

Angular and Metric Distance in Road Network Analysis: a nationwide correlational study

Miguel Serra^a , Bill Hillier^b

^a Citta – Research Centre for Territory, Transports and Environment, Faculty of Engineering, University of Porto, Portugal

^b Space Syntax Laboratory, The Bartlett School of Architecture, University College London, United Kingdom

ABSTRACT

There is an epistemological divide in the field of road network analysis, concerning the way network distance should be conceptualized. On one hand, the generality of studies in the field adopt metric distance as a self-evident choice. On the other, space syntax studies adopt a different distance concept, namely that of angular distance, which ignores actual physical lengths. Theoretically, these two distance concepts imply quite different assumptions. Analytically, they produce also different results.

In this paper we assess the differences between these two network distance concepts, using a model of the UK's complete road network (2,031,971 nodes) and a very large dataset of vehicular movement counts (20,752 locations). We describe the statistical associations between observed vehicular flows and the betweenness centrality of the road-network nodes where such flows were measured, the latter calculated with metric and angular distance functions, across a number of increasing radii. Relations to road capacity are also discussed in principal roads where this is known.

The geographical comprehensiveness of our model and the size of our movement sample allow us to state, with unprecedented statistical validity, the clear outperformance of angular distance over metric distance, in what concerns the strength of the studied correlations. This is true for all types of roads (urban and non-urban) and for all motorized vehicles, representing 99.05% of the total traffic; the remaining 0.95% correspond to pedal cycles, which yielded inconclusive results. We also demonstrate the existence of two distinct regimes of association between movement and centrality, conspicuous in cities but altogether absent in non-urban areas, which represent new evidence supporting space syntax's dual model of urban form.

1. INTRODUCTION

Node centrality assessment is a common procedure in the analysis of street and road networks. It allows the ranking of nodes according to how central they are, thereby revealing the network's underlying structural hierarchy. It has been shown by many studies, as for example (Hillier et al. 1993, Hillier and Iida 2005, Jiang 2009, Jayasinghe et al. 2015), that such centrality hierarchies are strongly associated with observed movement flows, both pedestrian and vehicular. This association is of great importance, because it demonstrates that centrality analysis indeed captures essential functional properties of street and road networks.

Centrality is quantified through graph theoretical measures, which are based on the preliminary calculation of the shortest paths between all pairs of nodes in the network. This calculation

total angular variations are not. Path a) is as simple as possible – a straight line between two points. Because it has minimal descriptive complexity, it also contains minimal information; angular distance measures its length as zero (details on the calculation of angular distance will be given later). Path b) is more complex, but still far from random. It has one evident regularity, namely the fact that all segments are joined at wide obtuse angles. It is closer to the sequence with lowest complexity than to any random sequence (i.e. it approximates linearity). Angular distance measures it as having length 0.95. Path c) is composed by angles of all amplitudes, without any apparent regularity. It is therefore the most complex, thus with the highest information content. Angular distance measures it as having length 6.64.

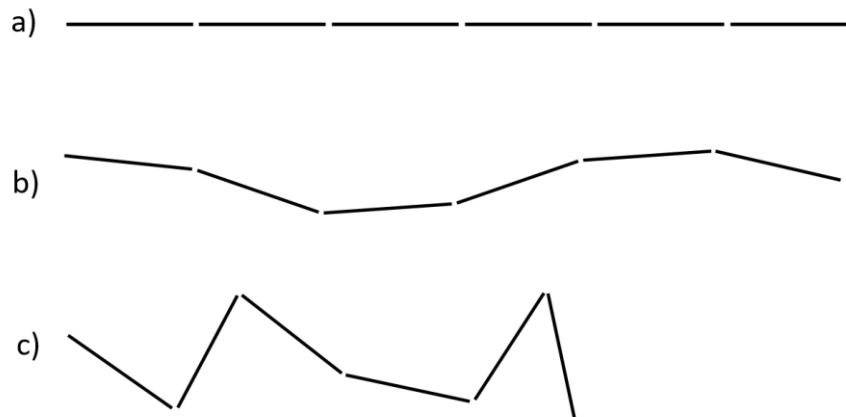


Figure 1. Three notional street network paths, with the same metric length but with increasingly higher angular length and descriptive complexity.

The difference between the three segment sequences in Figure 1 is obviously their descriptive complexity; i.e. the extent to which the presence (or the absence) of regularities in their geometrical variation, allows (or not) for the compression of the sequence into a simpler description. Therefore, to define the shortest paths as those with less amount of angular change (or, alternatively, with less information content), amounts to the following theoretical assumption: that the functionally meaningful hierarchy of street networks is to be determined by *information minimization principles* (i.e. that geometrically simpler paths will have a higher probability of being used and are therefore more important).

Space syntax's concept of angular distance arose from the close attention that the field has always granted to the *specific geometry* of street networks, in addition to their topology. Indeed, what separates space syntax's axial representation from other types of dual street network representations, is that the node-defining entities (i.e. axial lines) are determined exclusively by the *actual geometric constraints* that the built environment imposes on linear continuity. This way of defining the atoms of urban space is purely geometric, even if their relationships are modelled topologically.

The transition from the axial to the segment map (in which axial lines are broken into segments at intersections) and to its associated angular distance function, was a step towards increasing the discriminant power of space syntax's analytical model. As mentioned previously, the angular distance between perfectly aligned axial segments is zero – they are treated as a single spatial unit (i.e. as an axial line, see Figure 1 a). But if we look at sequences b) and c) in Figure 1 as also made of axial lines (i.e. ignoring their angular variations), we conclude that the axial model would be unable to distinguish them, because their unweighted graphs would be isomorphic. By taking into account the angles at which segments intersect, the segment model equipped with an angular distance function, is able to do so.

In their 2005 paper, Hillier and Iida provided strong evidence for the precedence of angular distance over metric distance, as revealed by its higher capability to predict pedestrian and vehicular movement flows, observed in four London local areas. This result was later confirmed by Turner (2009), through different analytical means and using a larger and more detailed movement sample of vehicular GPS traces. Turner (2007) also showed that angular distance could have an additional advantage – that of allowing the integration of the standard GIS street network representation, namely road-centre line (RCL) maps, into space syntax analysis.

Axial lines acquire their structural relevance from the fact of being *minimal geometric descriptions*. When they are converted into segments, these retain the original geometric parsimony. In contrast, RCL maps often reproduce irrelevant geometric details of the geographic features from which they derive. Thus, the dual graphs of the two representations are often of quite different orders. However, their structural differences may become insignificant if the edges of both graphs are weighted with angular distance, because the total amount of angular change is bound to be approximately the same in both representations (Turner 2007).

Beyond these methodological advantages, the concept of angular distance contributed also to the development of the *dual network model* of urban form, enunciated in (Hillier 1999) and detailed in (Hillier et al. 2010, Hillier 2012). This model describes the *generic form of the city* as a dual structure, composed of: i) a *foreground network* of main paths, made-up of a small number of long lines joined at wide obtuse angles, forming a web of highly central, multi-directional alignments; embedded in a ii) *background network* of local streets, made-up of a much larger number of shorter lines, forming grid-like patterns with lower centrality. Because of their different centrality levels, these two fundamental network tiers have also different *movement potentials*, with the foreground network carrying the bulk of urban movement and the background network being relatively secluded. This difference in movement potentials causes the foreground network to attract all types of public functions, while the background network remains mostly residential.

Such model assumes that urban space is hierarchized through subtle geometric regularities (Hillier 1999), making the descriptive complexity of higher order paths smaller than that of lower order ones. Indeed, angular centrality without radius restriction systematically identifies the foreground network (i.e. the most important roads), whereas global metric centrality often fails to do so (Hillier et al. 2010). This strongly suggests that the foreground network is indeed a web of *simplest paths* (i.e. paths with small angular change, thus with low descriptive complexity), rather than one of shortest paths (in the sense of those with least metric length).

There is therefore an epistemological divide in the field of street network analysis, concerning the question of how network distance should be conceptualized. On one hand, space syntax claims that this should be done through angular distance, because that is the way urban space is actually hierarchized. On the other, the majority of the studies from outside the space syntax field assume that, insofar as urban space has a structural hierarchy, such a hierarchy must be the outcome of metric distance relationships.

In this paper we readdress this epistemological 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 and angular and metric distance concepts (applied to the calculation of betweenness centrality), in order to assess their empirical and theoretical value. However, due to the sizes of both our street network model (the UK's entire road network) and vehicular movement sample (20,752 count points), the results reported here have a far superior validity than those reported in (Hillier and Iida 2005).

Also in contrast with (Hillier and Iida 2005), we will not interpret our results from the point of view of the individual's cognitive reading of the network's properties. Such inferential step has been questioned (Omer and Jiang 2015) and indeed it remains open to debate (Omer and Kaplan 2018). Instead, we focus our attention on the relevance of the physical properties of 'metric length' and 'information content', for the description of the hierarchical structure of street networks. We assess such relevance by comparing the network hierarchies induced by the two properties, with movement data describing the actual functioning of the network. But our object of interest is the structural nature of the network itself, and not how that structure is cognitively appropriated at the individual's level. Specifically, in this paper we ask:

- Which physical description of the UK's road network better reflects its actual use: one based on the metric length of network paths, or one based on their geometric descriptive complexity?
- Assuming there are significant differences between the correlations of the two physical descriptions with observed movement: how does that vary across geographic contexts, spatial scales and vehicle types?

In what follows, we first describe in detail our research methodology, the source datasets and the adopted analytical procedures. Next, we report the results of the main correlation exercise and we discuss their implications to our research questions. We conclude by summarizing our findings and by considering their relevance to the understanding of the generic structure of street and road networks.

2. DATASETS AND METHODS

2.1 The road network model

Our road network model is based on the Meridian 2 dataset (OS 2015), which represents the full hierarchy of Great Britain's road network, but not its absolute geometric constitution. Road representation is skeletal, collapsed into single RCLs, 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. For details on the generalization process, please see (OS 2015).

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

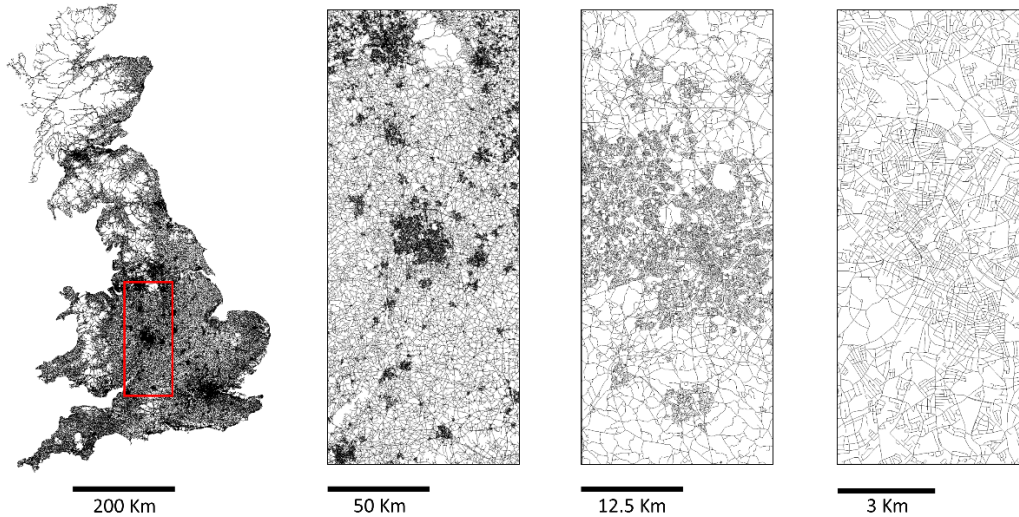


Figure 2. The road network model. Each of the images on the right depict sequential 400% zoom-ins of the red rectangle on the full map.

As in any syntactic segment model, individual line segments are encoded as the nodes $V = \{1, \dots, N\}$ 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 – angular and metric – denoted here respectively as w_a and w_m . The angular distance between two adjacent nodes, $w_a(i, j)$, is proportional to the angle of incidence θ between 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, $w_a(i, j)$ is defined as,

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

The metric distance between two adjacent nodes, $w_m(i, j)$, is the sum of the metric lengths of the segments encoded by i and j , denoted l_i and l_j , divided by 2; i.e. the actual length between the segments' mid-points, measured along the segments in metric units. Formally, $w_m(i, j)$ is defined as,

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

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 metric 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. We will use the following set of radii,

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

Ranging from the local scale (i.e. 1 Km), through the city scale (e.g. 10 Km) and up to the supra-regional scale (i.e. 150 Km). The two distance concepts are applied to the calculation of 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 as,

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

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 . Betweenness centrality quantifies how often a given node lies on the shortest paths between other nodes. From the point of view of vehicular movement, it may be seen as a direct indicator of the flow potential of a given node. The road-network model was processed in the network analysis software UCL DepthmapX, for each of the two distance concepts, at each of the network radii mentioned above.

2.2 The vehicular movement sample

Our vehicular movement sample is based on a publically available dataset 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 (DfT 2018). An AADF is the average over a full year of the number of vehicles passing a point in the road network each day (in both directions); for details on the generation of AADF values, please see (DfT 2018). After several pre-processing operations², we have validated a study sample of 20,752 count points (91% of the original dataset).

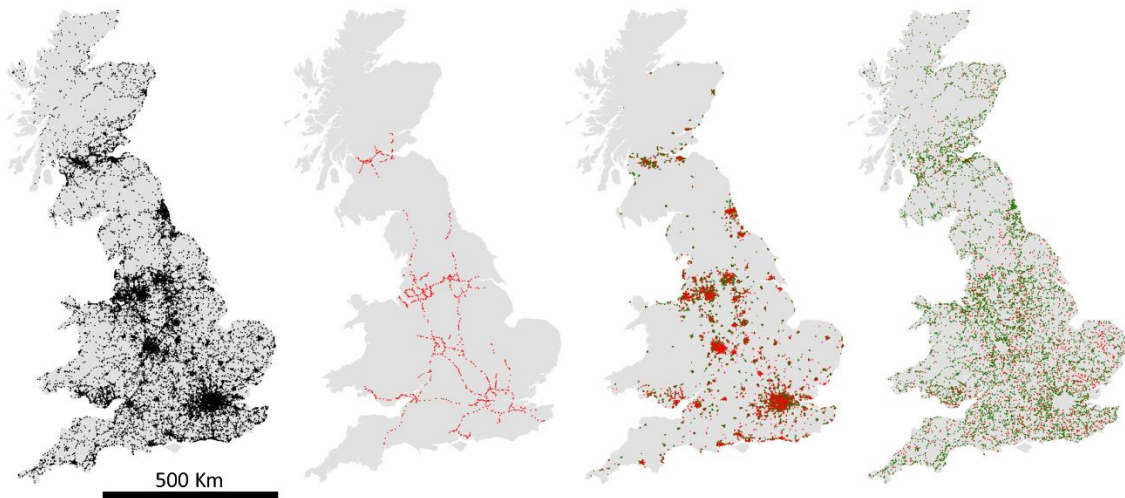


Fig.3 – Geographical distributions of count points. From left to right: ‘all’, ‘motorways’, ‘urban roads’ and ‘rural roads’ (‘principal’ in red, ‘minor’ in green).

The count points are geographically evenly distributed (Figure 3), but their frequencies per road class are very different. There are 12 road classes in the original dataset (Table 1), but the large majority of points (66.8%) are located on principal urban roads (PU, 39.6%) and on principal rural roads (PR, 27.2%), with all other 10 road classes representing only 33.2% of the occurrences. Thus, as it stands, the sample has a strong bias towards principal roads.

² Each count point is located precisely by its easting and northing coordinates and by a ‘road name’ attribute. However, because the geometry of Meridian 2 is simplified, some points do not fall exactly on the model’s lines, or are indeed located on lines that are not represented in the model (e.g. on motorways access ramps). We first discarded all points located more than 10 meters from the nearest line. We then assigned the remaining points to the nearest line with the same road name. We further discarded the points without road name match.

Original road class	Name	Frequency	Description	Aggregated road class	Name	Frequency
PM	Principal motorway	N = 18 (0.1%)	Double-carriageway roads of strategic importance, maintained by the local authorities.	MW	Motorways	N = 796 (3.9%)
TM	Trunk motorway	N = 778 (3.7%)	Double-carriageway roads of strategic importance, maintained by the central government.			
PU	Principal urban road	N = 8213 (39.6%)	A-roads in urban areas (pop. >= 10,000), maintained by the local authorities.	PU	Principal urban roads	N = 8474 (41.4%)
TU	Trunk urban road	N = 261 (1.3%)	A-roads in urban areas (pop. >= 10,000), maintained by the central government.			
PR	Principal rural road	N = 5637 (27.2%)	A-roads outside urban areas, maintained by the local authorities.	PR	Principal rural roads	N = 7203 (35.2%)
TR	Trunk rural road	N = 1566 (7.5%)	A-roads outside urban areas, maintained by the central government.			
BU	Urban B-road	N = 485 (2.3%)	B-roads in urban areas (pop.>= 10,000).	MU	Minor urban roads	N = 2742 (13.4%)
CU	Urban C-road	N = 541 (2.6%)	C-roads in urban areas (pop. >=10,000).			
UU	Unclassified urban road	N = 1716 (8.3%)	Local roads (mainly residential) in urban areas (pop. >=10,000).			
BR	Rural B-road	N = 670 (3.2%)	B-roads outside urban areas.	MR	Minor rural roads	N = 1537 (6%)
CR	Rural C-road	N = 492 (2.4%)	C-roads outside urban areas.			
UR	Unclassified rural road	N = 375 (1.8%)	Local roads (mainly residential but including rural lanes) outside urban areas.			

Tab.1 – Original and aggregated road classification schemes.

The original 12 road classes were first aggregated into a simpler scheme of just five classes (see table 1), namely: motorways, principal urban and rural roads, minor urban and rural roads; this is also the classification scheme adopted by DfT on their annual transport statistics reports (DfT 2018). Then, in order to mitigate the bias towards principal roads, the sample was studied under progressive levels of disaggregation, starting with the full sample and ending on individual road classes (see Figure 5).

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 aggregated, except bicycles). In order to decide which types of vehicles to study, we first want to see if their frequencies at the count locations are correlated or not. Logically, vehicle types whose frequencies are highly correlated must have similar behaviours in terms of road network use and are therefore interchangeable.

Such correlations are in general high, albeit with some variability. We use principal component analysis (PCA) to identify the dominant collinearity trends, extracting two principal components with eigenvalues higher than 1. These describe 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 vehicle type; see Figure 4).

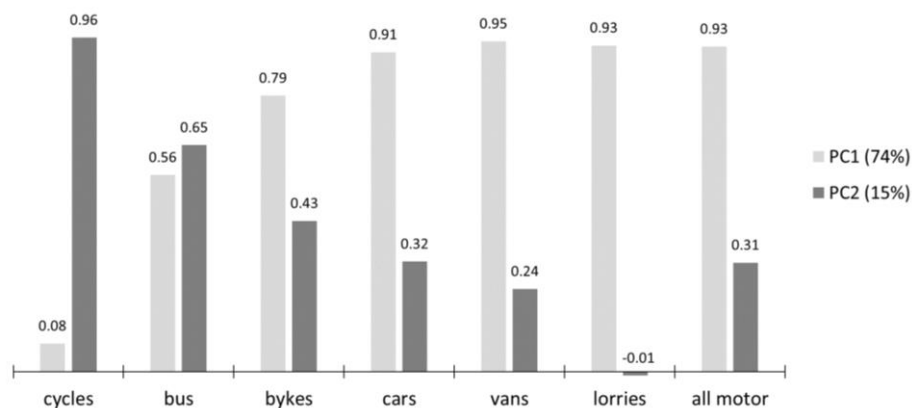


Fig.4 – Loadings of the two PCs describing the associations between the frequencies of vehicles types.

The first component (PC1), which is responsible for almost all variance explained (74%), is highly correlated [$r > 0.9$] with the frequencies of “cars”, “vans”, “lorries” and “all motor vehicles”; “bikes” are also strongly correlated [$r=0.79$] with PC1. The second component (PC2), explaining only residual variance (15%), is highly correlated with the frequencies of “cycles” [$r=0.96$] and, to a lesser extent, with those of “bus” [$r=0.65$].

We thus observe two distinct frequency trends, among the several types of vehicles: 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 classes against betweenness centrality values. It is important to bear in mind, however, that the “all motor vehicles” class corresponds to 99.05% of the counted vehicles, while the “cycles” class to just 0.95%. Therefore, the “all motor” class represents the overwhelming majority of the traffic flowing in the network, whereas “cycles” represent a very small (but highly differentiated) minority.

2.3 Analytical and statistical procedures

The sample was studied under progressive levels of desegregation accordingly to an analysis matrix (Figure 5), 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 matrix entry corresponds to a different sized sub-sample, corresponding to specific hierarchical and geographical contexts. When reading the matrix vertically, one can also get a picture of the results by socio-demographic context (i.e. for the whole country, only on cities or on rural areas). And horizontally, one may see how the results relate to the previously described foreground/background network model (i.e. on the whole network, on the foreground or background networks). The sizes of ‘all motor vehicles’ and ‘cycles’ sub-samples differ slightly, because in the ‘cycles’ case we only consider the counts higher than zero. Motorways [$n=752$], because they are not classified either as urban or rural road-infrastructures, are left out of the analysis matrix and their results will be presented apart.

Both movement and centrality variables strongly deviate from bivariate normality, with many outliers. We use therefore a non-parametric correlation method, namely Spearman’s rank correlation coefficient, denoted as ρ (rho), which does not assume normality and is not affected by the presence of outliers. Spearman’s ρ measures the association between the ranks of two variables, instead of their actual values. But this is particularly fit for our research subject, because we are not interested in the specific values of either movement or centrality; rather, we want to know which type of centrality hierarchy (angular or metric) is more associated with the relative magnitudes of observed traffic flows.

Given the large size of our sample and the large effects encountered in this study, the significance level of the reported correlation coefficients will always be $p < 0.001$ (except on very few, identified cases). For all correlations, we also produce 95% confidence intervals (CI). They indicate the interval around the sample’s correlation coefficient, where there is a 95% probability of finding the correlation coefficient of the *entire population of the correlated variables* (i.e. of all streets and roads in the UK and all vehicles circulating therein).

		Road geographical context					
		All roads	Urban roads	Rural roads			
Road hierarchical 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	Space syntax theoretical context	
	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.
		The Whole Country	The Cities	The Countryside			
		Socio-demographic context					

Fig.5 – Analysis matrix.

For each entry of the analysis matrix (i.e for each sub-sample), we will test the null hypothesis that the maximum correlation coefficients of angular and metric centrality with movement, 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 the two definitions of centrality will always be different (i.e. that their difference will not be zero). Let $\rho(A)$ be the Spearman's correlation coefficient between observed vehicular movement and angular-defined centrality, and $\rho(M)$ the coefficient between movement and metric-defined centrality. We can formally state our null (H_0) and alternative (H_1) hypotheses as,

$$H_0: \max|\rho(A)| - \max|\rho(M)| = 0$$

$$H_1: \max|\rho(A)| - \max|\rho(M)| \neq 0$$

The significance level for rejecting H_0 will be $\alpha=0.05$. In order to ascertain the significance of the difference $\max|\rho(A)| - \max|\rho(M)|$, we will perform a specific Z-test (Steiger 1980) for the difference between two different correlations obtained from the same sample, with one variable in common; i.e. $\rho(A)$ and $\rho(M)$, with the common variable AADF. 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 $p < 0.05$.

3. RESULTS

We start by studying the correlations between the values of angular and metric-defined betweenness, of the nodes where movement was observed (Figure 6). 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 was 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.

We note that the values of angular and metric betweenness centrality are strongly positively correlated – very much so at local radii ($\rho = 0.96$, $R = 2$ 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. We should therefore expect the correlations with movement to be similar at local scales, but different at global scales.

An important qualitative difference is noticeable between urban and rural roads. In cities, the angular/metric 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 sharp structural differentiation between those two road-classes. In contrast, in rural contexts, principal and minor roads follow rather close correlation curves, implying a lesser structural differentiation. 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.

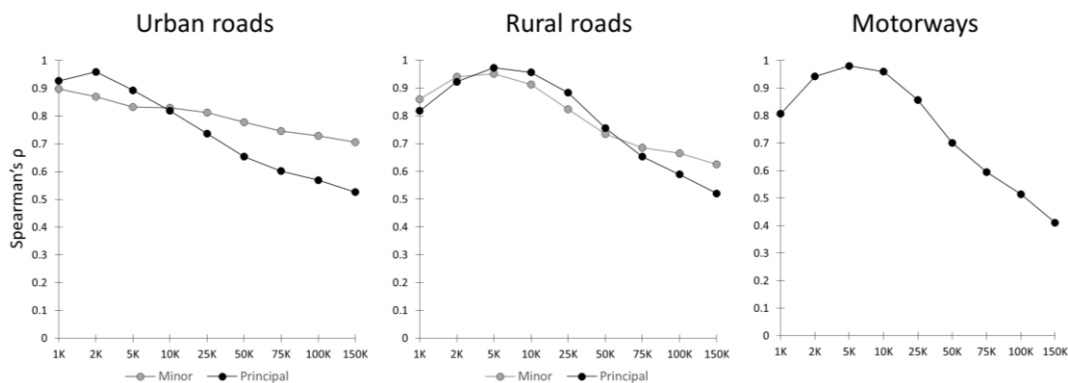


Fig.6 – Spearman’s correlation coefficients of angular and metric-defined betweenness centrality (y axis), across the several radii of analysis (x axis).

Figures 7 and 10 display on bar charts the results of the main correlation exercise, organized according to the analysis matrix described before. 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 metric 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. Furthermore, the differences between the maximal correlations are in general quite large (i.e. they have also practical significance). The mean of the differences $\max|\rho(A)| - \max|\rho(M)|$ 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.

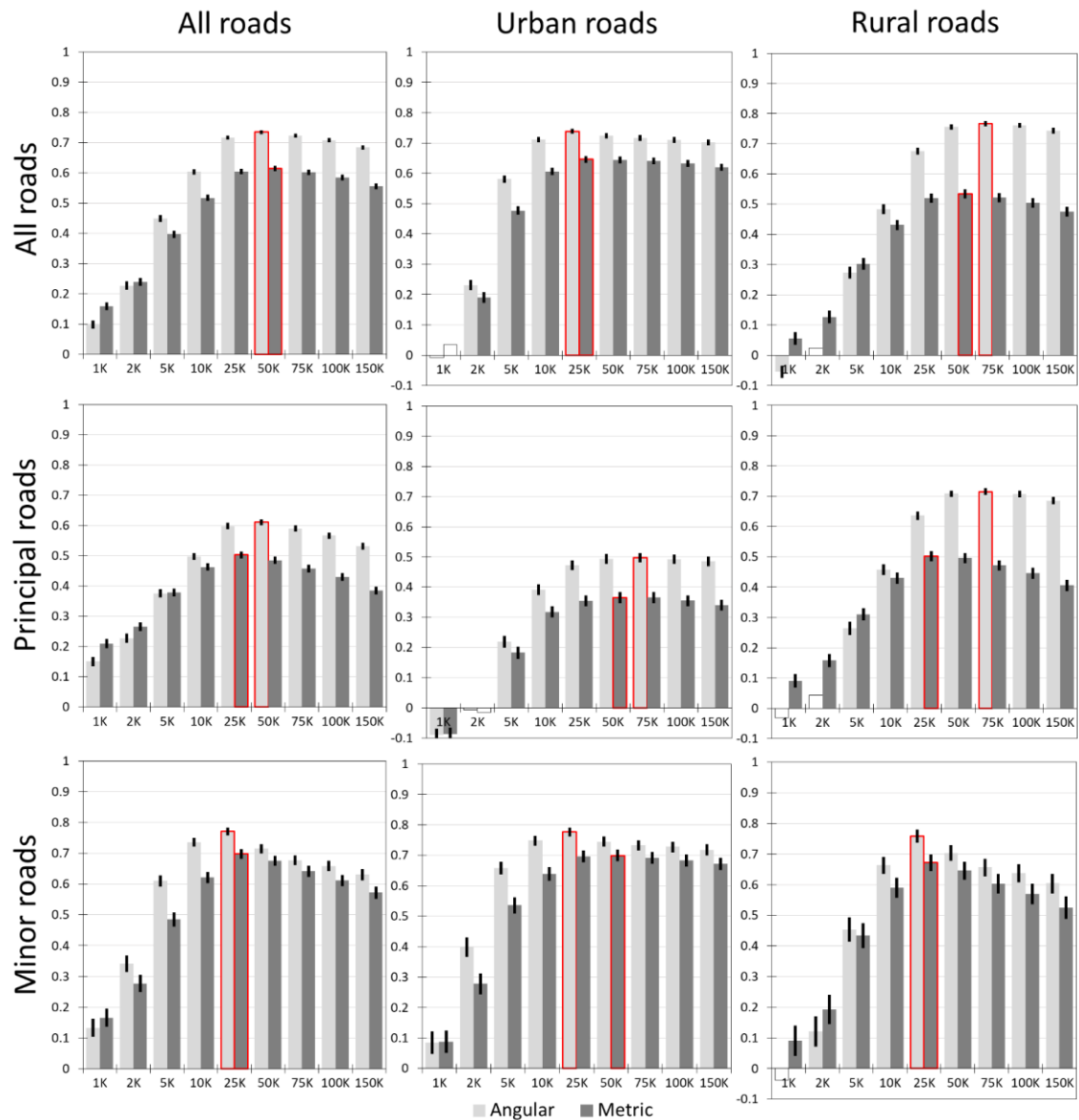


Fig.7 – Correlation results for ‘all motor vehicles’.

Figure 8 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 respectively the large differences observed and their high statistical significance.

The ‘motorways’ sub-sample produces similar results (Figure 9). 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 ($\rho=0.6$) that is more than twice that of metric-defined centrality ($\rho=0.237$). 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.

	All roads	Urban roads	Rural roads	
All roads	N = 19,952 max $\rho(A)$ = 0.735, radius = 50 Km $p < 0.001$, 95% CI [0.728, 0.741] max $\rho(M)$ = 0.614, radius = 50 Km $p < 0.001$, 95% CI [0.605, 0.623] max $\rho(A) - \max \rho(M)$ = +0.121 Z = +36.324, $p < 0.0001$	N = 11,216 max $\rho(A)$ = 0.739, radius = 25 Km $p < 0.001$, 95% CI [0.731, 0.747] max $\rho(M)$ = 0.646, radius = 25 Km $p < 0.001$, 95% CI [0.633, 0.655] max $\rho(A) - \max \rho(M)$ = +0.093 Z = +20.928, $p < 0.0001$	N = 8,740 max $\rho(A)$ = 0.767, radius = 75 Km $p < 0.001$, 95% CI [0.758, 0.775] max $\rho(M)$ = 0.534, radius = 50 Km $p < 0.001$, 95% CI [0.519, 0.549] max $\rho(A) - \max \rho(M)$ = +0.233 Z = +38.959, $p < 0.0001$	
	Principal roads	N = 15,677 max $\rho(A)$ = 0.611, radius = 50 Km $p < 0.001$, 95% CI [0.709, 0.725] max $\rho(M)$ = 0.503, radius = 25 Km $p < 0.001$, 95% CI [0.491, 0.514] max $\rho(A) - \max \rho(M)$ = +0.108 Z = +20.248, $p < 0.0001$	N = 8,474 max $\rho(A)$ = 0.498, radius = 75 Km $p < 0.001$, 95% CI [0.482, 0.513] max $\rho(M)$ = 0.365, radius = 50 Km $p < 0.001$, 95% CI [0.347, 0.383] max $\rho(A) - \max \rho(M)$ = +0.133 Z = +15.425, $p < 0.0001$	N = 7,203 max $\rho(A)$ = 0.715, radius = 75 Km $p < 0.001$, 95% CI [0.704, 0.726] max $\rho(M)$ = 0.502, radius = 25 Km $p < 0.001$, 95% CI [0.485, 0.519] max $\rho(A) - \max \rho(M)$ = +0.213 Z = +25.837, $p < 0.0001$
		Minor roads	N = 4,077 max $\rho(A)$ = 0.771, radius = 25 Km $p < 0.001$, 95% CI [0.498, 0.543] max $\rho(M)$ = 0.698, radius = 25 Km $p < 0.001$, 95% CI [0.476, 0.522] max $\rho(A) - \max \rho(M)$ = +0.073 Z = +11.908, $p < 0.0001$	N = 2,742 max $\rho(A)$ = 0.777, radius = 25 Km $p < 0.001$, 95% CI [0.762, 0.791] max $\rho(M)$ = 0.700, radius = 50 Km $p < 0.001$, 95% CI [0.680, 0.718] max $\rho(A) - \max \rho(M)$ = +0.077 Z = +10.139, $p < 0.0001$

Reject H_0
 Significant positive difference

Fig.8 – Hypothesis testing of the results for ‘all motor vehicles’.

Motorways

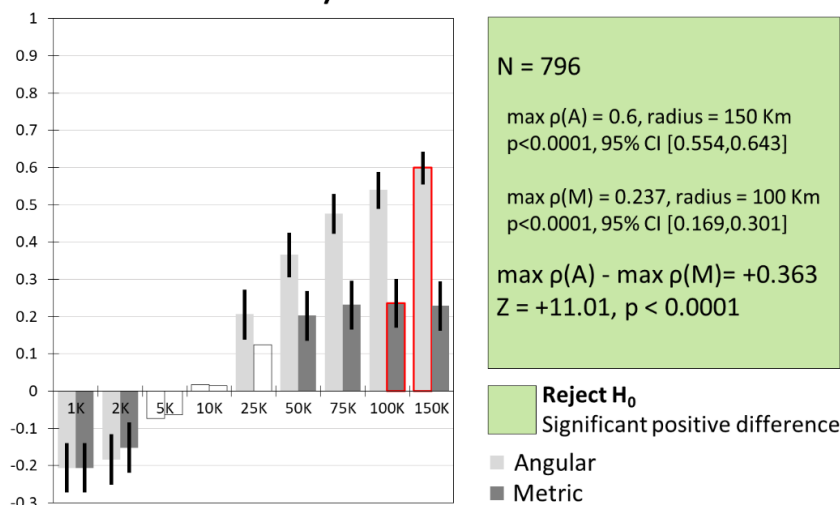


Fig.9 – Correlation results and hypothesis testing for ‘motorways’.

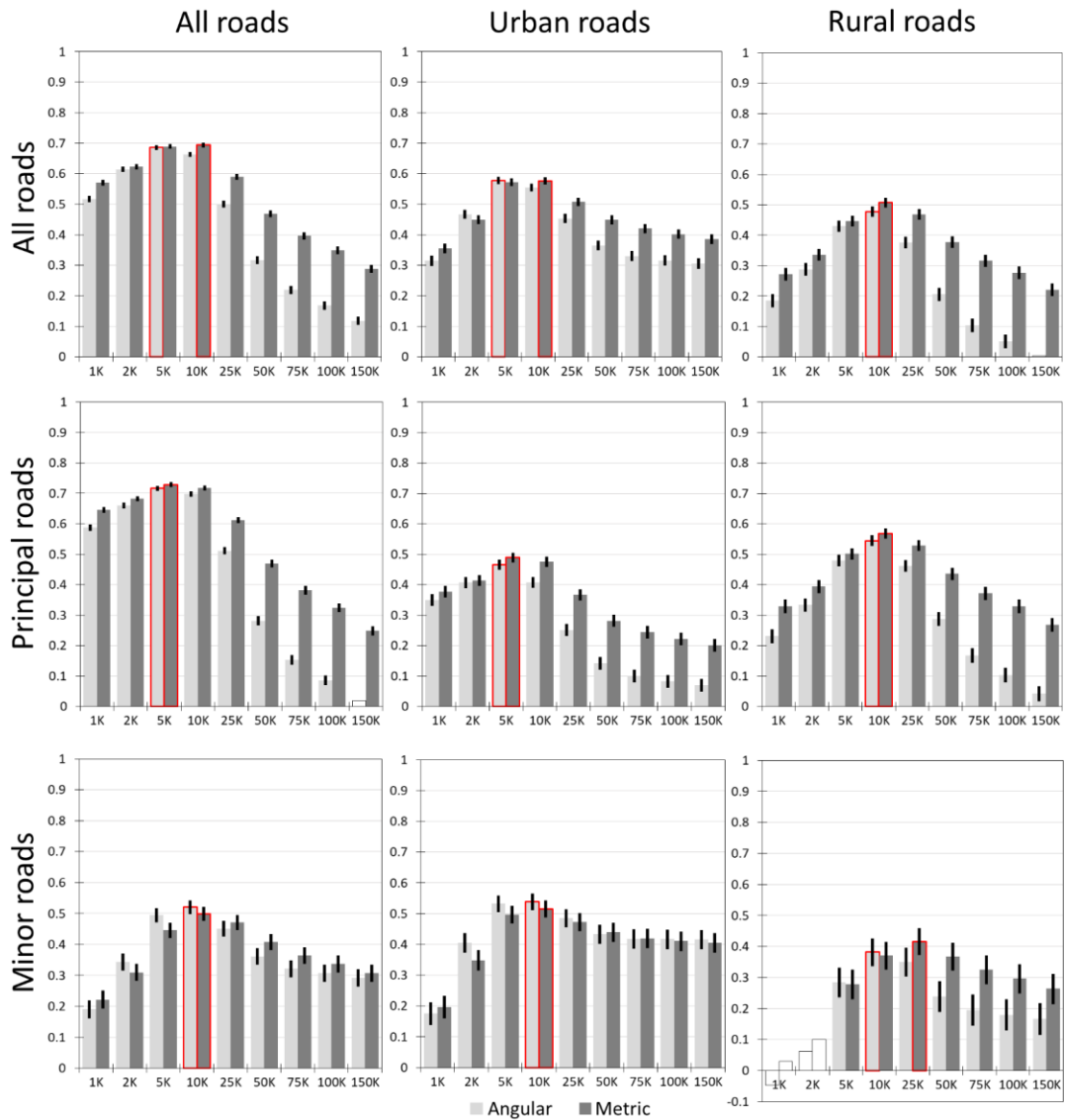


Fig.10 – Correlation results for ‘cycles’.

Finally, we look at the correlation results for the ‘cycles’ class of vehicles (Figure 10), which show a very different pattern. The first obvious observation, is that the previous large gap between angular and metric correlations has vanished. In several sub-samples the maximal metric 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, so we rely on the Z test for the assessment of the difference between maximal correlations (Figure 11).

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 centrality is higher), but in the remaining cases the differences are actually negative (i.e. the correlation with metric centrality is higher). Nevertheless, all the differences are small; they are significant only because the samples are large. However, statistical significance is not the same as practical significance. A difference of just -0.032 between correlation coefficients (the largest difference for ‘cycles’, in ‘all rural roads’) has very little meaning for the assessment of their relative postdictive validity.

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

Reject H_0 Significant positive difference	Fail to reject H_0 No significant difference	Reject H_0 Significant negative difference
---	---	---

Fig.11 – Hypothesis testing of the results for ‘cycles’.

Figure 12 summarizes the results of the overall correlation exercise. The chart on the top shows the values of the difference $max|\rho(A)| - max|\rho(M)|$ for each of the studied sub-samples, bounded by 95% confidence intervals, computed according to (Zou 2007). The ‘all motor vehicles’ class (99.05% of the observed traffic) produced unambiguous positive differences (mean of +0.15). The differences obtained for the ‘cycles’ class (0.95% of the observed traffic) were less stable and much weaker, oscillating around zero (mean of -0.007).

In order to make these effects comparable and easier to grasp, we also provide on Figure 12 shows the relative percentage difference in the correlation coefficients (i.e. their absolute difference, divided by the mean and multiplied by 100). Thus, for ‘all motor’ vehicles and in all road-classes excluding motorways, angular distance entails an average relative increase of 21% (s.d. 11%) in the correlation coefficients, reaching a staggering relative increase of 87% in motorways. In contrast, for the ‘pedal cycles’ vehicle class the relative changes in correlation coefficients are much lower (mean -1%, s.d. 5%, discarding non-significant differences).

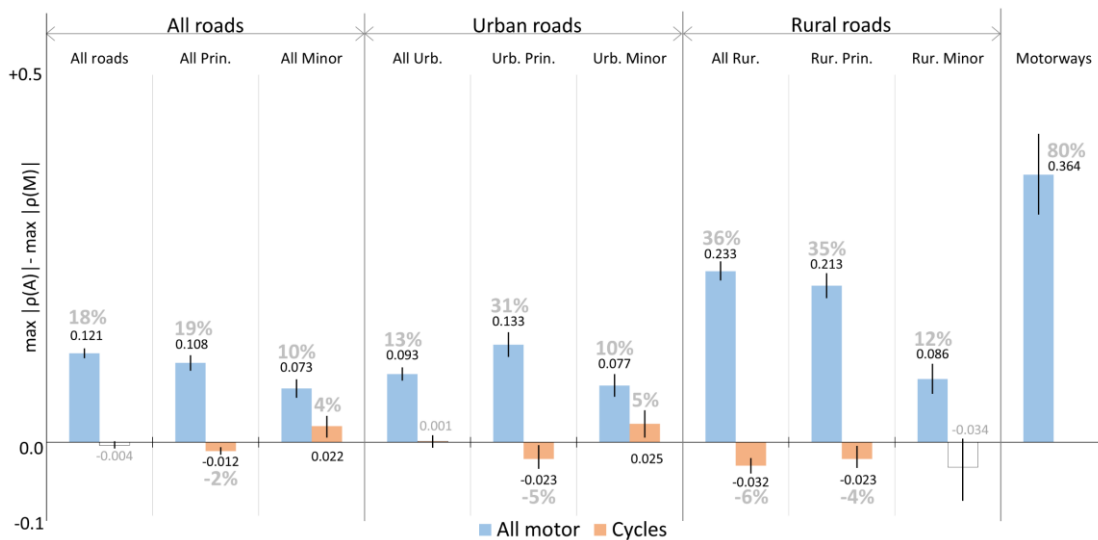


Fig.12 – Values of $\max |\rho(A)| - \max |\rho(M)|$ (in black typeface) and relative percentage change of correlation coefficients (in grey typeface) on all sub-samples, for 'all motor' and 'cycles' vehicle classes.

4. DISCUSSION

4.1 All motor vehicles

The 'all motor vehicles' class corresponds to 99.05% of all occurrences and represents the overwhelming majority of the traffic flowing in UK's road network. For such a majority, the correlations with angular-defined centrality proved to be always higher than those obtained with metric-defined centrality. This was found to be true on all geographical and road-hierarchical contexts represented by our analysis matrix (i.e. from the whole UK road-network to individual road-classes of urban and rural areas), at the radii where the maximal correlations were attained. Angular distance represents also a significant accuracy improvement over metric distance, namely an average relative increase of 21% in the predictive power of betweenness centrality.

But this does not imply that metric properties are simply irrelevant. Firstly, as we have seen in Figure 6, at low radii the hierarchies of angular and metric centralities are virtually equal. Secondly, if we look again at Figure 7, we will see that at the most local radius (1 Km) metric centrality is in general stronger, even though both correlations are low. The divergence between the two correlations becomes unquestionable only after 10 Km, when angular becomes clearly superior. Returning to Figure 6, one finds that 10 Km is also the radius at which the values of angular and metric centrality start to clearly diverge, after their strong initial correlation. Thus, even if the higher relevance of angular centrality at the city-scale and beyond (i.e. $R > 10$ Km) seems undisputable, our results also show that metric properties are important at the local scale, a fact already acknowledged in (Hillier et al. 2010).

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 geometric properties of each of those generic networks, as described by the dual network model of urban form. Given the angular-minimizing morphology of the foreground network, we should expect the prevalence of angular distance to be particularly expressive there.

And indeed this is what happens. However, this effect is actually more pronounced in *rural than in urban roads* (see Figure 7). This in turn demands an explanation, because one would also

expect the differences between the geometries of the foreground and the background networks to be clearer in cities, where they were identified in the first place. In Figure 13 we show two scatterplots, of the 'all urban' and 'all rural' sub-samples, with angular betweenness ($R=75\text{Km}$) on the x-axis, and 'all motor' 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.

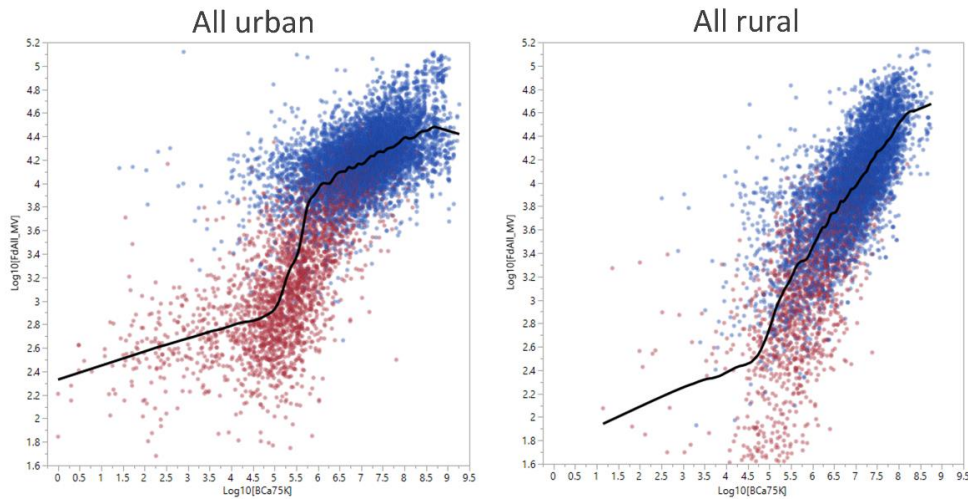


Fig.13 – Scatterplots of angular-defined betweenness centrality against vehicular movement, in urban and rural roads (minor roads in red and principal roads in blue).

What we see when looking at Figure 13 is that there is a striking qualitative difference between urban and 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 slope of the fitted curve is not the same for minor and principal roads, being clearly lower on the latter case. 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.

The saturation pattern for urban roads shown on Figure 13 implies that, in cities, 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 on (thus, least correlated with) centrality variance – that is, the *foreground network*. The pattern itself (formed by observations made in all Britain's cities) may indeed be seen as a signature, both structural and functional, of the dual network model of urban form proposed by space syntax. Moreover, this effect is entirely absent in rural areas. Thus, Figure 13 also provides new and suggestive evidence of the existence of intrinsic differences (both structural and functional), 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 7). 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 dataset³ (DfT 2017), containing the width of the space available for vehicular circulation of principal roads (both urban and rural), at each count location. 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 14). These three factors are: metric betweenness centrality at radius 75 Km (noted as BCm75k, in Figure 13), angular betweenness centrality at radius 75 Km (noted as BCa75k) and local road capacity (noted as Width).

Figure 14 reports the results of eight OLS regression models, describing the impact of each movement predictor, in urban and rural principal roads. Each variable is inserted sequentially 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).

Metric centrality is always a worse movement predictor than road capacity in principal roads, both urban and rural (models 1.4 and 2.4, respectively). The same is not true for angular centrality. Although in principal rural roads angular is capable of explaining more variance than width (model 2.1), the situation is inverted in principal urban roads, with width explaining a larger portion of movement variance (model 1.2). This seems to support the hypothesis that, in the foreground network of cities (i.e. on urban principal roads), the relationship between spatial centrality and movement is constrained by extant road capacities.

Principal urban roads						Principal rural roads					
Model	Step	Variables	β	R2	ΔR^2	Model	Step	Variables	β	R2	ΔR^2
1.1	1	Width	0.58	0.33		2.1	1	Width	1.12	0.48	
	2	Width	0.46	0.44	0.11		2	Width	0.73	0.66	0.18
		BCa75k	0.34				BCa75k	0.93			
1.2	1	BCa75k	0.5	0.25		2.2	1	BCa75k	1.36	0.51	
	2	BCa75k	0.34	0.44	0.19		2	BCa75k	0.93	0.66	0.15
		Width	0.46				Width	0.73			
1.3	1	Width	0.58	0.33		2.3	1	Width	1.12	0.48	
	2	Width	0.52	0.39	0.06		2	Width	0.96	0.6	0.12
		BCm75k	0.25				BCm75k	0.62			
1.4	1	BCm75k	0.37	0.14		2.4	1	BCm75k	0.91	0.27	
	2	BCm75k	0.25	0.39	0.25		2	BCm75k	0.62	0.6	0.33
		Width	0.52				Width	0.96			

Fig.14 – 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.

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 with rural settings). However, urban space is by definition scarce and urban streets, when completely delimited by buildings, create very strong limits to further

³ This dataset was obtained by personal communication (Richard German, October 28, 2016), through the email address ROADTRAFF.STATS@dft.gsi.gov.uk.

increases in road capacity. We thus provisionally propose that the saturation pattern shown on Figure 13, 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 (0.95% of the counted vehicles), we have found that it yielded a very different correlation pattern with the two types of centrality. In contrast with the remaining observed traffic, cycles produced very small differences between the correlations with metric and angular 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 centrality indicators, but always in conjunction with other variables in multiple regression models. Studying cycles flows in two central London local areas, (Raford et al. 2007) report significant correlations of $R^2=0.67$ and $R^2=0.76$, with angular closeness centrality combined with segment length and a dummy variable representing the presence of cycling lanes. Also in a London local area, (Law et al. 2014) report a coefficient of $R^2=0.66$ for angular betweenness also combined with the presence of cycling lanes. However, these two studies did not contemplate the option of introducing metric-defined centrality measures in their models. Cooper (2017) uses a complex version of network distance, including metric 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. In this paper, the maximum effect sizes observed for the 'cycles' class of vehicles were $\rho(A) = 0.72$ and $\rho(M) = 0.73$, 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. But the main difference is that the above mentioned studies do not compare the performances of angular and metric 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 metric centralities (which were large, at any rate), but rather the fact that the differences between such effects were negligible.

Discrete choice modelling of cyclists' route preferences (Broach et al. 2012) shows that cyclist route choice is idiosyncratic and influenced by many factors. Metric 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 sensitive to turn frequency, preferring simple routes. Our results seem to be in line with these findings, with metric 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 factors mentioned above.

Moreover, the studies reviewed above covered only local urban areas (except Cooper 2017), where dense sampling (i.e. count locations on almost every street) is practicable. These enhanced sampling densities can produce results different from ours, for cyclists are not restricted to the spaces of motorized circulation and may thus follow less predictable routes. As it is the case of pedestrian movement, the study of cyclists' movement may depend on samples with high spatial resolution, which is not the case with ours. In this sense, the inconclusiveness of the 'cycles' results clearly points to the need to investigate the theme of cyclist movement more intensely, in order to understand the roles of angular and metric factors in it.

5. CONCLUSIONS

We started by identifying an epistemological divide within the field of street and road network analysis, concerning the way network distance should be conceptualized – through metric or angular means. We associated each type of network distance to two different physical properties of network paths, namely their energetic cost (for metric distance) and their information content (for angular distance), and we identified the theoretical assumptions they imply regarding street and road networks. We then formulated our research questions, asking: i) are there correlational differences between the network hierarchies induced by the two distance concepts and the actual functioning of street and road networks? ii) assuming such differences exist, how do they change across geographical and road hierarchical contexts, and among different types of vehicles?

In order to answer these questions, we carried out a comprehensive correlational study, using the full UK's road network and a very large movement dataset, with more than twenty thousand observations distributed throughout UK's mainland, covering both urban and rural areas.

Answering our first research question, we can now state with great confidence that the network hierarchy induced by angular distance emulates clearly better the actual usage of the UK's road network. In other words, that when it comes to describe the network hierarchy with real functional meaning, the geometric descriptive complexity of network paths (i.e. their information content) is more relevant than their metric length. This was observed at the scale of an entire country, leaving no reasonable doubt of the validity and generability of the result. The distribution of vehicular movement within road and street networks is indeed deeply related with the geometric properties of the network itself.

Besides providing strong empirical support to space syntax's theoretical and methodological proposals described in the introduction, this result has also implications for the modelling and analysis of street networks at large. We found that angular distance produced an average relative increase of 21% in the correlation coefficients, which is a significant accuracy improvement. Current or future analytical models who ignore this, are prone to include an unjustified amount of error.

The answer to the second research question is that there are indeed variations in the correlation coefficients observed in different geographies, types of roads and types of vehicles. Firstly, the results mentioned above are indisputable for "all motor vehicles" (representing 99.05% of the observed traffic), but not for the "bicycles" vehicle class. Even though bicycles represent only 0.95% of the observed traffic, they yielded either non-significant or very small differences (both positive and negative) between correlation coefficients. This indicates that bicycles behave

differently from all other vehicles, which is something that demands further scrutiny. Secondly, the prevalence of angular over metric distance is stronger in principal roads than in minor ones (both in urban and rural contexts). This in turn provides strong support to space syntax's prediction that road networks are hierarchized according to the principle that higher order paths will be geometrically simpler than lower order ones.

However, we also found that this last effect was less clear in urban than in rural roads. When trying to elucidate this fact, we found that the bivariate relationship between movement and centrality is intrinsically different in rural and urban contexts, being simply linear in the former case, but monotonic and with two different regimes of association in the latter case. These two regimes occur in the background and foreground networks of cities, with the transition from one network tier to the other corresponding to a sudden decrease in the rate of response of movement to increased centrality. This breakdown was interpreted as a saturation effect of the urban foreground network, most probably caused by spatial constraints to potential increases in road capacity. At any rate, the conspicuity of such pattern in cities and its total absence in rural areas, provides a new functional and structural distinction between urban and rural road networks, as well as further evidence supporting space syntax's dual network model of urban form.

Finally, we would like to mention what we believe is the main takeaway of this paper. Namely, that when it comes to understand the nature and the functioning of street and road networks, *geometry matters*. The reason it matters, seems to be its capability of modulating the hierarchy of the network according to the *information content* of network paths (by making the descriptive complexity of higher-order paths smaller). And because of this regularity, information about that global hierarchy becomes available also *locally*, through the generic geometric properties of each location on the grid. That these properties must have their value for the navigation of street networks, seems highly probable. But here we have shown that, prior to anything else, they are inscribed in the geometry of network itself.

ACKNOWLEDGMENTS

Miguel Serra is funded by the FCT post-doctoral grant with reference SFRH/BPD/111260/2015.

BIBLIOGRAPHY

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

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

Crucitti, P., V. Latora and S. Porta (2006). "Centrality measures in spatial networks of urban streets." Physical Review E 73(3): 036125.

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

DfT (2018). "Traffic Statistics Methodology Review", Department for Transport, UK Government.

Grunwald, P. and P. Vitány (2003). "Kolmogorov Complexity and Information Theory. With an interpretation in terms of questions and answers." Journal of Logic, Language and Information 12(4): 497-529.

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. (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. 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., A. Penn, J. Hanson, T. Grajewski and J. Xu (1993). "Natural Movement: or, configuration and attraction in urban pedestrian movement." Environment and Planning B: Planning and Design **20**(1): 29-66.

Hillier, B., A. Turner, T. Yang and H. Tae-Park (2010). "Metric and Topo-Geometric Properties of Urban Street Networks: some convergences, divergences and new results." Journal of Space Syntax **1**(2): 258-279.

Jayasinghe, A., K. Sano and H. Nishiuchi (2015). "Explaining traffic flow patterns using network centrality measures." International Journal for Traffic and Transport Engineering **5**(2): 134 - 149.

Jiang, B. (2009). "Street Hierarchies: a minority of streets account for a majority of traffic flow." International Journal of Geographical Information Science **23**(8): 1033 -1048.

Law, S., F. Sakr and M. Martinez (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., D. Smith, A. Crooks and M. Batty (2009). "Random planar graphs and the London street network." European Physical Journal B **71**(2): 259-271.

Omer, I. and B. Jiang (2015). "Can cognitive inferences be made from aggregate traffic flow data?" Computer, Environment and Urban Systems **54**: 219-229.

Omer, I. and N. Kaplan (2018). "Structural properties of the angular and metric street network's centralities and their implications for movement flows." Environment and Planning B: Urban Analytics and City Science **0**(0): <https://doi.org/10.1177%1172F2399808318760571>.

OS (2015). Meridian 2: User Guide and Technical Specification. Ordnance Survey. <https://www.ordnancesurvey.co.uk/docs/user-guides/meridian-2-user-guide.pdf>.

Porta, S., V. Latora, F. Wang, S. Rueda, E. Strano, S. Scellato, A. Cardillo, E. Belli, F. Cárdenas, B. CORMENZANA and L. Latora (2012). "Street Centrality and the Location of Economic Activities in Barcelona." Urban Studies **49**(7): 1471-1488.

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

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

Strano, E., M. Viana, L. Costa, A. Cardillo, S. Porta and V. Latora (2013). "Urban street networks, a comparative analysis of ten European cities." Environment and Planning B: Planning and Design **40**: 1071-1086.

Turner, A. (2007). "From Axial to Road-Centre Lines: a new representation for space syntax and a new model of route choice for transport network analysis." Environment and Planning B: Planning and Design **34**: 539 - 555.

Turner, A. (2009). The Role of Angularity in Route Choice: an analysis of motorcycle courier GPS traces. Spatial Information Theory. COSIT 2009. Lecture Notes in Computer Science. C. Claramunt, M. Denis and G. Ligozat. Berlin, Heidelberg, Springer. **5756**: 489 - 504.

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