Spatializing Shannon Entropy: A Gaussian Kernel Approach to Studying the Territorial Distribution of Selected Foreign Population Groups in Italy

Massimo Mucciardi[1], Federico Benassi[2]

[1] University of Messina - Department of Cognitive Science, Education and Cultural Studies, University of Messina, Messina, Italy. ORCID: 0000-0003-1216-493

[2] University of Naples Federico II (Federico.benassi@unina.it) - Department of Political Sciences, University of Naples Federico II, Naples, Italy. ORCID: 0000-0002-8861-9996

Abstract – The paper proposes an original extension of the Shannon Entropy index that incorporates the spatial dimension in measuring the territorial distribution of population groups. This methodological advancement is achieved through the use of a Gaussian Kernel approach, which enhances the utility of the Shannon Entropy index, particularly for processes that are inherently spatial, such as residential segregation and related phenomena. An empirical application is presented, focusing on the spatial distribution of selected foreign groups residing in Italy. The results highlight notable characteristics of the index and suggest steps toward new approaches for measuring the territorial distribution of populations.

Key words: spatial entropy, spatial demography, foreign populations, diversity.

Acknowledgement: The paper has been conceived and realized as a part of the PRIN 2022-PNRR research project "Foreign population and territory: integration processes, demographic imbalances, challenges and opportunities for the social and economic sustainability of the different local contexts (For.Pop.Ter)" [P2022 WNLM7], funded by European Union-Next Generation EU, Component M4C2, Investment 1.1. The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

This preprint **has not undergone peer review or any post-submission improvement or corrections.** The version of record of this article is published in Quality & Quantity and is available online at https://doi.org/10.1007/s11135-025-02241-4.

At this link https://rdcu.be/euy3O it is available a full-text view-only version of the published paper.

1. Introduction

The spatial distribution of human populations has long been a central focus in applied demography and social sciences (Duncan 1957; Duncan et al. 1963; Voss 2007; Cohen et al. 2013; White 2016). One prominent research area is residential segregation, which refers to the physical separation of population groups within a given environment (Massey and Denton 1988). Since the early work of the Chicago Ecological School (Park et al. 1925; Park 1926), scholars have refined methods, measures, and scales for studying segregation (Duncan and Duncan 1955; Massey and Denton 1988; Wong 1999; Reardon and Firebaugh 2002; Reardon and O'Sullivan 2004; Feitosa et al. 2007; Roberto 2018; Yue et al. 2024).
Ethnic diversity is recognized as a critical factor shaping segregation and settlement patterns (Benassi et al. 2020a). This is particularly evident in urban and metropolitan areas, where immigrant populations face higher levels of concentration, marginalization, and segregation (Tammaru et al. 2021; Benassi and Iglesias-Pascual 2023). Such contexts highlight the multi-ethnic nature of contemporary urban societies (Reardon and Firebaugh 2002), where lower diversity—fewer distinct groups—often correlates with higher segregation levels (Benassi et al. 2020a). The ethnic diversity is quite often associated with the concept of entropy (McCulloch 2007; Mora and Ruiz-Castillo 2011; Harris 2020; Zachary and Dobson 2021; Steele et al. 2022).
Entropy, introduced by Shannon (1948), has been widely adapted across disciplines, including segregation studies (Mora and Ruiz- Castillo 2011; Guevara et al. 2016; Kramer and Kramer 2019). However, its original formulation was non-spatial, limiting its application to inherently spatial phenomena such as residential segregation (Reardon and O'Sullivan 2004). Efforts to develop spatially explicit versions of entropy began with Batty in the 1970s (Batty 1974, 1976) and have been refined over the years (Batty et al. 2014; Altieri et al. 2018a, 2018b, 2019). Building on this tradition, our study introduces an innovative approach to spatializing Shannon entropy by employing a spatial neighborhood function based on a Gaussian kernel. This function prioritizes the influence of nearby spatial units while gradually reducing the impact of more distant ones. This method transforms entropy from an a-spatial measure into one that integrates spatial interactions, offering a more nuanced framework for analyzing segregation. A key contribution of this study is the concept of "Lagged Spatial Entropy" (*LSE*), which compares the entropy of a specific territorial unit (a-spatial entropy) with that of its surrounding areas. This dual perspective captures spatial heterogeneity, revealing patterns of variability and information diffusion across spatial systems. The enriched framework proves particularly valuable for applications in demography, geography, and urban planning, where spatial dependencies are critical.
The empirical focus of this study is the diversity of foreign population in Italy in 2021. The analysis, conducted at the municipal level, spans approximately 7,900 administrative units—the most detailed scale provided by the Italian Institute of Statistics and takes into consideration the 10 most numerous foreign communities resident in Italy.
While diversity is fundamentally an urban phenomenon best analyzed at an intra-urban scale (Feitosa et al. 2007), this study adopts the municipal scale to address challenges such as small population sizes and the Modifiable Areal Unit Problem (MAUP), which often complicate sub-municipal analyses (Wong 2024). The municipal scale balances granularity with analytical feasibility, aligning with best practices in spatial analysis (Fotheringham and Sachdeva 2022). The remainder of this paper is organized as follows.
The next section details the methodology for spatializing Shannon entropy. Section 3 presents the empirical application to Italy, and the final section discusses the findings and their implications for segregation studies.

2. Spatialize Shannon entropy using a Gaussian kernel approach

Shannon entropy (1948) quantifies the randomness or uncertainty inherent in a set of probabilities or relative frequencies, making it a versatile tool for analyzing variability across diverse contexts. High entropy signifies a more uniform or random distribution, whereas low entropy indicates concentrated patterns. The entropy formula employed in this paper is a modified version of Shannon's original formula and can be described as a-spatial. Consider k populations or sub-population groups residing within a spatial unit i (such as a census tract, district, municipality, or region). For the *i-th* spatial unit, the entropy formula can be expressed as follows:

$$H_i = -\sum_{g=1}^{k} p_g log(p_g) \tag{1}$$

($i = 1 \dots . n$ and $g = 1 \dots . k$)

with $\sum_{g=1}^{k} p_g = 1$, where p_g is the probability (or relative frequency) of the g-th population[1], H_i reaches its maximum when all occurrences occur with equal frequency, quantified as $log(k)$. This measure can be standardized by dividing by it by $log(k)$, yielding an entropy value ranging between 0 and 1.

Our objective is to measure the entropy within the units j that are 'near' (i.e., in the neighborhood of) unit i. To achieve this, we apply an appropriate spatial neighbourhood function to summarize the distribution of the K populations residing outside unit i. Spatial relations between units j outside i are modelled using a spatial weigth matrix $\boldsymbol{W}(h)$, based on the Gaussian kernel function (Otranto et al., 2016).

This matrix incorporates the distance d_{ij} between each pair of spatial units i and j. The weight w_{ij} is computed using the following function:

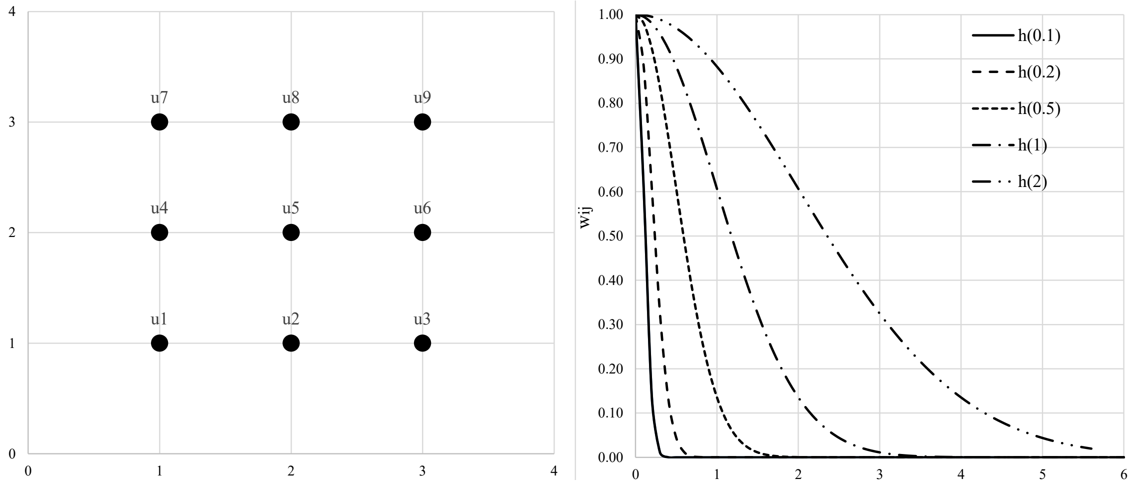
$$w_{ij} = exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right] \tag{2}$$

where d_{ij} represents the distance between the generic point units i and j, and h, known as the bandwidth, is a non-negative parameter that controls the rate of decay of influence with distance. When i and j coincide, the weight w_{ij} takes how rapidly the weights decrease over space (Figure 1). Specifically, for spatial units j located far from i the weight w_{ij} approaches to zero. To ensure consistency, the spatial weighted matrix is standardized so that all rows sum to one[2]. The use of the Gaussian kernel function addresses a fundamental issue in geographical data analysis, known as topological invariance. This problem occurs when the same spatial weight matrix is derived for different topological configurations of spatial units (Dacey 1965). The Gaussian kernel function, however, is sensitive to topological changes, making it a robust tool for capturing spatial variability.

[1] When the relative frequency of a group in a given spatial unit is equal to 0, then the log is set to 0 because $\lim_{x \to 0}(x\ logx) = 0$. All logarithmic calculations use the natural log.

[2] We employed row-standardized weights and excluded the diagonal elements from the calculation.

Figure 1. Example of Gaussian kernel functions for bandwidth values 0.1, 0.2, 0.5, 1 and 2 (right panel) applied to a simulated lattice grid with nine spatial units (left panel).



Source: our elaboration

By applying the Gaussian kernel function as defined in Equation (2), we construct a $n \times n$. symmetric matrix $\boldsymbol{W}(h)$. Figure 2 illustrates an example of a matrix from this formula.

Figure 2. Example of matrix $\boldsymbol{W}(h)$ for $h = 0.1$ applied to a simulated lattice grid with nine spatial units

Unit	u1	u2	u3	u4	u5	u6	u7	u8	u9	Sum
u1	0	**0.298**	**0.067**	**0.298**	**0.181**	**0.040**	**0.067**	**0.040**	**0.009**	1
u2	**0.212**	0	**0.212**	**0.129**	**0.212**	**0.129**	**0.029**	**0.047**	**0.029**	1
u3	**0.067**	**0.298**	0	**0.040**	**0.181**	**0.298**	**0.009**	**0.040**	**0.067**	1
u4	**0.212**	**0.129**	**0.029**	0	**0.212**	**0.047**	**0.212**	**0.129**	**0.029**	1
u5	**0.094**	**0.156**	**0.094**	**0.156**	0	**0.156**	**0.094**	**0.156**	**0.094**	1
u6	**0.029**	**0.129**	**0.212**	**0.047**	**0.212**	0	**0.029**	**0.129**	**0.212**	1
u7	**0.067**	**0.040**	**0.009**	**0.298**	**0.181**	**0.040**	0	**0.298**	**0.067**	1
u8	**0.029**	**0.047**	**0.029**	**0.129**	**0.212**	**0.129**	**0.212**	0	**0.212**	1
u9	**0.009**	**0.040**	**0.067**	**0.040**	**0.181**	**0.298**	**0.067**	**0.298**	0	1

Source: our elaboration

By fixing a bandwidth h and considering n spatial units and k populations (or groups), we can readily compute the values of a matrix $\boldsymbol{L}(h)$ representing the lagged spatial values:

$$\boldsymbol{L}(h)_{nk} = \boldsymbol{W}(h)_{nn} * \boldsymbol{D}_{nk} \quad (3)$$

here $\boldsymbol{D}_{nk}$ denotes the data matrix comprising n spatial units and k populations. The generic row i of this matrix provides the spatially lagged value l_g of the k populations for the spatial unit i, enabling the calculation of probabilities (or relative frequencies) in accordance with Shannon's entropy expressed in Equation (1):

$$H_{w_i}^h = -\sum_{g=1}^{k} pw_g log(pw_g) \quad (4)$$

$(i = 1 \dots . n$ and $g = 1 \dots . k)$

where $pw_g = \frac{l_g}{\sum_{g=1}^{k} l g}$ is the lagged relative frequency of the $g\text{-}th$ populations in the spatial units i (with $\sum_{g=1}^{k} pw_g = 1$). In this paper we refer to Hw_i^h as "Lagged Spatial Entropy" ($LSE(h)$), calculated for a specific h-bandwidth. In this context, $LSE(h)$ reaches its maximum value when all occurrences are equally lagged, quantified as $log(k)$. This measure can be standardised by dividing it by $log(k)$, yielding a value of $LSE(h)$ that ranges between 0 and 1[3].

Given that population evenness/diversity ranges from a minimum of 0 (indicating dominance by a single group) to a maximum of 1 (indicating equal representation of all k-groups), spatial units can be categorized by entropy type, as outlined in Table 1.

Table 1. Entropy classification for i-spatial unit

Value of Entropy	Significance within the framework of population evenness distribution	Label
$H_i \geq 0.5$ *and* $LSE(h) \geq 0.5$	The population is more evenly distributed among the groups in the unit i and its neighborhood	HH
$H_i < 0.5$ *and* $LSE(h) < 0.5$	One or a few groups dominate the population in the unit i and its neighborhood	LL
$H_i \geq 0.5$ *and* $LSE(h) < 0.5$	The population is more evenly distributed among the groups in the unit i but one or a few groups dominate its neighborhood	HL
$H_i < 0.5$ *and* $LSE(h) \geq 0.5$	One or a few groups dominate the population in the unit i but the population is more evenly distributed among the groups in its neighbourhood	LH

Source: our elaboration

3. Empirical application

In this section, we present an original application of the spatial version of the Shannon entropy index (Hw_i^h). We begin by describing the data utilized, proceed with an explanation of the Gaussian kernel adopted, and conclude with a presentation of the results obtained.

3.1 Demographic and geographical data

We analyze the top ten foreign groups residing in Italy as of January 1, 2022, identified based on their country of citizenship. For each of the nearly 7,900 Italian municipalities, we examine the ten nationalities listed in Table 2. This dataset captures a highly diverse and heterogeneous set of foreign population groups, including EU communities (e.g., Romanian nationals), communities

[3] In this study we consider the two entropy measures (H_i*and* $LSE(h)$) normalized by dividing by their maximum. Therefore, to avoid complicating the writing, we shall leave the symbols unchanged.

with strong cultural and historical ties to Italy (e.g., Albanian nationals), and long-standing foreign groups with an established history of immigration to Italy, such as Moroccans, Filipinos, and Egyptians (Strozza and De Santis, 2017). The data are provided by the Italian National Institute of Statistics (ISTAT) and freely available for download on its official website.

Additionally, the dataset encompasses relatively recent immigrant communities, such as those from Pakistan, Bangladesh, and India. Lastly, the Ukrainian community, which belongs to Eastern Europe, stands out due to its distinctive distribution across Italy (Benassi et al., 2020b).

The selected communities demonstrate a wide range of settlement patterns, demographic profiles, and immigration histories in Italy (Conti et al., 2023). The Romanian community, the largest among them, accounts for 21.5% of the total foreign population in the country. As Romania is a member of the European Union, it represents the only EU-nationality group among the ten. Romanians exhibit a dispersed settlement pattern, residing throughout the entire Italian territory (Benassi et al., 2019).

The Moroccan community, ranking second, has a long-standing immigration history in Italy and is primarily concentrated in the northern regions. Similarly, the Egyptian community, another North African group, is predominantly located in major urban centers such as Milan and Rome (Conti et al., 2023).

The remaining groups consist primarily of Asian communities, including Chinese, Indian, Bangladeshi, Filipino, and Pakistani populations. Among these, some groups (e.g., Chinese and Filipinos) have a long tradition of immigration to Italy and exhibit predominantly urban and metropolitan settlement patterns (Bitonti et al., 2023). Others, having arrived more recently, display more varied settlement behaviors.

Finally, the Ukrainian community is particularly notable. While their presence is especially prominent in southern Italy, their settlement pattern can be characterized as pseudo-diffuse.

Figure 3 provides a detailed visualization of the geographical distribution of these ten major foreign communities.

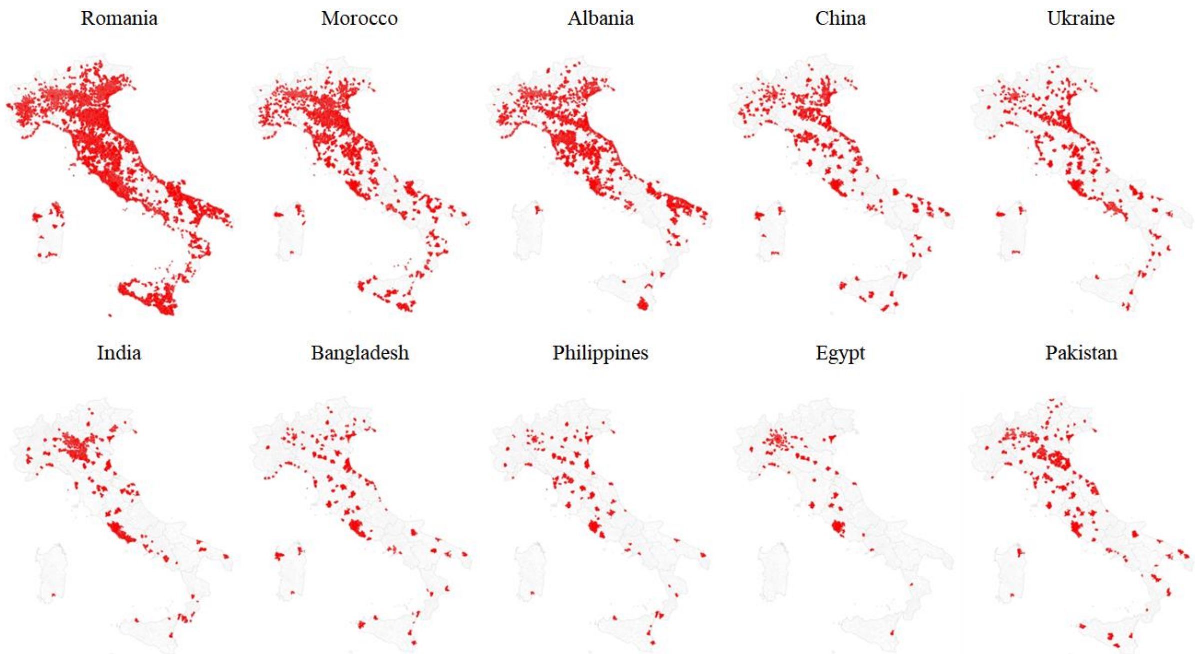
Geographical data refers to the shapefiles of municipalities, which are vector data provided by the Italian National Institute of Statistics (ISTAT) and freely available for download on its official website. Using these vector data, we computed the geographical barycenter for each municipality, defined as a point identified by its geographical coordinates. This step is essential for determining the bandwidth (see Section 1, Figure 1).

Table 2. Total foreign population and first ten foreign communities resident in Italy (January 1, 2022). Absolute values and percentage values.

Population groups	Absolute values	Percentage composition (%)
Romania	1,083,771	21.5%
Morocco	420,172	8.4%
Albania	419,987	8.4%
China	300,216	6.0%
Ukraine	225,307	4.5%
India	162,492	3.2%
Bangladesh	159,003	3.1%
Philippines	158,997	3.1%
Egypt	140,322	2.8%
Pakistan	134,182	2.7%
Other foreigners	*1,826,267*	*36.3%*
Total foreign population	*5,030,716*	

Source: our elaboration on ISTAT data

Figure 3 – Geographical distribution for the first ten foreign citizens throughout Italian municipalities (only municipalities with a minimum of 100 foreign residents are represented).

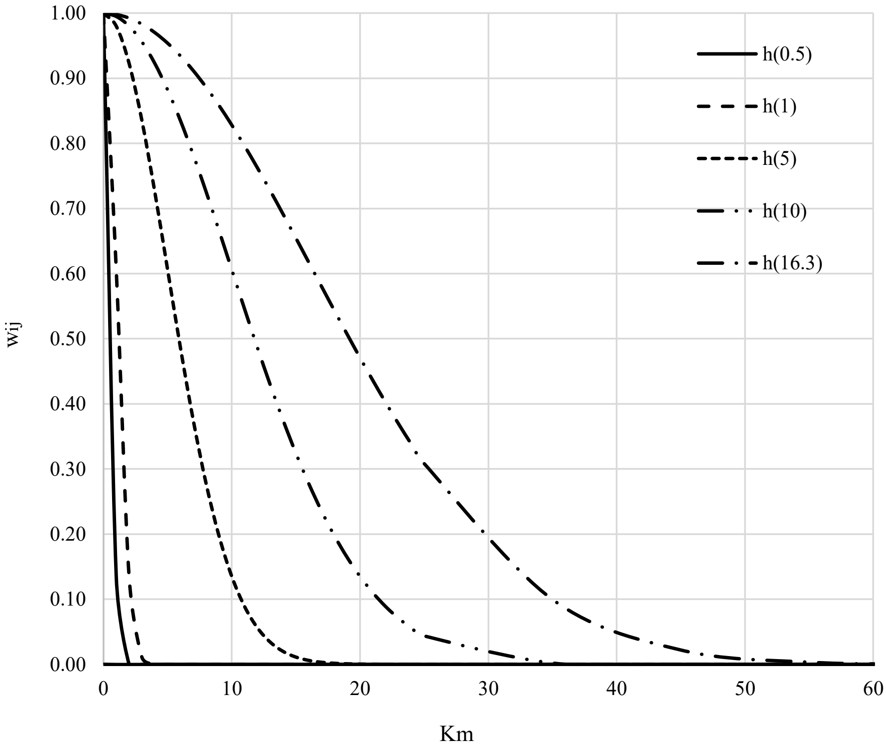


Source: our elaboration on ISTAT data

3.2 *Gaussian kernel configuration*

The distance d_{ij} between the generic units i and j in Equation 2 is computed based on the geographical coordinates (latitude and longitude) of the territorial barycenters of each municipality[4]. Regarding the bandwidth setup for the Gaussian kernel function, we consider a range of distances from the minimum inter-municipal distance (h_{min}=0.257 km) and the MaxMin distance (h_{Mm}=16.3 km), where the entire territory is fully interconnected and each municipality has at least one neighbor, ensuring no isolated units (Mucciardi, 2008). Among the different bandwidths tested, we select the following for the analysis: $h_{0.5}$ = 0.5 Km, h_1 = 1 Km, h_5 = 5 Km, h_{10} = 10 Km and h_{Mm}= 16.3 Km[5], as shown in Figure 4.

Figure 4. Gaussian kernel functions applied to the territorial barycentre of the Italian municipalities territory for bandwidth values 0.5, 1, 5, 10 and 16.3 km (in brackets values in kilometres)



Source: our elaboration on ISTAT data

3.3 *Results*

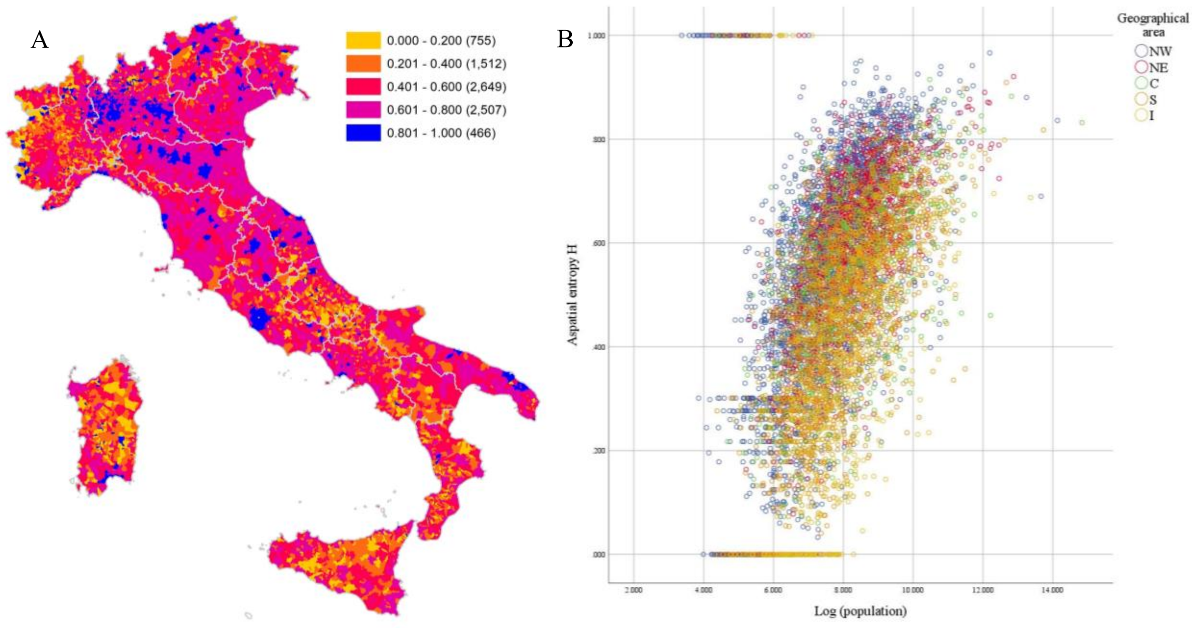
Now let's illustrate the principal findings of the empirical application. Figure 5 (panel A) displays the a-spatial entropy map (H_i). The normalised Shannon Entropy formula was applied to each municipality, yielding values ranging from 0 to 1. The values are classified into five categories to distinguish municipalities with low entropy (unevenness foreign population/low diversity) from those with high entropy (evenness foreign population/high diversity). Figure 5 (panel B), on the other hand, illustrates the relationship between a-spatial entropy and the logarithm of the population size of each municipality. The underlying hypothesis posits that the level of diversity increases with the demographic size of a municipality, as the latter serves as a proxy for the area's multi-ethnic composition. The scatterplot also displays the geographical areas of Italy to which belong each municipality. From a preliminary evaluation, 5.9% (466 out

[4] The barycenter of each municipality is calculated using ArcGIS, based on the shapefile provided by ISTAT (2024). The subsequent processing was carried out in Microsoft Excel and STATA.

[5] Beyond the distance h_{Mm} the LSE results do not vary significantly while we discard h_{min} since it generates a very narrow bandwidth. However, all additional bandwidth distance processing is available on request.

of 7,889) of the municipalities exhibit a non-spatial entropy level greater than 0.80, indicating a more evenly distributed foreign population, typically in proximity to the major urban areas. Conversely, 9.6% of municipalities (755 out of 7,889) exhibit low entropy values (less or equal than 0.20), suggesting a modest level of diversity among the foreign population. As shown in figure 5 (panel A), these municipalities are almost always located in Italy's mountainous areas. The remaining municipalities are distributed among medium-high entropy levels (31.8% between 0.60 and 0.80), primarily in the areas of central and northern Italy; medium entropy levels (33.6% between 0.40 and 0.60); and medium-low entropy levels (19.2% between 0.20 and 0.40), which are mostly found in central and southern Italy.Figure 5 (panel B), which relates non-spatial entropy to municipal population size, confirms corroborates the observations in Figure 5 (panel A). Overall, entropy increases with municipal population size, and this trend is more pronounced in municipalities in north-western (NW) and north-eastern (NE) Italy, where the resident foreign population is comparatively higher that in the Centre (C), South (S), and Islands (I). On this last point it is useful to recall that as of January 1, 2024, the resident foreign population in Italy totaled 5,307,598. Of this population: 1,815,107 foreigners resided in the North-West (34.2%), 1,293,774 in the North East (24.4%), 1,301,296 in the Centre (24.5%), 644,203 in the South (12.1%), and 253,218 on the Islands (4.8%).

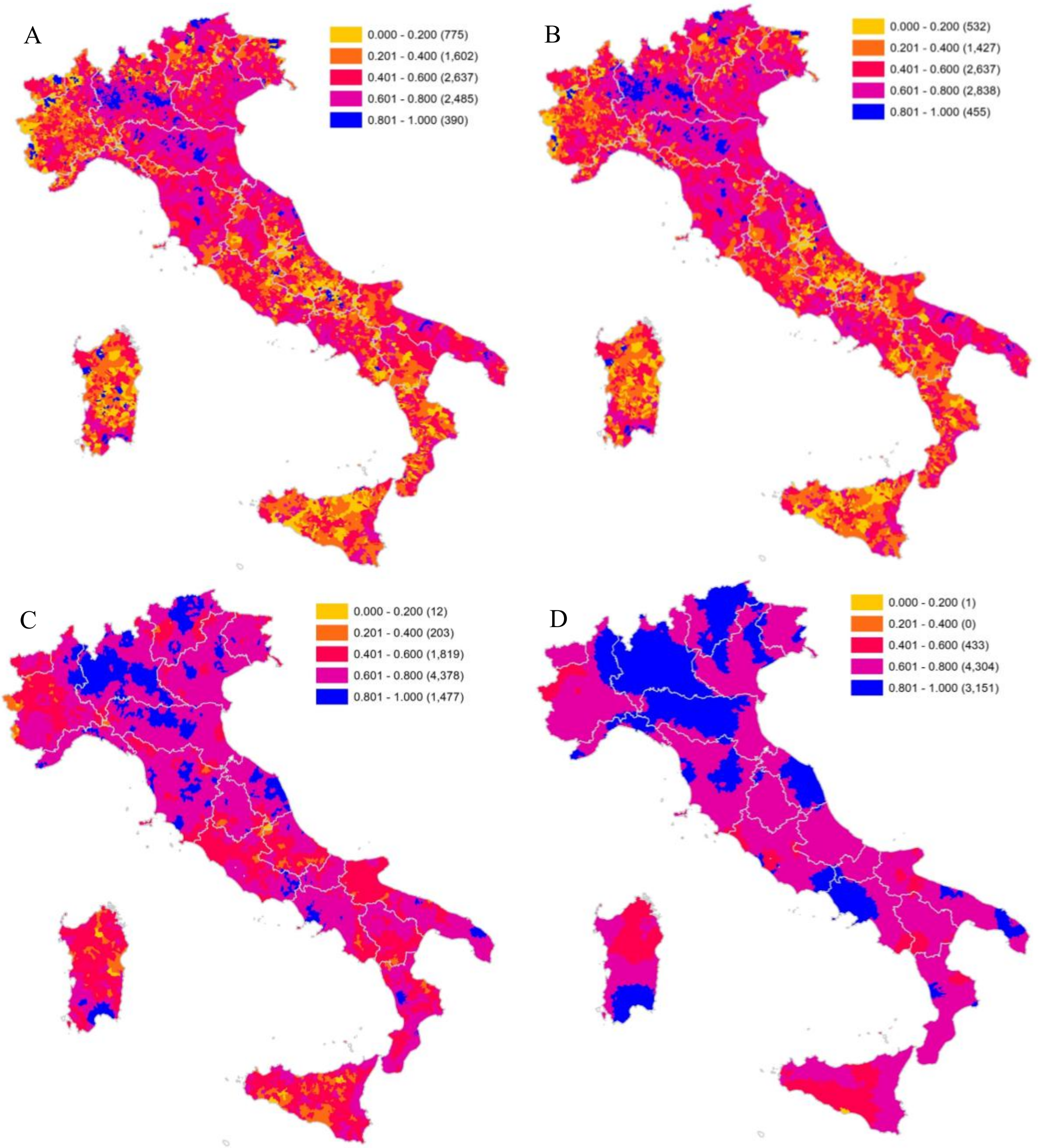
Figure 5. (A) Value (in classes) of non-spatial entropy H. In brackets the number of municipalities in each class. (B) Scatterplot between logarithm of resident population and non-spatial entropy H in each municipality by geographical macro-area. (NW= North-West; NE= North-Est; C= Centre; S=South; I=Islands).



Source: our elaboration on ISTAT data

Let us now examine the results when Local Spatial Entropy (LSE) is applied. Figure 6 presents the LSE values for the specified bandwidths. A global analysis of the maps reveals that entropy generally increases across municipalities as bandwidth widens. This trend can be attributed to the "smoothing effect" of the Gaussian kernel: at smaller bandwidths (e.g., 0.5 and 1 km), the influence of nearby municipalities is stronger, whereas at larger bandwidths (e.g., 5 and 16.3 km), the kernel incorporates more distant municipalities, diminishing the relative influence of closer ones. Larger bandwidths are particularly important for capturing the broader spatial trends in entropy.

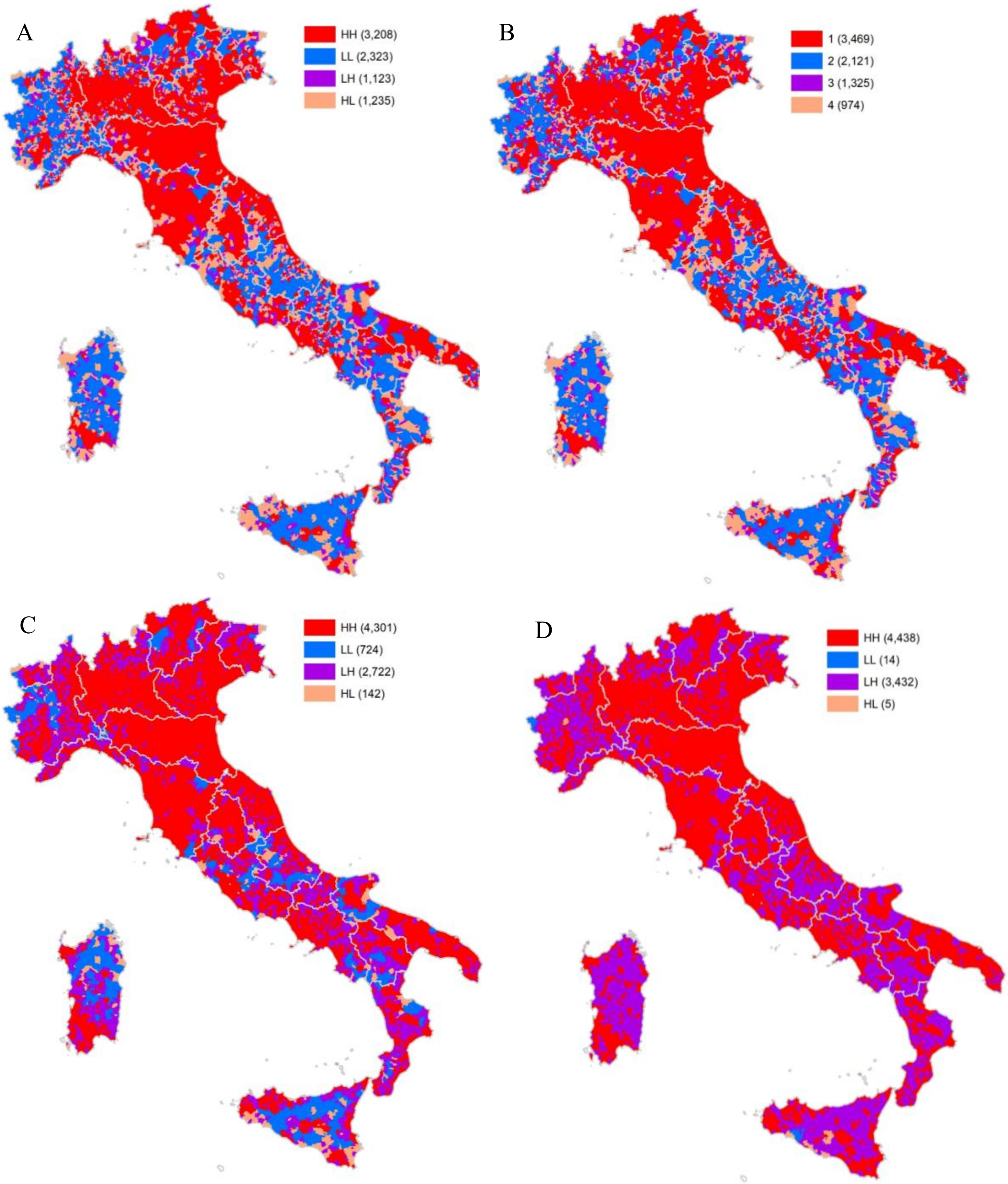
Figure 6. Value (in classes) of $LSE(h)$: (A) h= 0.5 Km; (B) h= 1 Km; (C) h= 5 Km; (D) h= 16.3 Km. In brackets the number of municipalities in each class.



Source: our elaboration on ISTAT data

In Figure 7 (panels A to D), an expansion of entropy values within the highest classes (violet and blue) is observed, particularly in Northern Italy and in some metropolitan areas of the Southern Italy like Cagliari, Naples, and Bari. This trend also reflected in the categorization presented in Table 1. As the bandwidth increases, the number of municipalities classified in the "HH" category rises, while those in the "LL" category decline.

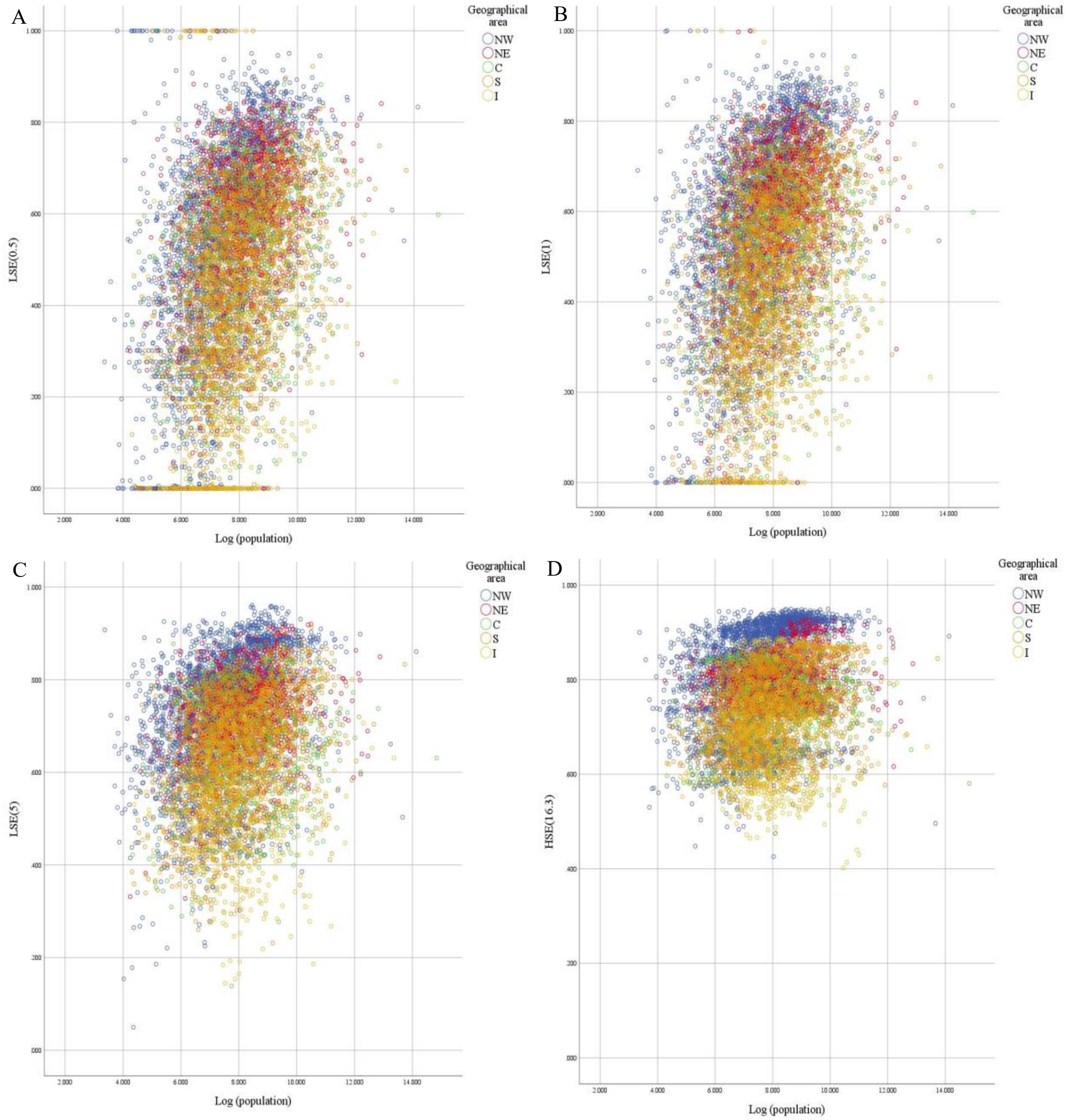
Figure 7. Entropy classification based on the thresholds in table 1 for the first ten foreign citizens on the 7889 municipalities: (A) h= 0.5 Km; (B) h= 1 Km; (C) h= 5 Km; (D) h= 16.3 Km. In brackets the number of municipalities in each class.



Source: our elaboration on ISTAT data

Figure 8 (panels A to D) further illustrates the relationship between $LSE(h)$ and the logarithm of municipal population size. The influence of the Gaussian kernel is clearly evident: at smaller bandwidths (e.g., 0.5 and 1 km), the scatterplot appears elongated, whereas at larger bandwidths (e.g., 5 and 16.3 km), it takes on a more rounded shape. Geographically, Northern Italy consistently displays higher entropy values compared to Southern Italy.

Figure 8. Scatterplot between logarithm of resident population and $LSE(h)$ in each municipality by geographical area: (A) h = 0.5 Km; (B) h = 1 Km; (C) h = 5 Km; (D) h = 16.3 Km.



Source: our elaboration on ISTAT data

5. Discussion and Conclusions

Contemporary societies in developed countries are increasingly characterized by multi-ethnic compositions (Reardon and Firebaugh, 2002). This trend is particularly pronounced in urban areas, where the concentration of foreign migrants is higher (Strozza et al., 2016). The complexity of these global trends —where the majority of the population lives in urban areas— poses significant new challenges for developing appropriate indicators, approaches, and tools. Among these, the measurement of diversity is a critical dimension (Vertovec, 2007).

This paper addresses these challenges by introducing a novel approach to studying Shannon entropy in the context of the territorial distribution of populations. The innovation lies in applying spatial kernels, constructed using carefully defined bandwidths inspired by locally weighted regression models (Fotheringham et al., 2002). While extending Shannon entropy to spatial dimensions is not entirely new (Batty, 1974, 1976; Altieri 2018a, 2018b, 2019), this approach distinguishes itself through its methodological clarity and step-by-step replicability.

The experimental application focuses on Italian municipalities and the 10 largest foreign communities as of 2021. Results align with theoretical expectations, demonstrating a correlation between diversity levels and the demographic size of municipalities (used as a proxy for multi-ethnic complexity). The findings also reveal a North-South gradient, reflecting the higher concentration of foreign populations in Northern Italy (Strozza et al., 2016).

The index proved sensitive to bandwidth variations, underscoring its flexibility and potential for future applications at various geographical scales, such as intra-urban analysis. This adaptability supports a multiscale approach (Oshan et al., 2019), enabling analyses at both regional and neighborhood levels. At broader scales, the index illuminates regional trends in diversity, while at finer scales, such as urban neighborhoods, it provides granular insights into segregation or integration processes. The study also highlighted significant differences between spatial and non-spatial indices, with the spatial approach more effectively capturing ethnic diversity. In light of the growing complexity of European and global metropolises and their characteristic *super-diversity* (Vertovec, 2007; Benassi et al., 2023), robust measures that accurately represent these realities are increasingly important. The proposed approach is notable for its ease of implementation, making it accessible to a wide range of researchers. In conclusion, this method's versatility makes it a valuable tool for demographers, urban planners, and sociologists aiming to understand spatial patterns of diversity in modern communities.

References

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

Altieri, L., Cocchi, D., Roli, G. (2018b). Spatentropy: Spatial entropy measures in r. arXiv preprint arXiv:1804.05521.

Altieri, L., Cocchi, D., Roli, G. (2019). *Advances in spatial entropy measures*. In: Stochastic Environmental Research and Risk Assessment, 33, 1223-1240.

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

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

Batty, M., Morphet, R., Masucci, P., & Stanilov, K. (2014). Entropy, complexity, and spatial information. *Journal of geographical systems*, 16, 363-385.

Benassi, F., & Iglesias-Pascual, R. (2023). Local-scale residential concentration and income inequalities of the main foreign-born population groups in the Spanish urban space. Reaffirming the model of a divided city. *Journal of Ethnic and Migration Studies*, *49*(3), 673-696.

Benassi, F., Casacchia, O., & Strozza, S. (2020b). Demographic and migratory characteristics and settlement models of the former Soviet Union citizens residing in Italy. *Migration from the Newly Independent States: 25 Years After the Collapse of the USSR*, 423-450.

Benassi, F., Iglesias-Pascual, R., & Salvati, L. (2020a). Residential segregation and social diversification: Exploring spatial settlement patterns of foreign population in Southern European cities. *Habitat International*, *101*, 102200.

Benassi, F., Lipizzi, F., & Strozza, S. (2019). Detecting foreigners' spatial residential patterns in urban contexts: Two tales from Italy. *Applied Spatial Analysis and Policy*, *12*, 301-319.

Benassi, F., Naccarato, A., Iglesias-Pascual, R., Salvati, L., & Strozza, S. (2023). Measuring residential segregation in multi-ethnic and unequal European cities. *International Migration*, *61*(2), 341-361.

Bitonti, F., Benassi, F., Mazza, A., & Strozza, S. (2023). Framing the Residential patterns of Asian communities in three Italian cities: Evidence from Milan, Rome, and Naples. *Social Sciences*, *12*(9), 480.

Cohen, J. E., Xu, M., & Brunborg, H. (2013). Taylor's law applies to spatial variation in a human population. *Genus*, 69(1), 25-60.

Conti, C., Mucciardi, M., & Simone, M. (2023). Exploring the settlement models of the main foreign communities residing in Italy (2003–2021). *Social Sciences*, *12*(9), 524.

Dacey, M. F. (1965). A review on measures of contiguity for two and k-color maps. In B. J. L. Berry & D. F. Marble (Eds.), *Spatial analysis: A reader in statistical geography* (pp. 479–495). Englewood Cliffs, NJ: Prentice-Hall.

Duncan, O. D. (1957). The measurement of population distribution. *Population Studies*, *11*(1), 27-45.

Feitosa, F. F., Camara, G., Monteiro, A. M. V., Koschitzki, T., & Silva, M. P. (2007). Global and local spatial indices of urban segregation. *International Journal of Geographical Information Science*, *21*(3), 299-323.

Fotheringham, A. S., & Sachdeva, M. (2022). On the importance of thinking locally for statistics and society. *Spatial Statistics*, 50, 100601.

Fotheringham, A. S., C. Brunsdon, and M. Charlton. (2002). Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Chichester: Wiley.

Guevara, M. R., Hartmann, D., & Mendoza, M. (2016). diverse: an R Package to Analyze Diversity in Complex Systems. The *R Journal*, *8*(2), 60-79.

Harris, R. (2020). Visualising the scales of ethnic diversity in London using a multilevel entropy index. *Environment and Planning A: Economy and Space*, *52*(7), 1239-1242.

ISTAT. (2022). *Popolazione residente per sesso, anno di nascita e stato civile al 1° gennaio*. Retrieved from http://dati.istat.it/Index.aspx?DataSetCode=DCIS_POPSTRRES1
ISTAT. (2024). *Confini delle unità amministrative a fini statistici al 1 gennaio 2018*. Retrieved from https://www.istat.it/notizia/confini-delle-unita-amministrative-a-fini-statistici-al-1-gennaio-2018-2
Kramer, R., & Kramer, P. (2019). Diversifying but not integrating: Entropic measures of local segregation. *Tijdschrift voor economische en sociale geografie*, *110*(3), 251-270.
Lee, B. A., Reardon, S. F., Firebaugh, G., Farrell, C. R., Matthews, S. A., & O'Sullivan, D. (2008). Beyond the census tract: Patterns and determinants of racial segregation at multiple geographic scales. *American sociological review*, *73*(5), 766-791.
Massey, D. S., & Denton, N. A. (1988). The dimensions of residential segregation. *Social forces*, *67*(2), 281-315.
Mazza, A., & Punzo, A. (2015). On the upward bias of the dissimilarity index and its corrections. *Sociological Methods & Research*, *44*(1), 80-107.
McCulloch, A. (2007). The changing structure of ethnic diversity and segregation in England, 1991–2001. *Environment and Planning A*, *39*(4), 909-927.
Mora, R., & Ruiz-Castillo, J. (2011). Entropy-based segregation indices. *Sociological Methodology*, *41*(1), 159-194.
Mucciardi, M. (2008). Use of a flexible weight matrix in a local spatial statistic. In *First joint meeting of the Société Francophone de Classification and the Classification and Data Analysis Group of the Italian Society of Statistics* (Vol. 5, pp. 385–388). Naples: Edizioni Scientifiche Italiane.
Oshan, T. M., Li, Z., Kang, W., Wolf, L. J., & Fotheringham, A. S. (2019). mgwr: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-Information*, *8*(6), 269.
Otranto, E., Mucciardi, M., & Bertuccelli, P. (2016). Spatial effects in dynamic conditional correlations. *Journal of Applied Statistics*, 43(3), 604–626. https://doi.org/10.1080/02664763.2015.1068263
Park, R. E. (1926, July). The concept of position in sociology. In *papers and proceedings of the American Sociological Society* (Vol. 20, No. 1926, pp. 1-14).
Park, R.E., Burgess, E.W., McKenzie, R.D. (1925). The City. University of Chicago Press, Chiago
Reardon, S. F., & Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological methodology*, *32*(1), 33-67.
Reardon, S. F., & O'Sullivan, D. (2004). Measures of spatial segregation. *Sociological methodology*, *34*(1), 121-162.
Reardon, S. F., Matthews, S. A., O'sullivan, D., Lee, B. A., Firebaugh, G., Farrell, C. R., & Bischoff, K. (2008). The geographic scale of metropolitan racial segregation. *Demography*, *45*(3), 489-514.
Roberto, E. (2018). The spatial proximity and connectivity method for measuring and analyzing residential segregation. *Sociological Methodology*, *48*(1), 182-224.
Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*(3), 379–423, 623–656.
Steele, L. G., Bostic, A., Lynch, S. M., & Abdelaaty, L. (2022). Measuring ethnic diversity. *Annual review of sociology*, *48*(1), 43-63.
Strozza, S., & De Santis, G. (2017). *Rapporto sulla popolazione. Le molte facce della presenza straniera in Italia*. Società editrice il Mulino spa.
Strozza, S., Benassi, F., Ferrara, R., & Gallo, G. (2016). Recent demographic trends in the major Italian urban agglomerations: The role of foreigners. *Spatial Demography*, *4*, 39-70.
Tammaru, T., Knapp, D., Silm, S., Van Ham, M., & Witlox, F. (2021). Spatial underpinnings of social inequalities: A vicious circles of segregation approach. *Social Inclusion*, *9*(2), 65-76.
Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and racial studies*, *30*(6), 1024-1054.

Voss, P. R. (2007). Demography as a spatial social science. *Population research and policy review*, *26*, 457-476.

White, M. J. (Ed.). (2016). *International handbook of migration and population distribution* (Vol. 6). Nueva York: Springer.

Wong, D. W. (1999). Geostatistics as measures of spatial segregation. *Urban geography*, *20*(7), 635-647.

Wong, D. W. (2004). The modifiable areal unit problem (MAUP). In *WorldMinds: geographical perspectives on 100 problems: commemorating the 100th anniversary of the association of American geographers 1904–2004* (pp. 571-575). Dordrecht: Springer Netherlands.

Yue, X., Wang, Y., Li, W., Wu, Y., Wang, Y., Zhang, H. O., & Ma, Z. (2024). Research Progress and Trends in Urban Residential Segregation. *Buildings*, *14*(7), 1962.

Zachary, D., & Dobson, S. (2021). Urban development and complexity: Shannon entropy as a measure of diversity. *Planning Practice & Research*, *36*(2), 157-173.