



HAL
open science

Urban Fabric with Bayesian LINCS. Empirical Evidence from the French Riviera

Giovanni Fusco, Alessandro Araldi

► **To cite this version:**

Giovanni Fusco, Alessandro Araldi. Urban Fabric with Bayesian LINCS. Empirical Evidence from the French Riviera. SAGEO 2016 Proceedings, GDR MAGIS / UMR ESPACE, 2016, 2-910545-11-3. hal-03533075

HAL Id: hal-03533075

<https://hal.science/hal-03533075>

Submitted on 26 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Urban Fabric with Bayesian LINCS. Empirical Evidence from the French Riviera.

Giovanni FUSCO¹, Alessandro ARALDI²

1. Université Côte d'Azur, CNRS, ESPACE
98 Bd Herriot, BP 3209, F-06204, Nice, France
giovanni.fusco@unice.fr

2. Université Côte d'Azur, CNRS, ESPACE
98 Bd Herriot, BP 3209, F-06204, Nice, France
alessandro.araldi@unice.fr

RESUME. Un ensemble d'indicateurs morphologiques est proposé pour identifier les tissus urbains par clustering spatial. L'approche LINCS est préférée aux plus classiques LISA pour mieux intégrer le point de vue du piéton dans la ville. Certains indicateurs morphologiques, comme ceux de couvertures du sol, sont des taux et posent un problème de variabilité avec la taille de la population mère. La correction empirique Bayésienne utilisée en épidémiologie, où la population mère serait la surface de l'unité spatiale, semble mal s'appliquer à la morphologie urbaine, car celle-ci influence le découpage spatial. Des nouvelles corrections Bayésiennes sont ainsi proposées et testées sur les paysages urbains de la Côte d'Azur. Une correction utilisant une fonction sublinéaire de la surface se montre plus apte à réduire l'hétéroscédasticité des taux et à identifier les concentrations de tissu pavillonnaire.

ABSTRACT. A set of morphological indicators is proposed to identify urban fabric using spatial clustering. The LINCS approach is preferred to classical LISA in order to better integrate the point of view of pedestrians moving in the city. Some morphological indicators like land coverage are rates and this poses the well-known problem of rate variability with the base population size. Classical empirical Bayesian correction used in epidemiology, with the spatial unit surface area as base population, seems unfit to the analysis of urban morphology, as spatial units depend from morphological phenomena. New empirical Bayesian corrections are thus proposed and tested on the case study of urban landscapes of the French Riviera. A new Bayesian correction which is a sublinear function of the unit surface proves better able to reduce rate heteroscedasticity and to identify hotspots of individual houses.

MOTS-CLES : Morphologie Urbaine, Taux de Couverture, Tissu Urbain, Réseau Viaire, LISA, LINCS, Correction Empirique Bayésienne, Côte d'Azur.

KEYWORDS: Urban Morphology, Coverage Ratio, Urban Fabric, Street Network, LISA, LINCS, Empirical Bayesian Correction, French Riviera

1. Introduction.

The classical schools of urban morphology (Caniggia and Maffei 1979, Castex et al. 1980, Conzen 1960, for an overview Pinon 1991) have developed a rich methodology to measure and evaluate urban morphology, and namely urban fabric, at the scale of a few blocks or a neighbourhood (Levy 2005). The analyses being carried out manually, their application to the whole urban space of a big city or of a conurbation becomes a daunting task. The use of geoprocessing and LISA/LINCS (Anselin 1995, Yamada and Thill 2010) has thus been proposed in the analysis of urban fabric within a large urban area (Araldi and Fusco 2016). The goal is to implement computer-aided chains of treatments to carry out analyses of urban fabric previously limited to small areas. The key to geo-processing is the use of morphological indicators well suited to identify specific traits of urban fabric. More particularly, the approach proposed by Araldi and Fusco (2016) is bottom-up: elementary morphological units are aggregated in order to identify through LISA/LINCS analysis of morphological indicators coherent patterns corresponding to particular urban fabrics.

However, when indicators of urban form are rates (for example a land coverage rate), the question arises of the influence of the base population size on rate variability among the spatial units of the analysis. If this variability is overlooked, LISA analysis can produce spurious spatial patterns. This problem is well known in epidemiology and has been addressed through the empirical Bayesian correction procedure (Marshall 1991, Assunção and Reis 1999). Straightforward application of solutions conceived in epidemiology seems nevertheless inappropriate for urban morphological analysis. The main specificity of urban morphology is the link between spatial unit definition and morphological characteristics, which makes the identification of a base population influencing rate variability much more complex.

The main aim of this paper is thus to present these specificities of urban morphological spatial analysis and to explore possible solutions of the Bayesian correction problem. The different solutions will be applied and tested on the case study of the French Riviera, allowing an expert-based feedback on the results of the analyses. The rest of the paper is organized as follows. Section 2 will present the main methodological choices of our research: spatial unit definition, morphological indicators and spatial analysis approach. Section 3 will explain why the classical Bayesian correction methods of epidemiology could be unfit for urban morphological analysis and will propose a few alternatives. Section 4 will present the main empirical results on the case of the French Riviera, when coverage rate of single-family houses are considered. Conclusions and future perspectives will put the research results in the broader context of computer-aided analysis of urban morphology.

2. Geoprocessing and LISA/LINCS for Urban Morphology Analysis: Methodological Choices.

2.1. Defining Elementary Spatial Units from a Pedestrian Point of View.

In the Italian, French and English traditional schools of urban morphology, the analysis of urban form, and more specifically the analysis of the urban fabric, has focused on three main aspects (Pinon 1991): (i) the identification of urban form components (urban network, buildings and parcels), (ii) their geometrical description and (iii) the analysis of their spatial relationships. The analyses of traditional urban morphology were normally carried out at the scale of a city neighbourhood, with manual calculations and a focus on the historical process behind observable urban forms. Geoprocessing of urban morphology within a GIS environment has become more widespread in the last twenty years, allowing for larger scales of analyses, but often losing the fine grain of the constituent elements of urban form (like in Berghauer Pont and Haupt 2010, or in Fusco 2016). Our research focuses in particular on the urban street network and the built-up space, which are the aspects of urban form more directly observable by pedestrians moving in urban space. Parcel structure plays a more important role in the historical production of urban form and has been omitted in our research.

In order to identify the elementary spatial units for the analysis of urban form, we consider a new division of urban space resulting from the combination of two elements: the urban street network, a connected set of segments allowing pedestrian movement, and the planar extension of urban space. A generalization of Thysen polygons is thus created around each street segment (topologically delimited by street junctions) to identify the portion of planar space conventionally served by the segment. For several morphological indicators, we only consider a double-sided proximity band within this polygon, in order to approximate visible space (Figure 1).

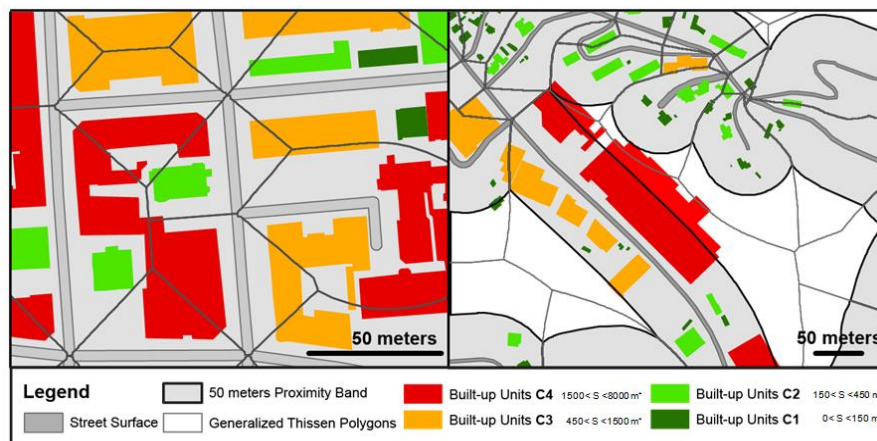


FIGURE 1. Street segment proximity bands and physical footprints of built-up units.

The rationale for this spatial unit definition is that a street segment should not be considered the limit, but rather the core of a fragment of urban fabric. This is often the case in European cities where discontinuities in urban fabric normally coincide with double carriageway boulevards (the latter produce two different spatial units in our model) but not with ordinary streets. Moreover, this approach is the most consistent with the pedestrian point of view: when standing in public space, people perceive the urban fabric on both sides of the street and not the elements within the four sides of a city block. In our study area, 99 562 elements were thus identified with 98% of street segment lengths varying between 5 m and 920 m (median value is 77.7 m). When considering the 50 m proximity band area, 98% of values range between 25 m² and 77 550 m² (median value being 3 700 m²).

2.2. A Set of Morphological Indicators

As anticipated, the street network morphology and the built-up forms are the main components of urban fabric considered in this research. Nine indicators, obtained through geoprocessing in GIS, were calculated for each spatial unit and cover five main aspects of urban fabric: network morphology, built-up morphology, network-building relationship, site morphology and site-network relationship. The indicators are presented in more detail in Araldi and Fusco (2016).

TABLE 1. A set of morphological indicators to characterise urban fabric.

Component of Urban Fabric Morphology	Indicator	Definition
Network Morphology	Length	Street segment length between two intersections
	Linearity/Windiness	Ratio between segment length and Euclidean distance
	Local connectivity	Number of degree 1 nodes
		Number of degree 4 nodes
Number of degree 3 or 5 and more nodes		
Built-up Morphology	Built-up Type Coverage Ratio	Ratio between 50 m proximity band surface and 0-150 m ² built-up units surface
		Ratio between 50 m proximity band surface and 150-450 m ² built-up units surface
		Ratio between 50 m proximity band surface and 450-1500 m ² built-up units surface
		Ratio between 50 m proximity band surface and 1500-8000 m ² built-up units surface
		Ratio between 50 m proximity band surface and >8000 m ² built-up units surface
Network-Building Relationship	Street corridor effect	Ratio between length of parallel façades and street length in the 10 m proximity band
	Proximity Band Coverage Ratio	Buildings coverage on the 20 m proximity band
	Proximity Band Building Height	Ratio between building vol. and surf. inside 20 m proximity band
Site Morphology	Surface Slope	Ratio between total and high sloped space-unit ($S > 30^\circ$) in 50m proximity band
Network Site Relationship	Street Acclivity	Computed as segment average of $\arctan(\text{slope})$



FIGURE 2. *Built-up types corresponding to physical footprint classes on the Riviera.*

Within this paper we will focus on the indicator describing the contribution to the coverage ratio of the 50 m proximity band around each street segment by the different classes of built-up elements. Built-up elements are either isolated buildings or sets of contiguous buildings. Empirical knowledge lets us establish multiple correspondences between built-up unit physical footprints and possible morphological types within our study area (Figure 2). 0-150 m² built-up footprints are mainly the fact of ordinary single-family independent houses (a) or of their annexes (b: sheds, detached garages). 150-450 m² footprints can correspond to big villas (c), to small series of row-houses (d) or to small block of flats (e). 450-1500 m² correspond mainly to four different typologies: big block of flats (f), small continuously built-up city blocks (i), small retail buildings (h) and long rows of townhouses (g), the latter being particularly rare in our study area. Footprints between 1500 and 8000 m² correspond to big modern housing estates (j), large continuously built-up city blocks (k) and big specialized buildings. Finally, class 5 are exceptional elements in the urban landscape of the study area, i.e. functionally specialized “giant” buildings (l: hospitals, stadiums, convention centres, shopping malls). The indicator is a rate measuring the part of the spatial unit 50 m wide around a street segment covered by each footprint class. The total coverage ratio of the spatial unit is thus the sum of the values calculated for each footprint class.

2.3. *Spatial Statistical Analysis of Morphological Indicators*

Once indicators are calculated for each spatial unit, local spatial clustering techniques are used to identify larger scale urban features (urban fabric). Spatial

clustering indicates where a phenomenon of interest has high/low incidence level, identifying hot/cold spots; many different methods have been developed, in different research fields and perspectives. Local Moran's I indicator of spatial association (LISA, Anselin 1995), based on Moran's I spatial correlation measure (Moran 1948), was identified as a valid geostatistical method. Despite its large application in other research fields, it has so far been used relatively little in the study of urban form (Tsai 2005, Musakwa and Niekerk 2014). In order to test the pedestrian perspective assumption, Araldi and Fusco (2016) analysed and compared the planar application of local Moran's I statistic LISA with the corresponding network-constrained ILINCS (I statistics in Local Indicator of Network-Constrained Clusters, Yamada and Thill 2007, 2010). Their work considered several network and planar neighbourhood depths, always following a topological queen contiguity approach. The case of the street corridor effect indicator was used in order to show that network-constrained LINCNS are better suited to capture the feature patterns from a pedestrian point of view. Following this indication, we will carry out spatial analysis of built-up type coverage through a LINCNS approach, a network extension of LISA. Adopting a network geometry, the GeoDa software (Anselin 2003) can be used to implement these statistics.

3. The Problem of Population Size in the Spatial Analysis of Rate Values

3.1. Bayesian Correction in Epidemiology

Epidemiology has been one of the first and is still one of the most important fields of application of spatial clustering techniques (Jacquez et al. 1996a, 1996b). Tackling the case of spatial clustering of rate values, Marshall (1991) and Assunção and Reis (1999) highlight that the size of the base population over which the rate is calculated has an impact on the variability of the rate. When dealing with spatial units having strong differences in population size, the variability of the prevalence rate of a given disease in large population units is lower than the variability measured in small population units. This is due to the limits of statistical estimators on small populations. When this dependence of variability on base population size is overlooked, spatial clustering techniques could detect spurious hot/cold spots, linked to the variability of estimations on small populations and not to spatial structure. The same authors thus propose an empirical Bayesian correction taking into account the knowledge of the base population size to correct the *a priori* assumptions on the variability of the rates. This approach, which is now commonly accepted in epidemiology, considers that spatial unit definition (normally administrative units like municipalities in a given region) is independent from the phenomena under investigation (here disease prevalence). Epidemics do not define spatial units and the epidemiological phenomenon (the prevalence rate) can be modelled as a Poisson stochastic process whose outcome (the number of infected people) is directly proportional to the population exposed in the spatial unit. Under this assumption, Empirical Bayesian correction of Moran local I (EBI) is defined as follows (Assunção and Reis 1999):

$$z_i = \frac{p_i - b}{\sqrt{a + b/x_i}} \quad (1)$$

$$EBI = \frac{m}{\sum w_{ij}} \frac{\sum w_{ij} z_i z_j}{\sum (z_i - E[z])^2} \quad (2)$$

Where m is the number of spatial units, w_{ij} are the neighbourhood coefficients for units i and j (normally 1 if j is considered neighbour of i , 0 otherwise), z_i are the new deviates for which the Moran local I is calculated and which are derived from the original rates p_i , a and b are Bayesian estimators described in Marshall (1991). The empirical Bayesian correction boils down to reducing the variance of rate values p_i *a posteriori* through standardization coefficients, which have a fixed part and a variable part which is inversely proportional to the population size x_i of each spatial unit.

LISA/LINCS clusters are aimed at verifying whether, on the one hand, the observed rates are the likely result of random outcomes of a spatially homogenous Poisson process with a unique rate or with different rates without spatial structure (null hypothesis) or, on the other hand, if they are the random outcome of a spatial structure of different underlying rates (alternative hypothesis). Bayesian corrections have a direct impact on the calculus of LISA/LINCS clusters: deviations from the study area average values tend to be more significant when they are calculated for units with large populations and less significant when they are calculated for units with small populations. Accordingly, the usual categories of High-High, Low-Low, High-Low and Low-High spatial patterns tend to be less significant when calculated on units having mainly smaller populations.

3.2. Why could the Analysis of Urban Morphology be Different from the Classical Epidemiological Framework.

At first view, it is easy to conceive a straightforward application of the epidemiological approach to urban morphological rates. The coverage ratio of a given built-up type within a spatial unit would have as base population the surface of the spatial unit. Every square meter of the 50 m proximity band around a street segment is the equivalent of an exposed person in epidemiological terms. In this context, small spatial units should exhibit higher variability than big spatial units, the latter corresponding to long street segments and the former to short street segments. The empirical Bayesian correction of these variabilities based on spatial unit surface amounts to considering as more reliable rates measured around long street segments. But this presupposes that spatial units are independent from the phenomenon under investigation.

Unfortunately, when dealing with urban morphology, spatial unit definition is directly linked to the phenomena under investigation. This is evident in our case as we define spatial units as proximity bands around street segments, which are themselves part of the urban morphology we want to study. Coarser (and often inappropriate) spatial unit definitions like urban neighbourhoods or street blocks (the natural urban grid) also depend on urban morphology. Neighbourhoods are in fact smaller where urbanization is denser and larger where it is less intense. Superposing

a conventional grid of decametric/hectometric grain on urban space is very popular in the analysis of land coverage, landscape ecology and, more recently, socioeconomic phenomena. Its use for the analysis of the morphology of urban fabric seems more problematic, as the spatial arrangements of form elements could be partially destroyed by the superimposed observation grid¹. Moreover, we should not see urban space as a coropleth phenomenon. Urban space is the result of the urban colonization of natural and agricultural land. The morphological process is here to be taken into account in order to understand the production of spatial units. As classical morphological analysis shows (Caniggia and Maffei 1979), when a spatial unit approximately defined as we do in our research (land close to a street segment) is relatively saturated (i.e. it has a high land coverage ratio), it becomes economically interesting to intensify the street network in order to develop adjacent natural or agricultural land. Unless strong topographic constraints forbid it, new street segments are added and the average length of the street segments diminishes locally, just as the average surface of the spatial unit (proximity band of 50 m around the street segment). Taking into account that spatial unit definition is linked to morphological phenomena, two alternative hypotheses could be considered against the epidemiological approach.

The first hypothesis considers that new urbanization is historically produced by the repetition of a given building type (with its typical open spaces like interior courts, front and back gardens, parking space, etc.), or of a very limited number of complementary building types. Even if these patterns can later evolve through destruction/reconstruction processes, the contribution of given building types to the land coverage around a street should tend to oscillate within given limits once the unit is totally urbanized. Variability could be much higher when a long street segment proximity band is only partially developed and natural or agricultural land can still be in it (even if variability should fall again for the longest segments, in completely undeveloped land). The understanding of the morphological process could thus suggest an approach that is opposite to the one used in epidemiology. Rates around short and medium-sized segments should be considered as more reliable because the morphological process has produced more stable urban features. Rates around longer segments should be corrected as higher variability characterizes a phase of the morphological process that could be considered as intermediate between natural/agricultural status and full urbanisation. Within this hypothesis of “inverse relation”, we can conceive an empirical Bayesian correction where the base population is not the surface of the area unit, but is inversely proportional to it.

The second alternative hypothesis is a milder revision of the epidemiological approach. The aforementioned link between spatial unit definition and the morphological phenomena would weaken the relationship between the morphological variability and population size (here spatial unit size), without being

¹ Nevertheless, when the focus of the analysis is not the relationship between buildings, street segments and parcels within urban fabric but more abstract aspects of urban form, like fractal dimension of the built-up area, pixel-based analyses are of paramount interest (like in Batty and Longley 1994 or in Frankhauser 1994).

able to inverse it. The statistical reduction of morphological rate variability within larger spatial units will be only partially offset by the higher heterogeneity associated with longer street segments in not fully urbanised areas. In this case, a sublinear variation of the total land coverage ratio and of the land coverage ratios by each built-up type could be considered, with an exponent less than one to be applied to the spatial unit surface in order to obtain a convenient EBI.

In the following section we will test these three different approaches to empirical Bayesian correction of ILINCS (EBI-LINCS) using data from the French Riviera.

4. Empirical Evidence from the French Riviera

4.1. *The French Riviera: an ideal test area for new morphological analyses*

The French Riviera is a conurbation in southern France stretching more than sixty kilometres from the Italian border to the Esterel mountains. Covering around 1500 km², the French Riviera has a population of over one million inhabitants. This space is a unique conjunction of natural and urban landscapes: firstly, the topography, with elevation ranging from the sea level up to 1700 meters of the pre-Alps (passing through hills and valleys differently sloped). Secondly the socio-political and historical influences on the urban planning. Traditional villages are also scattered around three high density urban areas. From east to west, we find: Monaco and its skyscrapers, the most densely populated sovereign nation in the world; the urban agglomeration of Nice with a regular meshed core inspired by the Turin model (Graff 2013), surrounded by hilly and less tightly planned areas. And finally the urban agglomeration of Cannes-Grasse-Antibes characterised by land irregularity together with old city-centres and the car-centred sprawl development of the last 50 years (Fusco 2016). The combination of all these elements produces a sequence of urban centres and peripheral areas of different size and different morphology. This study area gives the opportunity to test the validity of all the morphological indicators of Table 1, of the LISA/LINCS approaches (Araldi and Fusco 2016) and, as far as this paper is concerned, the proposed methodology of empirical Bayesian correction for morphological rates.

4.2. *Spatial Patterns of Single Family Houses: Results and Discussion*

Within this paper we will focus on the spatial clustering of coverage rates for the smallest footprint class (≤ 150 m²). This is the class which corresponds more strictly to a single building type: ordinary free-standing single-family houses and their annexes. Figure 3 shows the values of land coverage ratios for this building type within the 50 m proximity bands of street segments. The represented area is the central section of the French Riviera, around the city of Nice. The map clearly shows a strong under-representation of this building type in the city-centre of Nice. It is much harder to identify spatial patterns in more peripheral areas based on street connections reflecting the way pedestrians explore urban space. This fully justifies the use of spatial clustering analysis based on the LINCS approach.

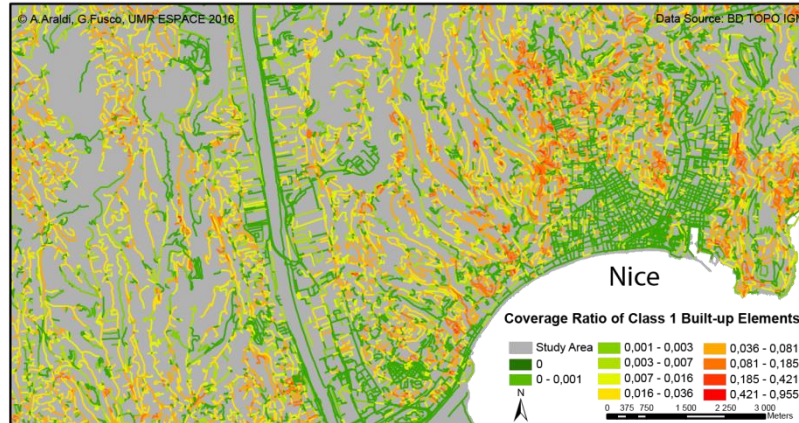


FIGURE 3. Land coverage ratios for single-family houses around the city of Nice.

For this first typo-morphological class, the interpretation of the hot/cold spots of the spatial clustering will be relatively easy. A good spatial clustering should identify as hotspots (High-High) the swaths of residential urban fabric mainly made of individual houses, within planned subdivisions or more spontaneous urbanisation; cold spots (Low-Low) should correspond either to urban fabric made up of other building types or to relatively undeveloped natural and agricultural land. Exceptions within these patterns or transition areas should be identified as Low-High or High-Low features. Finally, more mixed configurations without clear spatial structure should be classified as statistically non-significant in terms of spatial clustering (significance level of 0.05 will be used throughout our analyses). A few preliminary statistical analyses are nevertheless necessary before calculating spatial clustering indicators.

Figure 4.a shows the variation of the variance of single houses land coverage ratio with the increase of the spatial unit surface. The diagram clearly shows heteroscedasticity for the observed morphological rates, pointing to the need of empirical Bayesian correction: the ratio of the variances calculated in the ventiles with maximal and minimal variability is 32.1 and even considering only the 18 ventiles with most homogeneous variance, this ratio is still 7.0. Variance increases over the first three ventiles (with a maximum for the 3rd ventile, corresponding to an average unit surface of 312 m²) and steadily decreases afterwards. This trend invalidates the inverse relation hypothesis: like in epidemiology, and apart for extremely small spatial units (where the possibilities of urbanisation are technically limited to fewer options) rate variability steadily decreases with spatial unit size.

We still have to ascertain whether the classical epidemiological approach or the alternative hypothesis of a sublinear correction is best suited to our empirical data. Figure 4.b and 4.c answer this question. In the former, variance is calculated for empirically Bayesian corrected rates as in the epidemiological approach. Heteroscedasticity remains almost the same. On the contrary, through a Bayesian

correction which uses the cubic root of the spatial unit surface (Figure 4.c), we reduce heteroscedasticity considerably: the ratio between the highest and lowest variance is now 13 and even to 3.5 when only the 18 ventiles with most homogeneous variance are considered. This result points to the sublinear correction hypothesis: the morphological process can reduce but not invert the trend of coverage rate variability to decrease with spatial unit surface. Other sublinear corrections were tested within our research (exponents 0.667, 0.5, 0.25 and logarithmic transformation). The cubic root proved the most effective in reducing heteroscedasticity for single-family houses coverage ratio on the French Riviera.

We will however highlight the importance of the residual heteroscedasticity after the empirical Bayesian correction. Once the size of the spatial unit is taken into account (though a sublinear correction), intermediate spatial units still exhibit more rate variability than smaller and bigger ones. The link between the morphological phenomenon and the spatial unit definition is possibly more complex than what a monotonic Bayesian correction with the unit surface can handle. Intermediate spatial units are closely associated with urban fabric. High heterogeneity of urban fabrics in the study area could thus be a possible cause of observed residual heteroscedasticity.

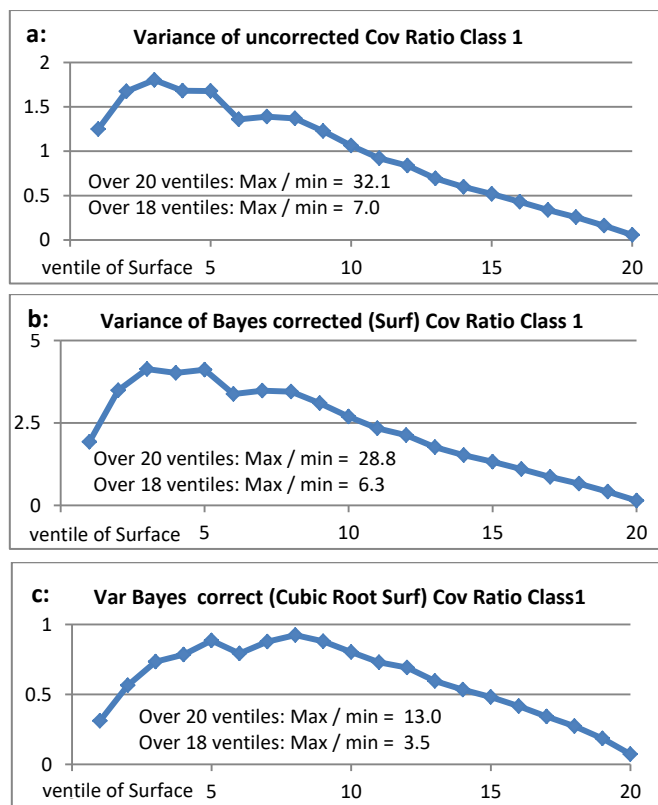


FIGURE 4. Variance of land coverage ratios for single-family houses.

Another important point to be considered before calculating ILINCS statistics is the lack of normality of the empirical distribution of our morphological rates. Griffith (2016) shows how pseudo-significance of Moran I is biased by skewness in data distribution. The same could be said for local Moran I and ILINCS. More particularly, single-family houses coverage ratio is extremely right-skewed on the French Riviera, with 48% of spatial units having the modal value 0, a median value of just 0.0004 and an average value of 0.02, whereas 71% of values are below average (this is not directly visible in Figure 3, where only the central area around the city of Nice is represented). The typical negative deviate from the average value is thus around 0.0004. Positive deviates can be much wider: 0.6% of deviates are two orders of magnitude larger than the typical negative deviate. As the average number of neighbours on each spatial units is around 20 (when a topological radius of 3 is used, see further), a single very large positive deviate can produce a higher than average value in the neighbourhood of a street segment. Now, a random sampling of our empirical distribution will easily obtain at least one very large deviate in 20 independent draws ($p = 20 \times 0.006 = 0.12$, well beyond the significance threshold of 0.05). Conversely, obtaining a lower than average neighbourhood from randomly sampling our empirical distribution is more difficult: almost all draws have to be below the average and no single draw must have a large positive deviate (here p is well below 0.05). As a consequence, when calculating pseudo-significance of ILINCS, HH or LH significance values are underestimated and LL or HL significance values are overestimated.

Given the particular empirical distribution of our data (48% of 0 values), no non-linear transformation of the data could produce a symmetric, almost-normal distribution, not even a rank transformation. Furthermore, EBI correction of rank data would not decrease data heteroscedasticity. As a consequence, ILINCS and EBI-LINCS will be calculated for the original coverage ratio data, and we will be forced to accept spatial clusters with different pseudo-significance levels.

ILINCS were thus calculated following topological contiguity on the street network with a neighbourhood depth of three (which corresponds well to the relatively local features of urban fabric). Values in Figure 5.1 have no Bayesian correction. They identify 32 000 units in Low-Low clusters. In the area represented on the map they correspond to the city-centre of Nice (a), to modern developments in the Var Valley or on the hills, like the IBM centre (b), to natural areas (c) or to old villages. High-High clusters involve more than 10 000 units, both in small subdivisions north of the city-centre (d) and in larger less regular areas on the hills (e). Nevertheless, differences in the variability of rates for different size of spatial units could bias the identification of significant and non-significant spatial clusters.

Figure 5.2 thus proposes a classical empirical Bayesian correction of coverage rates for individual houses, based on the straightforward epidemiological analogy. We observe a big increase of non-significant features. Spatially, significant patterns are severely reduced in urbanised areas (with small or medium-sized spatial units). Low-Low clusters are only marginally reduced: they resist well both in the very connected city-centre and in the undeveloped areas with long street-segments.

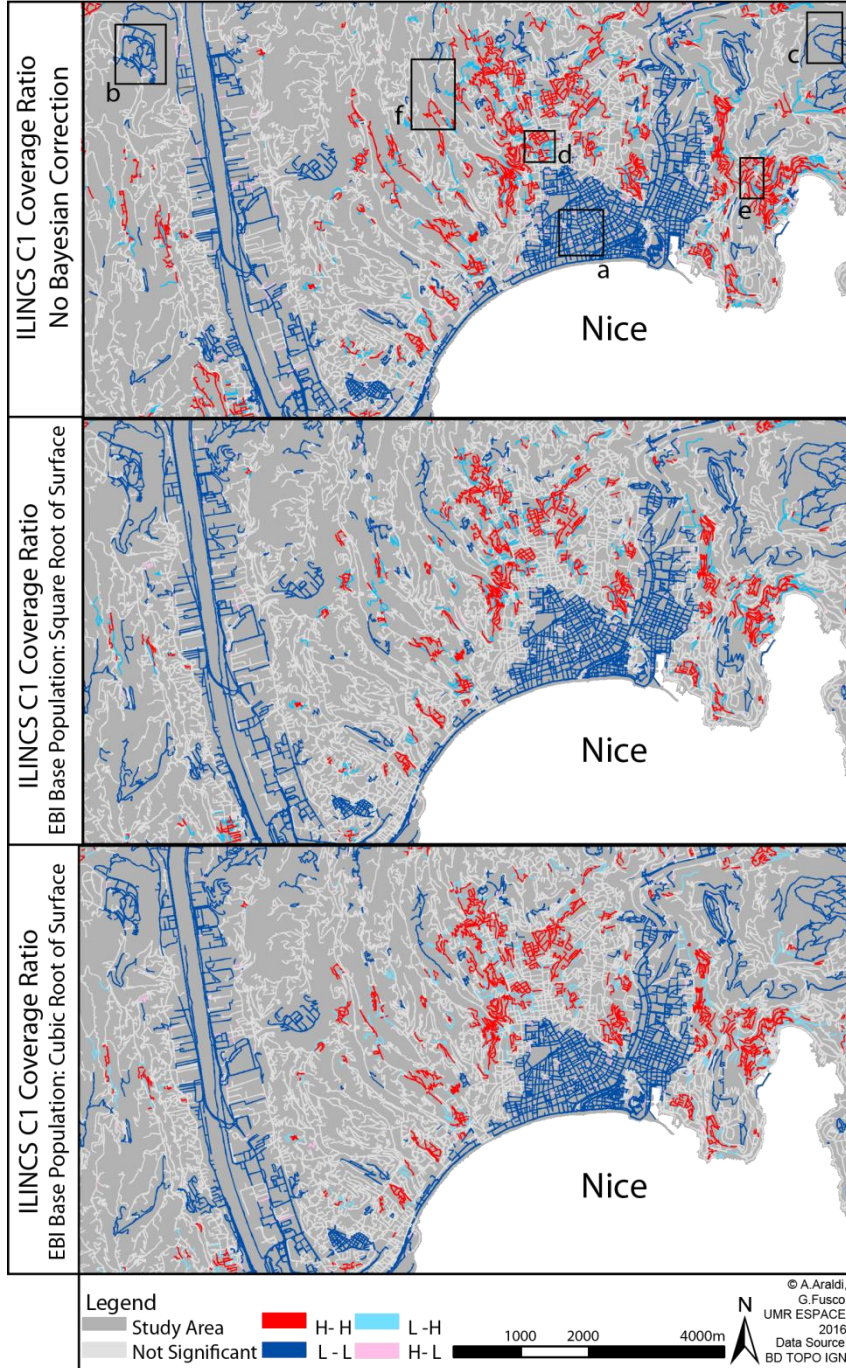


FIGURE 5. ILINCS of land coverage ratios for single-family houses.



FIGURE 6. Sample of Riviera landscapes within the areas identified in Fig.5.

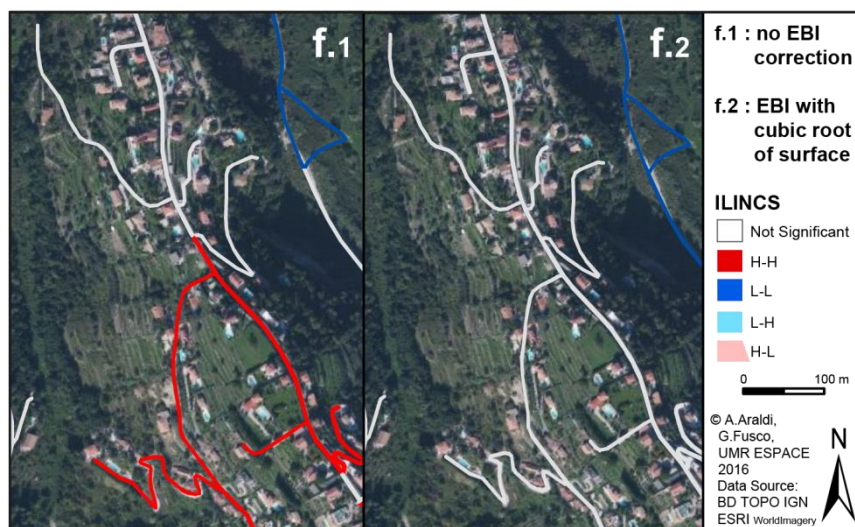


FIGURE 7. The case of a mixed area on the hills around Nice (f in Fig.5).

A few long Low-Low segments even appear and replace shorter ones which are now considered non-significant. The main loss of significant features is observed for the High-High clusters (they diminish of more than 60%): whenever street segments are not particularly well connected, the lower corrected values of short-segments deviates make whole High-High clusters become non-significant. Only the most typical urban subdivisions (d) or suburban concentrations of single-houses remain significant, many tree-like hotspots of individual houses (like e) disappear because of the Bayesian correction to the small branches of the main streets.

Figure 5.3 results from the application of our proposed sublinear empirical Bayesian correction. Overall, compared to Figures 5.1 and 5.2, there is a much slighter reduction of non-significant features. Spatially, most significant High-High and Low-Low clusters are maintained both in urbanised areas and in the less developed hilly peripheries. Sometimes, though, a slight penalization of the deviates of the smaller spatial units, reclassifies as non-significant a few structures that were previously considered as High-High. This is the case, for example, of relatively

heterogeneous areas on the hills around the city-centre of Nice (like f). The aerial image in Figure 7 shows no noticeable difference between the southern and the northern section of the area, with class 1 elements (single-family houses) intermingled with class 2 elements (big villas and a few semidetached houses). However, in the south many of the short street segments happen to lead to single-family houses only. In the north, they happen to connect more heterogeneous elements. The uncorrected ILINCS reveal the existence of a significant High-High structure in the south, whereas the corresponding EBI-LINCS, by overlooking the contribution of the smallest street segments, rightly identifies a seamless suburban area with non-significant overrepresentation of class 1 elements. The area in section f is thus clearly distinguished from areas of more pronounced homogeneous overrepresentation of single family houses like d and e.

5. Conclusions and Perspectives.

Specific knowledge of the morphological process led us to critically assess the classical empirical Bayesian correction of rate variability used for LISA/LINCS in epidemiology, because spatial unit definition is linked to the morphological process. We could thus formulate two new proposals: an EBI for morphological rates where the base population is the inverse of the spatial unit surface and an EBI where the base population is a monotonic sublinear function of the spatial unit surface.

The case study of urbanisation in the French Riviera let us verify empirically that the proposed sublinear EBI of individual house coverage rate is better able to reduce heteroscedasticity in the empirical data, to correctly detect typical suburban fabric as hotspots of single-family houses and central-city areas, functionally specialised areas, as well as scarcely developed areas as cold spots. Residual heteroscedasticity is nevertheless still important after the EBI correction. The link between morphological phenomena and the spatial unit definition is possibly more complex than what a monotonic Bayesian correction with the unit surface can handle, especially in the presence of high heterogeneity of urban fabrics, like in our study area.

The spatial analysis of the coverage rate of single houses is only a first step of the research. Other physical footprint classes, having multiple correspondences to morphological types, will have to be cross-analysed with single-family houses. Built-up volumes could then be used to better assess the correspondence between physical footprints and morphological types. The results presented in this paper make us however believe that the best way to take into account size-related variability of the coverage rates will be through an empirical Bayesian correction using as base population a sublinear function of the spatial unit surface.

A complete bottom-up identification of urban fabric types, coherently taking into consideration the pedestrian point of view, will of course be possible only through the combined spatial analysis of all indicators identified in our research design. Street network morphology and network-building relations could be essential to solve the ambiguous correspondences between physical footprints and morphological types. Araldi and Fusco (2016) showed the pertinence of the LINCS approach, already proposed by Yamada and Thill (2007, 2010) in other research

fields, for the spatial analysis of urban morphology. This paper proposes an alternative way of applying empirical Bayesian corrections to morphological rates. These two indications should thus guide the cross-analysis of the whole indicator set.

References

- Anselin L. (1995). Local Indicators of spatial association. *Geog. Analysis*, 27(2): p. 93-115.
- Anselin L. (2003). *GeoDa0.9users guide*. Phoenix, Spatial Analysis Laboratory.
- Anselin L., Kim W.Y., Syabri Y. (2003). Web based Analytical tools for the exploration of spatial data. *Journal of Geographical Systems*. 6(2): p. 197-218.
- Araldi A., Fusco G. (2016). Urban Form from the Pedestrian Point of View: Spatial Patterns on a Street Network. *INPUT 2016*, Turin, Sept. 14th-15th 2016, 6 p. (proceedings in press).
- Assunção R., Reis E. (1999). A New Proposal to Adjust Moran's I for Population Density. *Statistics in Medicine*, vol. 18, p. 2147-2162.
- Batty M., Longley P. (1994) *Fractal Cities: Geometry of Form and Function*, Academic Press
- Berghauser Pont M. Y., Haupt P. (2010). *SPACEMATRIX, Space, Density and Urban Form*. Rotterdam, NAI Publishers.
- Caniggia G., Maffei G.-L. (1979). *Lettura dell'edilizia di base*. Firenze, Alinea.
- Frankhauser P. (1994). *La fractalité des structures urbaines*, Economica, Paris.
- Fusco G. (2016). Beyond the Built-up Form / Mobility Relationship: Spatial Affordance and Lifestyles. *Computer, Environment and Urban Systems*. 60 (2016): p. 50-66
- Graff P. (2013). *Une ville d'exception. Nice dans l'effervescence du 20^e siècle*. Serre, Nice.
- Griffith D.A. (2010) The Moran coefficient for non-normal data, *Journal of Statistical Planning and Inference*. 140 (2010): p. 2980-2990
- Jacquez G M, Grimson R, Waller L, Wartenberg D (1996a). The analysis of disease clusters: Introduction to techniques. *Infection Control and Hospital Epidemiology*, 17, p. 385-397
- Jacquez G M, Waller L, Grimson R, and Wartenberg D (1996b). The analysis of disease clusters: State of the art. *Infection Control and Hospital Epidemiology*, 17, p. 319-327
- Lévy A. (2005). Formes urbaine et significations: revisiter la morphologie urbaine. *Espace et sociétés*, 122(4): p. 25-48.
- Marshall R. J. (1991). Mapping Disease and Mortality Rates using Empirical Bayes Estimators. *Applied Statistics*. 40(2): p. 283-294.
- Moran P. A. P. (1948). Interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*. 10(2): p. 243-251.
- Musakwa, W., Niekerk A. (2014). Monitoring urban sprawl and sustainable urban development using the Moran Index. *Int. Journ. Applied Geospatial Res.* 5(3): p. 1-20.
- Pinon P. (1991). *Lire et composer l'espace public*. Paris, STU.
- Tsai, Y.H. (2004). Quantifying urban form: Compactness versus sprawl. *Urban Studies*, 42(1): p. 141-161.
- Yamada, I. Thill, J.C. (2007). Local indicators of network-constrained clusters in spatial point patterns. *Geographical Analysis*, 39(3), p. 268-292.
- Yamada, I. Thill, J.C. (2010). Local indicators of network-constrained clusters in spatial patterns represented by a link attribute. *Annals Ass'n. Am. Geog.*, 100(2), p. 269-285.