#122

SPATIAL CONFIGURATION AND VEHICULAR MOVEMENT

A nationwide correlational study

MIGUEL SERRA
CITTA – Research Centre for Territory, Transports and Environment. Porto University.
mserra@fe.up.pt

BILL HILLIER
Space Syntax Laboratory, The Bartlett School of Architecture, University College London
bhillier@ucl.ac.uk

ABSTRACT

We provide the most comprehensive study to date on the correlation between network centrality measures and vehicular movement flows, using a model of the UK's entire road network (2,031,971 nodes) and a very large dataset of vehicular movement counts (20,752 instances, evenly distributed over the UK's territory). We describe the statistical associations between observed vehicular flows and the values of betweenness centrality of the road-network nodes where such flows were measured, the latter calculated using Euclidean and angular distance functions, across a number of increasing radii, from the local to the supra-regional scales. Relations to road capacity are also discussed in principal road networks where this is known.

The geographical comprehensiveness of our model and the size of our movement sample allows us to state, with unprecedented statistical validity, the clear outperformance of angular distance over Euclidean distance, on what concerns the effect sizes of the studied correlations. We also demonstrate the existence of two clearly different regimes of association between movement and centrality, occurring on the background and foreground networks of cities, which may be interpreted as new evidence of the dual structure of urban form, proposed by space syntax.

KEYWORDS

spatial configuration, vehicular movement, angular and metric distance, saturation

1. INTRODUCTION

The theory of cities that has emerged from space syntax studies (Hillier 2012, 2016), is grounded on two fundamental findings: the discovery of the deep relationship between the topological structure of urban spatial networks and the distribution of movement flows therein, which led to the concept of 'natural movement' (Hillier, Penn et al. 1993, Hillier 1996, Hillier and Iida 2005); and the identification of a number of geometric regularities specific of urban spatial networks that are directly related with their topological characteristics, which led to the proposal of the 'dual background / foreground network' model of urban form (Hillier 1999, Hillier 2002, Hillier, Turner et al. 2006).

The concept of natural movement states that, all other things being equal, the intensity of urban movement observed on a given space of the network will be proportional to the position of that space on the configurational hierarchy of the network; that is, proportional to the relative importance of that space in the web of connectivity relationships that the network creates.

Movement intensities, on their side, are determinant to the spatial location of urban functions, in that functions that benefit from public exposure (as commercial and other tertiary functions) tend to colonize movement-rich locations, while functions that do not (as the residential function) tend to occupy more secluded areas. The settling of movement-seeking functions in locations that are 'naturally' movement-rich (i.e. made so by their position on the network), generates a positive feedback loop, by the attraction that those functions exert on even more movement and on the further settling of similar functions. Cumulatively, the clustered functional pattern that we observe in cities emerges (Hillier 1996).

Space syntax has also dedicated close attention to the *specific geometry* of urban spatial networks. It was found that the lengths of axial lines possess self-similar properties, with the same proportion of few long lines to many short ones repeating itself at all scales (Hillier 2002, Carvalho and Penn 2004). But also, that long and short lines have distinct aggregation probabilities (Hillier 1999, Hillier 2002). Longer lines tend to aggregate sequentially, with each line linked to the following one at *wide obtuse angles*, forming a large-scaled web of multi-directional alignments. Shorter lines tend to form clusters in the interstices of that web, passing through or ending on each other at *nearly right angles* (Hillier 1999).

This probabilistic arrangement of urban grids has far-reaching topological implications. Indeed, it induces two strongly asymmetric *connectivity patterns*, because the length of axial lines is positively correlated with their connectivity (Hillier 1999). And because long lines are also much less frequent than short ones, the highly connected sequences that they form will tend to possess also relatively *higher centrality levels*, forming a large-scaled *foreground network* of main paths. While the much more numerous but less connected shorter lines, will have relatively *lower centrality levels*, creating a less differentiated background network. Most importantly, because of the asymmetry in their centrality levels, these two fundamental networks will have also different *movement potentials* – the foreground network carrying the bulk of urban movement, with the functional consequences mentioned before.

Through these two basic principles space syntax was able to propose a theory of urban form that links its topology, geometry and functioning, into a single explanatory framework. Such a theory is capable of making testable predictions, because it is based on specific morphological assumptions – such as the generic foreground / background model of urban form – which can be verified and potentially denied. As described in (Hillier, Turner et al 2006; Hillier 2012, 2016) the foreground network, which is the global structure that holds the city together and conveys the bulk of urban movement, has topo-geometric characteristics that make it a web of *simplest paths* (i.e. made-up of long lines with little angular variation), and not one of *shortest paths* (in the sense of those with less Euclidian length). A way of testing this proposition is to statistically compare the spatial hierarchies described by angular and Euclidean-defined centrality measures and the actual distribution of urban movement flows. If urban space is indeed globally hierarchized through topo-geometrical principles rather than by metric ones, the former hierarchy should describe better the actual distribution of urban movement than the latter.

Hillier and Iida (2005) conducted a correlational study of this kind, comparing the strength of the correlations between urban movement flows (vehicular and pedestrian) and two types of centrality measures (closeness and betweenness), defined under three types of distance functions – Euclidean (or metric), topological and a new angular function, which was designed to express the geometric properties described above. Although this work was carried out on localized urban areas and not at the scale of the entire city, the authors have provided convincing empirical evidence of the validity of the proposed angular distance function, as revealed by the stronger correlations between angular-defined centrality values and the observed movement flows. These results were followed closely by those obtained with topological-defined centrality values. However, in all of the studied cases, metric distance yielded the worst correlations both when applied to betweenness and closeness calculations, but particularly so in the latter case (Hillier and Iida 2005).

The adoption of the angular distance function for defining shortest paths when computing graph centrality measures, has since then become generalized in space syntax urban research. This is in strong contrast with other analytical approaches to urban spatial networks who do not rely on any specific morphological model and therefore assume that, insofar urban space has a hierarchy, such a hierarchy should be based solely on Euclidean distance relationships. Indeed, notwithstanding a few recent papers who acknowledge the relevance of the angular distance concept and apply it (Cooper, Fone et al. 2014; Gil 2014; Cooper 2015; Molinero, Murcio et al. 2015), most of the studies which resort to centrality indicators for describing urban spatial structure – as, for example (Crucitti, Latora et al. 2006, Scellato, Cardillo et al. 2006, Masucci, Smith et al. 2009, Porta, Strano et al. 2009, Porta, Latora et al. 2012, Strano, Nicosia et al. 2012, Strano, Viana et al. 2013) – still adopt Euclidean distance functions as a self-evident choice.

In this paper we readdress this disciplinary divide, through a correlational study similar to the one developed in (Hillier and Iida 2005). We will compare the strength of the statistical associations between observed vehicular movement flows and angular and metric distance concepts, in order to assess their methodological and theoretical value. However, due to the geographical comprehensiveness of the spatial network model employed here (the UK's entire road network) and the size of the studied vehicular movement sample (20,752 count points), the detection of potential differences between the correlations obtained with the two types of distance will have unprecedented statistical validity. We aim at providing robust empirical evidence, capable of validating or denying the topo-geometric spatial structuring that space syntax proposes and, consequently, the theoretical constructs described above. Moreover, because we will do this at the scale of an entire country and in several geographical contexts (urban and non-urban), we also will test the potential generalization of such theoretical constructs from the urban scale to that of regional and supra-regional road networks.

2. DATASETS AND METHODS

2.1 THE ROAD NETWORK MODEL

The road network model used in this study is based on the Meridian 2 dataset (OS 2014), representing the full hierarchy of Great Britain's road network, but not its absolute geometric constitution. Road representation is skeletal, collapsed into single road centre-lines (RCL) independently of the type of road or of its specific cross section (i.e. number of lanes or carriageways). All complex road junctions (e.g. roundabouts and motorway interchanges) are generalized as simple RCL intersections. The vector geometry of the RCLs themselves has been partially generalized through simplification, eliminating unnecessary detail while retaining their essential shape.

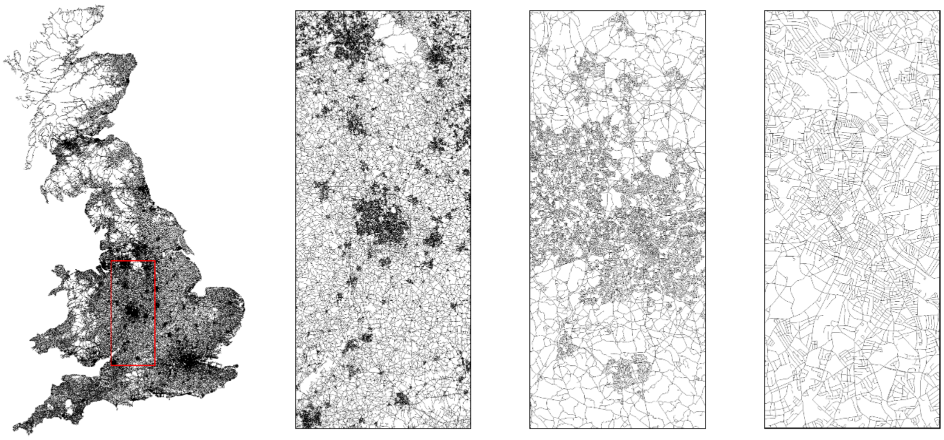

Figure 1 - The road network model. Each of the images on the right depict sequential 200% zoom-ins of the red rectangle on the full map.

These characteristics make this dataset particularly fit to serve as a basis for syntactic models, because its level of representation very much approximates that of a typical syntactic segment map. Given its geographic extent, the model used in this study should be seen as exhaustive, for it comprises the full national road hierarchy. However, at the level of the finer-grained network of local streets and lanes, the Meridian 2 dataset has a certain degree of incompleteness. Therefore, centrality measures calculated under short radii of analysis should be expected to contain some noise, induced by local inaccuracies of the model. In its final state (Figure 1), the road network model has 2,031,971 segments, corresponding to a total segment length of 341,588 Km.

As in any syntactic segment model, individual line segments are encoded as the nodes $V=\{1,\ldots,N\}$ of an undirected weighted graph $G(V,E)$, in which any pair of nodes $i\in V$ and $j\in V$ are held to be adjacent, $i\sim j$, when they correspond to segments that intersect on the segment map. The adjacency relations between nodes are encoded by edges $(i,j)\in E$, if and only if $i\sim j$.

Edges are weighted according to two types of distance cost – angular and Euclidean – denoted here respectively as w_a and w_e. The angular distance cost between two adjacent nodes, $w_a(i,j)$, is proportional to the angle of incidence θ defined by the two segments encoded by i and j, such that $w_a(i,j)=0$ when the two segments are aligned and $w_a(i,j)=1$ when the two segments make a right angle. Formally, the angular distance function may be defined as,

$$w_a(i,j)=\frac{2\theta}{\pi}, \theta\in[0,\pi[$$

The Euclidean distance cost between two adjacent nodes, $w_e(i,j)$, is the sum of the metric lengths of the segments encoded by i and j, denoted l_i and l, divided by 2; in other words, it is the actual length between the segments' mid-points, measured along the segments in metric units. Formally, $w_e(i,j)$ may be defined as,

$$w_e(i,j)=\frac{l_i+l_j}{2}, \{l_i, l_j\}\in\mathbb{R}^+$$

These two distance functions serve to define the shortest paths (or graph geodesics) between each pair of nodes, in two different ways. Angular distance defines geodesics as those paths with minimal sum of angular change, Euclidian distance defines geodesics as those paths with minimal sum of metric length. Due to the high computational cost of determining minimal paths in large graphs and to the nationwide size of our network model, angular and Euclidean geodesics are calculated here for a number of restricted network radii. A network radius, defined here in metric units, induces a sub-graph around each node containing the nodes that are reachable from the origin node within the radius distance. It may be seen as the maximum trip distance from the node under calculation. In order to calculate centrality values under the two definitions of distance described above, we will use the following set of radii,

$$R=\{1000,2000,5000,10000,25000,50000,100000,150000\}$$

Ranging from the local scale (i.e. 1Km), through the city scale (e.g. 10Km) and up to the supra-regional scale (i.e. 150 Km). The two distance concepts, angular and Euclidean, when applied to centrality measures, produce also different network centrality hierarchies in which a node may occupy quite different ordinal places. Here, we use the two types of angular and Euclidean defined geodesics, to compute the betweenness centrality (also called choice in space syntax) of each node $i \in V$, at each radii $r \in R$. The betweenness centrality of a given node i is defined by (Freeman 1977) as,

$$C_i^B = \sum_j \sum_k \frac{n_{jk}(i)}{n_{jk}} \; (j<k)$$

Where $n_{jk}(i)$ is the number of geodesics between nodes j and k that contain node i and n_{jk} is the number of all geodesics between j and k (there can be several). Betweenness centrality quantifies how often a given node lies on the shortest paths between other nodes. From the point of view of vehicular movement on road-networks, it may be seen as a direct indicator of the traffic flow potential of a given node or, in our case, of a given road or street segment. The road-network model was processed in the network analysis software UCL DepthmapX (Varoudis 2012), for each of the two centrality measures, at each of the network radii mentioned before.

2.2 THE VEHICULAR MOVEMENT SAMPLE

Our vehicular movement sample is based on a publically available dataset (DfT 2014) describing annual average daily flows (AADF) of different vehicles types, at 22,758 count locations on the UK's road network, distributed over the entire mainland territory. After several pre-processing operations, necessary for reasons of correct assignment of the count points to their respective locations on our road-network model, we have validated a study sample of 20,752 count points (91% of the original dataset).

These 20,752 count points are geographically evenly distributed (Figure 2), but their distribution per road class is neither random nor even. There are 12 road classes in the original dataset (Table 1), but the large majority of points (67%) are located on principal urban roads (PU, 40%) and on principal rural roads (PR, 27%), with all other 10 road classes representing only 33% of the occurrences. Thus, the sample has a strong bias towards principal roads and its direct use as such would certainly reflect that bias on the correlation results.

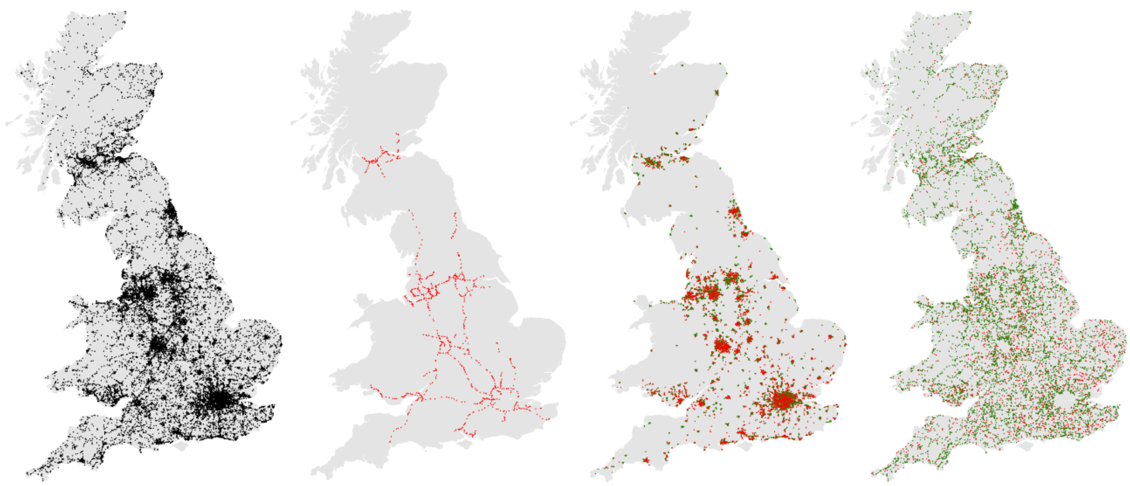

Figure 2 - Geographical distributions of count points. From left to right: 'all', 'motorways', 'urban roads' and 'rural roads' ('principal' in red, 'minor' in green)

Original road class	Name	Frequency	Description	Aggregated road class	Name	Frequency
PM	Principal motorway	N = 18	Double-carriageway roads of strategic importance, with two or more lanes in each direction, whose mantainance reponsability lies with the local authorities.	MW	Motorways	N = 796
TM	Trunk motorway	N = 778	Double-carriageway roads of strategic importance, with two or more lanes in each direction, whose mantainance reponsability lies with the central government.			
PU	Principal urban road	N = 8213	A-roads lying in urban areas with a population of 10,000 or more, whose mantainace responsibility lies with the local authorities.	PU	Principal urban roads	N = 8474
TU	Trunk urban road	N = 261	A-roads lying in urban areas with a population of 10,000 or more, whose mantainace responsibility lies with the central government.			
PR	Principal rural road	N = 5637	A-roads lying outside urban areas, whose mantainace responsibility lies with the local authorities.	PR	Principal rural roads	N = 7203
TR	Trunk rural road	N = 1566	A-roads lying outside urban areas, whose mantainace responsibility lies with thecentral government.			
BU	Urban B-road	N = 485	B-roads lying in urban areas with a population of 10,000 or more.	MU	Minor urban roads	N = 2742
CU	Urban C-road	N = 541	C-roads lying in urban areas with a population of 10,000 or more.			
UU	Unclassified urban road	N = 1716	Local roads (mainly residential) lying in urban areas.			
BR	Rural B-road	N = 670	B-roads lying outside urban areas.	MR	Minor rural roads	N = 1537
CR	Rural C-road	N = 492	C-roads lying outside urban areas.			
UR	Unclassified rural road	N = 375	Local roads (mainly residential but including rural lanes) lying in rural areas.			

Table 1 - Original and aggregated road classification schemes.

In order to mitigate that sample bias, we adopted the following re-sampling strategy. The original 12 road classes were aggregated into a simpler scheme of just 5 classes (motorways, principal urban and rural roads, minor urban and rural roads), which is the classification scheme adopted by DfT on their annual transport statistics reports (DfT 2016). After this operation, the sample was studied under progressive levels of disaggregation, starting with the full sample and ending on individual road classes (see Figure 4).

Besides the several types of road classes, the source dataset also provides AADF values for specific types of vehicles. These are "pedal cycles", "bus" (i.e. buses and coaches), "two-wheeled motor vehicles" (i.e. bikes), "cars" (i.e. cars and taxis), "light goods vehicles" (i.e. vans), "heavy goods vehicles" (i.e. lorries) and "all motor vehicles" (i.e. all motorized vehicles aggregated). Given all these different vehicle types, we first want to see if they all have similar route choice behaviours or if they show differences on that regard. With that purpose, we select only those points where all types of vehicles are present [N=17,283] and we correlate their respective frequencies at each count point (logically, vehicles whose frequencies are highly correlated should have similar behaviours in terms of road network use). Such correlations are in general high, but they also show some variability. We thus run a principal component analysis (PCA) on those correlations, in order to identify the most relevant collinearity trends. We extract two principal components with eigenvalues higher than 1, describing two groups of vehicles whose frequencies are strongly associated. We determine the members of each group, by inspecting the loadings of the two principal components (i.e. their correlations with the frequencies of each type of vehicle), displayed on Figure 3.

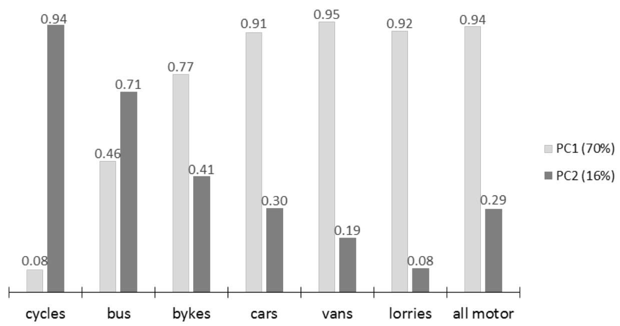

Figure 3 - Loadings of the two PCs describing the associations between the frequencies of vehicles types.

The first component (PC1), which is responsible for almost all variance explained (70%), is highly correlated [r > 0.9] with the frequencies of "cars", "vans", "lorries" and "all motor vehicles"; "bikes" are also strongly correlated [r=0.77] with PC1. The second component (PC2), explaining only residual variance (16%), is highly correlated with the frequencies of "cycles" [r=0.94] and, to a lesser extent, also with those of "bus" [r=0.71].

We thus observe two opposed behaviours among vehicles' route choices: one that may be represented by "all motor vehicles"; and another that may be represented by "cycles", with a weak relationship [r = 0.33] between their respective frequencies at each count point. We will therefore study the correlations between the AADF values of these two vehicle types and betweenness centrality values, while bearing in mind that "all motor vehicles" corresponds to 99.05% of the observed traffic, while "cycles" to only 0.95%.

2.3 ANALYTICAL AND STATISTICAL PROCEDURES

The sample was studied under progressive levels of desegregation accordingly to an analysis matrix (Figure 4), devised in such a way that each row corresponds to a specific hierarchical tier of the road-network (all, principal and minor roads) and each column to a specific geographical context (all, urban and rural). Each entry of the matrix corresponds to a different sized sub-sample, whose composition reflects its specific hierarchical and geographical contexts. When reading the matrix vertically, one can get a picture of the results by socio-demographic context (i.e. for the whole country, on cities and on rural areas). And horizontally, one may see how the results relate to the foreground/background-network model, previously identified and described in space syntax literature (Hillier, Turner et al. 2006, Hillier 2009, Hillier 2012). The sizes of 'all motor vehicles' and 'cycles' sub-samples differ slightly, because in the 'cycles' case

Road hierarchical context	Road geographical context →	**All roads**	**Urban roads**	**Rural roads**	Space syntax theoretical context
	All roads	**Road classes** {PU, PR, MU, MR} **Sub-sample size** All motor vehicles: N=19,952 Cycles: N=18,659 **Percentage of study sample** All motor vehicles: 96% Cycles: 90%	**Road classes** {PU, MU} **Sub-sample size** All motor vehicles: N=11,216 Cycles: N=11,000 **Percentage of study sample** All motor vehicles: 54% Cycles: 53%	**Road classes** {PR, MR} **Sub-sample size** All motor vehicles: N=8,740 Cycles: N=7,659 **Percentage of study sample** All motor vehicles: 42% Cycles: 37%	Whole Network
	Principal roads	**Road classes** {PU, PR} **Sub-sample size** All motor vehicles: N=15,677 Cycles: N=14,582 **Percentage of study sample** All motor vehicles: 76% Cycles: 70%	**Road classes** {PU} **Sub-sample size** All motor vehicles: N=8,474 Cycles: N=8,335 **Percentage of study sample** All motor vehicles: 41% Cycles: 40%	**Road classes** {PR} **Sub-sample size** All motor vehicles: N=7,203 Cycles: N=6,247 **Percentage of study sample** All motor vehicles: 35% Cycles: 30%	Foreground Net.
	Minor roads	**Road classes** {MU, MR} **Sub-sample size** All motor vehicles: N=4,077 Cycles: N=4,077 **Percentage of study sample** All motor vehicles: 20% Cycles: 20%	**Road classes** {MU} **Sub-sample size** All motor vehicles: N=2,742 Cycles: N=2,665 **Percentage of study sample** All motor vehicles: 13% Cycles: 13%	**Road classes** {MR} **Sub-sample size** All motor vehicles: N=1,537 Cycles: N=1,412 **Percentage of study sample** All motor vehicles: 7% Cycles: 7%	Background Net.
	Socio-demographic context →	The Whole Country	The Cities	The Countryside	

Figure 4 - Analysis matrix.

we only consider the count points with non-zero frequency. Motorways [n=752], because they are not classified either as urban or rural road-infrastructures, but also because they have very specific connectivity characteristics are left out of the analysis matrix and their results will be presented apart. For each of the sub-samples corresponding to each matrix entry (and for the motorways sub-sample), we will correlate the traffic flows and the network centrality values observed at each count location.

Both movement and centrality variables have highly right-skewed distributions, with many outliers, strongly deviating from bivariate normality. We will therefore use a robust non-parametric correlation method, namely Spearman's rank correlation coefficient, denoted as ρ (rho). Spearman's ρ is a measure of statistical association based on the ranks of two variables (i.e. on ordinal values indicating the relative magnitude of the actual values). Spearman's ρ is particularly fit for our research subject, because we are not interested in the specific values of either movement or centrality, but rather in knowing which type of network hierarchy (as described by angular and Euclidean defined betweenness) better emulates the observed relative magnitudes of vehicular traffic flows.

Given the large size of our sample and sub-samples and the large effects encountered in this study, the significance level of the reported correlation coefficients is always $p < 0.001$ (except on very few, identified cases). For all correlations, we also produce 95% confidence intervals (CI). They indicate the interval around the correlation coefficient of the sample, where there is a 95% probability of finding the correlation coefficient of the *entire population of the correlated variables.* CIs are important for visually comparing the differences between the obtained correlations (displayed in Figures 6, 8 and 9). When the confidence intervals of angular and Euclidean correlations don't overlap, we can be sure with 95% confidence that the difference between the two correlations is significant.

For each entry of the analysis matrix we will test the null hypothesis that the maximum correlation coefficients of angular-defined centrality and Euclidean-defined centrality with that movement sub-sample are equal (i.e. that the difference between the two maximum correlations will be zero). Our alternative hypothesis will state the opposite: that the maximum correlation coefficients of angular and Euclidean defined centrality will always be different (i.e. that the difference between both maximum correlation coefficients will not be zero). Although previous research (Hillier and Iida 2005) points clearly to the prevalence of angular over Euclidean results, we do not specify a direction for our alternative hypothesis, because the large size of our sample dispenses the added power of one-tailed tests and we do not wish to make *a priori* assumptions which may result in the non-detection of negative differences.

Let ρ(A) be the Spearman's correlation coefficient between observed vehicular movement and angular-defined centrality, and ρ(E) the coefficient between movement and Euclidean-defined centrality. We can formally define our null (H_0) and alternative (H_1) hypotheses as,

$$H_0: max|\rho(\mathrm{A})| - max|\rho(\mathrm{E})| = 0$$

$$H_1: max|\rho(\mathrm{A})| - max|\rho(\mathrm{E})| \neq 0$$

The significance level for rejecting H_0 will be $\alpha=0.05$. In order to ascertain the actual significance of the difference $max|\rho(\mathrm{A})| - max|\rho(\mathrm{E})|$, we will perform a specific test (Steiger 1980), implemented in the R package 'cocor' (Diedenhofen and Musch 2015), for the difference between two correlations obtained from the same sample (i.e. ρ(A) and ρ(E)) with one variable in common (i.e. vehicular movement flow). The result of the test is a z-score and H_0 is two-tailed, so the critical value will be $Z = \pm 1.96$ with $p < 0.05$. Except for that specific test, all other statistical procedures and calculations were carried out in JMP Pro (SAS 2015).

3. RESULTS

We start by studying the correlations between the values of angular and Euclidean-defined betweenness, of the nodes where movement was observed (Figure 5). The objective is to assess the degree of association between the network hierarchies induced by the two types of centrality, along the scale of radii defined above, before asking which one better emulates observed movement. This is done for 5 different sub-samples, namely 'urban roads' (principal [n=8,474] and minor [n=2,742]), 'rural roads' (principal [n=7,203] and minor [n=1,537]) and 'motorways' [n=796]. All correlations are significant at the $p < 0.001$ level. Figure 6 shows the results on line charts, with the correlation coefficient ρ on the y-axis and the several analysis radii on the x-axis.

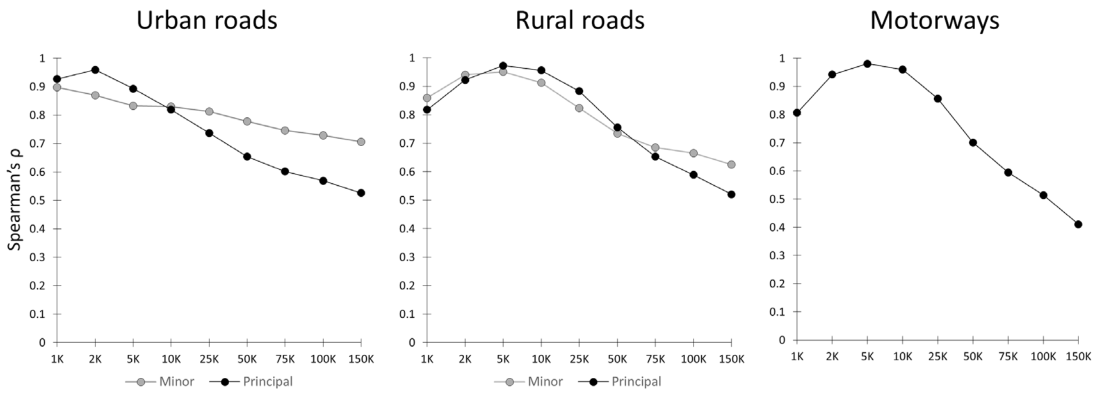

Figure 5 - Correlations for the values of angular and Euclidean-defined betweenness centrality.

We note that the values of angular and Euclidean betweenness centrality are strongly positively correlated – very much so at local radii (ρ = 0.96, R = 2 Km, on urban principal roads) and progressive less at larger radii. Thus, the network hierarchies induced by the two types of centrality are actually very similar when short distances are concerned, but they diverge as larger parts of the network are encompassed.

An important qualitative difference is noticeable between urban and rural roads. In cities, the correlations of principal roads (i.e. of the foreground network) decay faster than those of minor roads (i.e. of the background network), implying a clear structural differentiation between those two road-classes along spatial scales. In contrast, in rural contexts, principal and minor roads follow rather close correlation curves, implying a lesser structural differentiation between road-classes. Finally, the motorway's sub-sample shows a correlation curve that is similar to that of rural roads, but with an even stronger decay at large radii.

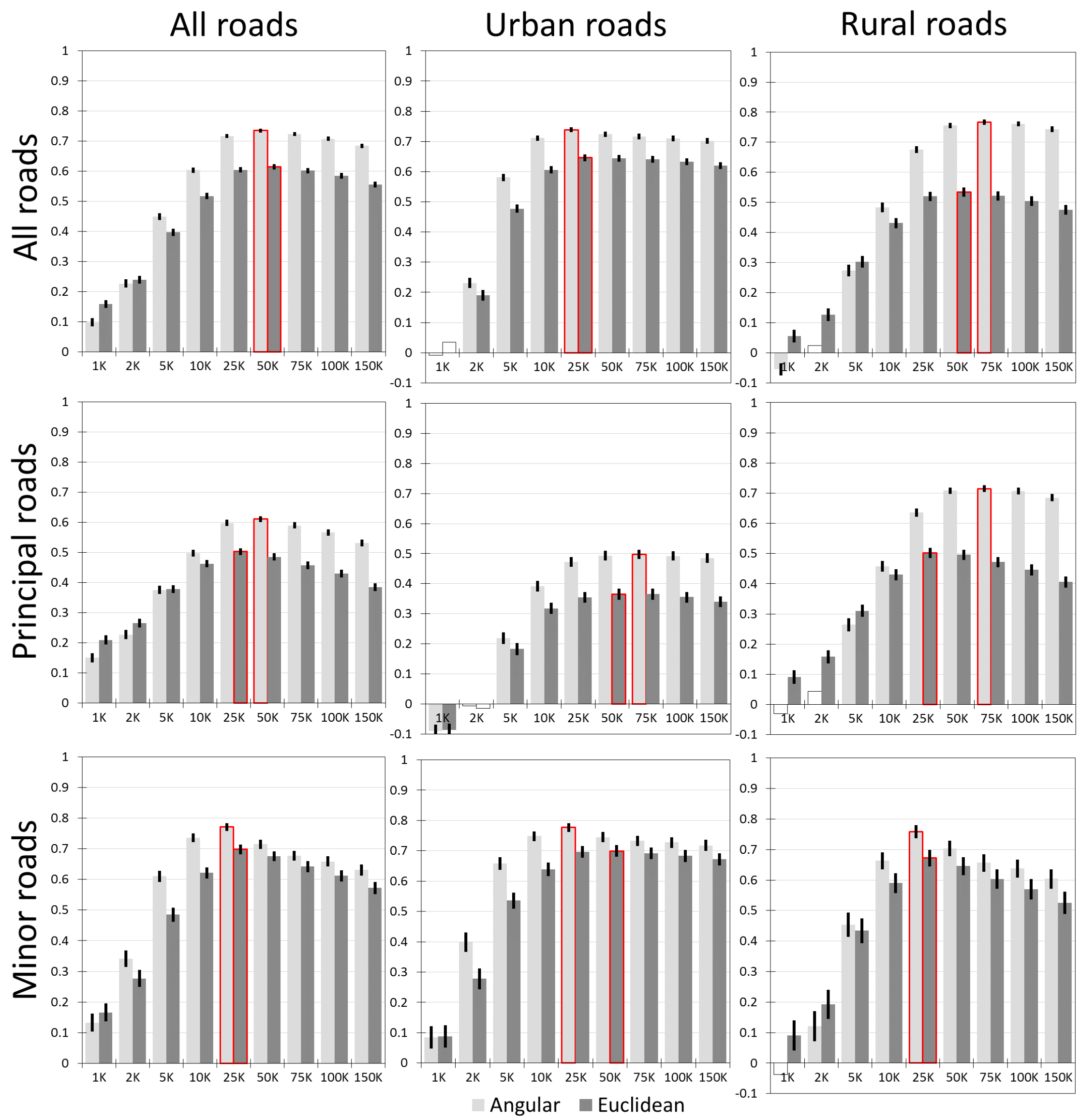

Figure 6 - Correlation results for 'all motor vehicles'.

Figures 6 and 9 display on bar charts the results of the main correlation exercise carried out in this study, organized according to the analysis matrix described before. On these charts, light grey bars represent the correlation coefficients obtained with angular-defined betweenness centrality and observed movement at each radius, or ρ(A); dark grey bars represent the correlations obtained with Euclidean-defined centrality and movement, or ρ(E). The maximal correlations in each sub-sample are highlighted in red. All correlations are significant at the p<0.001 level, except for very few cases, identified by non-coloured bars. Error bars represent 95% confidence intervals.

We start by looking at the results of 'all motor vehicles' (Figure 6). The first thing we should note is that, for all sub-samples, the maximal angular correlations are always higher than the maximal Euclidean correlations (both highlighted in red) and well beyond the limits of confidence intervals. We can thus immediately state that, for the 'all motor vehicles' class (which, we recall, represents 99.05% of all observed traffic), H_0 is rejected for all sub-samples at the p<0.05 level, with all the tested differences being positive; i.e. $\rho(A) > \rho(E)$. Beyond their statistical significance, the actual differences between the maximal correlations are in general quite large (i.e. they have also practical significance). The mean of the differences $max|\rho(A)|$ - $max|\rho(E)|$ for all samples is 0.126, with a maximal difference of 0.233 attained in 'all rural roads'.

Also, the observed effect sizes are considerable, with max $\rho(A) > 0.7$ in 7 of the 9 matrix entries.

	All roads	Urban roads	Rural roads
All roads	N = 19,952 max ρ(A) = 0.735, radius = 50 Km p<0.001, 95% CI [0.728,0.741] max ρ(E) = 0.614, radius = 50 Km p<0.001, 95% CI [0.605,0.623] max ρ(A) - max ρ(E)= +0.121 Z = +36.608, p < 0.0001	N = 11,216 max ρ(A) = 0.739, radius = 25 Km p<0.001, 95% CI [0.731,0.747] max ρ(E) = 0.646, radius = 25 Km p<0.001, 95% CI [0.633,0.655] max ρ(A) - max ρ(E)= +0.093 Z = +20.928, p < 0.0001	N = 8,740 max ρ(A) = 0.767, radius = 75 Km p<0.001, 95% CI [0.758,0.775] max ρ(E) = 0.534, radius = 50 Km p<0.001, 95% CI [0.519,0.549] max ρ(A) - max ρ(E)= +0.233 Z = +38.959, p < 0.0001
Principal roads	N = 15,677 max ρ(A) = 0.611, radius = 50 Km p<0.001, 95% CI [0.709,0.725] max ρ(E) = 0.503, radius = 25 Km p<0.001, 95% CI [0.491,0.514] max ρ(A) - max ρ(E)= +0.108 Z = +20.299, p < 0.0001	N = 8,474 max ρ(A) = 0.498, radius = 75 Km p<0.001, 95% CI [0.482,0.513] max ρ(E) = 0.365, radius = 50 Km p<0.001, 95% CI [0.347,0.383] max ρ(A) - max ρ(E)= +0.133 Z = +15.425, p < 0.0001	N = 7,203 max ρ(A) = 0.715, radius = 75 Km p<0.001, 95% CI [0.704,0.726] max ρ(E) = 0.502, radius = 25 Km p<0.001, 95% CI [0.485,0.519] max ρ(A) - max ρ(E)= +0.213 Z = +25.837, p < 0.0001
Minor roads	N = 4,077 max ρ(A) = 0.771, radius = 25 Km p<0.001, 95% CI [0.498,0.543] max ρ(E) = 0.698, radius = 25 Km p<0.001, 95% CI [0.476,0.522] max ρ(A) - max ρ(E)= +0.073 Z = +11.961, p < 0.0001	N = 2,742 max ρ(A) = 0.777, radius = 25 Km p<0.001, 95% CI [0.762,0.791] max ρ(E) = 0.700, radius = 50 Km p<0.001, 95% CI [0.680,0.718] max ρ(A) - max ρ(E)= +0.077 Z = +10.139, p < 0.0001	N = 1,537 max ρ(A) = 0.759, radius = 25 Km p<0.001, 95% CI [0.737,0.780] max ρ(E) = 0.673, radius = 25 Km p<0.001, 95% CI [0.644,0.699] max ρ(A) - max ρ(E)= +0.086 Z = +8.629, p < 0.0001

Reject H_0
Significant positive difference

Figure 7 - Hypothesis testing of the results for 'all motor vehicles'.

Figure 7 shows the actual z-scores and p-values of the Z test mentioned before (Steiger 1980); note the extreme positive values of Z (much higher than the critical value of Z=1.96) and the p-values always less than 0.0001, indicating the large differences observed and their high statistical significance. The consistency of these results across all sub-samples leaves no possible doubt that, at the scales where maximum correlations are attained (25Km, 50Km or 75Km), the network hierarchy described by angular-defined betweenness centrality emulates clearly better the actual usage of the road-network. These results strongly corroborate the findings of (Hillier and Iida 2005), validating them at the level of an entire country and within several of its geographical contexts.

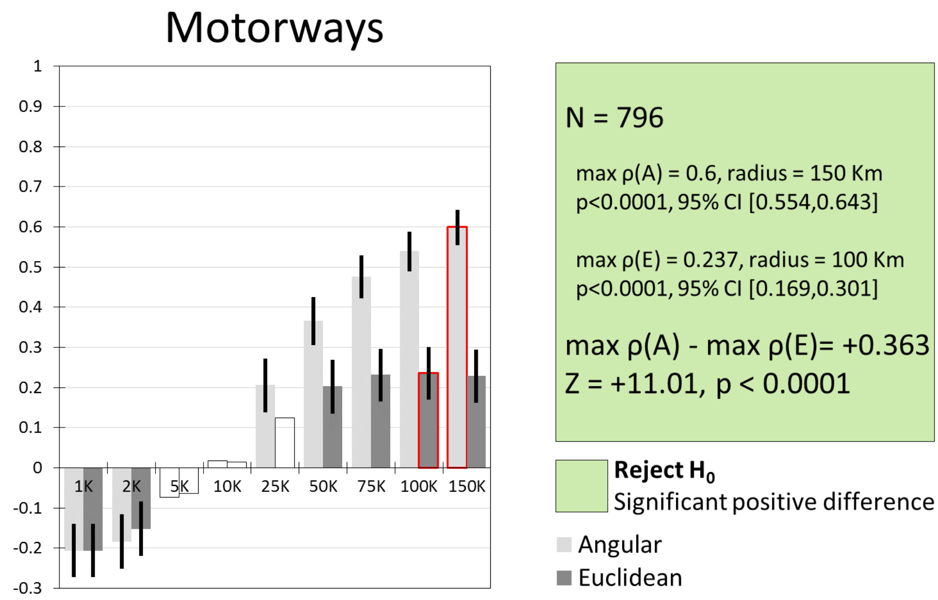

Figure 8 - Correlation results and hypothesis testing for 'motorways'.

The 'motorways' sub-sample produces similar results (Figure 8). The gap between the maximum correlations obtained with the two types of centrality is now even more clear, with angular-defined centrality attaining a coefficient (ρ=0.6) that is more than twice that of Euclidean-defined centrality (ρ=0.237). Therefore H_0 is again rejected without ambiguity. The radii at which these maximal correlations are attained (150 Km and 100 Km, respectively), as well as the clear negative correlations at local radii (1 Km and 2 Km), are consistent with the long-distance vehicular movement that motorways convey.

Finally, we look at the correlation results for the 'cycles' class of vehicles (Figure 9), which show a very different pattern. The first obvious observation, is that the previous large gap between angular and Euclidean correlations has vanished. In several sub-samples, the maximal Euclidean correlations are now slightly higher than the angular ones. Maxima are now attained at 5Km and 10Km (25Km in just one case), with correlations decaying fast afterwards (especially angular ones), reflecting the more localized range of cyclists trips. From the most local scale (1Km) until the scales at which maxima are attained (5-10Km), the differences between correlations are very small and with a general overlap of confidence intervals. Their significance cannot be assessed visually on the graphs, so we rely on the Z test for the difference between maximal correlations (Figure 10).

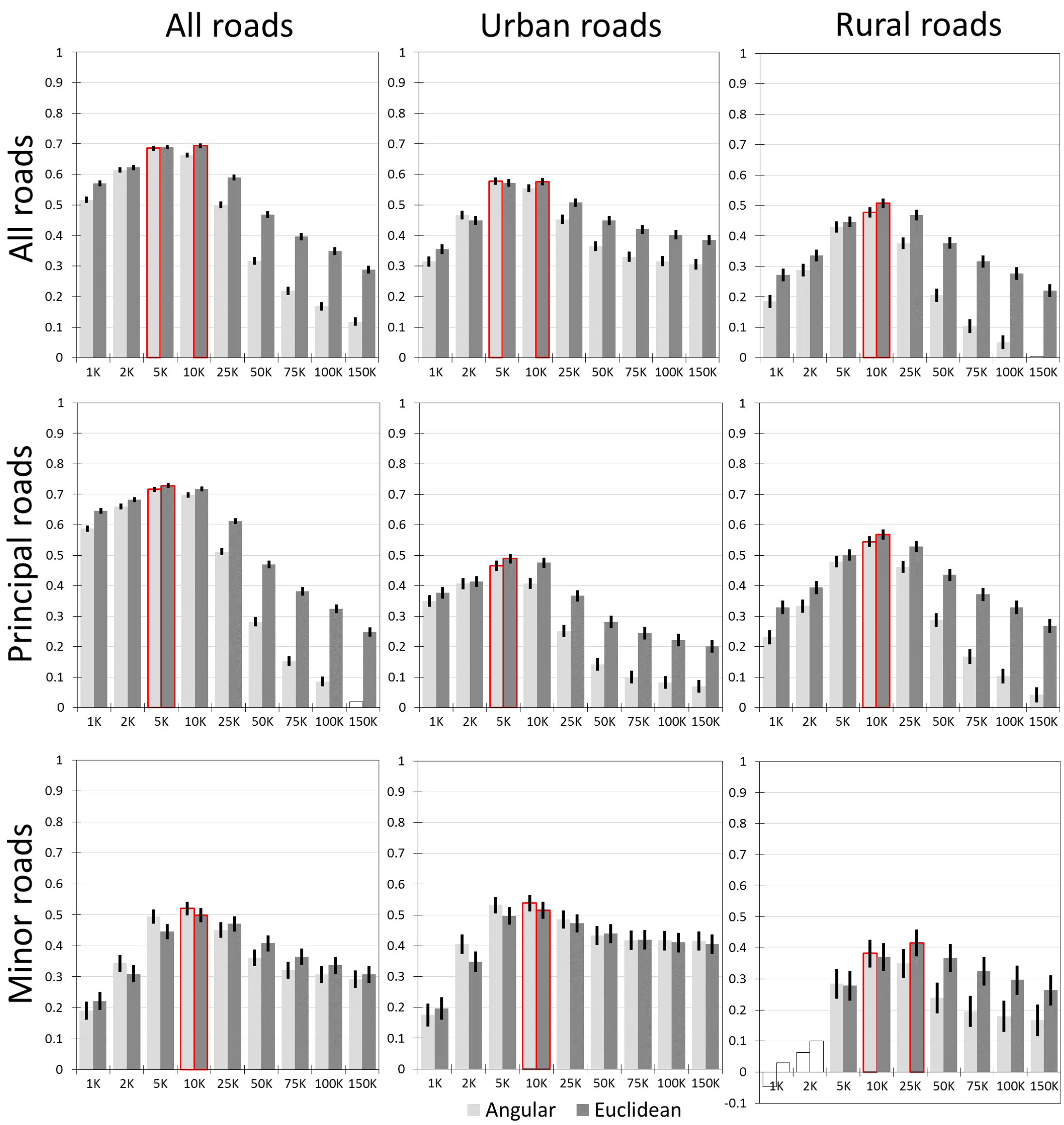

Figure 9 - Correlation results for 'cycles'.

We fail to reject H_0 in three cases, because there is no significant difference between the maximal correlations. We can reject H_0 in six cases; of these, two have significant positive differences (i.e. the correlation with angular-defined centrality is higher), but in the remaining four cases the differences are actually negative (i.e. the correlation with Euclidean-defined centrality is higher). Nevertheless, all the differences are small; it is the large sizes of our sub-samples that are capable of producing statistically significant results. However, statistical significance is not the same as practical significance. A difference of just | 0.032 | between correlation coefficients (the maximal significant observed difference for 'cycles', in 'all rural roads') has little or no relevance to the relative descriptive power of the two centrality measures. Thus, in this case, we cannot conclude for the superiority of any of the two tested definitions of betweenness centrality. Rather, we must note that the observed differences between the correlations of the two centralities types with the frequencies of 'cycles', have little practical significance and may be considered inconclusive.

	All roads	Urban roads	Rural roads
All roads	N = 18,659 max ρ(A) = 0.690, radius = 5 Km p<0.0001, 95% CI [0.678,0.694] max ρ(E) = 0.694, radius = 10 Km p<0.0001, 95% CI [0.686,0.701] max ρ(A) - max ρ(E)= -0.004 Z = -1.754, p = 0.079	N = 11,000 max ρ(A) = 0.578, radius = 5 Km p<0.0001, 95% CI [0.565,0.590] max ρ(E) = 0.577, radius = 10 Km p<0.0001, 95% CI [0.564,0.589] max ρ(A) - max ρ(E)= +0.001 Z = +0.234, p = 0.815	N = 7,659 max ρ(A) = 0.478, radius = 10 Km p<0.0001, 95% CI [0.461,0.495] max ρ(E) = 0.510, radius = 10 Km p<0.0001, 95% CI [0.491,0.524] max ρ(A) - max ρ(E)= -0.032 Z = -10.228, p < 0.0001
Principal roads	N = 14,582 max ρ(A) = 0.717, radius = 5 Km p<0.0001, 95% CI [0.709,0.725] max ρ(E) = 0.729, radius = 5 Km p<0.0001, 95% CI [0.721,0.736] max ρ(A) - max ρ(E)= -0.012 Z = -8.634, p < 0.0001	N = 8,335 max ρ(A) = 0.466, radius = 5 Km p<0.0001, 95% CI [0.449,0.483] max ρ(E) = 0.489, radius = 5 Km p<0.0001, 95% CI [0.473,0.510] max ρ(A) - max ρ(E)= -0.023 Z = -5.214, p < 0.0001	N = 6,247 max ρ(A) = 0.545, radius = 10 Km p<0.0001, 95% CI [0.523,0.562] max ρ(E) = 0.568, radius = 10 Km p<0.0001, 95% CI [0.551,0.585] max ρ(A) - max ρ(E)= -0.023 Z = -6.359, p < 0.0001
Minor roads	N = 4,077 max ρ(A) = 0.521, radius = 10 Km p<0.0001, 95% CI [0.498,0.543] max ρ(E) = 0.499, radius = 10 Km p<0.0001, 95% CI [0.476,0.522] max ρ(A) - max ρ(E)= +0.022 Z = +2.853, p = 0.003	N = 2,665 max ρ(A) = 0.540, radius = 10 Km p<0.0001, 95% CI [0.512,0.566] max ρ(E) = 0.515, radius = 10 Km p<0.0001, 95% CI [0.487,0.543] max ρ(A) - max ρ(E)= +0.025 Z = +2.637, p = 0.008	N = 1,412 max ρ(A) = 0.382, radius = 10 Km p<0.0001, 95% CI [0.337,0.426] max ρ(E) = 0.416, radius = 25 Km p<0.0001, 95% CI [0.372,0.458] max ρ(A) - max ρ(E)= -0.034 Z = -1.6669, p = 0.095

Reject H_0 Significant positive difference | **Fail to reject H_0** No significant difference | **Reject H_0** Significant negative difference

Figure 10 - Hypothesis testing of the results for 'cycles'.

Figure 11 summarizes the results of the overall correlation exercise, expressed as the values of the difference $\max|\rho(\mathrm{A})|$-$\max|\rho(\mathrm{E})|$, for each of the studied sub-samples; it also shows 95% confidence intervals for those differences, computed according to the procedure proposed by (Zou 2007) and implemented in (Diedenhofen and Musch 2015). The 'all motor vehicles' class (99.05% of the observed traffic) produced unambiguous positive correlation differences, for all sub-samples (mean of +0.15); whereas the same differences for the 'cycles' class (0.05% of the observed traffic) were less stable and much weaker, oscillating around zero (mean of -0.007) on the several sub-samples.

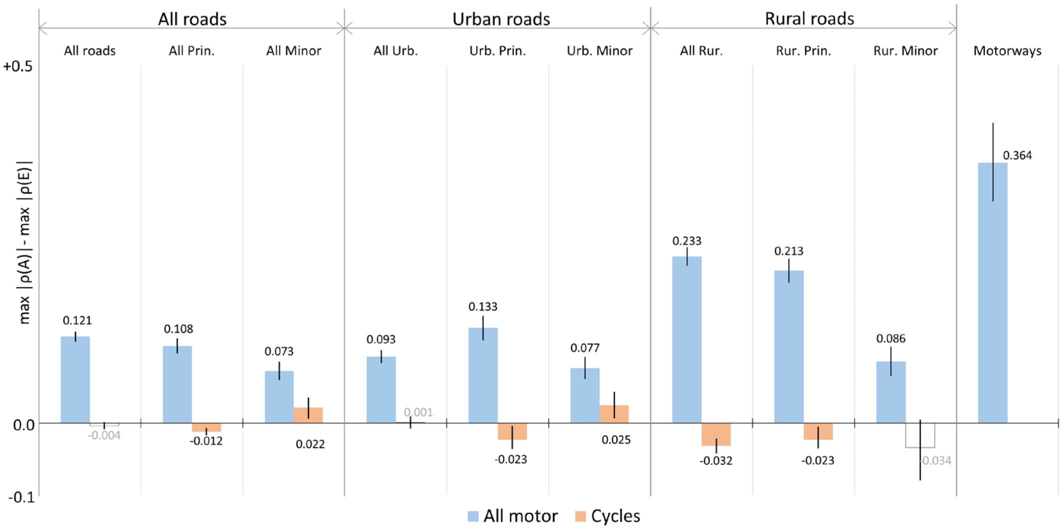

Figure 11 - Observed values of on all sub-samples, for 'all motor' and 'cycles'.

4. DISCUSSION AND CONCLUSIONS

4.1 ALL MOTOR VEHICLES

In all geographical and road-hierarchical contexts represented by our analysis matrix (i.e. from the whole UK's road-network to individual road-classes of urban and rural areas) and at the radii where the maximal correlations were attained, the network hierarchy induced by angular-defined betweenness centrality did emulate clearly better the actual usage of the network itself. However, the interpretation of this in terms of the drivers' cognitive reading of the geometric properties of the network, as proposed in (Hillier and Iida 2005), remains uncertain. It is however clear that angular shortest paths are also the *simplest paths*, i.e. those encoding the minimum amount of information (Rosvall et al. 2005). Thus, what we can state with renewed confidence, is that the space syntax prediction that road networks are hierarchized by *information minimization principles* (i.e. that simpler, thus straighter, network paths will correspond to more used streets and roads), is now supported by strong evidence. And that the network hierarchy induced by prioritizing paths by their least Euclidean lengths, has a clear weaker association with the observed movement intensities. The *informational content* of road-networks, whose morphological manifestations have since long been identified by space syntax (Hillier 1999, 2002, 2005), seems therefore to be an unavoidable and fundamental property of these objects, which can no longer be sidelined by ignoring the relevance of angular network distance.

But our results also shown that Euclidean properties should not be sidelined either. If we look again at Figure 6, focusing on the trajectories of the correlations' values along the analysis radii, we note that at the most local radius (1 Km) and in almost all sub-samples, and even though both correlations are low, Euclidean-defined centrality is in general stronger. The divergence between the correlations of the two types of centrality becomes unquestionable only after R=10 Km, when angular yields always higher correlations. Looking at Figure 5, one can also see that that is the radius at which the values of angular and Euclidean-defined betweenness centrality start to clearly diverge, after their strong initial correlation. Thus, even if the higher relevance of angular-defined centrality at the city-scale and beyond (i.e. R>10 Km) seems undisputable, our results also show a clear (albeit weak) signal that Euclidean structural principles are important at local scales; a fact already acknowledged in (Hillier 2006).

We also note that the differences between the values of the two correlations are clearly larger in 'principal roads' (which correspond to the foreground network) and narrower in 'minor roads' (corresponding to the background network). This is in strong accordance with the specific topo-geometrical properties of each of those generic networks, as described in (Hillier 2006); given its angular-minimizing morphology, we should expect the results on the foreground network to be particularly expressive, regarding the superiority of angular over Euclidean distance.

And indeed this is what happens. However, this effect is actually more pronounced in *rural than in urban roads* (see Figures 6 and 11). This needs an explanation, because one would also expect the differences between the geometries of the foreground and the background networks to be stronger in cities, where they were identified in the first place. In Figure 12 we show two scatterplots, of the 'all urban' and 'all rural' sub-samples, with angular-defined betweenness centrality values (R=75Km) on the x-axis, and 'all motor' vehicles AADF values on the y-axis (values are logged on both axes); minor roads are represented by red points and principal roads by blue points. Because of the noise in data, we fit a local kernel smoother (black curves) to each plot, in order to highlight the main trends in the clouds of points.

What we see when looking at Figure 12 is that there is a striking qualitative difference between urban *minor and principal* roads, that is not present in rural roads. In *rural roads*, the bivariate relationship between centrality and movement is linear. In other words, in all rural roads (minor and principal), more centrality means on average always more movement. But in *urban roads*, the average slope of the fitted curve is not the same for minor and principal roads (it is clearly lower on the latter group). This means that, in cities, from a certain threshold on, further gains in centrality will result only in marginal gains in movement. This is a clear sign of a saturation effect – a sudden and sustained decrease in the rate of response of one variable regarding the other. And the saturation threshold coincides with the minimum centrality level of principal roads; or, in other words, of the *foreground network.*

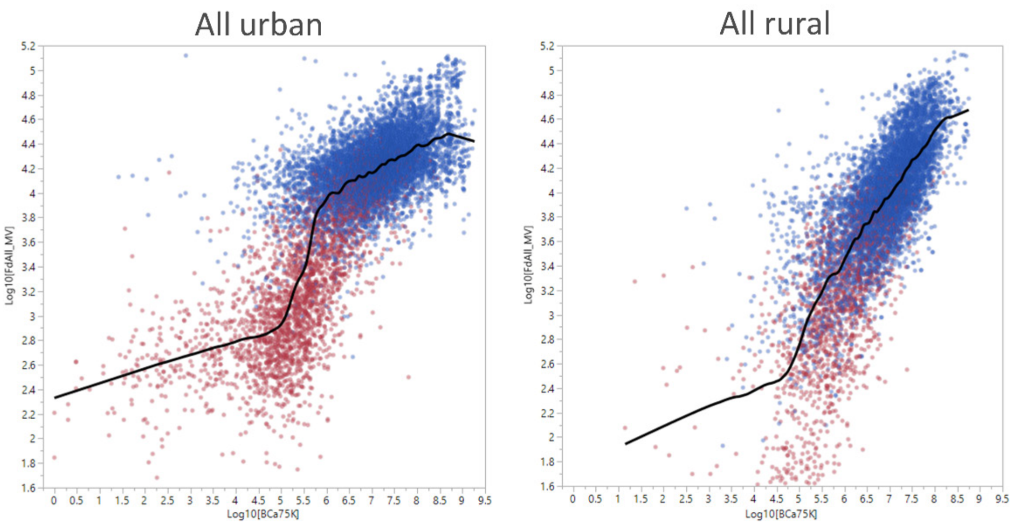

Figure 12 - Observed values of on all sub-samples, for 'all motor' and 'cycles'.

The saturation pattern for urban roads shown on Figure 12, may be seen as a signature (both structural and functional) of the dual generic morphological model of cities, proposed by space syntax. Such pattern implies that there is a sudden change from a system where low movement intensities increase gradually with centrality – that is, the *background network*; to another system where centrality is high, but where there is always lots of movement, with a more uniform intensity and least dependent of (thus, least correlated with) centrality variance – that is, the *foreground network*. Moreover, this effect is entirely absent in rural areas, providing also a very suggestive image of the intrinsic structural and functional differences between urban and rural road networks.

The lower dependence between the variances of movement and centrality in the foreground network of cities explains the lower correlations detected in the 'principal urban' roads sub-sample (see Figure 6). Indeed, notwithstanding the high movement intensities observed on those roads, the direct relationship between movement and centrality partially breaks down there, as if another variable was constraining it. We suggest this to be an effect of the spatial constraints that exist on cities, namely regarding existing road capacities and their potential increase.

We further explore this hypothesis with another dataset (DfT 2017), containing the width of the space available for vehicular circulation of principal roads (both urban and rural), at each count location of the main dataset used in this study (see endnote 1). We use multiple regression to study the inter-dependencies and relative importance of three factors, for predicting observed movement in urban and rural principal roads (Figure 13). These three factors are: Euclidean-defined betweenness centrality at radius 75 Km (noted as BCm75k, in Figure 13), angular-defined betweenness centrality at radius 75 Km (noted as BCa75k) and local road capacity (noted as Width).

Figure 13 reports the results of eight hierarchical OLS regression models, describing the impact of each movement predictor, in urban and rural principal roads. Each variable is inserted sequentially (i.e. hierarchically) into the models (see column 'Step' on Figure 13), in order to observe the change in two parameters: the standardized β coefficient (measuring the effect of each predictor on the dependent variable); and the change in R^2 (ΔR^2) when a variable is inserted last in the model (corresponding to its individual contribution in terms of explained movement variance, while controlling for the variable inserted first).

Principal **urban** roads

Model	Step	Variables	β	R2	ΔR2
1.1	1	Width	0.58	0.33	
	2	Width	0.46	0.44	0.11
		BCa75k	0.34		
1.2	1	BCa75k	0.5	0.25	
	2	BCa75k	0.34	0.44	**0.19**
		Width	**0.46**		
1.3	1	Width	0.58	0.33	
	2	Width	0.52	0.39	0.06
		BCm75k	0.25		
1.4	1	BCm75k	0.37	0.14	
	2	BCm75k	0.25	0.39	**0.25**
		Width	**0.52**		

Principal **rural** roads

Model	Step	Variables	β	R2	ΔR2
2.1	1	Width	1.12	0.48	
	2	Width	0.73	0.66	**0.18**
		BCa75k	**0.93**		
2.2	1	BCa75k	1.36	0.51	
	2	BCa75k	0.93	0.66	0.15
		Width	0.73		
2.3	1	Width	1.12	0.48	
	2	Width	0.96	0.6	0.12
		BCm75k	0.62		
2.4	1	BCm75k	0.91	0.27	
	2	BCm75k	0.62	0.6	**0.33**
		Width	**0.96**		

Figure 13 - Multiple-regression models, exploring the variances explained (ΔR^2) by each of the movement predictor variables (BCa75k, BCm75k and Width), while controlling for the others.

Euclidean-defined centrality is always a worse movement predictor than road capacity, both in urban and rural principal roads (models 1.4 and 2.4, respectively). The same is not true for angular-defined centrality. Although in principal rural roads angular is capable of explaining more variance than width (model 2.1), the situation is inverted in urban principal roads, with width playing a more important role (model 1.2). We see then, that in the foreground network of cities (i.e. on urban principal roads), the relationship between spatial centrality and movement is constrained by existing road capacity.

We must bear in mind that movement potential, as expressed by network centrality, is a more primitive and more fundamental characteristic than road capacity. Intuitively, one would expect the latter factor to be determined by the former and this must indeed be so, if no other spatial constraints are present (as it is the case of rural settings). However, urban space is by definition scarce and urban streets, when completely delimited by buildings, create very strong limits to increases in road capacity. We thus provisionally propose that the saturation pattern shown

on Figure 12, is the product of the spatial constraints characteristic of cities, which impose restrictions on the direct centrality / movement relationship.

This new insight, which is only touched upon here, will be theme for further research. But the finding of the foreground network's saturation regime sheds new empirical light on the dual model of urban form proposed by space syntax. It shows that the foreground network, more than just a main web of movement, may be seen as a whole phenomenon on its own right, highly differentiated from the rest of the city, both functionally and structurally.

4.2 BICYCLES

Despite the much smaller representativeness of the 'cycles' vehicular class regarding the overall observed traffic (0.05%), we have found that that class of vehicles produced a very different correlation pattern with the two types of centrality studied. In contrast with the remaining observed traffic, cycles yield very small differences between the correlations with Euclidean and angular-defined centrality, which in some cases were actually non-significant. Such an undifferentiated behaviour demands of course some reflexion.

Previous space syntax studies addressing cyclist flows have found significant correlations with angular-defined centrality indicators, but always in conjunction with other variables in multiple regression models. Studying cycles flows in two central London local areas, (Raford, Chiaradia et al. 2007) report significant correlations of R^2=0.67 and R^2=0.76, with angular-defined closeness centrality combined with segment length and a dummy variable representing the presence of cycling lanes. Also in a London local area, Law, Sakr et al. (2014) report a coefficient of R^2=0.66 for normalized angular-defined betweenness also combined with the presence of cycling lanes. However, these two studies do not contemplate the option of introducing Euclidean-defined centrality measures in their models. Cooper (2017) uses a complex version of network distance, including Euclidean and angular distance factors mixed with road slope and traffic volumes, for calculating betweenness centrality on Cardiff's entire street network. The author reports a maximum association of r=0.78 (R^2=0.61), between the composite betweenness centrality measure and observed cyclists flows.

Although the results of these studies are hardly comparable in numerical terms, we note that the range of the detected effect sizes is similar ($R^2 \approx 0.7$). In this paper, the maximum effect sizes observed for the 'cycles' class of vehicles were ρ(A) = 0.72 and ρ(E) = 0.73, at radius 5Km, in the 'all principal' roads sub-sample. This coefficients are lower than the ones cited before (as they are were not squared), but our sample is also much larger. Also, we use simple bi-variate correlations and not multiple regression models (which naturally yield higher correlations, due to the presence of multiple factors). But the main difference is that the above mentioned studies do not compare the performances of angular and Euclidean-defined centrality and thus do not provide information on that regard. Our main finding regarding 'cycles' does not concern the size of the maximal effects obtained with angular and Euclidean-defined centralities (which were large, at any rate), but rather the fact that the differences between such effects were negligible. Actually, there is no obvious reason to assume that cyclists would behave exactly in the same way as the generality of motorized vehicles.

Indeed, discrete choice modelling of cyclists' route preferences (Menghini, Carrasco et al. 2010; Broach, Dill et al. 2012) shows that cyclist route choice is highly idiosyncratic and influenced by many factors. Euclidean distance seems to be by far the most important negative factor, followed by a clear aversion for high traffic volumes and strong slopes. However, cyclists are also quite sensitive to turn frequency, preferring simple routes. Our results seem to be in line with these findings, with Euclidean distance postdicting marginally better the observed cycles flows, but being followed very closely by angular distance. We suggest that the minimal differences observed between the two distance types should reflect the overlap of the negative and positive route choice factors mentioned above.

However, given the relative incompleteness of our network model at the most fine-grained scale, and also given the relative spatial sparseness of the count points of the movement sample used here, we cannot deem our results conclusive for the 'cycles' vehicular class. Indeed, in the space syntax studies reviewed above (except Cooper 2017), the spatial distributions of the count locations were highly dense (as they covered local urban areas), with count locations on almost every street segment. These enhanced sampling densities can produce results different from ours, because cyclist movement (in contrast with that of motorized vehicles) is prone to follow less stable (or more unpredictable) routes, in the sense of not being altogether constrained to the spaces of motorized vehicular circulation. As it is the case of pedestrian movement, the study of cyclists' movement might depend on high spatial resolution samples, which is not the case of the sample used here. In this sense, the inconclusiveness of our results regarding cyclists' movement, clearly points to the need to investigate this theme more intensely, in order to understand the true roles of angular and Euclidean distance, in the movement patterns and route choice strategies of cyclists.

ENDNOTES

1 This dataset was obtained by personal communication (Richard German, October 28, 2016), through the email address ROADTRAFF.STATS@dft.gsi.gov.uk.

ACKNOWLEDGEMENTS

M. Serra is thankful to Dr. Fernanda Sousa for her advices regarding statistical methods and procedures.

M. Serra is supported by the FCT grant SFRH/BPD/111260/2015

REFERENCES

Broach, J., et al. (2012). "Where do cyclists ride? A route choice model developed with revealed preference GPS data." Transportation Research Part A(46): 1720-1740.

Carvalho, R. and A. Penn (2004). "Scalling and universality in the micro-structure of urban space." Physica A(332): 539-547.

Cooper, C. (2017). "Using spatial network analysis to model pedal cycle flows, risk and mode choice." Journal of Transport Geography 58: 157-165.

Cooper, C. H. V. (2015). "Spatial localization of closeness and betweenness measures: a self-contradictory but useful form of network analysis." International Journal of Geographical Information Science 29(8): 1293-1309.

Cooper, C. H. V., et al. (2014). "Measuring the impact of spatial network layout on community social cohesion: a cross-sectional study." International Journal of Health Geographics 13(1): 11.

Crucitti, P., et al. (2006). "Centrality measures in spatial networks of urban streets." Physical Review E 73(3): 036125.

DfT (2014). Traffic Counts [data set]. Department for Transports, UK Government. Retrived from https://www.dft.gov.uk/traffic-counts/.

DfT (2016). Road Use Statistics, Great Britain 2016, Department for Transport, UK Government.

DfT (2017). Road widths at the vehicular movement count locations in principal roads [data file], Department for Transports, UK Government. Unpublished dataset, cited with permission.

Diedenhofen, B. and J. Musch (2015). "cocor: A Comprehensive Solution for the Statistical Comparison of Correlations." PLoS ONE 10(4): e0121945. doi: 0121910.0121371/journal.pone.0121945.

Freeman, L. (1977). "A set of measures of centrality based on betweenness." Sociometry 40: 35-41.

Gil, J. (2014). "Analysing the Configuration of Multimodal Urban Networks." Geographical Analysis 46: 368-391.

Hillier, B. (1996). "Cities as Movement Economies." Urban Design International 1(1): 41-60.

Hillier, B. (1999). "The Hidden Geometry of Deformed Grids: or, why space syntax works, when it looks as though it shouldn't." Environment and Planning B: Planning and Design 26(2): 169-191.

Hillier, B. (2002). "A Theory of the City as Object, or, how spatial laws mediate the social construction of urbans space." Urban Design International 7: 153-179.

Hillier, B. (2009). Spatial Sustainability in Cities: organic patterns and sustainable forms 7th International Space Syntax Symposium. KTH Faculty of Architecture, Stockholm.

Hillier, B. (2012). The Genetic Code for Cities: Is It Simpler than We Think? Complexity Theories of Cities Have Come of Age: An Overview with Implications to Urban Planning and Design. J. Portugali, H. Meyer, E. Stolk and E. Tan. Berlin, Heidelberg, Springer Berlin Heidelberg: 129-152.

Hillier, B. (2016). "What are cities for? And how does this relate to their spatial form?" The Journal of Space Syntax 6(2): 199-212.

Hillier, B. and S. Iida (2005). Network and psychological effects in urban movement. Proceedings of the 2005 international conference on Spatial Information Theory. Ellicottville, NY, Springer-Verlag: 475-490.

Hillier, B., et al. (1993). "Natural Movement: or, configuration and attraction in urban pedestrian movement." Environment and Planning B: Planning and Design 20(1): 29-66.

Hillier, B., et al. (2006). Metric and Topo-Geometric Properties of Urban Street Networks: some convergences, divergences and new results. 6th International Space Syntax Symposium, ITU Faculty of Architecture, Instambul.

Law, S., et al. (2014). "Measuring the Changes in Aggregate Cycling Patterns between 2003 and 2012 from a Space Syntax Perspective." Behavioral Sciences 4(3): 278-300.

Masucci, A. P., et al. (2009). "Random planar graphs and the London street network." Eur. Phys. J. B 71(2): 259-271.

Menghini, G., et al. (2010). "Route choice of cyclists in Zurich." Transportation Research Part A(44): 754-765.

Molinero, C., et al. (2015) The angular nature of road networks. ArXiv: 1512.05659v1 [physics.soc-ph]

OS (2014). Meridian 2 [data set]. Ordnance Survey Limited. Retrived from https://www.ordnancesurvey.co.uk/business-and-government/products/meridian2.html.

Penn, A., et al. (1998). "Configurational Modelling of Urban Movement Networks." Environment and Planning B: Planning and Design 25(1): 59-84.

Porta, S., et al. (2012). "Street Centrality and the Location of Economic Activities in Barcelona." Urban Studies 49(7): 1471-1488.

Porta, S., et al. (2009). "Street Centrality and Densities of Retail and Services in Bologna, Italy." Environment and Planning B: Planning and Design 36(3): 450-465.

Raford, N., et al. (2007). Space Syntax: the role of urban form in cyclist route choice in central London. Transportation Research Board 86th Annual Meeting. Washington DC, USA.

Rosvall, M., et al. (2005). "Networks and Cities: an Information Perspective." Physical Review Letters 94, 028701.

SAS (2015). JMP Pro 12.0.1, SAS Institute Inc., Cary, NC.

Scellato, S., et al. (2006). "The backbone of a city." Eur. Phys. J. B 50(1-2): 221-225.

Steiger, J. H. (1980). "Tests for comparing elements of a correlation matrix." Psychological Bulletin 87: 241-251.

Strano, E., et al. (2012). "Elementary processes governing the evolution of road networks." Scientific Reports 2: 296.

Strano, E., et al. (2013). "Urban street networks, a comparative analysis of ten European cities." Environment and Planning B: Planning and Design 40: 1071-1086.

Varoudis, T. (2012). depthmapX - Multi-platform spatial network analysis software, The Bartlett School of Architecture, UCL, http://varoudis.github.io/depthmapX/.

Zou, G. Y. (2007). "Toward Using Confidence Intervals to Compare Correlations." Psychological Methods 12(4): 399-413.