405 | < **0.001** | 0.265 | 0.529 | 0.525 | < **0.001** | 0.401 | 0.630 |
| $G$ | −0.519 | < **0.001** | −0.625 | −0.394 | −0.576 | < **0.001** | −0.672 | −0.460 |
| $\eta'$ | **0.579** | < **0.001** | 0.464 | 0.675 | **0.636** | < **0.001** | 0.533 | 0.721 |
| $I$ | 0.538 | < **0.001** | 0.416 | 0.641 | 0.491 | < **0.001** | 0.361 | 0.601 |
| $B$ | 0.395 | < **0.001** | 0.253 | 0.520 | 0.492 | < **0.001** | 0.363 | 0.602 |

**Table 6**
Results obtained by the linear models that predict $\overline{d}$.

| | M. | $R^2$ | adj-$R^2$ | $F$ | $\beta_0$ | $\beta 1$ | Residuals min | Residuals Q1 | Residuals median | Residuals Q3 | Residuals max |
|---|---|---|---|---|---|---|---|---|---|---|---|
| car | $L$ | 0.276 | 0.271 | 58.63 | 1105 | 802 | −570 | −206 | −36 | 167 | 1142 |
| | $G$ | 0.331 | 0.327 | 76.3 | 3924 | −2678 | −629 | −205 | −45 | 151 | 1102 |
| | $\eta'$ | **0.405** | **0.401** | 104.88 | −117 | 2323 | −433 | −210 | −51 | 141 | 1188 |
| | $I$ | 0.241 | 0.236 | 48.8 | 892 | 3422 | −964 | −175 | 23 | 221 | 735 |
| | $B$ | 0.242 | 0.237 | 49.13 | 635 | 1238 | −645 | −225 | −34 | 188 | 1118 |
| bicycle | $L$ | 0.164 | 0.159 | 30.26 | 4241 | 2473 | −2786 | −993 | 54 | 893 | 3252 |
| | $G$ | 0.270 | 0.265 | 56.85 | 14137 | −9647 | −2804 | −903 | −73 | 878 | 3139 |
| | $\eta'$ | **0.335** | **0.331** | 77.7 | −473 | 8438 | −2330 | −913 | −99 | 847 | 3211 |
| | $I$ | 0.289 | 0.285 | 62.75 | 2661 | 14989 | −4022 | −634 | 169 | 685 | 3540 |
| | $B$ | 0.156 | 0.150 | 28.45 | 2670 | 3969 | −2653 | −1064 | 30 | 917 | 2851 |

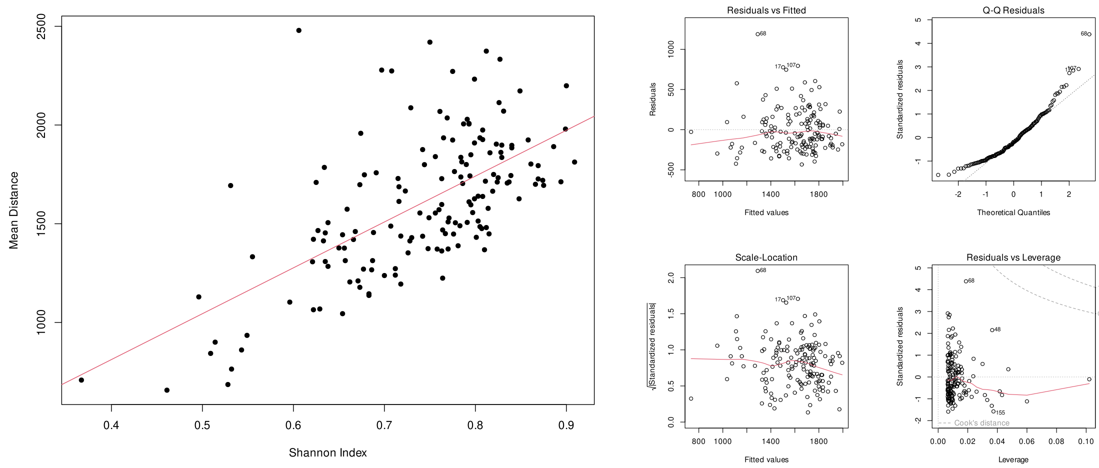
Out of the ten cities with the fastest routes, eight are in Spain, the remaining in Greece and Romania, and out of the ten cities with the lowest Shannon entropy, nine are in Spain and one in Southern Italy. Many Spanish cities, especially those in the country's interior, remain exceptionally compact compared to their European peers. Their compactness is evident in the density heatmaps (Figs. A.10 to A.14) in Appendix A (look, for instance, at Albacete, Burgos, Zaragoza, Vitoria-Gasteiz). Jaeger et al. (2010b), define urban sprawl as "a phenomenon that can be visually perceived in the landscape", and indeed, the attentive beholder venturing at the heart of the Iberian peninsula, will notice at sight a compactness of cities and enjoy large open spaces without traces of human settlements as it is barely possible elsewhere in Western Europe. Fig. 3b clarifies the compactness of Spanish cities, the sprawl of German ones, and the North-South divide in Italy. Among the whole sample, the ten cities with the longest routes are in Germany (3), Italy (3), Portugal (2), Austria (1) and Belgium (1), while the cities with the highest Shannon entropy are in Belgium (3), Germany (2), Italy (2), France (1), United Kingdom (1) and Portugal (1).

The reasons for these differences in Europe are a mix of historical, geographical, cultural, and socioeconomic factors. From a historical point of view, Mediterranean Europe has a prolonged urban tradition that traces back to the Greek and Roman settlements. Geographically speaking, Southern Europe is more mountainous than Northern and Central Europe. Barriers like the sea and the mountains and the lack of extensive fertile plains forced Mediterranean cities to develop more compactly. This contrast is evident in Italy, where the Northern Po Valley, the country's largest and most fertile plain, is also among the most sprawled areas in Europe. Recent socioeconomic reasons might be even more prominent, considering that most European cities experienced sustained population and surface growth in the last two centuries. Dissimilar economic prosperity and cultural perspectives led to diverging car ownership rates, access to detached housing, infrastructure development, and investments in public transportation across European countries. Additionally, Eastern Europe experienced state-led urban planning, while the development in Western Europe was mainly guided by private initiative, with a few exceptions (e.g., Vienna).

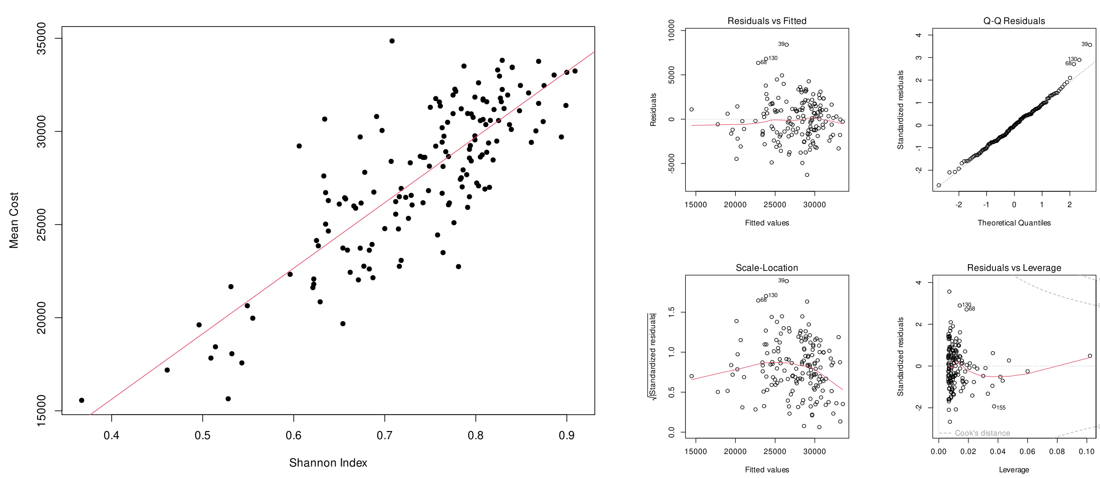
As a result of these factors, we observe varying degrees of housing compactness. We use data from Eurostat about the percentage of inhabitants living in buildings with more than ten dwellings, which we

**Table 7**
Results obtained by the linear models that predict $\overline{C}(\mathscr{G}_{ATc})$.

| $c$ | T. | M. | $R^2$ | adj-$R^2$ | $F$ | $\beta_0$ | $\beta 1$ | Residuals min | Q1 | median | Q3 | max |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 50 | car | $L$ | 0.503 | 0.500 | 156.13 | 19717 | 12709 | −7638 | −1859 | 162 | 2053 | 9968 |
| | | $G$ | 0.561 | 0.558 | 196.84 | 63007 | −40847 | −7247 | −1733 | 66 | 1874 | 9307 |
| | | $\eta'$ | **0.677** | **0.675** | 322.31 | 1550 | 35182 | −6283 | −1642 | −16 | 1444 | 8399 |
| | | $I$ | 0.054 | 0.048 | 8.8 | 23647 | 19013 | −12153 | −2085 | 907 | 3002 | 6761 |
| | | $B$ | 0.429 | 0.425 | 115.57 | 12492 | 19319 | −8364 | −2119 | 48 | 2056 | 9828 |
| | bicycle | $L$ | 0.265 | 0.260 | 55.41 | 64600 | 23934 | −26554 | −5779 | 38 | 5602 | 23818 |
| | | $G$ | 0.369 | 0.364 | 89.88 | 153992 | −85998 | −27599 | −5676 | −398 | 5164 | 23028 |
| | | $\eta'$ | **0.475** | **0.472** | 139.34 | 22750 | 76572 | −18241 | −4376 | −888 | 4060 | 23930 |
| | | $I$ | 0.113 | 0.107 | 19.6 | 64615 | 71375 | −38134 | −4398 | 1900 | 7463 | 24261 |
| | | $B$ | 0.233 | 0.228 | 46.78 | 50515 | 36994 | −21532 | −6363 | −429 | 5487 | 23635 |
| 10 | car | $L$ | 0.284 | 0.279 | 61.1 | 7831 | 4741 | −4503 | −880 | −70 | 982 | 6011 |
| | | $G$ | 0.313 | 0.309 | 70.29 | 23915 | −15163 | −4931 | −1053 | −136 | 1143 | 5761 |
| | | $\eta'$ | **0.418** | **0.415** | 110.81 | 598 | 13741 | −3124 | −1052 | −109 | 853 | 5445 |
| | | $I$ | 0.177 | 0.172 | 33.18 | 7220 | 17099 | −6812 | −1122 | 301 | 1331 | 3406 |
| | | $B$ | 0.276 | 0.272 | 58.81 | 4748 | 7703 | −4157 | −1091 | 57 | 940 | 6024 |
| | bicycle | $L$ | 0.163 | 0.157 | 29.96 | 28112 | 12424 | −15994 | −4020 | 141 | 4241 | 17214 |
| | | $G$ | 0.254 | 0.249 | 52.44 | 76763 | −47236 | −16230 | −3568 | −180 | 3690 | 19199 |
| | | $\eta'$ | **0.316** | **0.312** | 71.27 | 5205 | 41345 | −12539 | −3464 | −302 | 3239 | 18494 |
| | | $I$ | 0.220 | 0.215 | 43.35 | 22134 | 65868 | −22158 | −3176 | 810 | 4058 | 20616 |
| | | $B$ | 0.134 | 0.128 | 23.81 | 21305 | 18557 | −15239 | −4231 | 528 | 3938 | 16879 |




(a) $\bar{d}$.




(b) $\bar{C}(\mathcal{G})$.

**Fig. 6.** Linear model for the prediction of $\overline{d}$ and $\overline{C}(\mathscr{G}_{ATc})$ based on the Shannon Entropy $\eta'$.

**Table 8**
Countries in the sample, from left to right: country code, name, number of cities in the sample, average population of the cities in the sample, route lengths: min, median and max among all the generated road graphs, the percentage of inhabitants living in buildings with more than ten dwellings (% compact housing).

| Country | | Cities | | Route ($h$) | | | % Compact |
|---|---|---|---|---|---|---|---|
| Code | Name | # | Avg Pop. | Min | Median | Max | housing |
| AT | Austria | 5 | 886551 | 5.66 | 8.54 | 10.47 | 27.4 |
| BE | Belgium | 4 | 872799 | 7.09 | 8.59 | 9.96 | 7.3 |
| BG | Bulgaria | 3 | 832931 | 5.03 | 7.48 | 9.42 | 39.4 |
| CH | Switzerland | 2 | 1565620 | 7.56 | 8.80 | 10.42 | 25.0 |
| CY | Cyprus | 1 | 468500 | 6.53 | 6.75 | 7.61 | 8.4 |
| CZ | Czech Republic | 3 | 992805 | 6.68 | 8.22 | 9.44 | 38.1 |
| DE | Germany | 18 | 1570627 | 6.25 | 8.38 | 10.53 | 16.7 |
| DK | Denmark | 4 | 776939 | 5.99 | 7.17 | 8.87 | 26.3 |
| EE | Estonia | 1 | 573041 | 5.27 | 6.20 | 7.96 | 53.6 |
| EL | Greece | 5 | 1095925 | 3.56 | 6.61 | 9.56 | 23.2 |
| ES | Spain | 22 | 1162152 | 3.20 | 6.19 | 9.18 | 44.8 |
| FI | Finland | 3 | 654918 | 5.74 | 7.26 | 8.30 | 32.7 |
| FR | France | 13 | 910385 | 5.40 | 8.03 | 11.00 | 23.6 |
| HR | Croatia | 2 | 736749 | 5.69 | 7.92 | 10.88 | 15.7 |
| HU | Hungary | 2 | 1465965 | 6.58 | 8.19 | 8.95 | 23.0 |
| IE | Ireland | 3 | 739778 | 6.31 | 7.68 | 9.39 | 4.2 |
| IT | Italy | 18 | 1363203 | 5.38 | 8.13 | 13.02 | 25.5 |
| LT | Lithuania | 2 | 540661 | 4.65 | 6.50 | 8.12 | 51.7 |
| LU | Luxembourg | 1 | 737146 | 6.86 | 7.98 | 8.68 | 13.5 |
| LV | Latvia | 1 | 855892 | 6.87 | 7.81 | 8.69 | 58.7 |
| NL | Netherlands | 5 | 1794434 | 6.99 | 8.47 | 10.45 | 14.2 |
| PL | Poland | 7 | 989698 | 6.37 | 8.34 | 10.27 | 34.7 |
| PT | Portugal | 7 | 899932 | 5.76 | 8.66 | 11.30 | 19.6 |
| RO | Romania | 9 | 703521 | 3.88 | 6.67 | 9.23 | 29.9 |
| SE | Sweden | 3 | 1100291 | 6.32 | 7.77 | 10.55 | 35.2 |
| SI | Slovenia | 1 | 481233 | 6.76 | 8.29 | 9.45 | 19.5 |
| SK | Slovakia | 1 | 653203 | 7.48 | 9.09 | 10.30 | 40.8 |
| UK | United Kingdom | 10 | 1381130 | 5.58 | 7.97 | 9.81 | 6.3 |

define as *compact housing*[4]. Table 8 shows the share of inhabitants living in compact housing, together with a summary of the results for the countries in the sample. We see that high rates of compact housing ($>30\%$) are registered in the Baltic republics (Estonia, Latvia, Lithuania), Spain, Finland, Sweden, and Eastern Europe. Many cities from these countries registered low routing times. Exiguous rates of compact housing are recorded in the UK, Ireland, and Belgium. In order to visualize these data on the map, we extend Fig. 4 into Fig. 7 by assigning a gradient color to the countries according to their level of compact housing. Fig. 8 shows examples of compact housing.

### 5.2. *Winners and losers of sprawl*

The correlation between total population and routing times is 0.265. While this fact does not affect the results, which are based on relative density, it does allow us to derive some considerations.

In Fig. 9, we show the distribution of the cities in the sample according to their average optimal route length $\overline{C}(\mathscr{G}_{ATc})$ and their population. The population is on a logarithmic scale and we have normalized the variables in the interval $[0,1]$. Looking at Fig. 9a, we identify four quadrants. Clockwise, starting from the bottom left, we have (i) cities with low population and low sprawl, (ii) cities with low population and high sprawl, (iii) cities with high population and high sprawl, and (iv) cities with high population and low sprawl.

Cities in quadrants (i) and (iii) are where we expect them to be: the city sprawls as the population grows. Cities in quadrant (iv) are the *winners* of this sample. They remained route-efficient despite their high population. The list comprises, among others, Madrid, Athens, Bucharest, Vienna, Palermo, Valencia, Zaragoza, Barcelona and Seville. On the other hand, quadrant (ii) includes cities with high sprawl despite their low population, which are the *losers* in the sprawl challenge: inhabitants face longer distances and longer routes than in much larger cities, without enjoying the benefits that a higher population brings in terms of mobility. Here, we find cities like Coimbra, Klagenfurt, Graz, Udine, Piacenza, Perugia, and Jena. The names of some of these cities are shown in Fig. 9b.

### 5.3. *Other variables of interest*

Other control variables, such as infrastructure development and topology, might affect sprawl and routing. However, the choice of 156 cities across Europe mitigates their influence.

In contrast with many studies comparing cities based on their administrative borders, we have chosen to compare urban areas of the same size. Using administrative borders produces unfair comparisons and can lead to entirely distorted conclusions. On the contrary, if the size is the same, the deviation in the route lengths is attributable solely to the sprawl of the population. It would be possible, in principle, to compare cities with different sizes and to normalize the distances in the $[0,1]$ range at the price of losing the information of the actual expected routing time in the real city. Therefore, we decided to use a single size for all identical cities.

Our model does not consider parking times, which may negatively affect the performance of cars. Conversely, a possible reason for bicycle efficiency in compact cities is that routes rarely include points distant from the city center.

## 6. Conclusions and future work

This study highlights the underestimated importance of addressing urban sprawl to reduce transportation's economic, social, and environmental costs. Compact cities facilitate efficient transportation. Conversely, if cities continue sprawling, no matter how good we become at solving complex optimization routing problems, their spatial distribution will prevent them from achieving sustainable and efficient urban transportation.

These conclusions are rooted in a novel methodology that integrates

[4] EUROSTAT, dataset ILC_LVHO01, year 2018.

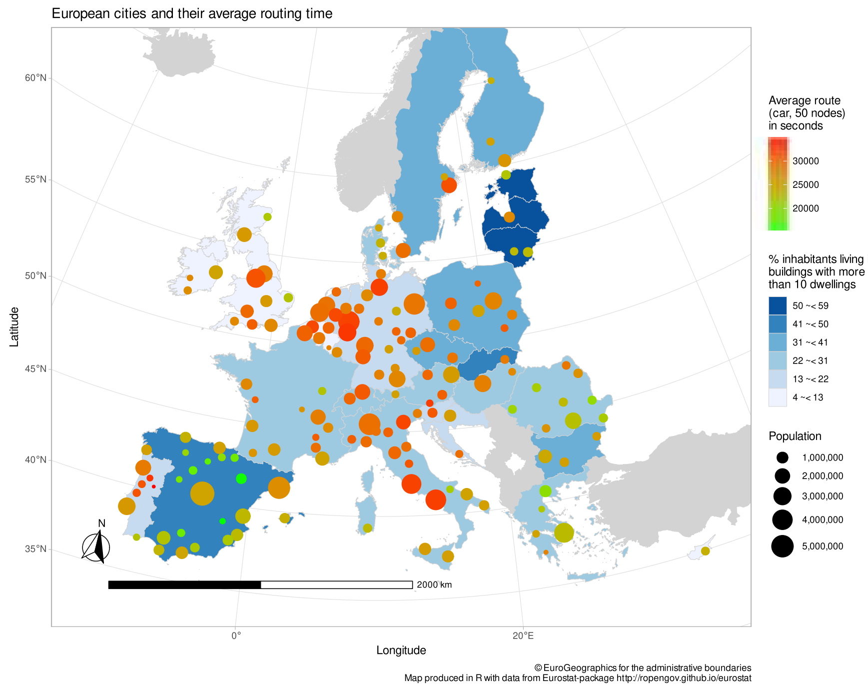


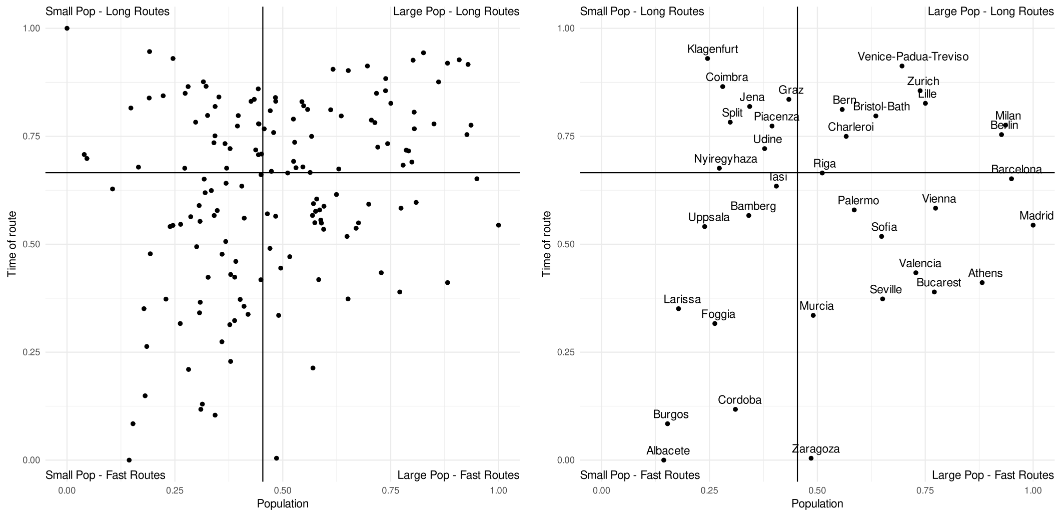

**Fig. 7.** The same information of Fig. 4, enriched with the indication of the percentage of inhabitants living in dense housing (defined as buildings with more than ten dwellings).

**Fig. 8.** Compact urban development in cities from the sample: Vienna, Austria (top left), Naples, Italy (top right), Zaragoza, Spain (bottom left), and Bilbao, Spain (bottom right). Photos by the author.

knowledge from operations research into urban studies, applied to 6240 realistic graphs from 156 European cities, which offer great spatial heterogeneity. We found a significant correlation between all indicators of sprawl (land usage, Gini coefficient, Shannon entropy, Moran I index, and Bribiesca index) and travel distances and routing times for both cars and bicycles. Shannon entropy is the best predictor of route efficiency. Therefore, it is a suitable indicator when the cities have the same size.

Many geographical insights also emerge from the study. Due to their compactness, many Spanish cities are the least sprawled in the sample and the most efficient when it comes to routing. Mediterranean and Eastern Europe host the majority of compact cities, while many Western and Central European countries are on the high end of sprawl.

In future work, we will extend the sample by including cities from other continents to validate our results on a global scale. We also plan to include more specific indicators, such as Weighted Urban Proliferation, proposed by Jaeger and Schwick (2014). The challenge is the




(a) Population vs Routes (b) Population vs Routes (Selected Cities)

**Fig. 9.** Distribution of cities by population and routes.

availability and reliability of data across several countries.

Regarding the routing model, we plan to consider parking to enhance realism. We will also consider public transportation and the possibility of choosing between multiple modes en route. Additionally, we intend to investigate the relationship between population distribution and the effectiveness and efficiency of railway networks.

Finally, as an ambitious long-term plan, we aim to establish a new stream of research that combines operations research and urban studies, providing a comprehensive framework to guide optimal decisions in urban planning.

**Declaration of Competing Interest**

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

The author acknowledges the CINECA IA4EVRP HP10CE285L and MATHLSCP HP10CDK4DW awards under the ISCRA initiative, for the availability of high-performance computing resources and support.

Furthermore, the author acknowledges the European Commission (Eurostat, Joint Research Centre and DG Regional Policy - REGIO-GIS) for providing the GEOSTAT 1 km$^2$ population grid.

The constructive and detailed feedback by three anonymous reviewers is gratefully acknowledged.

Finally, the author thanks Hải Yến Lưu and Eugenio Macor for the fruitful discussions.

## Appendix A. City data

We included 156 areas of size 2809 km$^2$ in the study, listed in Tables A.9 to A.11. Each area takes the name of a city (or the cities) that lie at the center of the circle or is most representative of the area, but other important settlements might lie within the 30 km. Besides the countries and the name of the city, we report the population, the density, the lived density, which is the density that is computed only on lived cells, and the maximum density, the five indicators of sprawl, the average distances by bicycle and car, and the optimal routes by bicycle and car for 10 and 50 points. The density heatmaps for all cities are also shown in Figs. A.10 to A.14. For all cities the density goes from white (minimum) to red (maximum in the city), and unpopulated cells are in gray. All figures have a radius of 30 km and are directed with the north up.

All EU countries are represented except Malta (which surface is smaller than 2809 km$^2$), plus Switzerland and the United Kingdom, for a total of 28 countries. The total area involved in the study is 438204 km$^2$.

The density of the area in the study is $398 inh./km^2$, higher than the density of the EU. The densest cell reported in this study lies in Barcelona, with a density of 52767. Very dense cells (over 40000) are also found in Madrid (44006), Zaragoza (42833), Bilbao (41389), and Oviedo-Gijón (40083). Despite their high total population, the densest cells in Duisburg-Essen-Dortmund and Manchester are just 15030 and 13313, respectively. The area Libramont-Chevigny has the lowest maximum density, $2357 inh./km^2$.

The most populated urban area is Madrid, with a population of 5940359, while the least populated one is Covilhã-Fundã, with 112731 inhabitants. We point out that we did not intentionally include the extensive urban areas of London and Paris, as they are of a different order of magnitude in terms of population and extension.

The areas with the highest percentage of cells with human presence ($L$) are Florence (IT, 91.5%), Milan (IT, 92.3%), Nantes (FR, 92.5%), Hengelo (NL, 93.2%), Venice-Padua-Treviso (IT, 93.3%), Duisburg-Essen-Dortmund (DE, 94.7%), Linz (AT, 95.1%), Lille (FR, 96.2%), Ghent (BE, 96.5%) and Toulouse (FR, 96.7%). Cities with the least human human cell occupation are Albacete (ES, 3.5%), Zaragoza (ES, 12.2%), Córdoba (ES, 13.6%), Valladolid (ES, 14.6%), Burgos (ES, 15.2%), Galaţi-Brăila (RO, 18.3%), Salamanca (ES, 18.4%), Vitoria-Gasteiz (ES, 19.4%), Granada (ES, 20.5%) and Cádiz-Jerez (ES, 20.7%)

If we look at Shannon entropy, the lowest sprawl is registered in Albacete, Burgos, Vitoria-Gasteiz, Córdoba, Salamanca, Zaragoza (ES), Foggia (IT), Pamplona, Valladolid, León (ES), with values spanning from 0.367 (Albacete) to 0.549 (León). In contrast, the highest sprawl, in decreasing order, is recorded in Duisburg-Essen-Dortmund (DE), Venice-Padua-Treviso (IT), Beringen (BE), Porto (PT), Ghent (BE), Manchester (UK), Milan (IT), Cologne-Bonn (DE), Lille (FR), Charleroi (BE), with values spanning from 0.909 (Duisburg-Essen-Dortmund) to 0.866 (Charleroi).

According to $\overline{C}(\mathscr{G}_{ATc})$ for the case $c = 50$ and $T = car$, the ten areas with the fastest routes are, in order, Albacete, Zaragoza, Burgos, Valladolid, Córdoba, Pamplona, Salamanca, Vitoria-Gasteiz (ES), Thessaloniki (EL) and Galaţi-Brăila (RO), with values spanning from 15563*s* for Albacete to 19973*s* for Galaţi-Brăila. Conversely, the cities with the longest routes are, in decreasing order, Covilhã-Fundão, Viseu (PT), Cologne-Bonn (DE), Klagenfurt (AT), Naples (IT), Hamburg (DE), Rome (IT), Duisburg-Essen-Dortmund (DE), Venice-Padua-Treviso (IT) and Ghent (BE), with values spanning from 34857*s* (Covilhã-Fundão) to 33027*s* (Ghent).

Data comes from the GEOSTAT-JRC database. EUROSTAT also provides information on data quality, which varies by country or by region. Belgium, Switzerland, Finland, Lithuania, Norway, and Slovenia are classified as “very high”, meaning data has a sub-kilometer resolution. Ireland, Denmark, Sweden, Estonia, Latvia, and the Netherlands are classified as “high”, meaning data has a resolution of one kilometer. The quality of data for Spain is classified from “medium” to “very high” (Andalucia and Canary Islands). Most remaining countries are classified as “medium”, meaning data was available at a lower resolution than 1 km$^2$. Finally, very few countries are classified as “low”. Most of them are extra-EU and outside our study. In the case of regions with “high” or “very high” quality, data collection methodology is unlikely to have influenced our results. In the case of regions with “medium” and “low” quality, data is interpolated, which might have introduced an approximation.

Table A.9: List of cities included in the study, part I

| | Area | | Density | | | Sprawl Indicators | | | | | $\overline{d}$ (*s*) | | $\overline{C}(\mathscr{G}_{ATc})$ (*s*) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | City | Pop. | Density | Lived | Max | *L* | *G* | $\eta'$ | *I* | *B* | bic. | car | b10 | b50 | c10 | c50 |
| AT | Graz | 629669 | 224 | 246 | 12074 | 91.0 | 0.780 | 0.808 | 0.123 | 0.949 | 5049 | 1638 | 35343 | 79893 | 10480 | 31682 |
| AT | Innsbruck | 351261 | 125 | 486 | 11968 | 25.7 | 0.947 | 0.656 | 0.220 | 0.738 | 4999 | 1376 | 30484 | 67269 | 11490 | 26438 |
| AT | Klagenfurt | 298412 | 106 | 146 | 6641 | 72.6 | 0.829 | 0.787 | 0.207 | 0.866 | 6516 | 2006 | 40373 | 91194 | 12255 | 33508 |
| AT | Linz | 729382 | 260 | 273 | 9835 | 95.1 | 0.782 | 0.820 | 0.208 | 0.961 | 6340 | 1750 | 36645 | 86060 | 12083 | 31175 |
| AT | Vienna | 2424033 | 863 | 1409 | 31813 | 61.2 | 0.892 | 0.748 | 0.116 | 0.771 | 3938 | 1374 | 24176 | 69262 | 8925 | 26820 |
| BE | Beringen | 991314 | 353 | 393 | 5480 | 89.8 | 0.661 | 0.899 | 0.291 | 0.942 | 7593 | 1980 | 41548 | 90418 | 12672 | 31394 |
| BE | Charleroi | 1066356 | 380 | 430 | 6079 | 88.2 | 0.738 | 0.866 | 0.220 | 0.903 | 6597 | 1700 | 36228 | 83352 | 11090 | 30028 |
| BE | Ghent | 1298194 | 462 | 479 | 10947 | 96.5 | 0.680 | 0.886 | 0.197 | 0.972 | 6920 | 1891 | 42821 | 87782 | 12264 | 33027 |
| BE | Libramont-Chevigny | 135335 | 48 | 98 | 2357 | 48.9 | 0.847 | 0.793 | 0.360 | 0.655 | 8044 | 2005 | 43387 | 90652 | 12294 | 29039 |
| BG | Plovdiv | 574258 | 204 | 585 | 19471 | 34.9 | 0.945 | 0.657 | 0.187 | 0.664 | 4310 | 1313 | 31217 | 66940 | 9638 | 26375 |
| BG | Sofia | 1476407 | 526 | 1208 | 16432 | 43.5 | 0.924 | 0.712 | 0.148 | 0.732 | 3596 | 1240 | 22790 | 59029 | 8846 | 25558 |
| BG | Varna | 448130 | 160 | 442 | 13677 | 36.1 | 0.955 | 0.635 | 0.183 | 0.622 | 5868 | 1454 | 38167 | 71578 | 10026 | 26715 |
| CH | Bern | 1028000 | 366 | 408 | 10861 | 89.7 | 0.789 | 0.831 | 0.239 | 0.942 | 7538 | 2070 | 45834 | 86490 | 12337 | 31230 |
| CH | Zürich | 2103241 | 749 | 833 | 14962 | 89.9 | 0.758 | 0.858 | 0.243 | 0.920 | 6837 | 1924 | 42311 | 89730 | 12189 | 32068 |
| CY | Larnaca | 468500 | 167 | 462 | 6175 | 36.1 | 0.923 | 0.715 | 0.139 | 0.791 | 7080 | 1729 | 37563 | 71868 | 11500 | 24764 |
| CZ | Brno | 752231 | 268 | 468 | 11248 | 57.2 | 0.875 | 0.763 | 0.231 | 0.676 | 5392 | 1597 | 35728 | 84282 | 11815 | 30198 |
| CZ | Pilsen | 379535 | 135 | 243 | 12797 | 55.7 | 0.901 | 0.718 | 0.279 | 0.648 | 5247 | 1438 | 35116 | 78866 | 10344 | 26938 |
| CZ | Prague | 1846651 | 657 | 833 | 25868 | 78.9 | 0.837 | 0.797 | 0.213 | 0.839 | 5201 | 1557 | 37451 | 81581 | 10734 | 30755 |
| DE | Bamberg | 435726 | 155 | 276 | 6736 | 56.2 | 0.846 | 0.793 | 0.258 | 0.665 | 6825 | 1611 | 40582 | 90998 | 10474 | 26493 |
| DE | Berlin | 4442380 | 1581 | 2153 | 26155 | 73.4 | 0.793 | 0.839 | 0.137 | 0.850 | 5392 | 1744 | 34177 | 81955 | 11847 | 30107 |
| DE | Bremen | 1052268 | 375 | 474 | 13548 | 79.0 | 0.847 | 0.795 | 0.194 | 0.866 | 5600 | 1596 | 31558 | 75232 | 10471 | 28416 |
| DE | Chemnitz | 766866 | 273 | 357 | 10064 | 76.5 | 0.793 | 0.828 | 0.241 | 0.813 | 6872 | 1861 | 40149 | 90174 | 13295 | 31596 |
| DE | Duisburg-Essen-Dortmund | 4488029 | 1598 | 1688 | 15030 | 94.7 | 0.635 | 0.909 | 0.250 | 0.966 | 7656 | 1812 | 45255 | 93743 | 11882 | 33242 |
| DE | Frankfurt am Mein | 2734674 | 974 | 1409 | 21317 | 69.1 | 0.775 | 0.848 | 0.250 | 0.815 | 6701 | 1626 | 43297 | 92466 | 11731 | 31111 |
| DE | Göttingen | 434706 | 155 | 337 | 6995 | 45.9 | 0.877 | 0.765 | 0.257 | 0.608 | 7083 | 1935 | 45034 | 91135 | 12270 | 29743 |
| DE | Hamburg | 2708264 | 964 | 1180 | 21384 | 81.7 | 0.789 | 0.840 | 0.159 | 0.878 | 5818 | 1884 | 37046 | 86215 | 11649 | 33432 |
| DE | Hof-Plauen | 410031 | 146 | 217 | 7880 | 67.2 | 0.843 | 0.791 | 0.271 | 0.763 | 8111 | 2029 | 45136 | 92564 | 13424 | 30962 |
| DE | Ingolstadt | 485841 | 173 | 300 | 11020 | 57.7 | 0.854 | 0.786 | 0.258 | 0.695 | 6355 | 1704 | 40151 | 84117 | 10303 | 27934 |
| DE | Jena | 439404 | 156 | 283 | 7174 | 55.3 | 0.868 | 0.761 | 0.260 | 0.654 | 7436 | 2069 | 45771 | 91231 | 11804 | 31364 |
| DE | Kiel | 638130 | 227 | 311 | 12589 | 73.1 | 0.873 | 0.763 | 0.242 | 0.808 | 6570 | 1728 | 42740 | 85010 | 12000 | 29423 |
| DE | Cologne-Bonn | 2980915 | 1061 | 1273 | 17790 | 83.3 | 0.735 | 0.869 | 0.220 | 0.889 | 6910 | 1727 | 40690 | 91492 | 11016 | 33762 |
| DE | Ludwigshafen-Heidelberg | 1935680 | 689 | 986 | 14313 | 69.9 | 0.799 | 0.835 | 0.289 | 0.813 | 7308 | 1706 | 40992 | 93021 | 11125 | 31952 |
| DE | Magdeburg | 484425 | 172 | 423 | 10745 | 40.8 | 0.909 | 0.726 | 0.225 | 0.585 | 5387 | 1353 | 35688 | 77312 | 11017 | 25332 |
| DE | Munich | 2471285 | 880 | 1205 | 23853 | 73.0 | 0.837 | 0.807 | 0.158 | 0.816 | 5008 | 1476 | 31079 | 78207 | 10441 | 28742 |
| DE | Osnabrück | 689838 | 246 | 269 | 8897 | 91.4 | 0.780 | 0.837 | 0.244 | 0.939 | 6389 | 1712 | 36352 | 83300 | 12341 | 30364 |
| DE | Ulm | 672830 | 240 | 432 | 9033 | 55.4 | 0.850 | 0.794 | 0.277 | 0.683 | 6815 | 1742 | 42223 | 88553 | 11362 | 29240 |
| DK | Aalborg | 298232 | 106 | 134 | 10876 | 79.2 | 0.890 | 0.730 | 0.206 | 0.889 | 5506 | 1430 | 35051 | 75898 | 10887 | 26052 |
| DK | Aarhus | 531476 | 189 | 225 | 15015 | 84.0 | 0.873 | 0.758 | 0.192 | 0.929 | 5172 | 1372 | 35512 | 76722 | 8725 | 24441 |
| DK | Copenhagen | 1907112 | 679 | 909 | 23136 | 74.7 | 0.831 | 0.808 | 0.180 | 0.921 | 8898 | 1975 | 54606 | 101723 | 12202 | 30643 |
| DK | Odense | 370936 | 132 | 158 | 6707 | 83.6 | 0.845 | 0.776 | 0.168 | 0.949 | 5083 | 1448 | 35668 | 73618 | 10234 | 25099 |
| EE | Tallin | 573041 | 204 | 281 | 16861 | 72.6 | 0.938 | 0.662 | 0.211 | 0.868 | 3638 | 1205 | 24062 | 61093 | 7655 | 22436 |
| EL | Athens | 3719119 | 1324 | 2060 | 28622 | 64.3 | 0.885 | 0.764 | 0.078 | 0.873 | 4144 | 1224 | 24253 | 67264 | 8109 | 23494 |
| EL | Kalamata-Sparta | 132094 | 47 | 163 | 12457 | 28.9 | 0.951 | 0.606 | 0.249 | 0.587 | 7852 | 2479 | 43018 | 81110 | 13690 | 29217 |
| EL | Larissa | 228574 | 81 | 207 | 12599 | 39.4 | 0.958 | 0.596 | 0.163 | 0.647 | 3672 | 1103 | 29004 | 66911 | 8791 | 22329 |
| EL | Patras | 320131 | 114 | 267 | 12522 | 42.8 | 0.940 | 0.650 | 0.173 | 0.673 | 6616 | 1377 | 40163 | 92788 | 9963 | 26100 |
| EL | Thessaloniki | 1079711 | 384 | 643 | 28771 | 59.8 | 0.939 | 0.654 | 0.132 | 0.805 | 3566 | 1045 | 23750 | 61366 | 8461 | 19679 |
| ES | Albacete | 199564 | 71 | 2057 | 38176 | 3.5 | 0.995 | 0.367 | 0.214 | 0.356 | 2428 | 708 | 16674 | 49230 | 5328 | 15563 |
| ES | Alicante-Elche | 1138319 | 405 | 1097 | 39240 | 37.0 | 0.944 | 0.659 | 0.258 | 0.663 | 7147 | 1574 | 35849 | 75997 | 10650 | 23628 |
| ES | Barcelona | 4868660 | 1733 | 3113 | 52767 | 55.7 | 0.895 | 0.749 | 0.155 | 0.754 | 5653 | 1530 | 36399 | 75148 | 9977 | 28134 |

(*continued on next page*)

(*continued*)

| | Area | | Density | | | Sprawl Indicators | | | | | $\bar{d}$ ($s$) | | $\bar{C}(\mathscr{G}_{ATc})$ ($s$) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | City | Pop. | Density | Lived | Max | $L$ | $G$ | $\eta'$ | $I$ | $B$ | bic. | car | b10 | b50 | c10 | c50 |
| ES | Bilbao | 1160264 | 413 | 1006 | 41389 | 41.0 | 0.953 | 0.638 | 0.224 | 0.687 | 4743 | 1284 | 28820 | 75236 | 9279 | 26288 |
| ES | Burgos | 206626 | 74 | 485 | 21257 | 15.2 | 0.985 | 0.461 | 0.213 | 0.318 | 1831 | 657 | 14006 | 41684 | 4049 | 17189 |
| ES | Cádiz-Jerez | 803008 | 286 | 1380 | 25519 | 20.7 | 0.961 | 0.625 | 0.262 | 0.632 | 8005 | 1709 | 46356 | 80675 | 10589 | 24143 |

Table A.10: List of cities included in the study, part II

| | Area | | Density | | | Sprawl Indicators | | | | | $\bar{d}$ ($s$) | | $\bar{C}(\mathscr{G}_{ATc})$ ($s$) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | City | Pop. | Density | Lived | Max | $L$ | $G$ | $\eta'$ | $I$ | $B$ | b50 | c50 | b10 | b50 | c10 | c50 |
| ES | Córdoba | 385415 | 137 | 1006 | 23951 | 13.6 | 0.984 | 0.509 | 0.168 | 0.592 | 3800 | 843 | 27054 | 64720 | 6890 | 17832 |
| ES | Granada | 594751 | 212 | 1033 | 24975 | 20.5 | 0.961 | 0.622 | 0.228 | 0.605 | 3660 | 1064 | 24936 | 62553 | 8090 | 22077 |
| ES | León | 234543 | 84 | 365 | 18801 | 22.9 | 0.967 | 0.549 | 0.188 | 0.474 | 2961 | 935 | 30495 | 57113 | 6185 | 20641 |
| ES | Madrid | 5940359 | 2115 | 5391 | 44006 | 39.2 | 0.881 | 0.770 | 0.177 | 0.707 | 5710 | 1372 | 36988 | 86544 | 8847 | 26062 |
| ES | Malaga | 1191134 | 424 | 808 | 39457 | 52.5 | 0.943 | 0.668 | 0.213 | 0.754 | 6478 | 1461 | 37298 | 81352 | 11489 | 25878 |
| ES | Mallorca | 727258 | 259 | 887 | 30393 | 29.2 | 0.954 | 0.635 | 0.186 | 0.579 | 5010 | 1309 | 31653 | 71997 | 8399 | 25024 |
| ES | Murcia | 787783 | 280 | 1084 | 22777 | 25.9 | 0.942 | 0.671 | 0.236 | 0.606 | 5159 | 1211 | 36326 | 72310 | 8971 | 22030 |
| ES | Oviedo-Gijón | 872589 | 311 | 743 | 40083 | 41.8 | 0.945 | 0.638 | 0.233 | 0.635 | 6684 | 1506 | 36797 | 79740 | 9034 | 24652 |
| ES | Pamplona | 390882 | 139 | 623 | 22937 | 22.3 | 0.978 | 0.532 | 0.203 | 0.407 | 2426 | 764 | 21142 | 53569 | 4784 | 18066 |
| ES | Salamanca | 231019 | 82 | 446 | 25603 | 18.4 | 0.978 | 0.514 | 0.236 | 0.408 | 2604 | 900 | 18955 | 51028 | 5744 | 18438 |
| ES | Seville | 1490259 | 531 | 2169 | 28221 | 24.5 | 0.943 | 0.677 | 0.197 | 0.707 | 4689 | 1270 | 29993 | 71456 | 8975 | 22765 |
| ES | Valencia | 2022195 | 720 | 2053 | 37447 | 35.1 | 0.937 | 0.686 | 0.200 | 0.628 | 4586 | 1267 | 30567 | 72001 | 9294 | 23934 |
| ES | Valladolid | 439393 | 156 | 1074 | 30431 | 14.6 | 0.978 | 0.543 | 0.171 | 0.481 | 3024 | 861 | 21934 | 57684 | 5874 | 17573 |
| ES | Vigo-Pontevedra | 766424 | 273 | 572 | 25755 | 47.7 | 0.886 | 0.723 | 0.229 | 0.736 | 6506 | 1666 | 34828 | 77377 | 12356 | 26460 |
| ES | Vitoria-Gasteiz | 344328 | 123 | 631 | 31788 | 19.4 | 0.982 | 0.496 | 0.169 | 0.353 | 4671 | 1130 | 32923 | 59809 | 8417 | 19616 |
| ES | Zaragoza | 772591 | 275 | 2246 | 42833 | 12.2 | 0.982 | 0.528 | 0.144 | 0.562 | 2248 | 686 | 18295 | 50398 | 4920 | 15648 |
| FI | Helsinki-Espoo | 1339937 | 477 | 681 | 15219 | 70.1 | 0.866 | 0.783 | 0.188 | 0.913 | 5134 | 1490 | 27471 | 70645 | 9910 | 27430 |
| FI | Oulu | 242293 | 86 | 217 | 8515 | 39.8 | 0.927 | 0.700 | 0.197 | 0.774 | 4688 | 1238 | 31953 | 67952 | 9185 | 24782 |
| FI | Tampere | 382525 | 136 | 206 | 10762 | 66.0 | 0.914 | 0.712 | 0.241 | 0.797 | 4119 | 1273 | 24986 | 62960 | 8649 | 26232 |
| FR | Avignon | 654892 | 233 | 284 | 7848 | 82.1 | 0.812 | 0.816 | 0.248 | 0.904 | 6520 | 1859 | 41586 | 82512 | 12649 | 30591 |
| FR | Bordeaux | 1117407 | 398 | 517 | 13679 | 76.9 | 0.832 | 0.801 | 0.112 | 0.894 | 4792 | 1431 | 29665 | 72245 | 10051 | 27230 |
| FR | Dijon | 383338 | 136 | 321 | 10709 | 42.5 | 0.927 | 0.683 | 0.141 | 0.602 | 3938 | 1137 | 24905 | 67996 | 8368 | 22616 |
| FR | Fontenay-Niort | 239872 | 85 | 108 | 5366 | 78.7 | 0.811 | 0.808 | 0.201 | 0.826 | 6526 | 1926 | 37701 | 83426 | 12817 | 31742 |
| FR | Grenoble | 670014 | 239 | 420 | 17071 | 56.8 | 0.896 | 0.728 | 0.138 | 0.803 | 4938 | 1413 | 32266 | 78132 | 11484 | 28316 |
| FR | Lille | 2207594 | 786 | 817 | 14881 | 96.2 | 0.730 | 0.869 | 0.151 | 0.966 | 6315 | 1794 | 39752 | 83232 | 11124 | 31505 |
| FR | Lyon | 1957863 | 697 | 768 | 24904 | 90.7 | 0.827 | 0.799 | 0.118 | 0.933 | 4992 | 1626 | 33464 | 76399 | 10706 | 29547 |
| FR | Marseille | 1640833 | 584 | 926 | 26682 | 63.0 | 0.870 | 0.771 | 0.115 | 0.855 | 6059 | 1529 | 33601 | 81560 | 10773 | 26163 |
| FR | Montelimar | 272736 | 97 | 124 | 4794 | 78.5 | 0.826 | 0.799 | 0.272 | 0.860 | 7843 | 2232 | 49561 | 92693 | 15008 | 31842 |
| FR | Nantes | 923746 | 329 | 356 | 12906 | 92.5 | 0.809 | 0.805 | 0.119 | 0.959 | 4734 | 1485 | 34896 | 78360 | 10372 | 28627 |
| FR | Pau-Tarbes | 401021 | 143 | 179 | 9618 | 79.9 | 0.830 | 0.784 | 0.152 | 0.901 | 6659 | 1836 | 35334 | 77500 | 12146 | 27510 |
| FR | Toulouse | 1194391 | 425 | 440 | 16059 | 96.7 | 0.815 | 0.810 | 0.116 | 0.971 | 4546 | 1368 | 32166 | 74201 | 10178 | 26906 |
| FR | Vichy | 171298 | 61 | 72 | 5628 | 84.3 | 0.810 | 0.790 | 0.223 | 0.864 | 5870 | 1800 | 34748 | 80689 | 10957 | 27678 |
| HR | Split | 367441 | 131 | 320 | 18979 | 40.8 | 0.945 | 0.634 | 0.196 | 0.667 | 5397 | 1785 | 31020 | 81628 | 13921 | 30662 |
| HR | Zagreb | 1106057 | 394 | 558 | 18069 | 70.6 | 0.861 | 0.763 | 0.155 | 0.839 | 4687 | 1362 | 28988 | 75797 | 8270 | 26679 |
| HU | Budapest | 2599452 | 925 | 1552 | 32186 | 59.6 | 0.822 | 0.814 | 0.163 | 0.803 | 4770 | 1578 | 34696 | 74943 | 11160 | 29378 |
| HU | Nyíregyháza | 332478 | 118 | 310 | 10485 | 38.1 | 0.893 | 0.742 | 0.246 | 0.630 | 6789 | 1876 | 40997 | 86590 | 12288 | 28605 |
| IE | Cork | 397437 | 141 | 156 | 7254 | 90.6 | 0.837 | 0.764 | 0.172 | 0.949 | 5371 | 1468 | 39525 | 80838 | 11849 | 28119 |
| IE | Dublin | 1604562 | 571 | 720 | 13082 | 79.4 | 0.855 | 0.791 | 0.125 | 0.934 | 5298 | 1507 | 28422 | 79343 | 9950 | 25924 |
| IE | Limerick | 217337 | 77 | 93 | 5228 | 83.2 | 0.809 | 0.770 | 0.177 | 0.907 | 5282 | 1509 | 34979 | 84406 | 11459 | 28654 |
| IT | Bari | 1167192 | 416 | 863 | 28559 | 48.2 | 0.943 | 0.674 | 0.309 | 0.752 | 7130 | 1958 | 41056 | 77752 | 12627 | 26155 |
| IT | Cagliari | 526055 | 187 | 398 | 12563 | 47.0 | 0.950 | 0.654 | 0.225 | 0.761 | 4906 | 1444 | 28354 | 69698 | 9778 | 23738 |
| IT | Catania | 1075165 | 383 | 737 | 17117 | 51.9 | 0.921 | 0.716 | 0.208 | 0.826 | 5969 | 1687 | 32792 | 81750 | 11356 | 26496 |
| IT | Forlì-Cesena-Rimini | 657477 | 234 | 289 | 8484 | 81.0 | 0.829 | 0.805 | 0.207 | 0.887 | 7156 | 1935 | 40451 | 81705 | 12661 | 30582 |
| IT | Cuneo-Fossano | 455033 | 162 | 183 | 8924 | 88.7 | 0.782 | 0.827 | 0.312 | 0.906 | 7790 | 2333 | 43453 | 86855 | 14364 | 31788 |
| IT | Florence | 1270537 | 452 | 495 | 16742 | 91.5 | 0.854 | 0.784 | 0.183 | 0.952 | 5723 | 1737 | 36881 | 85107 | 12807 | 31216 |
| IT | Foggia | 319019 | 114 | 291 | 19921 | 39.1 | 0.979 | 0.531 | 0.231 | 0.625 | 5785 | 1693 | 37135 | 61151 | 10490 | 21666 |
| IT | Genoa | 902277 | 321 | 536 | 25504 | 59.9 | 0.925 | 0.691 | 0.213 | 0.759 | 6555 | 1758 | 52269 | 90840 | 13427 | 30800 |
| IT | Lecce | 710333 | 253 | 460 | 10813 | 55.0 | 0.912 | 0.729 | 0.311 | 0.775 | 7037 | 2087 | 39275 | 75570 | 12109 | 26568 |
| IT | Milan | 4611600 | 1642 | 1778 | 25608 | 92.3 | 0.721 | 0.874 | 0.172 | 0.942 | 5919 | 1720 | 35092 | 88433 | 10565 | 30528 |
| IT | Naples | 4138032 | 1473 | 1876 | 31458 | 78.5 | 0.789 | 0.840 | 0.190 | 0.910 | 6206 | 1897 | 38816 | 83690 | 12782 | 33449 |
| IT | Palermo | 1148158 | 409 | 894 | 23289 | 45.7 | 0.936 | 0.688 | 0.157 | 0.765 | 5150 | 1455 | 34743 | 71173 | 9653 | 26741 |
| IT | Perugia | 404417 | 144 | 190 | 7655 | 75.8 | 0.860 | 0.777 | 0.286 | 0.833 | 6493 | 1764 | 42583 | 89187 | 12335 | 32265 |

Table A.11: List of cities included in the study, part III

| Area | | | Density | | | Sprawl Indicators | | | | | $\overline{d}$ $(s)$ | | $\overline{C}(\mathscr{G}_{ATc})$ $(s)$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | City | Pop. | Density | Lived | Max | $L$ | $G$ | $\eta'$ | $I$ | $B$ | b50 | c50 | b10 | b50 | c10 | c50 |
| IT | Piacenza | 539535 | 192 | 223 | 13012 | 86.3 | 0.858 | 0.769 | 0.255 | 0.902 | 7199 | 2036 | 43720 | 83514 | 12068 | 30490 |
| IT | Rome | 3716610 | 1323 | 1457 | 37192 | 90.8 | 0.813 | 0.824 | 0.175 | 0.956 | 5277 | 1711 | 36006 | 84728 | 12088 | 33298 |
| IT | Suzzara | 611173 | 218 | 247 | 8193 | 88.3 | 0.814 | 0.812 | 0.314 | 0.909 | 8210 | 2374 | 43295 | 91257 | 13771 | 31592 |
| IT | Udine | 504090 | 179 | 244 | 6357 | 73.4 | 0.804 | 0.823 | 0.226 | 0.842 | 6301 | 1903 | 40425 | 84028 | 11454 | 29481 |
| IT | Venice-Padua-Treviso | 1780952 | 634 | 679 | 12197 | 93.3 | 0.645 | 0.900 | 0.237 | 0.979 | 7788 | 2199 | 43514 | 92136 | 13542 | 33172 |
| LT | Kaunas | 412127 | 147 | 241 | 12307 | 61.0 | 0.931 | 0.673 | 0.174 | 0.785 | 3761 | 1178 | 28928 | 73144 | 7381 | 23733 |
| LT | Vilnius | 669195 | 238 | 332 | 16796 | 71.7 | 0.921 | 0.683 | 0.203 | 0.825 | 3443 | 1145 | 21375 | 65921 | 8585 | 23622 |
| LU | Luxembourg | 737146 | 262 | 332 | 9813 | 79.1 | 0.813 | 0.819 | 0.252 | 0.836 | 6508 | 1665 | 40068 | 85685 | 10871 | 28468 |
| LV | Riga | 855892 | 305 | 460 | 15995 | 66.2 | 0.922 | 0.707 | 0.205 | 0.830 | 4878 | 1488 | 33431 | 86012 | 10564 | 28393 |
| NL | Amsterdam | 2739883 | 975 | 1354 | 23302 | 72.0 | 0.833 | 0.811 | 0.178 | 0.914 | 6982 | 1715 | 41928 | 87657 | 11094 | 30368 |
| NL | Hengelo | 912586 | 325 | 348 | 7462 | 93.2 | 0.836 | 0.799 | 0.232 | 0.953 | 6540 | 1909 | 39477 | 84429 | 11718 | 29764 |
| NL | Leeuwarden | 543669 | 194 | 266 | 7456 | 72.7 | 0.866 | 0.775 | 0.235 | 0.879 | 7497 | 1924 | 37456 | 86648 | 11603 | 30956 |
| NL | Nijmegen | 1493816 | 532 | 635 | 11979 | 83.8 | 0.812 | 0.826 | 0.254 | 0.919 | 7647 | 2113 | 46066 | 89341 | 13862 | 32968 |
| NL | Rotterdam-The Hague | 3282220 | 1168 | 1622 | 21567 | 72.0 | 0.814 | 0.825 | 0.185 | 0.912 | 6493 | 1733 | 39994 | 82812 | 12428 | 30584 |
| PL | Łódź | 1098305 | 391 | 444 | 18546 | 88.0 | 0.875 | 0.742 | 0.175 | 0.906 | 4569 | 1437 | 27847 | 68517 | 9036 | 26171 |
| PL | Lublin | 654363 | 233 | 275 | 12272 | 84.7 | 0.834 | 0.756 | 0.187 | 0.897 | 4964 | 1554 | 33479 | 77055 | 10800 | 29208 |
| PL | Nowe Miasto | 202686 | 72 | 105 | 7804 | 68.7 | 0.848 | 0.750 | 0.303 | 0.797 | 7831 | 2420 | 43248 | 84376 | 14036 | 31295 |
| PL | Poznań | 975515 | 347 | 499 | 13682 | 69.6 | 0.872 | 0.760 | 0.190 | 0.774 | 4811 | 1571 | 32133 | 80328 | 10927 | 31579 |
| PL | Tarnobrzeg | 333755 | 119 | 174 | 15373 | 68.2 | 0.805 | 0.775 | 0.298 | 0.831 | 7453 | 2271 | 40244 | 85663 | 15372 | 31951 |
| PL | Warsaw | 2678997 | 954 | 1101 | 21372 | 86.7 | 0.819 | 0.812 | 0.190 | 0.915 | 4978 | 1481 | 29884 | 81654 | 11243 | 28881 |
| PL | Wrocław | 984267 | 350 | 555 | 21205 | 63.2 | 0.885 | 0.739 | 0.190 | 0.727 | 4755 | 1554 | 30355 | 75220 | 10680 | 28663 |
| PT | Coimbra | 342907 | 122 | 170 | 5888 | 71.7 | 0.778 | 0.829 | 0.230 | 0.790 | 6585 | 1836 | 39916 | 89561 | 12202 | 32252 |
| PT | Covilhã-Fundão | 112731 | 40 | 99 | 3452 | 40.7 | 0.917 | 0.708 | 0.301 | 0.649 | 7651 | 2274 | 47447 | 98159 | 15772 | 34857 |
| PT | Faro | 279741 | 100 | 178 | 13085 | 56.1 | 0.894 | 0.716 | 0.316 | 0.775 | 6515 | 1613 | 37500 | 67298 | 11138 | 22758 |
| PT | Leiria | 394621 | 140 | 177 | 6993 | 79.6 | 0.738 | 0.849 | 0.272 | 0.865 | 7810 | 2173 | 41277 | 87218 | 14218 | 32467 |
| PT | Lisbon | 2785065 | 991 | 1483 | 22306 | 66.9 | 0.844 | 0.803 | 0.191 | 0.908 | 5999 | 1514 | 37483 | 81970 | 9856 | 27078 |
| PT | Porto | 2143826 | 763 | 869 | 12117 | 87.8 | 0.663 | 0.894 | 0.160 | 0.951 | 7087 | 1712 | 43532 | 90999 | 11687 | 29703 |
| PT | Viseu | 240639 | 86 | 128 | 6119 | 67.0 | 0.783 | 0.829 | 0.240 | 0.772 | 6668 | 1898 | 44681 | 91814 | 11380 | 33817 |
| RO | Botoşani-Suceava | 438875 | 156 | 318 | 22716 | 49.2 | 0.890 | 0.697 | 0.284 | 0.718 | 7965 | 2278 | 44920 | 100051 | 13542 | 30052 |
| RO | Braşov | 506999 | 180 | 609 | 22197 | 29.7 | 0.955 | 0.627 | 0.244 | 0.689 | 5757 | 1466 | 33381 | 81081 | 9351 | 23858 |
| RO | Bucharest | 2396642 | 853 | 1615 | 39551 | 52.8 | 0.913 | 0.718 | 0.156 | 0.788 | 3255 | 1194 | 23589 | 64839 | 7613 | 23076 |
| RO | Cluj-Napoca | 502940 | 179 | 449 | 26263 | 39.8 | 0.952 | 0.621 | 0.213 | 0.632 | 5286 | 1307 | 34036 | 92987 | 9840 | 21615 |
| RO | Constanţa | 525667 | 187 | 782 | 22778 | 23.9 | 0.959 | 0.622 | 0.227 | 0.634 | 6057 | 1421 | 36867 | 84703 | 10313 | 21798 |
| RO | Craiova | 424526 | 151 | 431 | 23668 | 35.1 | 0.939 | 0.633 | 0.183 | 0.625 | 4774 | 1413 | 36838 | 88636 | 8736 | 27607 |
| RO | Galaţi-Brăila | 506746 | 180 | 988 | 25447 | 18.3 | 0.975 | 0.555 | 0.250 | 0.645 | 4681 | 1333 | 27473 | 72733 | 8243 | 19973 |
| RO | Iaşi | 561494 | 200 | 432 | 26459 | 46.3 | 0.909 | 0.678 | 0.185 | 0.690 | 6202 | 1748 | 37034 | 92386 | 12553 | 27805 |
| RO | Timişoara | 467804 | 167 | 681 | 14729 | 24.5 | 0.956 | 0.629 | 0.164 | 0.598 | 3980 | 1068 | 30750 | 72005 | 8189 | 20851 |
| SE | Gothenburg | 903761 | 322 | 437 | 17778 | 73.6 | 0.876 | 0.767 | 0.181 | 0.889 | 4780 | 1450 | 30316 | 77470 | 11806 | 28911 |
| SE | Stockholm | 2106507 | 750 | 958 | 27054 | 78.2 | 0.840 | 0.803 | 0.188 | 0.872 | 5342 | 1639 | 38565 | 82729 | 11662 | 32610 |
| SE | Uppsala | 290607 | 103 | 136 | 10867 | 76.1 | 0.919 | 0.666 | 0.183 | 0.847 | 4376 | 1420 | 33092 | 69540 | 9600 | 25997 |
| SK | Košice | 481233 | 171 | 536 | 19059 | 32.0 | 0.933 | 0.673 | 0.323 | 0.588 | 6218 | 1698 | 41258 | 83578 | 11378 | 29704 |
| SI | Ljubljana | 653203 | 233 | 309 | 11077 | 75.2 | 0.847 | 0.778 | 0.225 | 0.854 | 5497 | 1505 | 33946 | 82090 | 12380 | 32150 |
| UK | Aberdeen | 380841 | 136 | 234 | 9721 | 57.9 | 0.927 | 0.687 | 0.161 | 0.726 | 4935 | 1313 | 30033 | 74400 | 8695 | 22145 |
| UK | Bournemouth | 765035 | 272 | 458 | 9806 | 59.5 | 0.890 | 0.756 | 0.169 | 0.743 | 6006 | 1840 | 35628 | 85307 | 11937 | 31762 |
| UK | Brighton | 1368623 | 487 | 625 | 16024 | 78.0 | 0.852 | 0.793 | 0.187 | 0.857 | 7796 | 2008 | 46416 | 82264 | 11960 | 28571 |
| UK | Bristol-Bath | 1400438 | 499 | 639 | 11827 | 78.0 | 0.852 | 0.795 | 0.151 | 0.852 | 6574 | 1849 | 42241 | 93915 | 11418 | 30940 |
| UK | Exeter | 488461 | 174 | 225 | 9540 | 77.4 | 0.884 | 0.744 | 0.220 | 0.834 | 6565 | 1799 | 37434 | 77593 | 11156 | 28607 |
| UK | Glasgow | 1802024 | 642 | 1047 | 13837 | 61.3 | 0.827 | 0.815 | 0.156 | 0.802 | 4907 | 1449 | 30796 | 73918 | 10643 | 26998 |
| UK | Leeds | 2543888 | 906 | 1070 | 15731 | 84.6 | 0.753 | 0.861 | 0.186 | 0.882 | 6393 | 1802 | 40389 | 90920 | 11758 | 29412 |
| UK | Manchester | 3424004 | 1219 | 1511 | 13313 | 80.7 | 0.721 | 0.875 | 0.172 | 0.927 | 6272 | 1694 | 39874 | 86356 | 11615 | 32461 |
| UK | Northampton | 1084779 | 386 | 571 | 10578 | 67.7 | 0.865 | 0.785 | 0.220 | 0.713 | 8298 | 1813 | 47516 | 93963 | 11939 | 27022 |
| UK | Norwich | 553212 | 197 | 244 | 9097 | 80.8 | 0.852 | 0.781 | 0.184 | 0.840 | 6012 | 1388 | 35812 | 85808 | 10020 | 22744 |

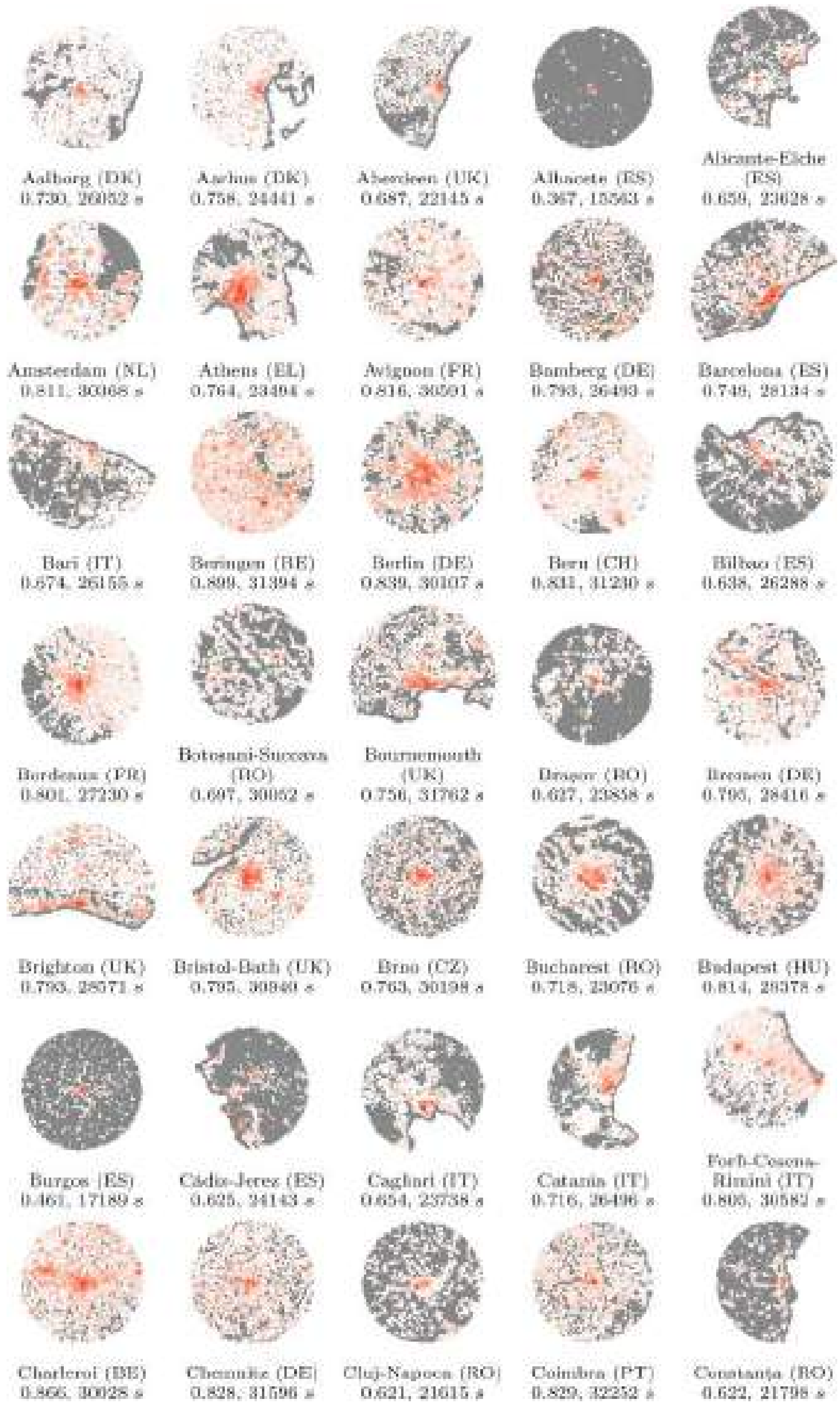



Fig. A.10: Density heatmaps, part I

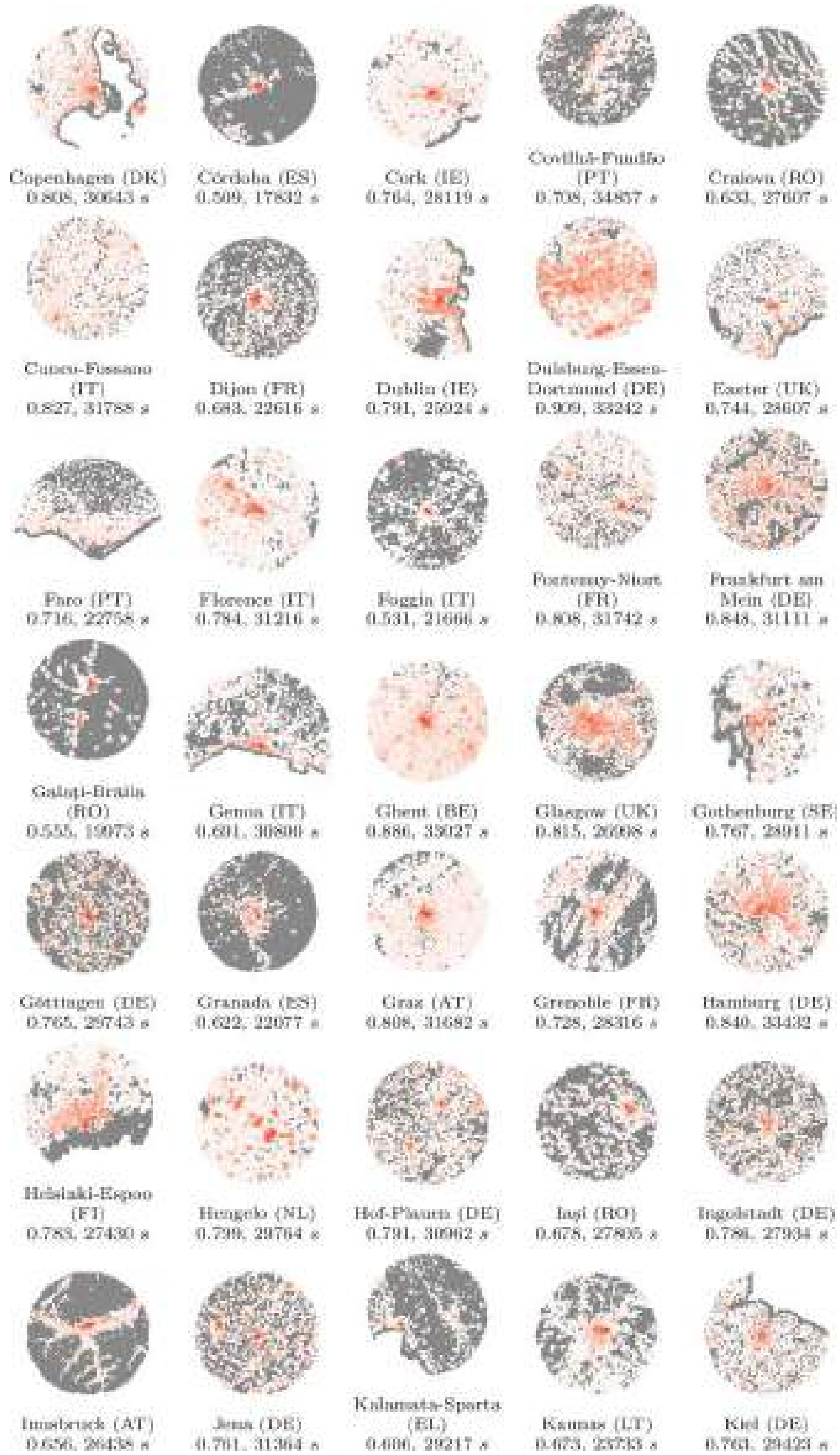



Fig. A.11: Density heatmaps, part II

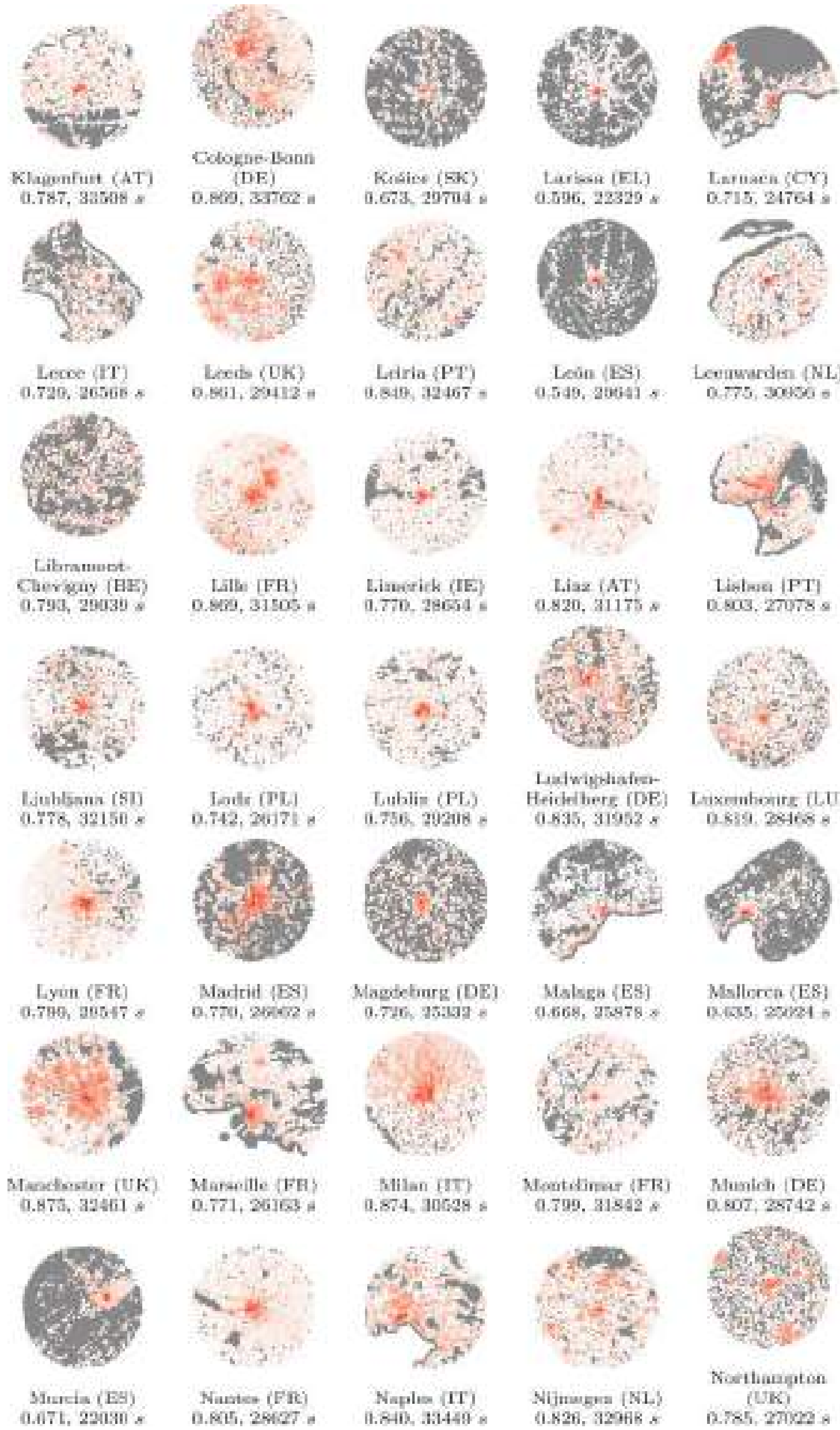



Fig. A.12: Density heatmaps, part III

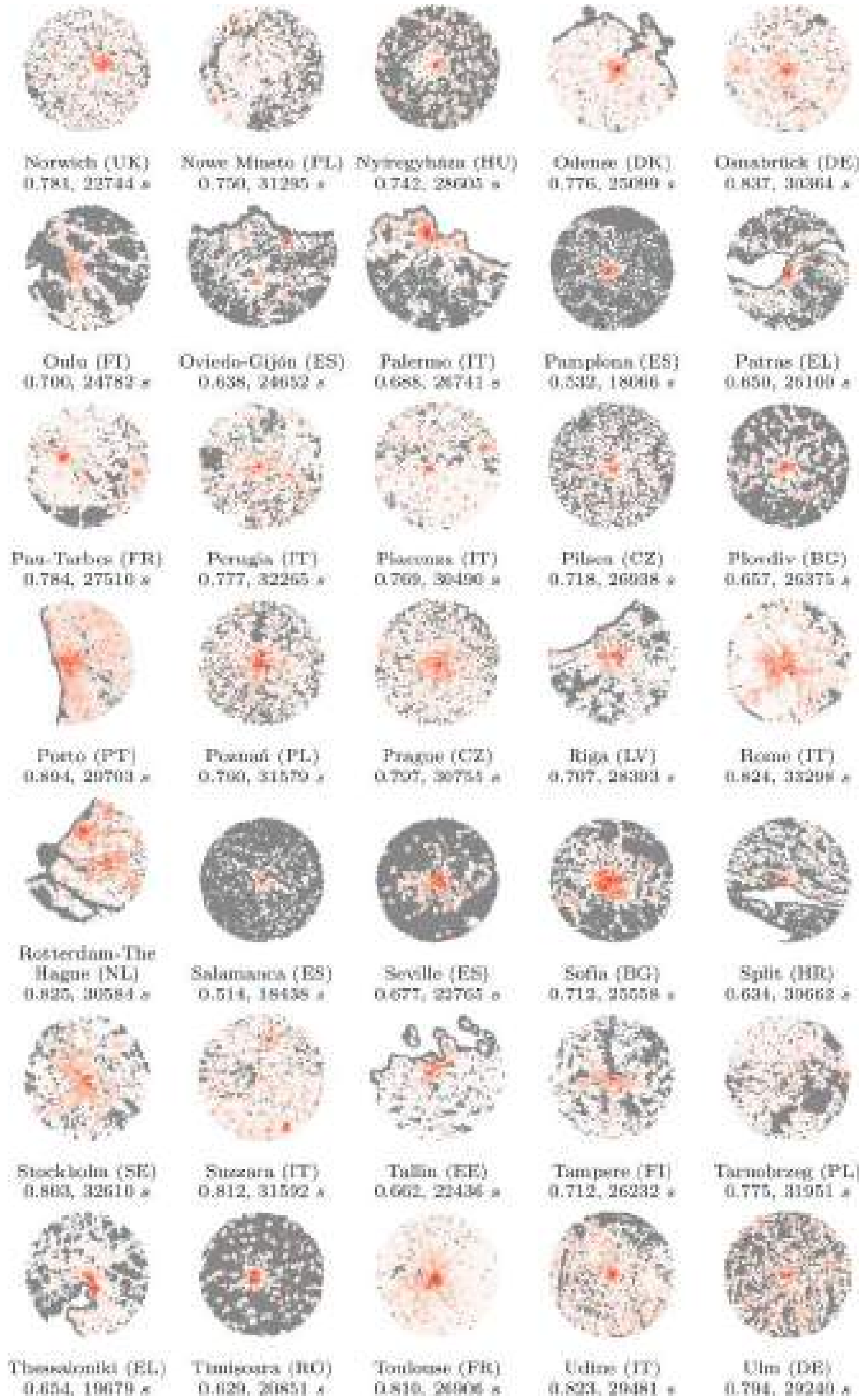



Fig. A.13: Density heatmaps, part IV

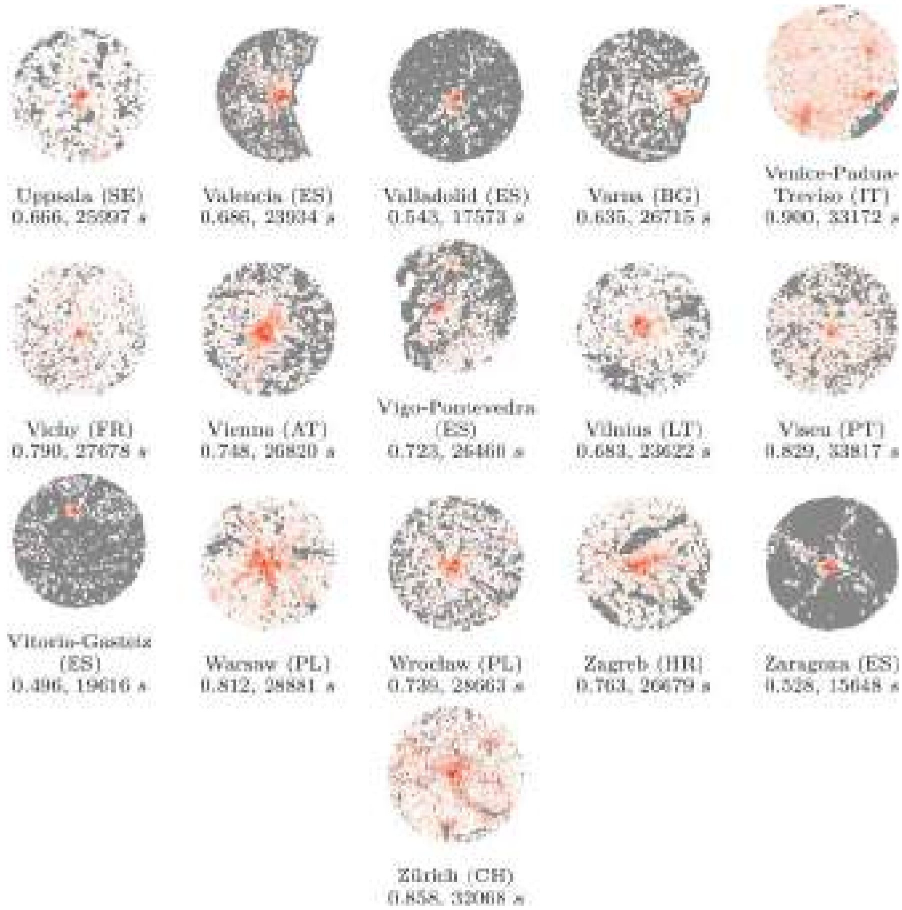



Fig. A.14: Density heatmaps, part V

## References


Applegate, D., Bixby, R., Chvatal, V., Cook, W. (2006). Concorde TSP solver.

Bas Waterhout, F. O., & Sykes, O. (2013). Neo-liberalization processes and spatial planning in France, Germany, and the Netherlands: an exploration. *Planning Practice & Research, 28*, 141–159.

Behnisch, M., Krüger, T., & Jaeger, J. A. (2022). Rapid rise in urban sprawl: Global hotspots and trends since 1990. *PLOS Sustainability and Transformation, 1*, e0000034.

Bribiesca, E. (1997). Measuring 2-D shape compactness using the contact perimeter. *Computers & Mathematics with Applications, 33*, 1–9.

Burckhardt, L., Frisch, M., & Kutter, M. (1955). *Achtung: die Schweiz. Ein Gespräch über unsere Lage und ein Vorschlag zur Tat* (volume 2). Basel Verlag F. Handschin.

Cattaruzza, D., Absi, N., Feillet, D., & González-Feliu, J. (2017). Vehicle routing problems for city logistics. *EURO Journal on Transportation and Logistics, 6*, 51–79.

Ceschia, S., Di Gaspero, L., Rosati, R. M., Schaerf, A. (2024). Multi-neighborhood simulated annealing for the home healthcare routing and scheduling problem. PREPRINT available at Research Square.

Colantoni, A., Grigoriadis, E., Sateriano, A., Venanzoni, G., & Salvati, L. (2016). Cities as selective land predators? A lesson on urban growth, deregulated planning and sprawl containment. *Science of the Total Environment, 545*, 329–339.

Crainic, T. G., Ricciardi, N., & Storchi, G. (2009). Models for evaluating and planning city logistics systems. *Transportation science, 43*, 432–454.

Dantzig, G., Fulkerson, R., & Johnson, S. (1954). Solution of a large-scale traveling-salesman problem. *Journal of the operations research society of America, 2*, 393–410.

Davis, C., & Schaub, T. (2005). A transboundary study of urban sprawl in the Pacific Coast region of North America: The benefits of multiple measurement methods. *International Journal of Applied Earth Observation and Geoinformation, 7*, 268–283.

De Vos, J., & Witlox, F. (2013). Transportation policy as spatial planning tool; reducing urban sprawl by increasing travel costs and clustering infrastructure and public transportation. *Journal of Transport Geography, 33*, 117–125.

European Environment Agency, 2006. Urban sprawl in Europe: the ignored challenge. Publications Office.

European Environment Agency, 2016. Urban sprawl in europe. Joint EEA-FOEN report.

Ewing, R., & Hamidi, S. (2015). Urban sprawl as a risk factor in motor vehicle occupant and pedestrian fatalities: Update and refinement. *Transportation research record, 2513*, 40–47.

Ewing, R., Hamidi, S., & Grace, J. B. (2016). Urban sprawl as a risk factor in motor vehicle crashes. *Urban Studies, 53*, 247–266.

Ewing, R., Schmid, T., Killingsworth, R., Zlot, A., & Raudenbush, S. (2003). Relationship between urban sprawl and physical activity, obesity, and morbidity. *American journal of health promotion, 18*, 47–57.

Ferm, J., Clifford, B., Canelas, P., & Livingstone, N. (2021). Emerging problematics of deregulating the urban: the case of permitted development in England. *Urban Studies, 58*, 2040–2058.

Frenkel, A., & Ashkenazi, M. (2008). Measuring urban sprawl: how can we deal with it? *Environment and Planning B: Planning and Design, 35*, 56–79.

Frumkin, H. (2002). Urban sprawl and public health. *Public Health Reports, 117*, 201–217.

Gini, C., 1912. Variabilità e mutabilità: contributo allo studio delle distribuzioni e delle relazioni statistiche.[Fasc. I.]. Tipogr. di P. Cuppini.

Hamidi, S., Ewing, R., Preuss, I., & Dodds, A. (2015). Measuring sprawl and its impacts: An update. *Journal of Planning Education and Research, 35*, 35–50.

Hennig, E. I., Schwick, C., Soukup, T., Orlitová, E., Kienast, F., & Jaeger, J. A. (2015). Multi-scale analysis of urban sprawl in Europe: towards a European de-sprawling strategy. *Land use policy, 49*, 483–498.

Hortas-Rico, M., & Solé-Ollé, A. (2010). Does urban sprawl increase the costs of providing local public services? Evidence from Spanish municipalities. *Urban studies, 47*, 1513–1540.

Intergovernmental Panel on Climate Change, 2022. Global Warming of 1.5C: IPCC Special Report on Impacts of Global Warming of 1.5C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty. Cambridge University Press.

Jaeger, J. A., Bertiller, R., Schwick, C., Cavens, D., & Kienast, F. (2010a). Urban permeation of landscapes and sprawl per capita: New measures of urban sprawl. *Ecological Indicators, 10*, 427–441.

Jaeger, J. A., Bertiller, R., Schwick, C., & Kienast, F. (2010b). Suitability criteria for measures of urban sprawl. *Ecological indicators, 10*, 397–406.

Jaeger, J. A., & Schwick, C. (2014). Improving the measurement of urban sprawl: Weighted urban proliferation (WUP) and its application to Switzerland. *Ecological indicators, 38*, 294–308.

Karp, R. M. (1972). Reducibility among combinatorial problems. *Complexity of computer computations. Springer*, 85–103.

Kasanko, M., Barredo, J. I., Lavalle, C., McCormick, N., Demicheli, L., Sagris, V., & Brezger, A. (2006). Are European cities becoming dispersed?: a comparative analysis of 15 European urban areas. *Landscape and urban planning, 77*, 111–130.

Lee, C. (2020). Impacts of two-scale urban form and their combined effects on commute modes in us metropolitan areas. *Journal of Transport Geography, 88*, 102821.

Lee, C. (2020). Metropolitan sprawl measurement and its impacts on commuting trips and road emissions. *Transportation Research Part D: Transport and Environment, 82*, 102329.

Makido, Y., Dhakal, S., & Yamagata, Y. (2012). Relationship between urban form and co2 emissions: Evidence from fifty Japanese cities. *Urban Climate, 2*, 55–67.

Marique, A. F., Dujardin, S., Teller, J., & Reiter, S. (2013). School commuting: the relationship between energy consumption and urban form. *Journal of transport Geography, 26*, 1–11.

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*, 17–23.

Nazarnia, N., Harding, C., & Jaeger, J. A. (2019). How suitable is entropy as a measure of urban sprawl? *Landscape and Urban Planning, 184*, 32–43.

Nilles, J. M. (1991). Telecommuting and urban sprawl: Mitigator or inciter? *Transportation, 18*, 411–432.

Oueslati, W., Alvanides, S., & Garrod, G. (2015). Determinants of urban sprawl in european cities. *Urban studies, 52*, 1594–1614.

Patacchini, E., Zenou, Y., Henderson, J. V., & Epple, D. (2009). Urban sprawl in europe. *Brookings-Wharton papers on urban affairs*, 125–149.

Pendall, R. (2000). Local land use regulation and the chain of exclusion. *Journal of the American Planning Association, 66*, 125–142.

Pourtaherian, P., & Jaeger, J. A. (2022). How effective are greenbelts at mitigating urban sprawl? a comparative study of 60 European cities. *Landscape and Urban Planning, 227*, 104532.

Power, A. (2001). Social exclusion and urban sprawl: Is the rescue of cities possible? *Regional Studies, 35*, 731–742.

Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. Simon and schuster.

Rosati, R. M., & Schaerf, A. (2024). Multi-neighborhood simulated annealing for the capacitated dispersion problem. *Expert Systems with Applications, 255*, 124484.

Salvati, L. (2015). Lost in complexity, found in dispersion: 'Peripheral' development and deregulated urban growth in Rome. *Cities, 47*, 73–80.

Schneider, A., & Woodcock, C. E. (2008). Compact, dispersed, fragmented, extensive? a comparison of urban growth in twenty-five global cities using remotely sensed data, pattern metrics and census information. *Urban Studies, 45*, 659–692.

Schwanen, T. (2002). Urban form and commuting behaviour: a cross-european perspective. *Tijdschrift voor economische en sociale geografie, 93*, 336–343.

Schweitzer, L., & Zhou, J. (2010). Neighborhood air quality, respiratory health, and vulnerable populations in compact and sprawled regions. *Journal of the American Planning Association, 76*, 363–371.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal, 27*, 379–423.

Batista e Silva, F., Dijkstra, L., Poelman, H., 2021. The JRC-GEOSTAT 2018 population grid. Technical Report. JRC Technical Report (forthcoming).

Steurer, M., & Bayr, C. (2020). Measuring urban sprawl using land use data. *Land Use Policy, 97*, 104799.

Tiznado-Aitken, I., Lucas, K., Munoz, J. C., & Hurtubia, R. (2022). Freedom of choice? Social and spatial disparities on combined housing and transport affordability. *Transport Policy, 122*, 39–53.

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

Travisi, C. M., Camagni, R., & Nijkamp, P. (2010). Impacts of urban sprawl and commuting: a modelling study for italy. *Journal of Transport Geography, 18*, 382–392.

Trowbridge, M. J., Gurka, M. J., & O'connor, R. E. (2009). Urban sprawl and delayed ambulance arrival in the US. *American Journal of Preventive Medicine, 37*, 428–432.

Tsai, Y. H. (2005). Quantifying urban form: Compactness versus 'sprawl'. *Urban studies, 42*, 141–161.

Turan, B., Hemmelmayr, V., Larsen, A., & Puchinger, J. (2024). Transition towards sustainable mobility: the role of transport optimization. *Central European Journal of Operations Research, 32*, 435–456.

Van Ommeren, J. N., & Gutiérrez-i Puigarnau, E. (2011). Are workers with a long commute less productive? An empirical analysis of absenteeism. *Regional Science and Urban Economics, 41*, 1–8.

Whyte, W. (1958). *The Exploding Metropolis*. Doubleday: Doubleday anchor books.

Wolman, H., Galster, G., Hanson, R., Ratcliffe, M., Furdell, K., & Sarzynski, A. (2005). The fundamental challenge in measuring sprawl: Which land should be considered? *The Professional Geographer, 57*, 94–105.

Yeh, A. G. O., & Li, X. (2001). Measurement and monitoring of urban sprawl in a rapidly growing region using entropy. *Photogrammetric Engineering and Remote Sensing*.

Zhao, Z., & Kaestner, R. (2010). Effects of urban sprawl on obesity. *Journal of Health Economics, 29*, 779–787.

Zolnik, E. J. (2011). The effects of sprawl on private-vehicle commuting distances and times. *Environment and Planning B: Planning and Design, 38*, 1071–1084.