A dynamical taxonomy of population density: moving around in the Moran scatterplot

Davide Fiaschi — University of Pisa

Angela Parenti — University of Pisa

Cristiano Ricci — University of Pisa

March 15, 2023

To respect the size limit in the submission, figures are available at
https://poisson.phc.dm.unipi.it/~cricci/fiaschi_parenti_ricci_LDMS.pdf

Abstract

In this paper, we analyse the spatial distribution dynamics of the population density of Italian municipalities over the period 1984-2019. Firstly, we refine the standard Moran-based classification, using as additional dividers the bisector and the estimated nonparametric Moran's I. The proposed taxonomy resembles other classifications of municipalities into urban and rural but has the advantage of being based on a very limited amount of information that is only population density and the definition of Local Labour Areas (LLA). This allows us to get a taxonomy for each year and study its dynamic over time. Moreover, we are also able to study the evolution of municipalities in continuous state space without relying on discrete taxonomy, therefore providing a more comprehensive understanding of the historic track that led to the current configuration. Our findings show the presence of three dynamic attractors, an *urban attractor*, a *suburban attractor* and a *rural attractor*, where all municipalities and their LLA are converging.

JEL Classification Numbers: C14, O18, R12, R23

Keywords: nonparametric methods, spatial dependence, spatial distribution dynamics, urban development

Contents

1 Introduction

The concentration of people and economic activities has been identified as evidence of agglomeration economies (Glaeser, 2010). Population dynamics is the result of competition between economic forces, i.e. centripetal vs centrifugal (Krugman, 1994). On the one hand, wages, higher local amenities, and knowledge spillovers are among the primary motivations which drive people toward urban areas. On the other hand, house prices and congestion lead to urban sprawl. The COVID-19 pandemic has accelerated the adoption of digital infrastructures and the use of smart working, further incentivising people to move out of denser metropolitan areas (Glaeser and Cutler, 2021, Peiser and Hugel, 2022). Although the urban dynamics are relatively similar among European countries, there is still a component of spatial heterogeneity that needs to be further investigated. This heterogeneity accounts not only for country-specific effects, like national policies but also for geographical characteristics and historical patterns (Accetturo and Mocetti, 2019).

Whilst it is true that population dynamics is a phenomenon that can only be observed over a long time period (low resolution), the same does not hold for geographical resolution. For example, urban development can be studied at sub-municipal levels as in the EU-OECD definition of urban centre and city, while a change in commuting patterns can be grasped by observing the evolution of the local labour areas (LLAs) or functional urban areas (FUAs) (Dijkstra et al., 2019 and Lamorgese and Petrella, 2019).

In this paper, we analyse the spatial dynamics of population density among Italian municipalities from 1984 to 2019. Although the temporal dynamics appears very stable, it is very heterogeneous in space, revealing different agglomeration patterns over the Italian territory. To shed light on this miscellaneous configuration we advance a taxonomy of Italian municipalities based on their position in the Moran scatterplot of population density (Anselin, 2019), taking into account their density and the average within their LLA. The classification that we propose is a refinement of the standard Moran-based classification in the four quadrants HH, HL, LH and LL which also uses the bisector and the estimated nonparametric Moran's I to further subdivide the space into eight regions. Our classification resembles those already present in the literature (see e.g., Barca et al., 2014 and ISTAT, 2017) that are based also on other sources of information such as the presence of primary services, infrastructure, and transport. However, our methodology has the remarkable advantage of being based only on the population density and the definition of neighbours. This data can be easily gathered for a long time series thus opening the possibility to study the evolution of the classification over time. The understanding of the development of the classification may constitute an important tool for policymakers to shed light on how the system has evolved towards the present configuration, therefore providing a piece of important evidence for future policies. This backward analysis is not always possible for other classifications that are based on other sources of administrative data. However, the definition of neighbours is crucial in our methodology and has to be selected in each case of study according to the phenomenon of interest. In our case is well established in the literature that the correct definition of neighbouring to study population dynamics is the membership to the same LLA (Lamorgese and Petrella, 2019, Manzoli and Mocetti, 2019). A comprehensive understanding of the different tracks of territorial development is reached using the Moran-based taxonomy in conjunction with the technique developed in Fiaschi et al. (2018) which allows estimating nonparametrically the joint dynamics of the population density of the municipality and the average level within its LLA. The outcome of the estimation is an expected arrow in every point of the Moran space, that best approximates the local direction in which municipalities are evolving.

When we apply the methodology to the case of Italian municipalities we find that the Moran-based taxonomy is able to distinguish rural from urban areas. Moreover, our classification identifies all the metropolitan cities and their surrounding municipalities characterized by high density. The classification is also able to capture the exceptions of Venice, Catania and Naples which are constrained by the geographical characteristics of their territories. Concerning rural areas, we can also distinguish the most important municipalities in the LLA, analogously identified as *poles* in Barca et al. (2014). The spatiotemporal analysis of the Moran scatterplot allows us to detect the presence of three attractor points, an *urban attractor*, a *suburban attractor* and a *rural attractor*, where municipalities and their LLA are converging. This highlights a tendency toward the formation of three settlement systems whose basin of attraction can be clearly identified. In particular, the basin of the urban attractor is composed of both the HH and HL quadrants as well as about half of the LH quadrant. The basin of the suburban attractor expands both in the LH quadrant and the LL above the nonparametric Moran's I. Finally, the basin of the rural attractor is composed of municipalities and LLAs with a density lower than the average, in the LL Moran quadrant below the nonparametric Moran's I.

The paper is structured as follows. Section 2 explains the data sources and the sample choice; Section 3 provides a first look at the dynamics of the population density for the Italian municipalities over the period 1984-2019; Section 4 presents the taxonomy based on the Moran scatterplot and its evolution over the period; Section 5 reports the complete spatiotemporal analysis for all the municipalities; finally, Section 6 concludes.

2 Data sources

The analysis of spatial agglomeration is conventionally carried out at functional urban areas (FUAs) or local labour areas (LLAs), i.e. using population density and travel-to-work flows (commuting zone) as key information. We, therefore, follow the literature and use the population density at municipal and LLA levels. Data on municipal population and definition of Italian LLAs come from ISTAT (Italian National Institute of Statistics).

Since 1861 the census of the Italian population is repeated every 10 years. From 1982 ISTAT also provides inter-census estimates of the population at a yearly frequency. These estimates are obtained by updating the census population using the data on births, deaths, and internal and external migrations collected by the municipalities. Since October 2018 instead, ISTAT has been yearly conducting a sample survey by collecting the main characteristics of the Italian resident population at the municipal level to be integrated with administrative sources to get a permanent census. Thus considering both the 10-year census data, the inter-census estimates, and the permanent census the yearly data are available over the period 1982-2021. To avoid anomalies induced by the COVID pandemic, we decide to end our analysis in 2019. Moreover, since in Section 5 we perform a space-time analysis of population density, it is more convenient to subdivide the time span of observations into subperiods of the same length. Therefore we consider 1984 as the first year.

In the 36 years considered the definition of some of the municipalities changed several times, either because the municipal territory changed, or because two or more municipalities merged together. To make it possible to compare municipalities over time we homogenized them to their definition in 2019.

Since 1981 ISTAT also classifies Italian municipalities into LLAs, areas where the bulk of the labour force lives and works. LLAs are usually considered a good unit of observation of spatial agglomeration since they contain both the place of residence and the place of work of almost all the residents. The initial classification of municipalities into LLAs has been updated every 10 years, up to 2001. In 2011 ISTAT changed the definition radically to be consistent with the European definition. For sake of comparability also the 2001 classification has been revised using the updated rule. No previous updates have been done for the classification of 1991 and 1981. Hence, we only restrict ourselves to use the classification into LLAs of 2001 and 2011, updated with the latest definition. However, since the time span of our observation starts before the first considered classification (2001) we decide to adopt the classification of 2001 for the years from 1984 to 2001, and the classification of 2011 for the years from 2002 to 2019. Aware that this time discrepancy induces a bias in the representation of commuting patterns in years previous to the classification, we decided to use only the classifications of 2001 and 2011 since they are the only ones comparable between themselves and based on the European definition. Moreover, in Section 5 we also control how much our estimates are affected by the change in the LLAs.

3 A first glance to population density in Italy

Figure 1 reports the maps of the (log) population density of about 8000 Italian municipalities for the years 1984 and 2019 and its absolute variation over the period. A high level of aggregation in the spatial distribution of population density can be detected which seems to be not increasing over time. The distribution of people is strongly affected by geographical characteristics: we observe a low density in the municipalities located in the mountainous regions (along the Alps and Apennines) and a higher density in the Po Valley and along the coasts. On the maps, we also report the 14 Italian metropolitan cities with yellow dots which exhibit some of the highest levels of population density (see list in Appendix 2). Looking at the variation over time, we observe high persistence with a tendency to increase the population density of the already high-density municipalities and an opposite trend for the low-density ones, confirming the presence of an agglomeration effect. Despite this agglomeration effect seems to be still present in some areas where we observe an increase in the density of the main urban centre associated with a decrease in its surrounding municipalities (e.g., around Grosseto, Matera, Lecce and L'Aquila), the opposite phenomenon can be observed in some metropolitan cities (e.g., Turin, Milan, Rome, Naples and Cagliari).

Evidence of agglomeration can also be grasped by looking at Figure 2a showing, on average, an increasing trend in the level of population density together with an increasing Moran's I index. This tendency however has reversed after 2013, when the mean density decreased and the coefficient of variation slightly increased. The reversion in the trend can be ascribable to the decline in the total population, as shown in Figure 2b. In Figure 2a a discontinuity in the time series can be observed between 1991 and 1992. This jump is likely induced by the inter-census estimates between the 1991 and 2001 censuses. The subsequent estimates do not seem to suffer from the same bias. The same effect cannot be found in the time series of the total population. This discrepancy is due to the fact that the reconstruction of residents at the municipal level is more difficult since it is more arduous to track inter-municipal migration than international migration.

Figure 2c shows how much of the total variance of the municipal population density is accounted for by the variance between LLAs and within LLAs. Over time slightly more than 50% of the total variance is due to the variability of population density of the municipalities within the same LLA. This fraction reduces over the years previous to the change in the definition of the LLAs. A discontinuity, in particular, can be observed between 2001 and 2002 due to the change in the definition of LLAs.[1]

[1] As stated in Section 2 we use the LLA definition of 2001 for the years 1984-2001 and that of 2011 for the years 2002-2019. The adoption of the LLAs in 2001 for the previous years induces an overestimation of the variance accounted for by the variability in the density among the municipalities that belong to the same LLA. The same applies to the years previous 2011 although with a lower magnitude.

Figures/map_popDensSLL_1984.pdf

(a) 1984

Figures/map_popDensSLL_2019.pdf

(b) 2019

Figures/mapDiffLog1984_2019.pdf

(c) Difference between 1984 and 2019

Figure 1: Map of population density of Italian Municipalities in 1984 and 2019.

Source: ISTAT (Italian National Institute of Statistics).

Figures/timeSeriesDensityMeanMoranI.pdf

(a) Mean and Moran'I index of municipal population density

Figures/timeSeriesTotPop.pdf

(b) Italian population (millions)

Figures/varianceDecomposition.pdf

(c) Between and within LLA variance of population density (log)

Figure 2: Population over time

4 Moran-based taxonomy

To further characterise the spatial distribution of population density of the Italian municipalities we look at Figure 3 showing the Moran scatterplot in 1984 and 2019. Each point represents the (log) population density of each municipality and the corresponding average value within its LLA[2]. The size of the bullets is proportional to the municipality's total population (see Figure 3). We also report in yellow the 14 Italian metropolitan cities together with their name, and the average values of both axes (dashed black lines). We observe that all the more populated municipalities (top 1%) show a density that is higher than the average (biggest brown bullets). In contrast, the least populated municipalities (bottom 50%) have a density lower than the average.

In the Moran scatterplot, we also report the nonparametric estimation of Moran's I index with its 95% confidence band (blue solid and dashed lines respectively). In 1984 we observe on average positive spatial dependence although almost zero in the part of the Moran space characterized by low-density municipalities belonging to LLAs with municipalities with low-density (LL). Moreover, on the HH quadrant, the nonparametric Moran is close but below the 45-degree line while on the LL is above: a municipality with a density higher than the average is surrounded by municipalities with high density but lower than its own. On the other hand, in the LL quadrant a municipality with a density lower than the average is surrounded by municipalities with a density lower than the average but higher than its own. We observe

[2]The spatial matrix corresponding to the definition of LLA is the row-standardise matrix of membership where all municipalities belonging to the same LLA are considered neighbours.

that almost all the metropolitan cities are in the HH quadrant, below the bisector and the nonparametric Moran. Exceptions are Venice, Catania, and Naples which are much closer to the bisector, meaning that they are closer in density to their neighbours. This is accountable to the geographical characteristic of the territory since all these cities cannot expand vertically too much (Venice is on a lagoon, and Catania and Naples are close to Vulcans, Etna, and Vesuvius respectively).[3]

Most of the discussed results also hold for 2019. However, some differences are present: Naples, Catania and Venice are closer to the bisector; the nonparametric estimated Moran is higher in the HH quadrant; Genoa and Turin decrease the average level of their LLAs. This latter effect is due to a change in the LLA definition and will be further discussed in the next section.

Figures/moranScatterplot_1984-eps-converted-to.pdf

(a) 1984

Figures/moranScatterplot_2019-eps-converted-to.pdf

(b) 2019

Figure 3: Moran scatterplot of population density of Italian Municipalities in 1984 and 2019.

Figures/LegendPicture.pdf

Figures/LegendLabels.pdf

Figure 4: Division of municipalities according to their position in the Moran Scatteplot

[3]Notice that in most cases LLAs are made of municipalities that lie on the same horizontal line. This is due to the fact the average population density of an LLA is roughly the same as the average of all the neighbours of each municipality that composes the LLA.

4.1 Refined classification

In this section, we further refine the standard classification of the Moran scatterplot in the four quadrants HH, LH, LL, and HL, using additionally the bisector and the estimated nonparametric Moran's I.[4] In particular, we separately consider eight classifications as reported in Figure 4:

1. High High-Top (**HH-T**, pink) : high-density municipalities with high-density neighbours above bisector;
2. High High-Middle (**HH-M**, red): high-density municipalities with high-density neighbours below bisector but above nonparametric Moran;
3. High High-Bottom (**HH-B**, wine): high-density municipalities with high-density neighbours below nonparametric Moran;
4. Low High (**LH**, yellow): low-density municipalities with high-density neighbours;
5. Low Low-Top (**LL-T**, light green): low-density municipalities with low-density neighbours below nonparametric Moran;
6. Low Low-Middle (**LL-M** , green): low-density municipalities with low-density neighbours above bisector but below nonparametric Moran;
7. Low Low-Bottom (**LL-B**, olive green): low-density municipalities with low-density neighbours below bisector;
8. High Low (**HL**, brown): high-density municipalities with low-density neighbours.

In Figure 5 we colour in the map the municipalities according to the above classification to understand the geographical pattern. This allows us to transpose the above classification into a taxonomy in the geographical space.

Both in 1984 and 2019 we observe that all the metropolitan cities belong to the HH-B (wine) except Venice, Catania and Naples which belong to the HH-M (red) together with municipalities surrounding the urban areas. The neighbours of municipalities surrounding

[4] The nonparametric Moran used for the taxonomy is estimated on the pooled data, i.e. considering all the observations in the initial and final years.

the urban areas usually belong to the HH-T (pink). Municipalities in the LH (yellow) and LL-T (light green) are characterised by almost zero spatial correlation with their neighbours as observed from the zero slope of the nonparametric Moran's I in these regions, and indeed appear to have no clear geographical pattern. Municipalities in the LL-M (green) and LL-B (olive green) are characterized by low density and are surrounded by low-density neighbours. These municipalities are located in rural areas mostly along the mountains (Alps, and Appennines). Finally, municipalities in the HL (brown) are the most important in the rural areas given that they have a higher density than their neighbours and higher than the average.

Our Moran-based classification, only based on the municipal and LLA population density, resembles the one of Barca et al. (2014) established on the presence of essential services provision and degree of remoteness. According to Barca et al. (2014)'s definition, Italian municipalities can be divided into: "Urban poles of attraction" and "Intermunicipal poles of attraction" which are either single municipalities or a group of neighbouring municipalities able to provide a full range of secondary education, at least one grade 1 emergency care hospital and at least one Silver category railway station; "Outlying areas", "Intermediate areas", "Peripheral areas" and "Ultra-peripheral areas" which have not all the essential services and are located within a distance of 20, 40, 75 or more minutes respectively to the nearest hub. When we compare Figure III.2 in Barca et al. (2014) with our map of municipalities coloured according to the Moran-based classification in Figure 5b we observe that it is possible to make a correspondence between the two taxonomies.[5] "Urban poles of attraction" and "Intermunicipal poles of attraction" mostly correspond to HH-T, HH-M, HH-B and HL, "Outlying areas" and "Intermediate areas" to LH and LL-T while "Peripheral areas" and "Ultra-peripheral areas" to LL-M and LL-B. Although is well-established in the literature that the demographic size of the municipality is not sufficient to qualify territories as the pole of attraction/peripheral areas (see also Perchinunno et al., 2019), our Moran-based classification which considers the size (in terms of density) of both the municipality and its neighbours jointly is effective.

Analogously, we try to classify the LLAs. LLAs that include only red and yellow municipalities are those whose densities are higher than the average density of all the municipalities. Those that include only green and brown municipalities are instead below the average. Hence we immediately observe that most of the LLAs made of green and brown municipalities are

[5]For sake of space we do not report our taxonomy in 2014 since is the same as in 2019.

found in remote areas, like mountains (Alpes and Apennines) and Sardinia. LLAs characterized by municipalities with the highest density (wine, red, and pink) are located in Po Valley (Milan, Brescia, Padua, Modena), along the Adriatic coast, and in Apulia. Only a few of them are on the Tirrenic coast (Naples). We can also identify big LLAs with the metropolitan city (Turin, Bologna, Rome) which includes also municipalities with low density (yellow). Along the mountains and in the most rural areas we observe LLAs characterized by municipalities with low density (light green, green and olive green). If the LLA includes a major city, the corresponding municipality is brown in colour.

Even in this case, we can relate our classification of LLAs with others advanced in the literature. Specifically, ISTAT (2017) proposes a definition of LLAs in "major urban areas", "medium-sized cities" and "other LLA", based on different classifications adopted at national and international levels, as the definition of metropolitan cities and that used in the National Rural Development Programme 2014-2020. Comparing Figure Cartogramma 2 in ISTAT (2017) with Figure 5b we notice that all the 21 major urban LLAs correspond to LLAs which are identified to be wine-red-pink or wine-red-pink-yellow, with the exception of the LLA of Genoa. No clear correspondence appears instead with the "medium-sized cities" LLA identified in ISTAT (2017). However, our Moran-based classification is able to classify rural LLAs (green-olive green) and LLA located in remote areas but with a major city (brown-green).

Figures/map_popDensSLL_UpperLower_Bisector_MoranNP_1984.pdf

Figures/map_popDensSLL_UpperLower_Bisector_Moran

(a) 1984

(b) 2019

Figure 5: Map of Moran scatterplot quadrants, bisector, and nonparametric Moran

4.2 From 1984 to 2019

A first comparison of the maps in 1984 and 2019 reported in Figure 5 allows us to recognise some phenotypic examples.

When we look at the map in 1984 of the previous Moran-based classification for the LLAs of Bolzano, Cagliari, L'Aquila and Ciriè, they are very similar (green with a brown centre, see Figure 6). However, in 2019 Bolzano, Cagliari and Ciriè exhibited a completely different taxonomy (yellow with wine centre, see Figures 6a-6b, 6c-6d) while L'Aquila kept the original classification (green with a brown centre, see Figures 6e-6f). This difference can be observed also in the absolute variation of municipal density as highlighted in Section 3. In particular, the decrease in the density of the central municipality (Bolzano, Cagliari and Ciriè) and the increase in the surroundings induce the change in the Moran-based classification. A possible explanation of the divergence in 2019 among these two groups of LLAs can be ascribed to the difference in transport infrastructures. In particular, both Bolzano and L'Aquila are in mountainous areas but Bolzano is better connected both by roads and high-speed rail. Even though L'Aquila has a distance to Rome which is lower in kilometres than the one between

Bolzano and Verona (117 km and 155 km respectively) the travel time by train is doubled (around 3 hours and 1.5 hours respectively). Analogously, Cagliari has the port and the airport which are the main hubs for tourism in Sardinia. Finally, the change in the taxonomy of Cirié is ascribable to the endogenous change in the LLA. In particular, the LLA of Cirié is almost completely absorbed by the LLA of Turin (see Figures 6g-6h).[6]

To detect and analyze all the possible transitions in the Moran-based taxonomy, without looking at each municipality one by one, we look at the transition matrix between the possible discrete states HH-T, HH-M, HH-B, HL, LH, LL-T, LL-M, LL-B. Namely, for every Moran state, we count which fraction of municipalities start in one state in 1984 and ends in another (including the initial state) in 2019. In Table 1 we report the transition probability matrix, while in Figure 7 the associated graph.[7] We first observe that there is high persistence of the initial state since the highest transition probability for each row is on the diagonal (most of the municipalities tend to remain in the state defined in 1984). Secondly, the matrix is almost block diagonal, since transitions from the HH-* and LL-* blocks are very rare and non-significant. The largest transition probabilities except those on the diagonal in the HH-* block are from HH-B to HH-M and from HH-T to HH-M. This implies an effect of convergence towards the state HH-M. Moreover, while the transition from HH-B to HH-M might indicate an increase in the average density within the LLAs, the transition from HH-T to HH-M is likely caused by the geographical expansion of LLAs. Differently, in the LL-* block the highest transition probabilities out of the diagonal are from LL-M to LL-T and vice versa. In this case, there is no clear evidence of convergence toward a specific state.

The transition matrix provides only a rough picture of the overall distribution dynamics, i.e. only for the changes in the discrete Moran-based classification. To have a complete picture we need to estimate the joint local dynamics of the population density of each municipality and its LLA average value in the Moran space without depending on the definition of the states. To solve this problem in the next section we use the Local Directional Moran Scatter Plot (LDMS) developed in Fiaschi et al. (2018).

[6]This change in LLA implies an upward shift in the Moran scatterplot since the average density within the LLA increases.

[7]In the graph we only report probabilities higher than 0.05.

Figures/Bolzano_1984.png

(a) Bolzano 1984

Figures/Bolzano_2019.png

(b) Bolzano 2019

Figures/Cagliari_1984.png

(c) Cagliari 1984

Figures/Cagliari_2019.png

(d) Cagliari 2019

Figures/Roma_LAquila_1984.png

(e) L'Aquila 1984

Figures/Roma_LAquila_2019.png

(f) L'Aquila 2019

Figures/Torino_Milano_1984.png

(g) Cirié 1984

Figures/Torino_Milano_2019.p

(h) Cirié 2019

Figure 6: Change in Moran-based classification 1984-2019

	HH-T	HH-M	HH-B	HL	LH	LL-T	LL-M	LL-B
HH-T	**0.71**	*0.19*	0.03	0.03	0.03	0	0	0.01
HH-M	0.15	**0.64**	0.16	0.03	0.01	0.01	0	0
HH-B	0.03	*0.21*	**0.66**	0.10	0	0	0	0
HL	0.03	0.03	0.09	**0.77**	0	0	0	0.07
LH	0.16	0.02	0.01	0.01	**0.64**	0.12	0.02	0.02
LL-T	0	0	0	0.01	0.14	**0.64**	0.16	0.05
LL-M	0	0	0	0.01	0.03	0.16	**0.66**	0.14
LL-B	0.01	0.01	0	0.12	0.05	0.04	0.11	**0.66**

Table 1: Transition Matrix 1984-2019

Figures/grafo-eps-converted-t

Figure 7: Graph representation

5 Comprehensive spatiotemporal dynamic

In this section, we improve the previous analysis by estimating the joint dynamics of the density of the municipality and of its LLA. Our methodology, the Local Directional Moran Scatterplot (LDMS), is based on the nonparametric estimation of a vector field generating the movement of municipalities in the Moran space over the period of observation. This allows us to get for each point in the Moran space an expected arrow resulting from the average of the movements around the point. The main advantages of this methodology are that the dynamic can be estimated at every point of the space and that the estimation is not affected by the definition of the discrete states as in the case of the transition matrix.

Figure 8a shows the estimated LDMS for the period 1984-2019.[8] In particular, we report in the Moran scatterplot both the observations in the initial year (yellow points) and in the final year (green points). Metropolitan cities are reported in the two years with orange (initial year) and brown points (final year). We also draw the nonparametric estimate of Morans' I and its confidence bands in both years (blue and purple lines for the initial and final year respectively). The estimated LDMS is represented as red arrows. We only report the arrows that are significant at 5% of significance level.[9] For each evaluation point, we also report in blue the estimated variance of the direction of the arrow (the darker the colour the higher the variance).

Looking at Figure 8a we first observe that the increase in the slope of the nonparametric Moran's I in the HH quadrant agrees with the direction of the estimated arrows, in particular those around Catania, Milan and Naples. Moreover, there is a tendency of convergence toward the nonparametric Moran's I, coherently with the evidence derived from the transition matrix of convergence to the HH-M state. All the arrows in the HH quadrant point to the right, i.e. towards an increase in the level of municipal density, with the exception of those around Naples, Milan and Turin highlighting an urban sprawl phenomenon. Finally, it is also possible to notice the presence of an attractor for the LDMS around (8.5,8).

In the HL quadrant we find a tendency to increase the LLA density, i.e. all the arrows point upwards. In particular, in Figure 9a we highlights the movement of the four cases analysed in Section 4, i.e. Bolzano, Calgiari, L'Aquila and Cirié. Notice that the direction of these movements is coherent with the estimated arrows. Moreover, while is evident that the displacement of L'Aquila is negligible, those of Cirié, Cagliari and Bolzano induce a change in their taxonomy. The special case of the complete change in the LLA definition of Cirié is clearly noticeable. On the contrary, in the LH quadrant, the significant arrows mostly point downwards and on the right, i.e. the average density in the LLA decreases while the density of the municipality increases. This might indicate an agglomeration effect. However, a large part of the quadrant has no significant arrows, indicating a non-clear direction for the municipalities

[8]The length of each arrow has been divided by a factor 5 for graphical purposes.

[9]The standard errors of the estimated arrows are derived from nonparametric block bootstrap.

located in that part of the Moran space, although some of the arrows indicate the presence of another attraction point.

In the LL quadrant, we observe a convergence along the nonparametric Moran as in the HH quadrant. We observe a decrease in the density of the municipalities (arrows point to the left) showing a tendency for depopulation. In particular, we may notice the presence of an attractor for the LDMS around (0,3). However, some divergent behaviours appear: points close to the average level of the x axis tend to escape from the quadrant. In particular, in the LL-B sector, it clearly appears a separating line where municipalities above a certain threshold of density move towards the HL quadrant. The same phenomenon is present close to the average level of the y axis where some arrows are pointing towards the LH quadrant.

To detect more precisely the presence of the attraction points, we allow the observations of 2019 in the Moran space to evolve according to the estimated arrows for a long time. This allows us to understand where each municipality will converge thus identifying both the attraction points and their basin of attraction.[10] In Figure 8b we colour the observations of 2019 according to the attractor point to which they converge and we map them in Figure 8c.[11] Moreover, in the Moran scatterplot attractor points are circled with the corresponding colour and all the estimated arrows are reported. Our methodology identifies three attractor points: i) a *urban attractor* around (8.5,8) in the HH quadrant; ii) a *suburban attractor* around (2,5) in the LH quadrant; and iii) a *rural attractor* around (0,3) in the LL quadrant. The basin of the urban attractor is composed of both the HH and HL quadrants as well as about half of the LH quadrant. The basin of the suburban attractor expands both in the LH quadrant and the LL above the nonparametric Moran's I. Finally, the basin of the rural attractor is totally included in the LL quadrant and almost coincides with the LL-B and LL-M sectors. We also observe that the nonparametric Moran's I is approximately the separating line between the basin of attraction of the suburban and rural attractors. Moreover, in Figure 8b is also possible to detect the other separating lines which discern observations in the LL-B sector (rural vs urban) as well as in the LH quadrant (suburban vs urban).

Overall there are three attractors where municipalities and their LLA are converging, one in the HH quadrant, one in the LL quadrant (both located around the nonparametric Moran's

[10] In this step we consider all the arrows including the non-significant ones in order to prevent the formation of artificial equilibrium points induced by the absence of an arrow.

[11] Missing values in the map are due to change of municipal boundaries from 1984 to 2019.

I) and one in the LH quadrant. This highlights a tendency toward the formation of three settlement systems. Notice that only the metropolitan attractor can be approximately detected using the transition matrix. In Section 4 we find the convergence toward the HH-M state while no clear evidence of convergence to a specific state is detected either in the LL or in the LH quadrants. Our methodology overcomes this limitation by estimating an arrow at every point of the Moran scatterplot.

Figures/RVF/RVF_1984-2019.pdf

(a) LDMS 1984-2019

Figures/RVFAttractors-eps-converted-to.pdf

(b) Basins of attractors

Figures/MapAttractors-eps-converted-to.pdf

(c) Map of attractors

Figure 8: Estimated LDMS

Figures/RVF/RVF_4Munic1984-2019.pdf

(a) 1984-2019

Figures/RVF/RVF_2001-2002.pdf

(b) 2001-2002

Figure 9: Estimated LDMS

The previous analysis has been carried out considering two different definitions of LLA. As explained in Section 2 we use the LLA definition of 2001 for the years 1984-2001 and that of 2011 for the years 2002-2019. To understand how much our estimations depend on this change in Figure 9b we report the LDMS estimated only on the adjacent years 2001 and 2002. This should minimise any effect of the change in municipal density due to the short time span, therefore isolating the effect of the change in LLA definition. In Figure 9b almost all the arrows are vertical (either upwards or downwards) reassuring us that the change in municipal density between 2001 and 2002 is negligible. Moreover, all the arrows point to the nonparametric Moran's I. This implies that the convergence effect described above arises from the change in the LLA definition.

5.1 Subperiod analysis

Figure 9b has shown us the impact that the endogenous change of the LLAs has on the full dynamics over the period 1984-2019. To disentangle the latter from the change in municipal density we also estimate the LDMS over subperiods in which there is no change in the LLA definition. Moreover, from Figure 2b we also notice that around the year 2013, there has been a trend inversion in the total population time series. This suggests analyzing the dynamic in the subperiods also to understand how much this change in trend is heterogeneous among geographical units. The choice of the sub-periods follows the following rationale: to use the most accurate data we chose as the final year every year of the census, except for 2019 where the continuous census was in place (1991,2001,2011,2019). The length of each subperiod is taken commonly to all of them so that in each of them there is no change in the LLA definition. Therefore the subperiods are 1984-1991, 1994-2001, 2004-2011, and 2012-2019.

Figures 10a-10d report the estimated LDMS in the subperiods. Differently from the estimated LDMS in Figure 8a there is no clear attractor point in any of the estimated LDMS in the subperiods. However, there is clearly the presence of a repellor point, which in figures 10a-10c is located in the intersection of the mean values along both axes (dashed lines), while in the last subperiods, Figure 10d, is located in the HH quadrant. Moreover, almost all the arrows are pointing radially from the repellor points. The change in the trend of the total population can also be grasped from this last LDMS since most of the arrows have their horizontal component pointing to the left (decrease in density). The absence of convergence towards the nonparametric Moran's I in all subperiods is in line with the fact that within each subperiod there is no change in the LLA definition.

The dynamic of each subperiod is strongly different from the dynamic of the full period, and in particular, the one of the last subperiod is the one that differs the most. Under the hypothesis that the dynamics of the last subperiod 2012-2019 will persist in the next years, we look at a forecasted municipal density distribution in 2033. Given that the subperiod lasts seven years, the forecasted distribution in 2033 is derived by applying the estimated (significant) arrows twice: first to the initial distribution in 2019 to compute that in 2026, then again to the latter to find the final distribution in 2033.

Looking at Figure 11a the geographical pattern is the same as the one in 2019. However, Figure 11b highlights in agreement with the above results on the LDMS that only a few municipalities show an increase in their density which are primarily located in Po Valley and around metropolitan cities. On the other hand, the majority of municipalities experience a decrease in their density, especially in the inner areas that decrease the most. If we look at the Moran taxonomy in 2033 in Figure 11c no change appears with very few exceptions (Cagliari and Bolzano are now brown instead of wine). This result is in line with the estimated LDMS in Figure 10d where arrows are very short indicating a very slow dynamic.

Figures/RVF/RVF_1984-1991.pdf

(a) 1984-1991

Figures/RVF/RVF_1994-2001.pdf

(b) 1994-2001

Figures/RVF/RVF_2004-2011.pdf

(c) 2004-2011

Figures/RVF/RVF_2012-2019.pdf

(d) 2012-2019

Figure 10: Estimated LDMS for subperiods

Figures/map_popDensSLL_FORECAST2033.pdf

(a) Forecasted ditribution in 2033

Figures/mapDiffLog2019_2033.pdf

(b) Difference between forecasted 2033 and observed 2019

Figures/map_popDensSLL_UpperLower_F

(c) Moran taxonomy forecast 2033

Figure 11: Forecasted distribution

6 Conclusions

The distribution of people across settlements of different sizes is well-established in the literature. Italy in particular shows a very peculiar structure due to historical and geographical constraints (Accetturo and Mocetti, 2019). Understanding the historical track that led to the present layout is crucial to improve social policies in the future. In this paper, we propose a classification of Italian municipalities which is only based on their population density and the definition of neighbours. This constitutes the main advantage of our methodology with respect to other classifications since it allows us to gather data for a long period to study the evolution of the taxonomy. Thanks to the estimation of the LDMS in Moran space we are also able to study the complete dynamics of the population density keeping into account spatial dependence. Combining it with the proposed taxonomy also allows us to predict the expected taxonomy in the future, therefore forecasting not only the population density of each municipality but also the correlation structure with the municipalities in the same LLA.

We show that our taxonomy resembles the one of Barca et al. (2014) which is based on other data and explicitly proposed for the national strategy for the inner areas. In particular, our taxonomy is able to distinguish between urban/rural areas but also to identify exceptions both within metropolitan cities (as for Venice, Naples and Catania) as well as the main poles of attraction of peripheral areas. We also identify LLAs characterised by municipalities with the highest density, big LLAs with metropolitan cities which also include municipalities with low density and rural LLAs characterised by municipalities with low density. When we analyse the changes in the taxonomy between 1984 and 2019 through the transition matrix, we find evidence of convergence towards the HH-M state from the two adjacent states HH-T and HH-B. This convergence seems to be driven by different mechanisms: while the convergence from HH-B is due to an increase in the average density within the LLAs, the transition from HH-T arises from the geographical expansion of the LLAs. Differently, in the case of LL-* block there is no clear evidence of convergence towards a specific state. Finally, the comprehensive spatiotemporal analysis through the LDMS in the Moran scatterplot reveals the presence of three attractors, an urban, a suburban and a rural one. The technique is also able to tell us which municipalities converge to which attractor and to forecast their future path.

The methodology can be easily applied to study other economic/social phenomena characterised by both spatial and temporal dynamics. In particular, one can use the Moran-based taxonomy to perform an exploratory analysis able to characterise the unit under observation into subgroups and study its evolution over time using the transition matrix. This spatiotemporal analysis can be subsequently refined through the estimation of the LDMS which allows studying the dynamics of the entire distribution without dividing it into discrete states. This innovative methodology can provide deeper insights into a broad range of economic and social phenomena for more effective decision-making and policy planning.

References

Accetturo, A. and Mocetti, S. (2019). Historical origins and developments of italian cities. *Italian Economic Journal*, 5(2):205–222.

Anselin, L. (2019). The moran scatterplot as an esda tool to assess local instability in spatial association. In *Spatial analytical perspectives on GIS*, pages 111–126. Routledge.

Barca, F., Casavola, P., and Lucatelli, S. (2014). A strategy for inner areas in italy: Definition, objectives, tools and governance. Materiali UVAL-Documenti: Vol. 31: UVAL.

Dijkstra, L., Poelman, H., and Veneri, P. (2019). The eu-oecd definition of a functional urban area.

Fiaschi, D., Gianmoena, L., and Parenti, A. (2018). Spatial club dynamics in european regions. *Regional Science and Urban Economics*, 72:115–130.

Glaeser, E. and Cutler, D. (2021). *Survival of the City: Living and Thriving in an Age of Isolation*. Penguin.

Glaeser, E. L. (2010). *Agglomeration economics*. University of Chicago Press.

ISTAT (2017). Livelli e dinamiche dell'urbanizzazione in italia.

Krugman, P. (1994). Complex landscapes in economic geography. *The American Economic Review*, 84(2):412–416.

Lamorgese, A. and Petrella, A. (2019). Stylized facts on italian cities. *Italian Economic Journal*, 5:223–249.

Manzoli, E. and Mocetti, S. (2019). The house price gradient: evidence from italian cities. *Italian Economic Journal*, 5:281–305.

Peiser, R. and Hugel, M. (2022). Is the pandemic causing a return to urban sprawl? *JCULP*, 5:26.

Perchinunno, P., d'Ovidio, F. D., and Rotondo, F. (2019). Identification of "hot spots" of inner areas in italy: Scan statistic for urban planning policies. *Social Indicators Research*, 143(3):1299–1317.

A List of metropolitan cities

Municipality	Density 1984	Population 1984	Density 2019	Population 2019
Rome	7.70	2844903	7.69	2820219
Milan	9.06	1560155	8.95	1395980
Naples	9.23	1207337	8.99	954318
Turin	9.04	1097355	8.80	860793
Palermo	8.40	711194	8.31	652720
Genoa	8.05	748634	7.77	569184
Bologna	8.06	447893	7.93	393248
Florence	8.37	442345	8.19	369885
Bari	8.05	368772	7.90	316491
Catania	7.64	380564	7.40	297752
Venice	6.70	339883	6.44	259961
Messina	7.08	254951	6.98	229280
Reggio di Calabria	6.61	177807	6.60	176299
Cagliari	7.93	236165	7.49	151504

Table 2: List of the 14 Italian metropolitan cities