
DISCREPANCIES IN CLOSENESS CENTRALITY FORMULATIONS: IMPLICATIONS FOR REPRODUCIBILITY IN URBAN NETWORK ANALYSIS.

January 7, 2026

The authorship and affiliations will be inserted after review.

ABSTRACT

Street network analysis is widely used to relate urban form to land-use intensity and travel behaviour, yet “closeness” centrality is inconsistently described, with cited formulas diverging from those implemented in software packages. This matters for contemporary analysis, where centralities are increasingly computed in localised (distance-threshold) forms to reduce edge effects and enable multi-scalar comparisons. In localised analysis, reachable node counts vary by origin. While *Closeness* and *Normalised Closeness* are effective for comparing nodes within a fixed graph, they behave counter-intuitively across differently sized sub-graphs because varying numbers of nodes cause them to scale in unanticipated ways. We clarify this mechanism, demonstrate reliable alternatives (*Improved* and *Harmonic Closeness*), and provide an openly reproducible workflow.

We test implications in Madrid, Spain, using street-network, land-use, and travel-survey data. We compute multiple formulations, including *Closeness*, *Normalised Closeness*, *Improved*, *Harmonic*, and *Gravity*, for metric and angular distances across multiple distance thresholds, and relate them to land-use accessibility and trip volumes using Spearman correlations with spatially robust (block-bootstrap) uncertainty estimates. *Closeness* shows negative associations with land-use intensity, while *Normalised Closeness* shows weak associations. In contrast, *Improved*, *Harmonic*, and *Gravity* remain stable and consistently positively associated across scales, consistent with their adoption in commonly used software packages. We conclude with practical guidance: prioritise formulations that scale appropriately with varying sub-graph sizes, and clearly document the precise formulations and workflows to support reproducibility and comparability of methods.

Keywords: closeness centrality; street networks; urban analytics; reproducibility; space syntax; accessibility

1 Introduction

Streets form connections between diverse activities and places. While their functional role is apparent, their emergent properties are less so. It can be observed that particular sidewalks are livelier than others, that certain landuses may have a proclivity for particular locations, or that walkable and vibrant streets can become public destinations in their own right [1]. Yet, without recourse to the incremental and evolutionary development typical of historical towns and cities, it is not straightforward to anticipate these dynamics for planned street networks [2]. Even if the benefits of compact and walkable forms of urban development are widely reported, new developments continue to prioritise vehicular movement to more distant places at the expense of pedestrians and connectivity to local street networks [3]. This is where network analytic methods are particularly relevant: by enabling the comparison of emergent properties of street configurations, they make it feasible to explore the emergent properties of street systems and how these relate to potential activity and use.

Graph network methods have been widely utilised across numerous disciplines, variously adopting network analytical methods for purposes ranging from early centrality methods in the study of telecommunication networks [4] to more

recent advances in social network analysis [5]. Street network analysis likewise emerged as a pivotal tool in urbanism, urban design, and urban planning. The utilisation of graph theories for spatial problems can be traced to 1736 when Euler proposed a graph-based solution to the Königsberg Seven Bridges problem [6]. In the 1950s and 1960s, Froshaug compared hierarchies of street networks by mapping them as simplified graphs with street intersections as nodes, notably the ULM4 project in 1959 [7]. Space syntax theory, developed by Bill Hillier and colleagues in the late 1970s and early 1980s [8], integrated a social theory of space with network measures to represent and analyse cities, prompting wider consideration on the usage of street network analysis for insights into urban design and planning.

Despite common foundations, the interpretation and application of centrality measures can vary across fields and traditions, leading to inconsistencies in terminology and implementation. Subtle differences in formulation or methodology can lead to significant changes in the mathematical behaviour of a measure, resulting in conflicting interpretations of the relevance of centralities to characteristics such as walkability or landuse accessibility. This may lead to spurious interpretations and widespread misattribution of results due to unacknowledged variations in the underlying metrics and methods. Previous authors have noted related challenges stemming from differences in model construction and representation [9, 10], variations in algorithmic implementations and metric calculations [11], and the lack of buffering of network boundaries to control for edge effects [12]. These inconsistencies present a significant barrier to the reproducibility, interpretation, and validity of results.

In the following analysis, we draw attention to the interpretation of closeness centralities in urban analytics and ask whether differences between formulations would lead to observable differences in behaviour for the now predominant localised (distance-threshold-based) form of network analysis. We provide open code and workflows based on open datasets for reproducibility.

What this paper adds: We identify a marked divergence between the closeness formulation widely cited in street network analysis literature (*Normalised Closeness*) and the variants actually implemented in computational packages such as Depthmap, Place Syntax Tool, and cityseer (*Improved Closeness*). We show theoretically and empirically that this distinction is consequential for localised network analysis because *Closeness* and its normalised variant can produce counter-intuitive behaviour when sub-graphs vary in node count. For localised street network analysis, practitioners should therefore prefer closeness formulations that scale appropriately with varying sub-graph sizes, such as *Improved Closeness*, *Harmonic Closeness*, or gravity-based accessibility measures.

Research questions and hypotheses: We pose two research questions. **RQ1:** Is there a meaningful distinction between *Normalised Closeness* and *Improved Closeness* (alongside other local variants such as *Harmonic Closeness*) in the context of localised street network analysis? **RQ2:** If so, does this distinction manifest in empirical associations with land-use accessibility and origin-destination trip counts? We hypothesise that *Closeness* will behave counter-intuitively under localised analysis because it conflates variations in sub-graph size with variations in average distance, and that standard *Normalised Closeness* does not correct for this. By contrast, *Improved Closeness* and *Harmonic Closeness* scale appropriately. Section 3 develops this hypothesis formally; Sections 4–6 test it empirically.

2 Street Network Analysis

The intuition of networks (graphs) is that nodes (vertices) are connected by links (edges) to other nodes. For example, a social network may consist of individuals connected through relationships with one another. Network analysis can then be used to answer how connected or ‘relatively central’ a particular person is, thus inferring their structural importance within the network. Two of the more common measures of importance are *closeness centrality* [13], how closely a node is located to other nodes, and *betweenness centrality* [14], how often a node provides the shortest path between other nodes. The computation of these measures involves the use of algorithms calculating *shortest-paths* through the network: in the basic case, distance is topological: the number of nodes between origin and destination pairs. It is also possible to represent distance by assigning a traversal cost to the links between the nodes. For example, in street networks, these traversal costs can correspond to the lengths of streets (links) between road intersections (nodes).

To understand the contemporary context of street network analysis in practice — and potential ramifications for closeness centrality — it is beneficial to be aware of the literature and distinctions related to *global* and *localised* forms of analysis; differences in street network representation; the use of *topological*, *metric*, and *geometric* distances as a heuristic for determining shortest paths through the network; and the impact of network topology on the robustness and comparability of results. These distinctions are sources of significant variation in closeness centralities and introduce barriers for interpretability and generalisability if not clearly explicated. For reproducibility, this section provides a detailed overview of these distinctions and related background literature prior to the next section, which introduces formulas for the calculation of closeness centralities.

2.1 Scales of Analysis

Network centrality methods can be applied to the entirety of a network, e.g., for a given city, which is typically referred to as ‘global’ analysis. In this case, the magnitude of the resultant network centrality is coupled to the size of the network with the implication that the larger the network, the larger the centrality values. Different boundary definitions therefore lead to fluctuating centrality values, raising some issues:

- It is difficult to rigorously and consistently define boundaries from city to city. This problem is exacerbated for large urban agglomerations or when the boundaries between urban areas are not clearly delineated. One solution is to apply a technique such as network percolation [15] to heuristically delineate boundaries;
- Global analysis is subject to significant “edge-effects”, where centrality values diminish towards the edge of the network because the algorithms are constrained by the boundary. This may be manageable when using rigorous boundary definitions, but otherwise creates significant issues for comparability and generalisability [16, 17];
- More localised phenomena within networks cannot be directly analysed or compared because local-scale properties are masked by the global-scale characteristics of the network [18, 17]. This makes it difficult to meaningfully compare localised properties of networks between locations, particularly from the point of view of urban design and planning interventions where it is necessary to gauge local-scale impacts of design decisions.

These issues are mitigated through the use of “localised” network analysis which works by iteratively visiting each node using a windowing methodology to define a local catchment area for analysis [16]. This intuition is conveyed in Figure 1: each node is visited in turn, all other reachable nodes within a selected distance threshold are then identified and isolated from the network at large before the analysis then proceeds (using only the locally extracted sub-graph). The algorithm then steps to the next node and repeats the process until the entirety of the network has been visited.

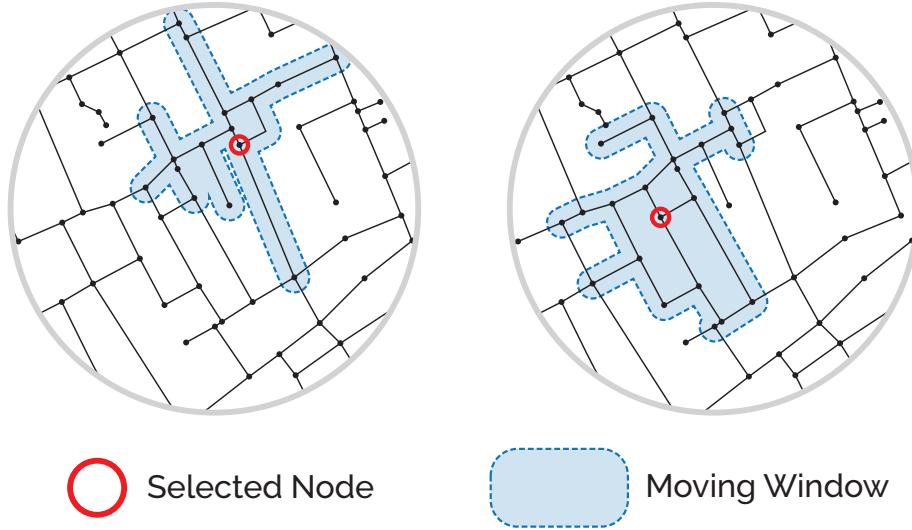


Figure 1: Moving window catchment area.

Localised methods confer several advantages:

- A measure computed for a given distance threshold on one network can be compared directly to the same measure and distance threshold computed for another location — even in a different city — because the localised boundary extents are defined in a methodologically consistent manner from location to location;
- Localised analysis substantially mitigates edge-effects as long as the network is adequately buffered relative to the considered distance threshold (for example, if using a 1km local catchment then a 1km buffer will eliminate edge effects);
- Centralities can be computed for a range of nearer or farther distance thresholds (commonly referred to as radii) to draw-out smaller or larger structures within the network. A number of different distance thresholds is ordinarily computed to understand the properties of the street network at different scales of analysis. Use

of smaller distances can yield information about local-scale walkability whereas the use of larger distances can offer larger-scale insights equivalent to global methods without the aforementioned drawbacks such as edge effects.

The use of localised network analysis is therefore conventional in current practice; the remainder of this discussion and the ensuing analysis is based on the localised form of analysis. In common parlance, localised measures computed for larger radii (e.g. 10km or 20km) are sometimes termed as ‘global’ analysis in the literature. To avoid terminological inconsistency, we refer to ‘localised’ analysis regardless of the local distance threshold used.

2.2 Model Representations

The *primal* representation is when intersections are represented by nodes and streets by links, thus corresponding to conceptions of streets embedded in Euclidean space: intersections adopt specified coordinates connected by streets to other intersections [19]. *Space Syntax*, on the other hand, emerged around the use of the *dual* representation [8]. This originated with the use of *axial lines* to generate a topological representation of urban space. An *axial line* is an uninterrupted longest line of sight connecting convex spaces (an area where any two points can be connected by a straight line), potentially spanning multiple contiguous street segments, and distances are topological, meaning the number of steps from axial line to axial line. This is a form of *dual* representation where the nodes represent street corridors and the links represent the topological steps between them. The techniques for the extrapolation of *axial lines* from the street network can be complex and variable, historically leading to debates on whether it is possible to do so in an algorithmically rigorous manner [20, 21, 19]. It is important to recognise that newer and more tractable forms of dual representation have since been widely adopted by the space syntax community: *fractional analysis* and the now prevalent *angular segment analysis* [22, 23, 21, 16] forego axial lines, instead deriving the network from road centre-lines through a direct inversion of the *primal* representation into its *dual*, as shown in Figure 2. Nodes therefore correspond to the mid-points of street segments and links typically correspond to the *geometric distance* (angular change) linking them, though can also represent *metric distance* (such as metres), as is typically the case for *primal* representations [24, 10, 25]. Forms of distance are discussed in the following sub-section.

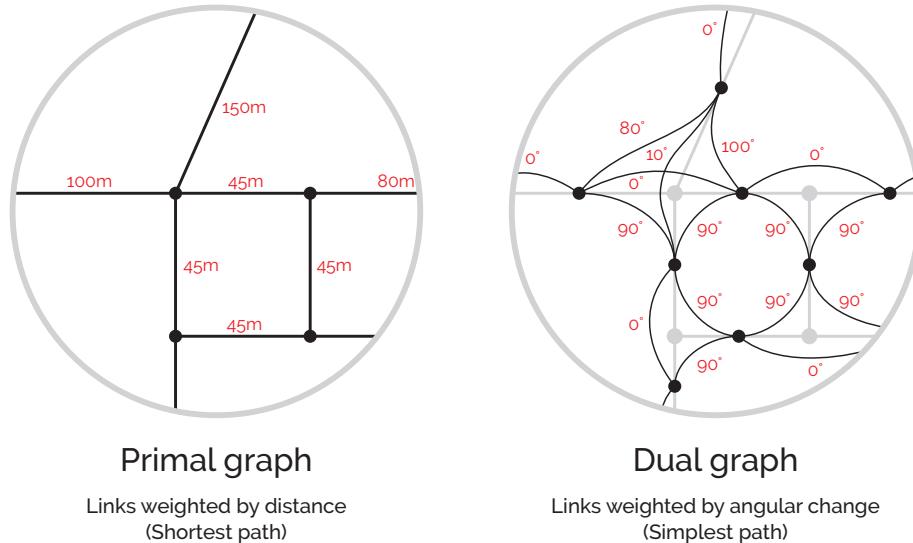


Figure 2: The *primal* street network representation with *metric distances* (left) and a corresponding example of a dual representation (right) with *geometric distances*. Note that the dual network can be used with *metric distances*.

There is a large variety of potential network representations [10], some of which can be used to analyse properties such as the connectivity of the network and its *small-world* and *scale-free* characteristics, emblematic of complex systems and network analysis more generally [26, 19]. However, contemporary forms of street network analysis tend to adopt a pragmatic approach which leverages the widespread availability of street network datasets, typically used either directly in the *primal* representation or its direct *dual* inversion.

2.3 Cost parameters

Network links (edges) can be assigned a cost parameter to reflect distances incurred by traversing a particular link. Historically, *topological distances* were used in the context of axial representations. Contemporary approaches more typically make use of *metric distances* or *geometric distances*. The former is ordinarily used with physical distance units such as metres whereas the latter approach is instead based on angular deviation, or in other terms the linearity of routes. Simpler routes with better lines-of-sight are therefore considered ‘shorter’ than more convoluted routes even if the Euclidean distances are greater; thus, routes requiring minimal geometric complexity are favoured to routes requiring minimal physical effort [27, 28]. This tends to be the preferred distance metric in space syntax, where geometric properties of the street network are seen as important determinants in the general evolution of landuses and the wider patterns of activities in cities [29].

The use of *geometric distance* introduces an implementation nuance in that shortest-path algorithms will ‘bypass’ sharp angular turns in cases where smaller combinations of adjacent angles can be combined instead (Figure 3), thus making it necessary to enforce directional constraints for algorithms as they pass-through nodes during graph traversals [16]. Note that unlike tailored network analysis tools such as *depthmapX* [30], *Place Syntax Tool* [31], or *cityseer* [32], generic network analysis packages typically do not take this into account and would also not necessarily return symmetrical routes if the direction of travel were reversed [33].

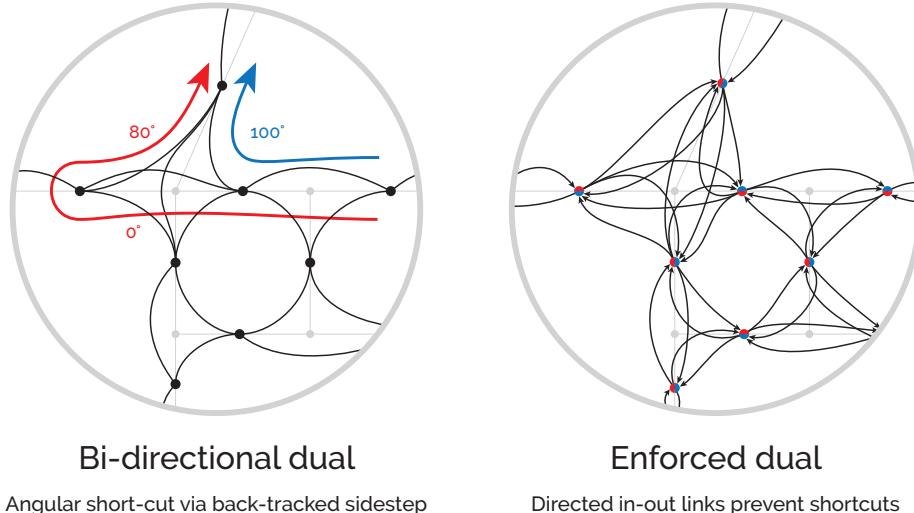


Figure 3: Shortest angular path side-stepping.

Metric distance and *geometric distance* as a cost parameter are not mutually exclusive. While expressing distances differently, both are valid approaches and their interrelation is complex [34]. Pedestrian route choices may vary from place to place and from person to person. Shortest routes are frequently the same for either cost parameter [35, 36] and strong correlations are observed between them for certain distances [27, 28]. Modal choice presents an additional nuance [28]; research has shown strong support for *geometric distances* within the context of vehicular travel, particularly at distances larger than 2000m. However, there are indications that *metric distances* retain relevance for smaller-scale analysis and for non-vehicular forms of transport such as cycling [28, 37].

Note that *metric distances* can be computed on the *dual* and that *geometric distances* can technically also be computed on the *primal*, and that it is also possible to simultaneously apply multiple representations [38].

2.4 Distortions related to topology and geometry

Network analysis algorithms are sensitive to the topological structure of the network, and low-quality network datasets will therefore lead to spurious results. For example, broken links will misroute shortest-path algorithms and unnecessarily complex representations of features such as intricate road intersections will introduce inflated centralities to the surrounding network.

A common problem is the conflation of the topological structure of the network with the geometrical trajectories of streets, which introduces different intensities of nodes for equivalent lengths of streets segments. For example, a straight street segment may be represented as a single link between two nodes whereas an arced street of equivalent

length is often represented with additional nodes to approximate the geometric curvature of the roadway. Each additional node results in more summations when calculating centralities, consequently skewing the outcome of measures. It is therefore important to use cleaned network representations which maintain the distinction between the geometric representation of roadways and the topological structure of the network. If using more topologically complex networks from sources such as OpenStreetMap, then it is generally necessary to first clean and simplify the network. This should be done while preserving the geometry of the original street segments so that accurate distances or angular changes in direction can be measured [39]. In this research, these techniques are facilitated by the `cityseer-api` package [32]. There is wider interest within the network analysis community in the formalisation and standardisation of network simplification methods [40].

Another manifestation of this phenomenon is the topological divergence between *primal* and *dual* representations (see Figure 4). Even though both derive from the same underlying structure, the outputs for the same network measures using the same cost parameters will yield different results; the *dual* representation generates higher equivalent metrics due to larger quantities of nodes and links, which becomes more pronounced as the complexity of the network increases. Comparative evaluation between *metric distance* measures on *primal* networks and *geometric distance* measures on *dual* networks should be avoided because the outcomes will also reflect representational differences.

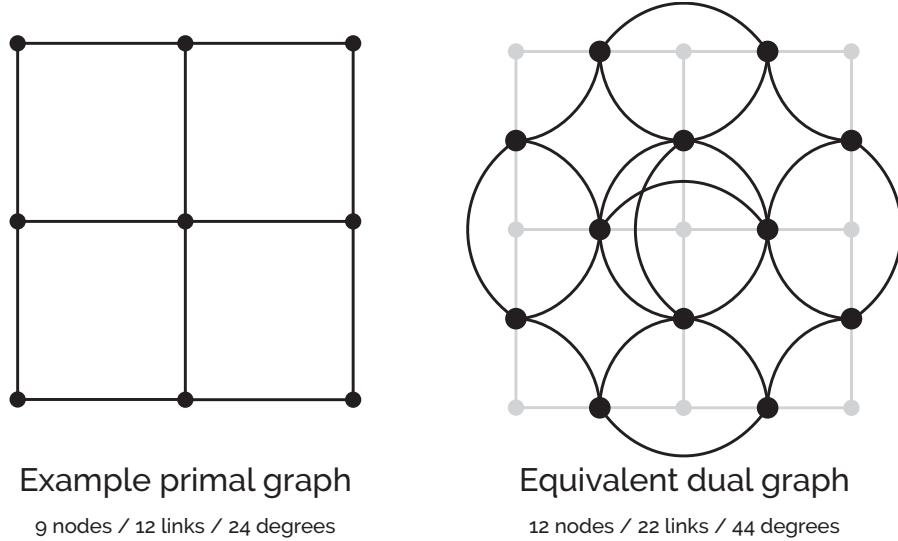


Figure 4: Topological divergence of primal and dual representations.

3 Closeness Centrality in Street Network Analysis

3.1 Varying descriptions for the concept of closeness

Application of the concept of closeness to street network analysis emerged with the development of space syntax in the late 1970s and early 1980s [8]. In this context, closeness is referred to as *Integration* and implies “to-movement” or “closeness” in the general sense of the words¹. In the case of the axial topological network representations used at the time, *Integration* is calculated through counting the number of topological steps to surrounding nodes, taking the average, “relativising” the result to the size of the network, then taking the reciprocal, thereby converting the measure from farness-like to closeness-like.

After the introduction and increasing shift to angular segment analysis in the 2000s, which used street segments based on road centrelines instead of axial lines, the method for calculating *Integration* was adapted to work with weighted graph edges by using summed angular deviation instead of counting topological steps. Whereas Turner at first suggested the use of *Normalised Closeness* (Equation 3) [16], he would later suggest taking the node count divided by the average distance to the nodes [41], which reflects the same intention as with the aforementioned axial integration: take *Normalised Farness* (average distance), re-normalise by the number of nodes in the network, then take the reciprocal. This is reinforced in a subsequent paper [42], where Hillier and Turner et al. state that the Depthmap

¹We here and throughout the paper refer to uncapitalised “closeness” in the general sense, as opposed to the capitalised form which we use to express the formal mathematical definition for *Closeness*, a specific form of closeness centrality.

software package calculates angular segment integration by “*dividing mean angular depth by node count... and taking the reciprocal*” (where *Mean Depth* refers to average distance). Closer inspection shows that this variant of closeness resembles a simplified form of the so-called *Improved Closeness* (see Equations 6, 7) proposed by Wasserman and Faust [5], which differs from *Normalised Closeness* in that it divides the node count by the **average** as opposed to **total** distance to the nodes. Notably, at least three street network analysis packages continue to use or provide this form of *Improved Closeness*, including Depthmap [42, 30], the Place Syntax Tool [31], and the cityseer Python package [32].

Concurrent with the adoption of road centreline representations in the mid-2000s, and the wider interest in street network analysis more generally, researchers both within and outside of space syntax have since consistently cited *Normalised Closeness* (Equation 3) when referring to the measure of closeness, even when these citations depend on packages such as Depthmap [36]. Citations for non-normalised mathematical *Closeness* (Equation 1) and even *Normalised Farness* (Equation 4) are also found. We present a selection of example references in Table 1.

<i>Closeness & Normalised Closeness</i>	Representative of literature citations on street network analysis in general, which typically mention or show the formula for <i>Normalised Closeness</i> [43, 18, 38, 44, 16, 45, 46] though also often discussed in the non-normalised form [47, 48, 49, 50].
<i>Improved Closeness</i>	Limited examples, which tend to come from technical literature or street network analysis computational packages [30, 31, 32]. This form divides the node count by the average as opposed to total distance to the nodes. “ <i>Hillier has suggested $\text{Integ} = NC/MD$ where NC is node count (i.e., the number of nodes within a network radius), and MD is mean depth of the nodes with respect to the root node.</i> ” [41] “ <i>Angular segment analysis: The integration solution. Hillier’s integration measure gives a solution that works both at low radius and radius n: $\text{Integ} = (NC * NC)/TD$</i> ” [41] (Algebraically rearranged form of $\text{Integ} = NC/MD$). “ <i>...the current angular integration measure in Depthmap. This is found by re-dividing mean angular depth by node count... and taking the reciprocal to have high values for high integration</i> ” [42]
<i>Normalised Farness</i>	Limited examples, which possibly unintentionally omit mention of taking the reciprocal (which would then give <i>Normalised Closeness</i>). “ <i>Segment integration / segment angular closeness of a segment is the mean of all the angles of all the shortest paths...</i> ” (C_θ) [51, 17] “ <i>Closeness...calculates the average distance from each node to all other nodes...</i> ” [12]

Table 1: Examples of different formulations attributed to the concept of closeness.

3.2 Research Question

The context of our research question stems from the varying cited definitions of closeness centralities for street network analysis. At first glance, this may simply be attributable to historical evolution or different traditions of analysis; however, closer scrutiny is warranted given the divergence between cited formulations and those used by several computational packages. This has potentially significant ramifications for the comparability of findings and the reproducibility of results.

We accordingly ask whether there is a distinction in the behaviour of *Improved Closeness* and *Normalised Closeness* in the context of contemporary street network analysis. By “contemporary”, we imply the use of either *metric distances* (in metres) or *geometric distances* (in angular deviation) on networks constructed from topologically cleaned road centreline representations, and where the measures are computed using localised methods to control for edge effects, thereby facilitating comparisons across locations and scales of analysis (see Section 2).

We proceed with a theoretical hypothesis for why a distinction between *Improved Closeness* and *Normalised Closeness* is potentially meaningful for closeness centralities in street network analysis. We then conduct an empirical evaluation where we compare the behaviour of different forms of closeness centrality in the context of landuse and trip data to assess whether the measures behave as anticipated by the hypothesis.

3.3 Hypothesis

As discussed in Section 2, complications with boundary edge effects means that street network analysis has shifted to using distance localised sub-graphs for comparability across locations and scales of analysis. Turner, an early developer of road centreline street network analysis methods, expresses in a presentation given in 2008 that they (including Bill Hillier) were looking for a closeness-like formulation that is compatible with localised methods of analysis, for which a form of *Improved Closeness* is proposed [41] (see quotations in Table 1) and subsequently adopted into depthmap [42]. Other than for developers of computational packages, we surmise that the distinction on workable formulations for localised closeness analysis appears not to have been broadly recognised by the wider street network analysis research community, where literature now almost universally refers to *Normalised Closeness* even in the context of the predominantly used localised methods. We consequently hypothesise that *Normalised Closeness* may produce counter-intuitive results for localised forms of network analysis and proceed with a theoretical description for why this may be the case.

When formally defined, the mathematical *Closeness* measure

$$\text{Closeness}_{(i)} = \frac{1}{\sum_{j \neq i} d_{(i,j)}} \quad (1)$$

is the reciprocal of *farness*

$$\text{Farness}_{(i)} = \sum_{j \neq i} d_{(i,j)}, \quad (2)$$

where *Farness* is the sum of distances d from node i to all reachable nodes j [13]. Mathematical *Closeness* conveys how proximate node i is to surrounding nodes, with the implication that high closeness centralities afford increased likelihood of access and interaction. *Closeness* can be normalised by the number of nodes in the graph N divided by the sum of the distances

$$\text{Normalised Closeness}_{(i)} = \frac{N-1}{\sum_{j \neq i} d_{(i,j)}}, \quad (3)$$

which is the inverse of the arithmetic mean (average) of *Farness*:

$$\text{Normalised Farness}_{(i)} = \frac{\sum_{j \neq i} d_{(i,j)}}{N-1}. \quad (4)$$

When normalising, it is common to use $N-1$ to imply that the origin node i is not technically counted (though this only has a notable impact on small graphs).

In localised analysis, reachable node counts vary by origin. While *Closeness* and *Normalised Closeness* are effective for comparing nodes within a fixed graph, they behave counter-intuitively across differently sized sub-graphs because they aggregate distances in ways that conflate the quantity and proximity of reachable nodes. In the broader network analysis field, it has been shown that *Harmonic Closeness* centrality [52, 53] scales more suitably for sub-graphs, with the difference being that the division happens prior to the summation

$$\text{Harmonic Closeness}_{(i)} = \sum_{j \neq i} \frac{1}{d_{(i,j)}}. \quad (5)$$

For the same reason, others have suggested *Improved Closeness* centrality [5]

$$\text{Improved Closeness}_{(i)} = \frac{N_i/g}{\sum_{j \neq i} d_{(i,j)}/N_i}, \quad (6)$$

which is intended for situations where a limited subset of nodes is reachable. It is defined as the ratio of the fraction of reachable nodes N_i/g to the average distance to those nodes $\sum_{j \neq i} d_{(i,j)}/N_i$. In the context of street network analysis, the global number of nodes g is ordinarily unknown; nevertheless, since this is effectively the worldwide street network and is constant for all sub-graphs, it can be proposed that the number of reachable nodes N_i in the numerator can forego the normalisation by g , thus giving

$$\text{Improved Closeness}_{(i)} = \frac{N_i}{\sum_{j \neq i} d_{(i,j)}/N_i} = \frac{N_i^2}{\sum_{j \neq i} d_{(i,j)}}, \quad (7)$$

which is synonymous with the formula proposed by Turner and Hillier and the variant found in several computational packages. (Note that when algebraically rearranged, the formulation can be expressed as the square of the number of reachable nodes divided by the sum of the distances to those nodes.)

The intuition of *Improved Closeness* can be contrasted to *Normalised Closeness* (Eq: 3): the number of reachable nodes is divided by the **average** distance instead of the **total** distance to the nodes. This distinction is important for localised graphs because *Improved Closeness* is consequently able to scale intuitively even if the number of nodes varies: it increases either for a greater number of locally accessible nodes or if the average distance to those nodes decreases.

Figure 5 illustrates the implications with a simple example: Scenario B is expected to have greater closeness within an urban context than Scenario A because it is equivalently close to a greater number of nodes. Scenario C contains a more distant node and should therefore have a lower closeness centrality than Scenario B but should still be greater than A.

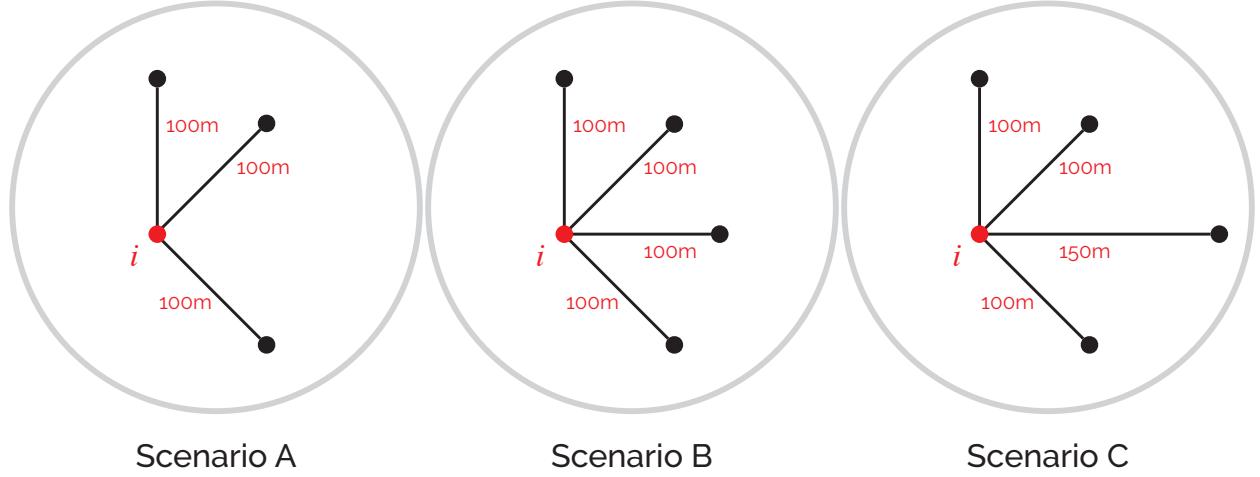


Figure 5: Simple comparative localised closeness scenarios.

As shown in Table 2, in the context of localised analysis, mathematical *Closeness* behaves opposite to its intended usage because it scales differently across different quantities of nodes, decreasing from Scenario A to B and from A to C. This means that whereas mathematical *Closeness* may be useful for comparing the centralities of nodes within the same graph, it behaves counter-intuitively when comparing centralities across graphs containing different numbers of nodes. *Normalised Closeness* is similarly problematic for localised analysis because it now neutralises meaningful variations, such as from Scenario A to B, while also continuing to demonstrate unexpected behaviour, such as counter-intuitively decreasing from Scenario A to C. On the other hand, and as indicated in the broader network analysis literature, *Harmonic Closeness* [53] and the simplified form of *Improved Closeness* [5] behave in a manner consistent with expectations across differently sized sub-graphs.

	Scenario A	Scenario B	Scenario C
<i>Closeness</i>	$\frac{1}{100+100+100} = 0.00\bar{3}$	$\frac{1}{100+100+100+100} = 0.0025$	$\frac{1}{100+100+150+100} = 0.00\bar{2}$
<i>Normalised</i>	$\frac{3}{100+100+100} = 0.01$	$\frac{4}{100+100+100+100} = 0.01$	$\frac{4}{100+100+150+100} = 0.00\bar{8}$
<i>Harmonic</i>	$\frac{1}{100} + \frac{1}{100} + \frac{1}{100} = 0.03$	$\frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100} = 0.04$	$\frac{1}{100} + \frac{1}{100} + \frac{1}{150} + \frac{1}{100} = 0.03\bar{6}$
<i>Improved</i>	$\frac{3}{(100+100+100)/3} = 0.03$	$\frac{4}{(100+100+100+100)/4} = 0.04$	$\frac{4}{(100+100+150+100)/4} = 0.03\bar{5}$

Table 2: Closeness comparisons.

Out of these, *Harmonic Closeness* offers a high degree of precision because it considers the inverse of the distances independently in contrast to *Improved Closeness*, which first averages the distances; yet, for the same reason, the averaging implicit with *Improved Closeness* may be advantageous when working with poorer quality or unsimplified representations of street networks where individual summations may otherwise approach infinity if a network segment approaches zero length.

4 Empirical Methodology

We now investigate whether the empirically observed behaviour of different closeness formulations matches that anticipated from the theoretical hypothesis. Network centralities are typically associated with landuse intensities [18] and travel volumes [27]; Madrid has been selected for the case study due to the availability of a high quality street network dataset accompanied by high resolution landuse premises information and an origin-destination travel dataset. For full reproducibility, access to the datasets, code workflow, and related data preparation notes is provided in an open code repository.

<https://anonymous.4open.science/r/paper-centrality-F121/>

The link to the permanent code repository will be inserted after review.

The analysis proceeds with a range of network centralities using localised analysis from 500m to 10km so that typical behaviour can be observed at both smaller and larger scales of analysis, as is typical in the broader literature. All measures are computed on a road-centreline street network using *metric* and *geometric* distances applied on a dual representation of the street segments. The network is a high-quality “cleaned” representation, with the geometrical curvature of streets separated from the topological structure of the network.

We then provide a visual and statistical comparison on the behaviour of the aforementioned centralities:

- The behaviour of the different forms is visually compared on the plotted maps;
- Spearman rank correlations are compared between network centralities for associations to landuse accessibility measures;
- Spearman rank correlations are compared between network centralities for associations to origin-destination trip data.

The following variants of closeness are included, and are denoted as follows in the plots:

1. *Closeness* represents the standard non-normalised mathematical formulation of closeness centrality.
2. *Closeness N*¹ indicates *Normalised Closeness*, where the node count is divided by *Farness*. This form is commonly cited in the street network analysis literature.
3. *Closeness N*^{1.2} raises the node count in the numerator to the power of 1.2. This is a form of closeness commonly used in the space syntax research community and is referred to as “normalised” least angular integration (NAIN) [42].
4. *Improved Closeness* (denoted *N*² in plots) squares the node count prior to the division by *Farness*, representing the version found in computational packages. (Mathematically equivalent to dividing the node count by *Normalised Farness*).

For additional comparative context, *Harmonic Closeness*, *Gravity Index*, *Betweenness*, and *length-weighted* variants of the centralities are also computed; these are provided in the Supplementary Materials, where we also introduce continuous forms of the length-weighted measures as derived from calculus.

Accessibility [54, 55, 56] to different landuses is calculated for Food & Beverage, Retail, Services, Creative & Entertainment, and Accommodation relative to each street segment for maximum street network distances of 100m, 200m, 500m, 1km, and 2km. A dimensionality reduction is performed on the landuse accessibility data using Principal Component Analysis (PCA), which expresses the variance contained in the datasets while removing collinearity between the variables. The input variables were preprocessed using the scikit-learn [57] package PowerTransformer (to reshape variables to a normal distribution) and StandardScaler (mean centering and variance scaling) prior to PCA, with loadings (correlations to input variables) shown in Supplementary Materials Figure S3. The centralities are then compared to stand-alone accessibilities for Food & Beverage and Retail as well as the first principal component of the PCA, which expresses 63.6% of the variance and is used as a proxy of more general landuse accessibility.

To cross-check the observations for landuse correlations, we examine whether similar patterns of observation persist when instead correlating against origin-destination travel survey data. The travel survey comprises trip-level records filtered to journeys with main purpose codes indicating non-home destinations (purposes 2–6: work, education, shopping, leisure, other). These trips are aggregated to origin-destination travel zones, yielding counts for both trip origins and destinations per zone. The zones are filtered to those overlapping with the street network, yielding approximately 625 zones for analysis. The counts of origin-destination travel volumes for these travel zones are correlated to the average centrality values for all street segments falling within the same zones.

The correlation plots make use of Spearman Rank ρ correlation coefficients (step-wise monotonicity of the data) as opposed to Pearson's r (linearity of the data) because heavily skewed datasets would otherwise require preprocessing steps (e.g. max-log optimised boxcox transformations).

4.1 Descriptive Statistics

Descriptive statistics are computed at both the segment and zone levels. At the segment level, we report $N = 42,167$ street segments, providing summary statistics (mean, median, quartiles, interquartile range, minimum, and maximum) for all centrality measures across the five analysed distance thresholds (500m, 1000m, 2000m, 5000m, 10000m). At the zone level, we report $N \approx 625$ travel zones for which street network and travel survey data overlap, with corresponding summary statistics for: (1) outcome variables—origin trip counts, destination trip counts, and normalised trip densities (trips per km^2)—and (2) aggregated centrality measures computed as simple means of segment-level values falling within each zone. These descriptive statistics are presented in the Supplementary Materials (Tables 5–8 for segments; Tables 9–11 for zones).

- Total street segments: 42,167
- Total survey records: 222,744
- Filtered trips (non-home destinations, purposes 2–6): 132,280
- Total zones in survey: 1,259
- Zones with both origin & destination trips: 1,220
- Final zones in analysis (overlapping with street network): 625
- Mean zone area: 1.25 km^2

4.2 Note on Correlation Magnitudes

Correlations tend to be larger at greater distance thresholds because aggregation smooths variance. This phenomenon is a characteristic of the Modifiable Areal Unit Problem [58] with the implication that correlations should only be directly compared between measures at the same scale of aggregation. A stronger correlation at a larger threshold is not necessarily "better" than a weaker correlation for a smaller distance, which retains a higher degree of local detail (and therefore variance).

4.3 Spatial Autocorrelation

Street network centrality measures are expected to exhibit positive spatial autocorrelation for closeness measures: nearby segments have similar values due to their shared network context. We quantify this using Moran's I [59], computed with k -nearest neighbour weights where k is set to the median network density at each distance threshold.

Spatial dependence reduces the effective sample size for statistical inference. We report an approximation $N_{\text{eff}} \approx N(1 - I)/(1 + I)$ to indicate the degree of redundancy introduced by autocorrelation. To provide robust confidence intervals for the correlation estimates, we employ a spatial block bootstrap: segments are grouped into spatial clusters using k -means on centroid coordinates, and blocks are resampled with replacement to preserve local spatial structure.

Given that this study is primarily comparative (examining how different closeness formulations behave relative to one another) rather than making causal claims, the comparative patterns across formulations remain valid, though individual correlation p -values should be interpreted with appropriate caution. Moran's I results, effective sample sizes, and block bootstrap confidence intervals are provided in the Supplementary Materials (Tables 12–14).

5 Data Sources

Network, landuse, and travel survey data are prepared for Madrid, Spain. Access to the input datasets and further detailed preparatory notes are provided in the data repository.

<https://anonymous.4open.science/r/ua-dataset-madrid-818B>

The link to the permanent data repository will be inserted after review.

5.1 Network Data

The network dataset consists of an open data street network for Madrid, which is a high-quality cleaned network structure based on street centrelines [60].

Street Map of the Community of Madrid: “Set of roads officially approved by the municipalities of the Community of Madrid, ordered by different characteristics.”

Source: Madrid Open Data

License: [Creative Commons Attribution License](#)

5.2 Land Use Data

Land uses are derived from an open dataset for Madrid, consisting of 153,953 premises classed according to a schema comprising approximately 80 landuse categories [61].

Census of premises and their activities: “Microdata file of the census of premises and activities of the Madrid City Council, classified according to their type of access (street door or grouped), situation (open, closed...) and indication of the economic activity exercised and the hospitality and restaurant terraces that appear registered in said census.”

Source: Madrid City Council

License: [License](#)

5.3 Origin-Destination Travel Survey Data

The Regional Transport Consortium of Madrid travel survey data contains approximately 200,000 origin-destination journeys encompassing approximately 1250 regions within Greater Madrid, including travel mode and travel reasons information [62, 63].

Regional Transport Consortium of Madrid Travel Survey for 2018: Powered by CRTM

Source: Regional Transport Consortium of Madrid

License: [License](#)

6 Results

Following the methodology outlined in Section 4, we present the results of the empirical analysis. A visual survey of the 1000m closeness centralities is plotted for *metric* distance measures in Figure 6, with *geometric* distance (angular) measures shown in Figure 7. The correlation matrix for the discussed closeness centralities is shown in Figure 8.

6.1 Visual Comparisons

As illustrated in Figure 6, *Closeness* exhibits notable challenges due to its mathematical behaviour, which scales differently depending on the number of nodes under consideration. The iterative process of isolating catchments for localised calculations can therefore lead *Closeness* to yield values opposite to its intended usage. *Normalised Closeness* and *NAIN* reduce the variance, but do not recover meaningful contrast across different street network intensities. The *Improved*, *Gravity*, and *Harmonic* formulations behave as intended and give highly comparable results, the main differences being how these handle regions of high intensities due to differences in how nearby distances are factored into summations; see, for example, the upper right corners of the plots.

Angular variants, as illustrated in Figure 7, exhibit the same pattern for *Closeness*. *Normalised Closeness* and *NAIN* again flatten the distribution of values, though due to the use of geometric (angular) distances, these now have a tendency to emphasise more orthogonal portions of the network with the implication that the more densely interconnected areas are less emphasised. The *Improved* and *Harmonic* formulations again behave as intended, once more yielding similar outputs to each other. Note that the *Gravity* formulation is not applicable to geometric distances.

The above described similarities and differences between the measures are reflected in the correlation grid (Figure 8), where *Normalised Closeness* differs markedly from the other forms and where *Closeness* shows negative correlations, particularly when compared for smaller distances.

6.2 Key Finding: Sign Reversal Across Formulations

Before presenting detailed correlation grids, we highlight the central empirical finding. Block-bootstrap confidence intervals (accounting for spatial autocorrelation; see Section 6.5 and Table 14) indicate a marked divergence in how different closeness formulations are associated with landuse intensity at localised scales:

- **Closeness** is associated with a strong *negative* correlation ($\rho \approx -0.70$, 95% CI: $[-0.76, -0.64]$ at 500m), a sign reversal relative to the intended interpretation.
- **Normalised Closeness** (N^1) shows correlations near zero ($\rho \approx -0.11$, 95% CI: $[-0.16, -0.07]$ at 500m), suggesting that simple normalisation does not address the underlying issue.
- **Improved Closeness** (N^2), **Gravity**, and **Harmonic Closeness** show consistently strong *positive* correlations ($\rho \approx +0.67$, 95% CI: $[+0.60, +0.72]$ at 500m for *Improved Closeness*).

These confidence intervals account for the high spatial autocorrelation ($I \approx 0.79\text{--}0.84$) present in well-behaved closeness measures, which reduces the effective sample size from $N = 42,167$ segments to approximately 3,700 independent observations. The pattern persists across distance thresholds (500m–10km), for both metric and geometric distances, and for origin-destination trip data (Figures 13, 14). The following subsections present detailed correlation grids and the spatial autocorrelation analysis underpinning these estimates.

6.3 Comparisons to Landuse Accessibility

Figures 9, 10, 11, 12 represent correlation grids comparing network centralities to landuse accessibilities. Each grid can be read as follows:

- The y axis labels correspond to a network centrality measure.
- The x axis labels correspond to the distance used for the localised catchment (so-called radii or moving windows) at which the network centrality measures shown on the y axis have been computed, ranging from 500m to 10km.
- The correlations shown in the individual cells correspond to the Spearman Rank correlation for a given network centrality measure (y axis) at a given distance (x axis) correlated to the landuse theme for a given correlation matrix (indicated in the title).

The results are broadly summarised as follows:

- The density measure represents a basic count of nodes, else of street lengths for streets inside the distance thresholds for the length-weighted case. Density shows strong associations for smaller distances in spite of its simplicity, potentially indicating that for smaller catchments the overriding issue is direct access to as much of the street network as possible. The *Cycles* measure, a basic count of network cycles, performs similarly.
- The *Farness* measure is positively associated and shows stronger associations for smaller distance thresholds for reasons likely similar to density, where simple direct access to the surrounding network affords greater access to landuses. For larger distances, it lags the more complex measures because it doesn't directly consider the effective closeness of the network as the network size expands. The implication of *Farness* being positively associated is that its inverse, *Closeness* correlates negatively. The normalised case of either measure is ineffective.
- *NAIN*, where the numerator is raised to the power of 1.2, behaves more suitably than *Normalised Closeness*, though lags the *Improved Closeness*, *Gravity*, and *Harmonic Closeness* variants which show the most consistently strong associations. The latter two appear to slightly outperform *Improved Closeness*, possibly because these factor distances directly for each summation instead of first summing and then averaging the distances prior to division. However, this trend is subtle and does not likely warrant favouring one over the other.
- In the context of Madrid, the betweenness variants are more weakly associated to landuses than closeness-like measures. This is to be expected given Madrid's high intensities of landuses in its walkable core, which follow a fractal or 'space-filling' logic utilising all available street-frontages even where not directly adjacent to the most heavily walked streets. For similar reasons, Madrid generally shows slightly weaker associations for geometric (angular) distance measures. Note that these observations may be different for other contexts (e.g. London's high streets) or for associations against pedestrian volumes as opposed to landuse intensities.
- The street-length weighted variants and their accompanying continuous variants exhibit similar behaviour, with generally slightly stronger associations compared to the unweighted versions.



Figure 6: Comparative plots of shortest metric distance closeness centralities for a 1000m catchment.



Figure 7: Comparative plots of shortest geometric distance (angular) closeness centralities for a 1000m catchment. Gravity is not applicable to geometric distances.

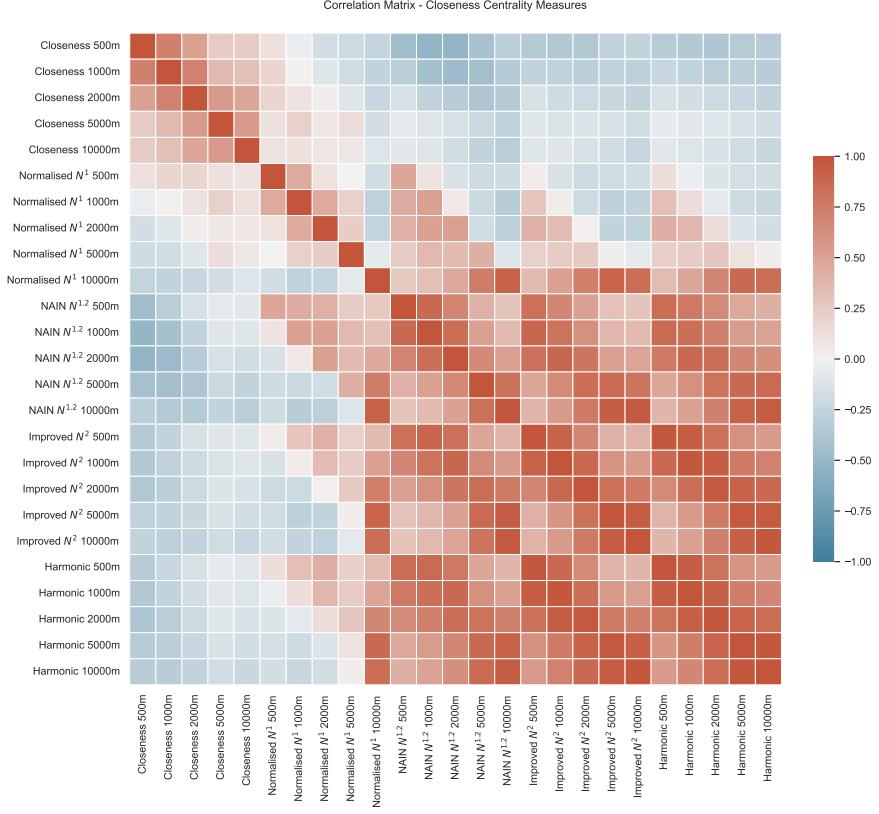


Figure 8: Correlation grid of different closeness centralities.

6.4 Comparisons to Origin-Destination Trip Data

Correlations of network centralities to origin-destination travel survey data, shown in Figures 13 and 14, largely reflect the aforementioned patterns, with *Closeness* and *Normalised Closeness* again showing negative or negligible associations whereas the *Improved*, *Gravity*, and *Harmonic Closeness* versions are more strongly associated.

6.5 Spatial Autocorrelation and Uncertainty Quantification

The spatial autocorrelation analysis (Table 12) indicates a further divergence between closeness formulations. Measures that scale as anticipated — *Improved Closeness*, *Gravity*, and *Harmonic Closeness* — exhibit high Moran's I values ($I \approx 0.79\text{--}0.84$), consistent with the smooth spatial clustering typically associated with closeness centrality. Nearby street segments sharing similar network context have similarly pronounced centrality values under these formulations. In contrast, *Closeness* and *Normalised Closeness* exhibit lower spatial autocorrelation ($I \approx 0.18\text{--}0.2$ at 500m), consistent with less coherent spatial patterning.

To account for this spatial dependence when estimating uncertainty, we compute effective sample sizes and block-bootstrap confidence intervals. The effective sample size varies from approximately 3,700 for measures with high spatial coherence ($I \approx 0.84$) to 27,000 for *Closeness* ($I \approx 0.21$), compared to the nominal $N = 42,167$ segments. The block-bootstrap procedure (Table 14) uses spatially contiguous blocks to preserve autocorrelation structure, yielding conservative confidence intervals consistent with the sign-reversal pattern summarised in Section 6: the intervals for *Closeness*, *Normalised Closeness*, and *Improved Closeness* do not overlap, indicating that the observed divergence is statistically robust.

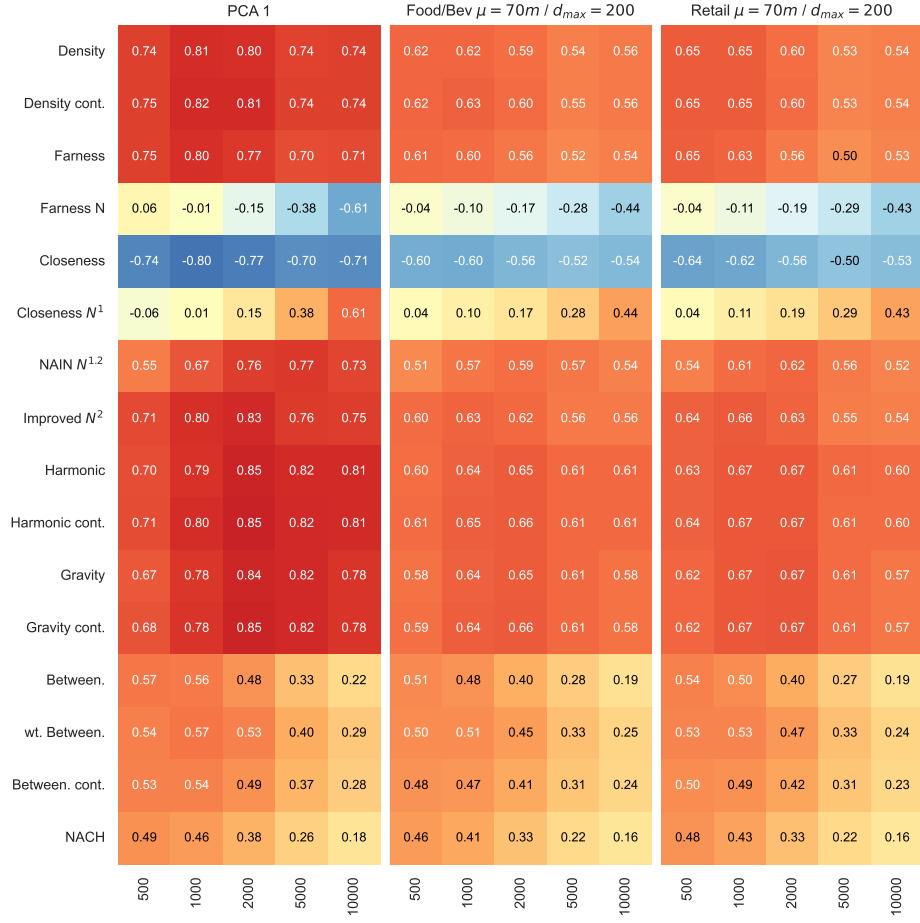
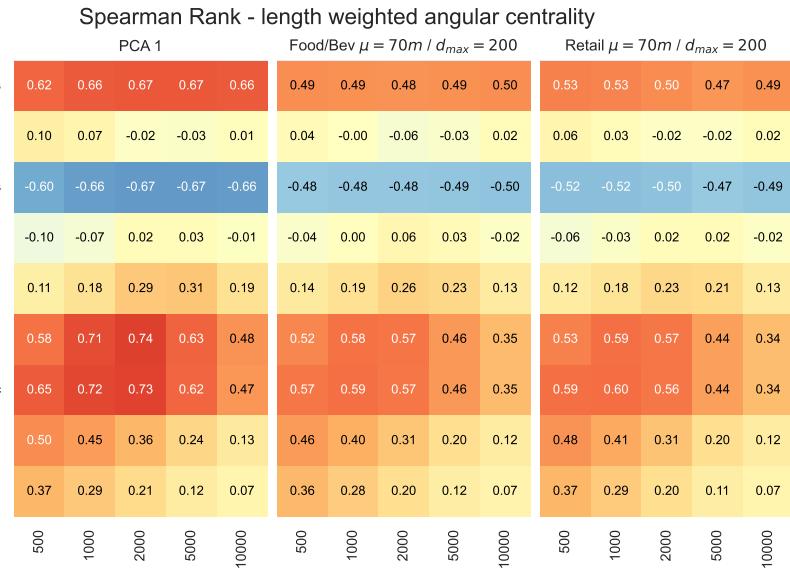
Spearman Rank - centrality													
	PCA 1					Food/Bev $\mu = 70m / d_{max} = 200$					Retail $\mu = 70m / d_{max} = 200$		
Density	0.70	0.78	0.81	0.74	0.74	0.58	0.60	0.60	0.55	0.56	0.62	0.64	0.61
Farness	0.72	0.78	0.77	0.70	0.69	0.58	0.59	0.57	0.52	0.53	0.62	0.62	0.57
Farness N	0.11	0.01	-0.15	-0.34	-0.66	0.02	-0.07	-0.16	-0.25	-0.48	0.02	-0.09	-0.19
Closeness	-0.70	-0.78	-0.77	-0.70	-0.69	-0.57	-0.59	-0.57	-0.52	-0.53	-0.61	-0.62	-0.57
Closeness N^1	-0.11	-0.01	0.15	0.34	0.66	-0.02	0.07	0.16	0.25	0.48	-0.02	0.09	0.19
NAIN $N^{1,2}$	0.50	0.61	0.69	0.77	0.74	0.45	0.52	0.55	0.57	0.55	0.49	0.56	0.58
Improved N^2	0.67	0.77	0.83	0.77	0.75	0.56	0.60	0.62	0.57	0.57	0.60	0.64	0.64
Harmonic	0.65	0.74	0.83	0.84	0.83	0.55	0.60	0.64	0.63	0.63	0.59	0.63	0.66
Gravity	0.63	0.73	0.82	0.84	0.80	0.54	0.60	0.63	0.63	0.59	0.58	0.63	0.63
Cycles	0.71	0.80	0.82	0.74	0.74	0.59	0.62	0.61	0.55	0.56	0.62	0.64	0.62
Between.	0.58	0.59	0.54	0.41	0.28	0.52	0.51	0.45	0.34	0.24	0.55	0.53	0.46
wt. Between.	0.55	0.59	0.58	0.47	0.37	0.50	0.52	0.49	0.39	0.31	0.53	0.55	0.51
NACH	0.52	0.50	0.44	0.33	0.24	0.48	0.45	0.38	0.29	0.21	0.50	0.47	0.39
	500	1000	2000	5000	10000	500	1000	2000	5000	10000	500	1000	2000
													10000

Figure 9: Correlation grids comparing *metric distance* network centralities to landuses.

Spearman Rank - angular centrality													
	PCA 1					Food/Bev $\mu = 70m / d_{max} = 200$					Retail $\mu = 70m / d_{max} = 200$		
ang. Farness	0.62	0.68	0.70	0.69	0.68	0.50	0.51	0.51	0.51	0.51	0.54	0.55	0.53
ang. Farness N	0.07	0.03	-0.07	-0.13	-0.10	0.01	-0.03	-0.11	-0.12	-0.06	0.03	-0.01	-0.08
ang. Closeness	-0.61	-0.68	-0.70	-0.69	-0.68	-0.48	-0.51	-0.50	-0.50	-0.51	-0.53	-0.55	-0.53
ang. Closeness N^1	-0.07	-0.03	0.07	0.13	0.10	-0.01	0.03	0.11	0.12	0.06	-0.03	0.01	0.08
ang. NAIN $N^{1,2}$	0.24	0.34	0.47	0.49	0.32	0.25	0.33	0.41	0.38	0.22	0.24	0.32	0.39
ang. Improved N^2	0.66	0.77	0.81	0.70	0.56	0.57	0.62	0.63	0.52	0.41	0.60	0.65	0.63
ang. Harmonic	0.68	0.77	0.81	0.72	0.58	0.59	0.63	0.63	0.54	0.43	0.62	0.65	0.64
ang. Between.	0.52	0.50	0.43	0.31	0.20	0.47	0.43	0.36	0.26	0.17	0.50	0.45	0.37
ang. NACH	0.39	0.32	0.26	0.18	0.12	0.38	0.31	0.24	0.16	0.11	0.39	0.31	0.24
	500	1000	2000	5000	10000	500	1000	2000	5000	10000	500	1000	2000
													10000

Figure 10: Correlation grids comparing *geometric distance* angular network centralities to landuses.

Spearman Rank - length weighted centrality

Figure 11: Correlation grids comparing length-weighted *metric distance* network centralities to landuses.Figure 12: Correlation grids comparing length-weighted *geometric distance* angular network centralities to landuses.

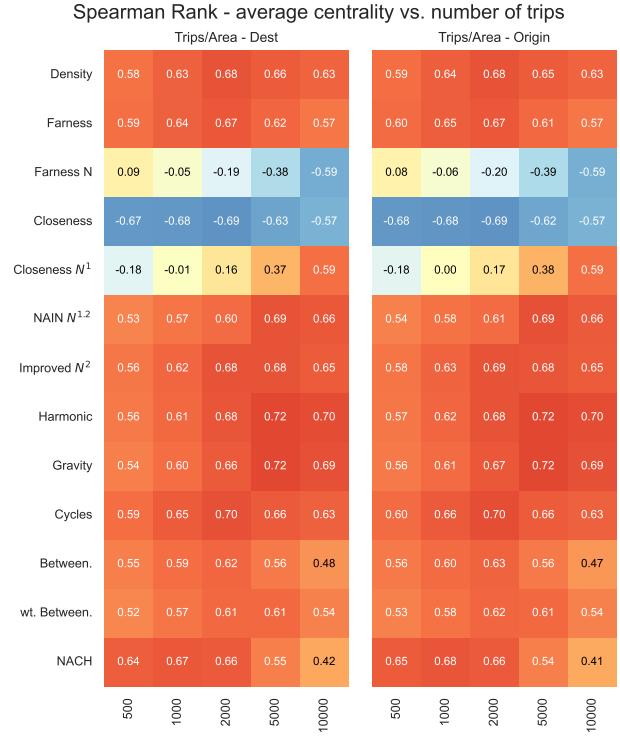


Figure 13: Correlation grids comparing average *metric distance* network centralities to trip origins and destinations for travel survey zones.

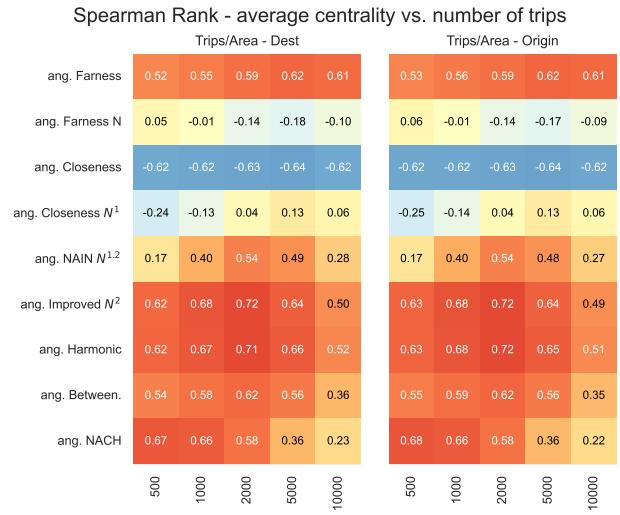


Figure 14: Correlation grids comparing average *geometric distance* angular network centralities to trip origins and destinations for travel survey zones.

6.6 Summary

The empirical results are consistent with the theoretical hypothesis presented in Section 3. *Normalised Closeness*, despite being widely cited in the street network analysis literature, shows weak or negative associations with landuse accessibility and origin-destination trip volumes when computed for localised analyses. In contrast, *Improved Closeness* and *Harmonic Closeness* — which scale appropriately across varying numbers of reachable nodes — show consistently stronger associations. These findings support our contention that the divergence between cited formulations and those implemented in computational packages reflects a meaningful methodological distinction rather than notational variation.

6.7 Limitations

This study is based on a single city case study, and the specific correlation magnitudes may differ for other urban contexts. Madrid's dense, walkable historic core with high landuse intensities distributed across the network may favour closeness-like measures over betweenness; contexts with more pronounced linear high streets (e.g., London) or studies based on detailed pedestrian counts might show different relative performance between measures and stronger associations for *geometric* distance variants. The analysis is cross-sectional and does not capture temporal dynamics in landuse or travel patterns.

7 Discussion and Conclusions

The results of the empirical observations for the Madrid case are consistent with the behaviour anticipated by the theoretical hypothesis. Regarding **RQ1**, we find a meaningful distinction: mathematically defined *Closeness* behaves opposite to its intended usage when applied to localised street network analysis, and when normalised in the form of *Normalised Closeness*, it performs poorly compared to localised variants such as *Improved Closeness* and *Harmonic Closeness*. Regarding **RQ2**, the distinction is reflected in empirical associations: the results hold for access to landuses as well as for origin-destination trip counts, regardless of whether closeness centralities are length weighted or not, and irrespective of whether *metric* or *geometric* distances are used.

Developers of domain-specific urban network analysis packages appear to be aware of this issue even if not reflected in wider discussion; the *DepthmapX* [30], *Place Syntax Toolkit* [31], and *cityseer* [32] packages all use or offer a form of *Improved Closeness* [5], and in some cases offer other suitable variants such as *Harmonic Closeness* and the *Gravity Index*.²

We suggest the following implications:

- There appears to be a misunderstanding in the general street network analysis literature that *Normalised Closeness* is being used for localised forms of street network analysis. Localised forms of analysis are used by convention and are considered best practice for reasons explained in Section 2.
- We suspect that one of the reasons this misattribution persists is that even though *Normalised Closeness* is typically cited by the authors of papers, these studies are often relying on computational packages which use forms of closeness that work for localised analysis. Consequently, outputs would typically perform as expected even if users were under mistaken impressions regarding the form of closeness being used.
- In cases where computational workflows relying on generic network analysis packages or non-open-source packages are used, this raises wider issues of reproducibility. On one hand, researchers trying to replicate the behaviour of existing studies are faced with a conundrum because they may be trying to replicate work using an ineffective form of closeness centrality. Secondly, where the code is not openly available, there may be issues regarding the reliability or interpretability of findings because it may not be clear whether the cited formulation matches that used for the calculation process.

These ambiguities warrant further investigation into the robustness of currently held assumptions regarding the relative performance of different centralities and forms of distance heuristic, which have been influenced by widely cited earlier studies. Some such studies cite *Normalised Closeness* but do not provide open code to verify the actual computational methods used, select small geographic areas for analysis, and do not buffer boundaries to control for edge effects (which would skew closeness centralities). The conclusions drawn in such cases regarding the relative performance of closeness measures may warrant re-examination [16, 46].

²Some packages also provide *Normalised Closeness* as an option.

We consequently encourage the use of openly available reference datasets and reproducible workflows, as we have made available in this study, so that different measures can be benchmarked openly and collaboratively against common points of reference. This would facilitate the evaluation of newly available methods against established variants while accounting for differences which might otherwise be attributable to variations in urban context or data representation. It would also help uncover situations such as the widespread misconceptions on closeness centralities discussed in this study.

In conclusion, our findings for RQ1 and RQ2 support a clear practical recommendation: researchers working with localised street network analysis should use *Improved Closeness*, *Harmonic Closeness*, or similar variants rather than *Normalised Closeness*, and should clearly document and share the methods and precise computational formulations used in their analyses.

8 Funding Statement

9 Declaration of generative AI and AI-assisted technologies in the manuscript preparation process.

The majority of this work was completed without the assistance of generative AI. During the final stages of preparation the authors used Claude Opus 4.5 for code, grammar, and style. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

- [1] Jane Jacobs. *The Death and Life of Great American Cities*. Random House, New York, 1961.
- [2] Christopher Alexander. A City is Not a Tree. *Ekistics*, 23(139):344–348, 1967.
- [3] Transport for New Homes. What is being built in 2025 - In search of the station. Technical report, 2025. URL <https://www.transportfornewhomes.org.uk/wp-content/uploads/2025/03/What-is-being-built-in-2025-In-search-of-the-station.pdf>.
- [4] Alfonso Shimbel. Structural Parameters of Communication Networks. *Bulletin of Mathematical Physics*, 15: 501–507, 1953.
- [5] Stanley Wasserman and Katherine Faust. *Social Network Analysis*. Cambridge University Press, Cambridge, 1994.
- [6] Leonhard Euler. Solutio problematis ad geometriam situs pertinentis. *Euler Archive - All Works*, 53, 1741. URL <https://scholarlycommons.pacific.edu/euler-works/53>.
- [7] Anthony Froshaug. Visuelle Methodik. 1959. doi: 10.11588/DIGLIT.60957.2. URL <https://digi.ub.uni-heidelberg.de/diglit/ulm4/0001>. Publisher: : NN.
- [8] Bill Hillier and Julienne Hanson. *The Social Logic of Space*. Cambridge University Press, Cambridge, 1984.
- [9] Chen Feng, Daniel Koch, and Ann Legeby. Accessibility patterns based on steps, direction changes, and angular deviation: Are they consistent? In *Proceedings: 13th International Space Syntax Symposium*, pages 534:1–20, Western Norway University of Applied sciences, 2022.
- [10] Stephen Marshall, Jorge Gil, Karl Kropf, Martin Tomko, and Lucas Figueiredo. Street Network Studies: from Networks to Models and their Representations. *Networks and Spatial Economics*, 2018. ISSN 1572-9427. doi: 10.1007/s11067-018-9427-9. URL <https://doi.org/10.1007/s11067-018-9427-9>.
- [11] Krenz, Kimon. Developments in Space Syntax: Past, Present and Future. *Urbanism*, 4(NO 042):80, 2022. URL http://urbandesign.tsinghua.org/CN/abstract/article_154236.shtml.
- [12] Jorge Gil. Street network analysis 'edge effects': Examining the sensitivity of centrality measures to boundary conditions. *Environment and Planning B: Urban Analytics and City Science*, 44(5), 2017. ISSN 23998091. doi: 10.1177/0265813516650678.
- [13] Gert Sabidussi. The Centrality Index of a Graph. *Psychometrika*, 31(4):581–603, 1966.
- [14] Linton C Freeman. A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40(1):35–41, 1977. URL <http://www.jstor.org/stable/3033543>.

The funding information will be inserted after review.

- [15] Elsa Arcaute, Carlos Molinero, Erez Hatna, Roberto Murcio, Camilo Vargas-Ruiz, A Paolo Masucci, and Michael Batty. Cities and regions in Britain through hierarchical percolation. *Royal Society Open Science*, 3, 2016. doi: 10.1098/rsos.150691. URL <http://dx.doi.org/10.1098/rsos.150691>.
- [16] Alasdair Turner. 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, 2007. doi: 10.1068/b32067.
- [17] Akkelies Van Nes and Claudia Yamu. *Introduction to Space Syntax in Urban Studies*. Springer International Publishing, Cham, 2021. ISBN 978-3-030-59139-7. doi: 10.1007/978-3-030-59140-3. URL <https://link.springer.com/10.1007/978-3-030-59140-3>.
- [18] Sergio Porta, Emanuele Strano, Valentino Iacoviello, Roberto Messora, Vito Latora, Alessio Cardillo, Fahui Wang, and Salvatore Scellato. Street Centrality and Densities of Retail and Services in Bologna, Italy. *Environment and Planning B: Planning and Design*, 36(3):450–465, January 2009. ISSN 0265-8135. doi: 10.1068/b34098. URL <https://doi.org/10.1068/b34098>. Publisher: SAGE Publications Ltd STM.
- [19] Sergio Porta, Paolo Crucitti, and Vito Latora. The network analysis of urban streets: A dual approach. *Physica A*, 369:853–866, 2006. doi: 10.1016/j.physa.2005.12.063.
- [20] Carlo Ratti. Space syntax: some inconsistencies. *Environment and Planning B: Planning and Design*, 31: 487–499, 2004. doi: 10.1068/b3019.
- [21] Alasdair Turner and Nick Dalton. A Simplified Route Choice Model Using the Shortest Angular Path Assumption. Technical report, Bartlett School of Graduate Studies, University College London, London, 2005.
- [22] Alasdair Turner. Angular Analysis: A Method for the Quantification of Space. Technical report, UCL Centre For Advanced Spatial Analysis, London, 2000. URL <http://www.casa.ucl.ac.uk>.
- [23] Nick Dalton. Fractional Configurational Analysis And a solution to the Manhattan problem. In *3rd International Space Syntax Symposium*, Atlanta, 2001.
- [24] M Rosvall, A Trusina, P Minnhagen, and K Sneppen. Networks and Cities: An Information Perspective. *Physical Review Letters*, 94, 2005. doi: 10.1103/PhysRevLett.94.028701.
- [25] Michael Batty. Distance in Space Syntax. *Working Paper Series*, 04(04):32, 2004. URL www.casa.ucl.ac.uk. Place: London ISBN: 1467-1298 Publisher: UCL Centre For Advanced Spatial Analysis.
- [26] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *REVIEWS OF MODERN PHYSICS*, 74(January):47–97, 2002.
- [27] Bill Hillier, Alasdair Turner, Tao Yang, and Hoon-Tae Park. METRIC AND TOPO-GEOMETRIC PROPERTIES OF URBAN STREET NETWORKS. In *Proceedings, 6th International Space Syntax Symposium*, Istanbul, 2007.
- [28] Miguel Serra and Bill Hillier. Angular and Metric Distance in Road Network Analysis: A nationwide correlation study. *Computers, Environment and Urban Systems*, 2019. ISSN 01989715. doi: 10.1016/j.compenvurbssys.2018.11.003.
- [29] A Penn, B Hillier, D Banister, and J Xu. Configurational modelling of urban movement networks. *Environment and Planning B: Planning and Design*, 25:59–84, 1998.
- [30] Alasdair Turner, Eva Friedrich, Tasos Varoudis, Christian Sailer, and Petros Koutsolampros. depthmapX, 2020. URL <https://github.com/SpaceGroupUCL/depthmapX>.
- [31] Alexander Stahle, Lars Marcus, Daniel Koch, Martin Fitger, Ann Legeby, Gianna Stavroulaki, Pont Berghauser, Anders Karlstrom, Pablo Miranda Carranza, and Tobias Nordstrom. Place Syntax Tool: PST Documentation, February 2023.
- [32] Gareth Simons. The cityseer Python package for pedestrian-scale network-based urban analysis. *Environment and Planning B: Urban Analytics and City Science*, 50(5):1328–1344, June 2023. ISSN 2399-8083, 2399-8091. doi: 10.1177/23998083221133827. URL <http://journals.sagepub.com/doi/10.1177/23998083221133827>.
- [33] Andrea Banino, Caswell Barry, Benigno Urias, Charles Blundell, Timothy Lillicrap, Piotr Mirowski, Alexander Pritzel, Martin J. Chadwick, Thomas Degris, Joseph Modayil, Greg Wayne, Hubert Soyer, Fabio Viola, Brian Zhang, Ross Goroshin, Neil Rabinowitz, Razvan Pascanu, Charlie Beattie, Stig Petersen, Amir Sadik, Stephen Gaffney, Helen King, Koray Kavukcuoglu, Demis Hassabis, Raia Hadsell, and Dharshan Kumaran. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, May 2018. ISSN 0028-0836, 1476-4687. doi: 10.1038/s41586-018-0102-6. URL <https://www.nature.com/articles/s41586-018-0102-6>.

- [34] Marc Barthelemy. From paths to blocks: New measures for street patterns. *Environment and Planning B: Urban Analytics and City Science*, 44(2):256–271, August 2015. ISSN 2399-8083. doi: 10.1177/0265813515599982. URL <https://doi.org/10.1177/0265813515599982>. Publisher: SAGE Publications Ltd STM.
- [35] Matheus P Viana, Emanuele Strano, Patricia Bordin, and Marc Barthelemy. The simplicity of planar networks. *Scientific Reports*, 3(3495), 2013. doi: 10.1038/srep03495.
- [36] Itzhak Omer and Nir Kaplan. 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*, 46(6): 1182–1200, February 2018. ISSN 2399-8083. doi: 10.1177/2399808318760571. URL <https://doi.org/10.1177/2399808318760571>. Publisher: SAGE Publications Ltd STM.
- [37] Samia Sharmin and Md. Kamruzzaman. Meta-analysis of the relationships between space syntax measures and pedestrian movement. *Transport Reviews*, 38(4):524–550, July 2018. ISSN 0144-1647, 1464-5327. doi: 10.1080/01441647.2017.1365101. URL <https://www.tandfonline.com/doi/full/10.1080/01441647.2017.1365101>.
- [38] A Paolo Masucci and Carlos Molinero. Robustness and closeness centrality for self-organized and planned cities. *Eur. Phys. J. B*, 89(53), 2016. doi: 10.1140/epjb/e2016-60431-2.
- [39] Jorge Gil, I Kolovou, Stephen Law, and L Versluis. Road centre line simplification principles for angular segment analysis. In *Proceedings of 11th International Space Syntax Symposium.*, Lisbon, Portugal.
- [40] Kimon Krenz. Employing volunteered geographic information in space syntax analysis. In *Proceedings of 11th International Space Syntax Symposium.*, Lisbon, Portugal.
- [41] Alasdair Turner. Getting Serious with Depthmap; Segment Analysis and Scripting, January 2008. URL <https://archtech.gr/varoudis/depthmapX/LearningMaterial/advanceddepthmap.pdf>.
- [42] Bill Hillier, Tao Yang, and Alasdair Turner. Normalising least angle choice in Depthmap and how it opens up new perspectives on the global and local analysis of city space. *The Journal of Space Syntax*, 3(2):155–193, 2012. URL <http://www.journalofspacesyntax.org/>.
- [43] Paolo Crucitti, Vito Latora, and Sergio Porta. Centrality in networks of urban streets. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 16(1):015113, March 2006. ISSN 1054-1500. doi: 10.1063/1.2150162. URL <https://doi.org/10.1063/1.2150162>.
- [44] Laura Vaughan. Glossary of Space Syntax. 2025.
- [45] Segment angular integration Space Syntax – Online Training Platform. URL <https://www.spacesyntax.online/term/segmental-integration/>.
- [46] Sergio Porta, Paolo Crucitti, and Vito Latora. The Network Analysis of Urban Streets: A Primal Approach. *Environment and Planning B: Planning and Design*, 33:705–725, 2006. doi: 10.1068/b32045.
- [47] Bill Hillier and Shinichi Iida. Network and Psychological Effects in Urban Movement. In David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Dough Tygar, Moshe Y. Vardi, Gerhard Weikum, Anthony G. Cohn, and David M. Mark, editors, *Spatial Information Theory*, volume 3693, pages 475–490. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005. doi: 10.1007/11556114_30. URL http://link.springer.com/10.1007/11556114_30. Series Title: Lecture Notes in Computer Science.
- [48] Andres Sevtsuk and Michael Mekonnen. Urban network analysis: A new toolbox for ArcGIS. *RIG*, 22:287–305, 2012.
- [49] Crispin H V Cooper. 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, 2015. doi: 10.1080/13658816.2015.1018834.
- [50] Michael. Batty. *The New Science of Cities*. MIT Press, Cambridge, MA, 2013.
- [51] Mahbub Rashid. *The Geometry of Urban Layouts*. Springer International Publishing, Cham, 2017. ISBN 978-3-319-30748-0. doi: 10.1007/978-3-319-30750-3. URL <http://link.springer.com/10.1007/978-3-319-30750-3>.
- [52] Massimo Marchiori and Vito Latora. Harmony in the small-world. *Physica A*, 285(89):539–546, 2000. URL www.elsevier.com/locate/physa.
- [53] Yannick Rochat. Closeness Centrality Extended To Unconnected Graphs : The Harmonic Centrality Index. Technical report, Institute of Applied Mathematics University of Lausanne, Switzerland, Zurich, 2009.

- [54] Walter G Hansen. How Accessibility Shapes Land Use. *JOURNAL OF THE AMERICAN INSTITUTE OF PLANNERS*, 25(2):73–76, 1959. doi: 10.1080/01944365908978307. URL <http://www.tandfonline.com/action/journalInformation?journalCode=rjpa19>.
- [55] S. L. Handy and D. A. Niemeier. Measuring accessibility: An exploration of issues and alternatives. *Environment and Planning A*, 1997. ISSN 0308518X. doi: 10.1068/a291175. ISBN: 0308-518X.
- [56] Michael Iacono, Kevin Krizek, Ahmed El-Geneidy, and Hubert H Humphrey. Access to Destinations: How Close is Close Enough? Estimating Accurate Distance Decay Functions for Multiple Modes and Different Purposes. 2008.
- [57] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Edouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [58] A S Fotheringham and D W S Wong. The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. *Environment and Planning A: Economy and Space*, 23(7):1025–1044, July 1991. ISSN 0308-518X. doi: 10.1068/a231025. URL <https://doi.org/10.1068/a231025>. Publisher: SAGE Publications Ltd.
- [59] P. A. P. Moran. Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1/2):17, June 1950. ISSN 00063444. doi: 10.2307/2332142. URL <https://www.jstor.org/stable/2332142?origin=crossref>.
- [60] Community of Madrid. Callejero de la comunidad de madrid. Dataset, 2019. URL https://datos.comunidad.madrid/catalogo/dataset/spacm_callescm. Open data. Set of roads officially approved by the municipalities of the Community of Madrid. Licensed under Creative Commons Attribution 4.0 (CC BY 4.0). License terms: <https://creativecommons.org/licenses/by/4.0/legalcode.es>. Original dataset link since removed: https://datos.comunidad.madrid/catalogo/dataset/spacm_callescm Possible replacements from https://data.europa.eu/data/datasets/https-idem-madrid-org-catalogocartografia-srv-resources-datasets-spacm_callescm?locale=en https://gestiona.comunidad.madrid/iestadis/fijas/estructu/general/territorio/estructu_descargas.htm https://gestiona.comunidad.madrid/nomecalles_web/#/inicio via Download Calejero.
- [61] Madrid City Council. Census of premises and activities of the madrid city council. Dataset, 2014. URL <https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=66665cd99be2410VgnVCM1000000b205a0aRCRD>. Open data. Origin of the data: Madrid City Council. Licensed under Spanish Law 37/2007 on Reuse of Public Sector Information. License terms: <https://datos.madrid.es/egob/catalogo/aviso-legal>.
- [62] Consorcio Regional de Transportes de Madrid. Encuesta Domiciliaria de Movilidad (EDM) 2018. Microsoft Excel Spreadsheet, 2020. URL <https://datos.crtm.es/documents/6af74db8175d4902ada0803f08ccf50e/about>. Travel survey data for the Madrid metropolitan region. File size: 22.88 MB. Published: 27 February 2020. Open data license from the Madrid Regional Transport Consortium, allowing commercial and non-commercial reuse with attribution.
- [63] Consorcio Regional de Transportes de Madrid. ZonificacionZT1259. Feature Layer Dataset, 2020. URL <https://datos.crtm.es/search?q=1259>. Transport zones (ZT1259) from EDM2018. 1,259 geographic zones for the Community of Madrid at a territorial scale between neighborhood and census section. Published: 29 August 2019. Updated: 28 February 2020. Open data license from the Madrid Regional Transport Consortium, allowing commercial and non-commercial reuse with attribution.
- [64] Scott Rutherford. Use of the Gravity Model for Pedestrian Travel Distribution. *Transportation Research Record*, (728):53–59, 1979.
- [65] David S Vale and Mauro Pereira. The influence of the impedance function on gravity-based pedestrian accessibility measures: A comparative analysis. *Environment and Planning B: Urban Analytics and City Science*, 44(4):740–763, July 2017. ISSN 2399-8083, 2399-8091. doi: 10.1177/0265813516641685. URL <http://journals.sagepub.com/doi/10.1177/0265813516641685>.
- [66] Andres Sevtsuk, Raul Kalvo, and Onur Ekmekci. Pedestrian accessibility in grid layouts: The role of block, plot and street dimensions. *Urban Morphology*, 2016. ISSN 10274278. doi: 10.1080/10464883.2012.714912. ISBN: 0870702823.
- [67] Maren Reyer, Stefan Fina, Stefan Siedentop, and Wolfgang Schlicht. Walkability is only part of the story: Walking for transportation in Stuttgart, Germany. *International Journal of Environmental Research and Public Health*, 2014. ISSN 16604601. doi: 10.3390/ijerph110605849. ISBN: 1660-4601 (Electronic)\n1660-4601 (Linking).

- [68] Siamak Baradaran and Farideh Ramjerdi. Performance of Accessibility Measures in Europe. *Journal of Transportation and Statistics*, 4(2/3):31–48, 2001. URL https://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/journal_of_transportation_and_statistics/volume_04_number_23/paper_03/index.html.
- [69] Jan Scheurer and Carey Curtis. Accessibility Measures: Overview and Practical Applications. Technical report, Curtin University, 2007. URL <http://urbanet.curtin.edu.au/>. Publication Title: Accessibility Measures.
- [70] Dustin T Duncan, Jared Aldstadt, John Whalen, Steven J Melly, and Steven L Gortmaker. Validation of Walk Score ® for Estimating Neighborhood Walkability: An Analysis of Four US Metropolitan Areas. *Int. J. Environ. Res. Public Health*, 8:4160–4179, 2011. doi: 10.3390/ijerph8114160. URL www.mdpi.com/journal/ijerph.
- [71] Susan Handy and Kelly Clifton. Evaluating Neighborhood Accessibility: Possibilities and Practicalities. *Journal of Transportation and Statistics*, 4(2/3):67–78, 2001.
- [72] Britton Harris. Accessibility: Concepts and Applications. *Journal of Transportation and Statistics*, 4(2/3):15–30, 2001.
- [73] John Bates. History of Demand Modelling. In David Hensher and Kenneth Button, editors, *Handbook of Transport Modelling*, pages 11–34. Elsevier Science, 2007. URL <https://doi.org/10.1108/9780857245670-002/>.

10 Supplementary Materials

10.1 Additional Centralities

10.1.1 The *Gravity Index*

The use of spatial impedance as an accessibility measure, often referred to as the *Gravity Index*, shifts the emphasis to the potential flow of interactions over the street network [50, 54, 64]. Use of a spatial impedance function introduces a more configurable approach explicitly modelling spatial impedances; it reflects the potential for spatial interaction from nodes j to node i and — like gravity — how the potential for this interaction decays with distance. This typically takes the form of the negative exponential

$$\text{Gravity}_{(i)} = \sum_{j \neq i} \exp(-\beta \cdot d) \quad (8)$$

where the rate of decay β (in the negative exponential) can be set to model specific trip-purposes or transportation modes to reflect people's willingness to travel a given distance d to particular types of locations [48, 55, 56]. A variety of different impedance functions can be used, though these tend to behave similarly [65].

The term *Gravity Index* can be somewhat misleading: full-fledged gravity models form the basis of broader land-use and transportation modelling where the attraction of the origins and destinations are, as per gravity, taken into account. However, the *Gravity Index* generally assumes equal attractions from each node to every other node and, as such, is simply a means to provide distance-weighted counts of accessible locations j proximate to i at the given impedance β . This approach works well for quantifying access to specific land-uses or, in the case of the street network structure, quantifying physical access from nodes j to node i . In the context of network analysis, the strength of the decay parameter can be varied to emphasise smaller or larger structures in the network.

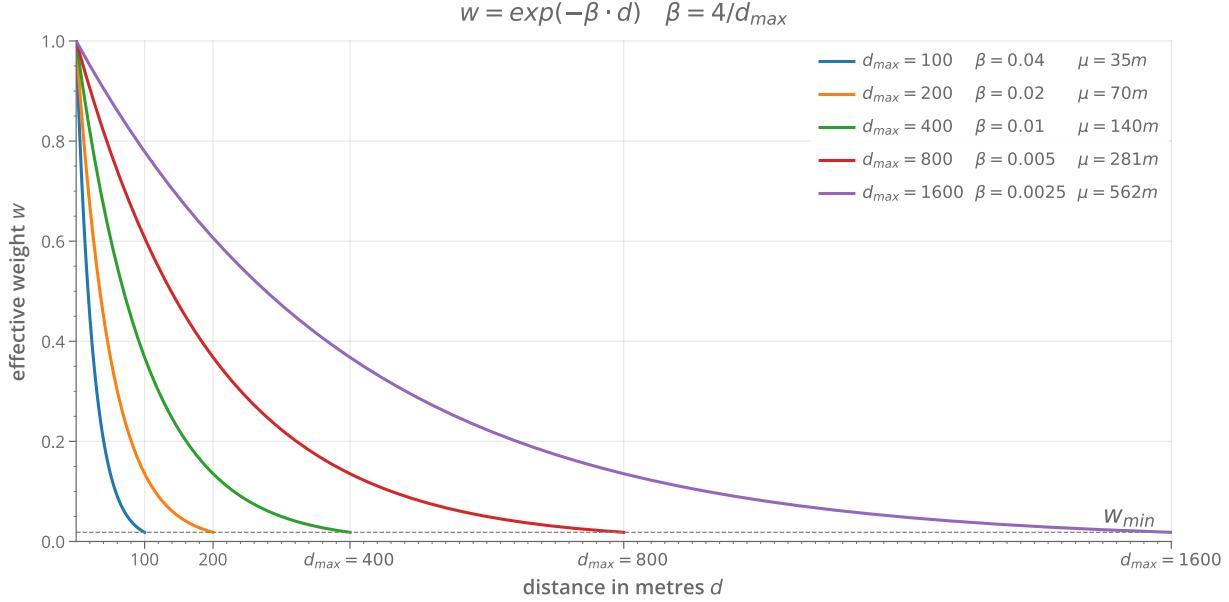


Figure S1: Spatial impedance curves for different β parameters. Nearer locations can be weighted more heavily than farther locations through use of the negative exponential decay function.

Gravity measures are inherently localised and do not strictly require distance cutoffs. It is nevertheless computationally advantageous to retain thresholds commensurate with distances at which the decay renders additional computation sufficiently negligible (Figure S1). For the proceeding discussion, a range of impedances is applied with the respective β parameters anchored to distance thresholds d_{max} through $\beta = 4/d_{max}$. For example, a 100m distance threshold d_{max} corresponds to $\beta = 0.04$ and gives an average trip distance of 35m. The distances and the corresponding impedances are selected with pedestrians in mind, where, for example, $\beta = 0.005$ (800m d_{max}) may represent relatively typical walking distances to bus stops [66, 55]. Realistically, these values vary significantly based on the purpose of the trip because pedestrians may be willing to walk much farther for purposes such as recreation and fitness than for purposes

such as shopping [56]. These parameters also vary based on location: it can be argued that North American contexts are less supportive of walking [67] than European equivalents. Regardless, pedestrians tend to be unwilling to walk distances greater than a mile (1600m, $\beta \approx 0.0025$) for non-recreational purposes and, per the exponential decay function, are more likely to walk to nearer locations than those farther away [68, 69, 70, 71, 72, 73].

10.1.2 Betweenness centrality

The shortest path from any node j to any other node k will pass through an assortment of nodes i , that is, unless j and k are directly adjacent. *Betweenness*

$$\text{Betweenness}_{(i)} = \sum_{j \neq i} \sum_{k \neq j \neq i} n(j, k) i \quad (9)$$

is the summation of shortest paths between all (j, k) pairs of nodes passing through a given node i [14], and in the case of street networks conveys how likely a street is to be traversed by people travelling between other locations. Note that this measure is typically referred to as *Choice* by the Space Syntax community. Localised *Betweenness* only considers other (j, k) node pairs within the threshold cutoff distances.

Betweenness can be weighted by distances:

$$\text{Betweenness}_{(i)} = \sum_{j \neq i} \sum_{k \neq j \neq i} n(j, k) i \cdot \exp(-\beta \cdot d), \quad (10)$$

in which case the negative exponential (see 8) can be used to reflect the notion that trips between closely located (j, k) node-pairs are more likely to occur than those located farther apart, with d in this case representing the corresponding trip distance for a given (j, k) pair of nodes passing through node i .

Note that the Space Syntax methods also include the normalised least angular choice (*NACH*) measure, which is described as a normalised form of *Betweenness* (*Choice*) achieved through division by *Farness* (Total Depth)

$$\text{NACH}(i) = \frac{\log(\text{Betweenness} + 1)}{\log(\text{Farness} + 3)}. \quad (11)$$

This is better described as a hybrid measure (a betweenness weighted closeness measure) rather than a normalisation, with the originators of the method stating that “*it seems to combine our two measures – depth and choice, to and through-movement*” [42]. The measure is ordinarily shown with the addition of constants in both the numerator and denominator to guard against situations such as taking the *log* of a number less than 1.

10.1.3 Length-weighting

Topological distortions due to varying intensities of nodes can exaggerate the outcome of centrality measures. It is therefore common to see the nodes weighted by factors such as street lengths [21, 16] or the number of adjacent buildings [48]. The implication is that greater exposure to street lengths offers a greater potential for interaction, with the contribution of unusually high concentrations of nodes tempered by correspondingly shorter street segments and vice-versa. In practical terms, assuming the use of the dual representation, the length of the primal street segment is assigned to the dual node, with the nodes weighted accordingly during the calculation of the centrality measure. For example, in the case of *Harmonic Closeness* the 1 in the numerator is replaced by the length of the segment l :

$$\text{HC}_{(i)} = \sum_{j \neq i} \frac{l}{d_{(i,j)}}. \quad (12)$$

An alternative is to use the integral of *Harmonic Closeness*. The integral for $f(x) = 1/d$ takes the form

$$\int_a^b f(x) dx = \ln(b) - \ln(a) \quad (13)$$

and can be applied to sum the ‘area under the (spatial impedance) curve’ for the respective lower and upper segment bounds a and b for all reachable segments S :

$$\text{HC}_{(i)} = \sum_{(a,b)}^S \int_a^b f(x) dx = \sum_{(a,b)}^S \ln(b) - \ln(a). \quad (14)$$

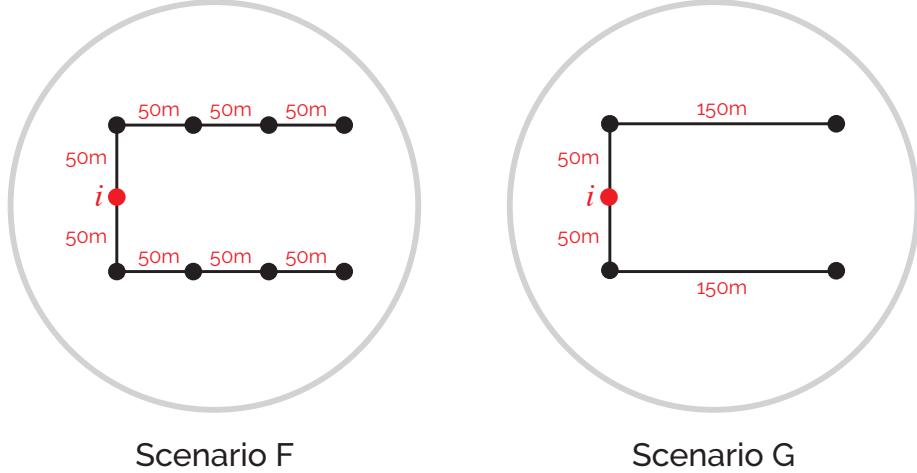


Figure S2: Length weighted closeness formulations comparing for varied node intensities.

	$length\ wt.\ HC_{(i)} = \sum_{j \neq i} \frac{l}{d_{(i,j)}}$	$Harmonic\ C_{(i)}(0, x) = \int_0^x \ln(x) dx$
Scenario F	$(\frac{50}{50} + \frac{50}{100} + \frac{50}{150} + \frac{25}{200}) * 2 = 3.92$	(simplified) $\ln(200) * 2 = 10.60$
Scenario G	$(\frac{100}{50} + \frac{75}{200}) * 2 = 4.75$	(simplified) $\ln(200) * 2 = 10.60$

Table 3: Length weighted closeness formulations comparing street-length weighted *Harmonic Closeness* and a continuous form of *Harmonic Closeness*. Continuous forms remain consistent regardless of the number of subdivisions.

This allows spatial impedances to increase continuously. For example, the contribution of a 10m street segment adjacent to the origin is now found as $\ln(10) = 2.303$ and a segment from 10m to 20m distant is found as $\ln(20) - \ln(10) = 0.693$. As shown in Figure S2 and Table 3, the continuous form remains consistent regardless of how many times street lengths are split at intervening nodes.

Note that the continuous form of *Harmonic Closeness* suggested per Equation 14 cannot be used with *geometric distance* (angular) impedances, which do not increase continuously.

As with closeness centrality, it may be preferable to use the gravity index in a continuous form. In this case, gravity is computed as the area under the curve for all reachable segments S for the respective lower and upper segment bounds a and b at the specified impedance β :

$$G_{(i)} = \sum_{(a,b)}^S \int_a^b f(x) dx = \sum_{(a,b)}^S \frac{\exp(-\beta \cdot b) - \exp(-\beta \cdot a)}{-\beta}. \quad (15)$$

10.2 Additional Figures and Tables

10.2.1 Descriptive Statistics Tables

Note on closeness centrality values: The closeness centrality measures exhibit small absolute values because they are computed as the inverse of *Farness* or *Normalised Farness* where the number of reachable nodes is divided by the total cumulative distance (in meters). For example, with a 500m catchment containing 100 nodes and 28,000 meters of cumulative farness, *Closeness* gives $1/28,000 = 0.000036$ and *Normalised Closeness* $100/28,000 = 0.00357$. These small values are mathematically correct and do not affect statistical analyses because Spearman rank correlations are based on rank order.

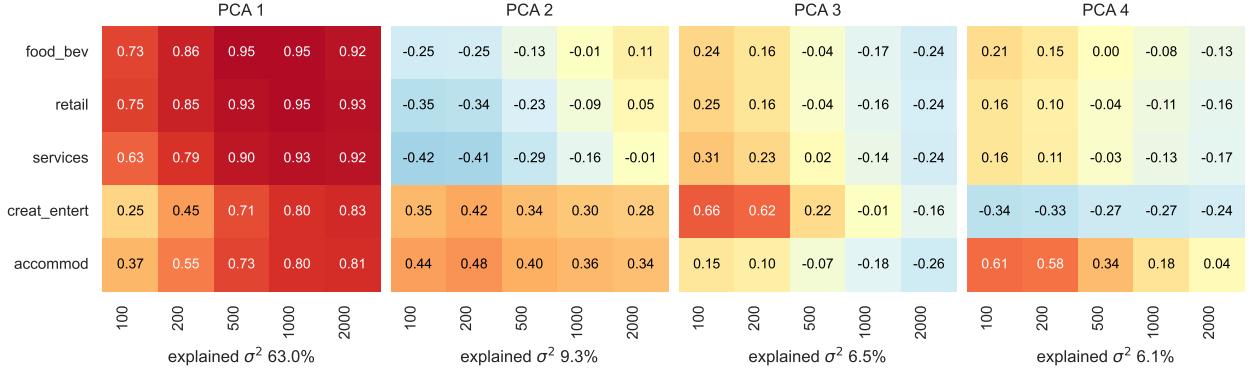


Figure S3: PCA loadings matrix.

Table 4: Descriptive statistics for land use variables.

	N	Mean	Median	Q1	Q3	IQR	Min	Max
pca_1	42167	-6.89e-04	-0.439	-3.177	2.809	5.986	-5.900	10.000
food_bev_200	42167	0.925	0.265	0.00e+00	1.134	1.134	0.00e+00	21
retail_200	42167	1.997	0.461	0.00e+00	2.328	2.328	0.00e+00	88

10.2.2 Zone-Level Descriptive Statistics

The following tables present descriptive statistics for the zone-level analysis, where segment centralities have been averaged to origin-destination travel zones. Trip counts represent the number of trips originating from or destined to each zone; trip numbers are normalised by zone area (km^2).

10.2.3 Spatial Autocorrelation and Bootstrap Results

The following tables present results from the spatial autocorrelation analysis. Moran's I is computed using k -nearest neighbour weights, where k is derived from the median network density at each distance threshold. The effective sample size (N_{eff}) accounts for spatial dependence using the approximation $N_{\text{eff}} \approx N(1 - I)/(1 + I)$. Block bootstrap confidence intervals (95%) are computed by resampling 100 spatial clusters with replacement (1,000 iterations) to preserve local spatial structure.

Table 5: Descriptive statistics for centrality measures.

	N	Mean	Median	Q1	Q3	IQR	Min	Max
density_500	42167	100	86	49	144	95	0.00e+00	367
density_1000	42167	352	330	192	496	304	0.00e+00	1092
density_2000	42167	1242	1199	796	1771	975	0.00e+00	2614
density_5000	42167	7078	7161	4790	9549	4759	0.00e+00	13191
density_10000	42167	25667	26803	21574	30862	9288	0.00e+00	36199
far_500	42167	32418	28121	15814	47338	31523	0.00e+00	122029
far_1000	42167	227738	216367	123304	316214	192910	0.00e+00	721591
far_2000	42167	1603791	1554475	978723	2329708	1350985	0.00e+00	3296895
far_5000	42167	23277712	23682352	15904454	30917964	15013510	0.00e+00	44062764
far_10000	42167	165974864	174153936	144394816	194249440	49854624	0.00e+00	223477152
far_norm_500	41537	326	328	310	344	34	79	500
far_norm_1000	42008	647	652	613	684	72	223	1000
far_norm_2000	42147	1289	1296	1213	1378	165	532	1987
far_norm_5000	42163	3295	3283	3168	3421	253	1422	4627
far_norm_10000	42163	6552	6510	6204	6817	614	5290	9101
closeness_500	41537	8.05e-05	3.49e-05	2.10e-05	6.12e-05	4.02e-05	8.19e-06	6.97e-03
closeness_1000	42167	1.40e-05	4.59e-06	3.14e-06	8.01e-06	4.86e-06	0.00e+00	2.67e-03
closeness_2000	42147	1.95e-06	6.43e-07	4.29e-07	1.02e-06	5.92e-07	3.03e-07	5.77e-04
closeness_5000	42163	9.91e-08	4.22e-08	3.23e-08	6.29e-08	3.05e-08	2.27e-08	1.05e-04
closeness_10000	42163	7.50e-09	5.74e-09	5.15e-09	6.92e-09	1.78e-09	4.47e-09	1.60e-06
close_N1_500	41537	3.10e-03	3.05e-03	2.91e-03	3.23e-03	3.17e-04	2.00e-03	1.27e-02
close_N1_1000	42167	1.56e-03	1.53e-03	1.46e-03	1.63e-03	1.71e-04	0.00e+00	4.49e-03
close_N1_2000	42147	7.85e-04	7.72e-04	7.26e-04	8.25e-04	9.90e-05	5.03e-04	1.88e-03
close_N1_5000	42163	3.06e-04	3.05e-04	2.92e-04	3.16e-04	2.34e-05	2.16e-04	7.03e-04
close_N1_10000	42163	1.53e-04	1.54e-04	1.47e-04	1.61e-04	1.45e-05	1.10e-04	1.89e-04
close_N1_2_500	41537	7.48e-03	7.59e-03	6.72e-03	8.41e-03	1.69e-03	2.00e-03	1.46e-02
close_N1_2_1000	42167	4.85e-03	4.92e-03	4.38e-03	5.47e-03	1.09e-03	0.00e+00	1.01e-02
close_N1_2_2000	42147	3.16e-03	3.21e-03	2.90e-03	3.52e-03	6.23e-04	5.03e-04	4.98e-03
close_N1_2_5000	42163	1.76e-03	1.81e-03	1.65e-03	1.94e-03	2.91e-04	3.05e-04	2.15e-03
close_N1_2_10000	42163	1.16e-03	1.18e-03	1.09e-03	1.27e-03	1.80e-04	3.14e-04	1.37e-03
close_N2_500	41537	0.312	0.272	0.158	0.447	0.289	2.00e-03	1.172
close_N2_1000	42167	0.549	0.509	0.298	0.777	0.479	0.00e+00	1.728
close_N2_2000	42147	0.969	0.946	0.609	1.336	0.726	5.03e-04	2.264
close_N2_5000	42163	2.159	2.180	1.451	2.964	1.513	9.25e-04	3.954
close_N2_10000	42163	3.982	4.137	3.208	4.905	1.697	1.47e-02	6.007
harmonic_500	42167	0.407	0.354	0.201	0.583	0.382	0.00e+00	1.721
harmonic_1000	42167	0.747	0.685	0.408	1.059	0.651	0.00e+00	2.532
harmonic_2000	42167	1.343	1.315	0.834	1.838	1.004	0.00e+00	3.564
harmonic_5000	42167	3.011	3.025	2.071	4.110	2.038	0.00e+00	6.155
harmonic_10000	42167	5.524	5.730	4.291	6.952	2.661	0.00e+00	9.433
gravity_500	42167	12	10	5.901	17	11	0.00e+00	47
gravity_1000	42167	44	40	23	63	40	0.00e+00	154
gravity_2000	42167	157	152	93	216	124	0.00e+00	436
gravity_5000	42167	856	840	569	1189	621	0.00e+00	1688
gravity_10000	42167	3169	3252	2312	4135	1823	0.00e+00	5272

Table 6: Descriptive statistics for length-weighted centrality measures.

	N	Mean	Median	Q1	Q3	IQR	Min	Max
lw_density_500	42167	7348	7184	4856	9915	5059	0.00e+00	18311
lw_density_1000	42167	28119	28248	19274	36788	17514	0.00e+00	64442
lw_density_2000	42167	108908	108452	79721	146356	66635	0.00e+00	195076
lw_density_5000	42167	690422	713402	531238	875962	344724	0.00e+00	1141834
lw_density_10000	42167	2659032	2759148	2332424	3079680	747256	0.00e+00	3581994
lw_far_500	42167	2446879	2390021	1592962	3301552	1708590	0.00e+00	6266275
lw_far_1000	42167	18561794	18738480	12520390	24216358	11695968	0.00e+00	43358996
lw_far_2000	42167	144247104	144698080	104013600	193948736	89935136	0.00e+00	257134432
lw_far_5000	42167	2311166464	2394943744	1786027776	2902958592	1116930816	0.00e+00	3842761216
lw_far_10000	42167	17492131840	18184536064	15647161344	19928051712	4280890368	0.00e+00	22876850176
lw_far_norm_500	41537	334	334	321	346	24	79	500
lw_far_norm_1000	42008	661	662	635	685	49	345	1000
lw_far_norm_2000	42147	1322	1324	1274	1375	101	764	1987
lw_far_norm_5000	42163	3356	3342	3277	3434	157	2420	4625
lw_far_norm_10000	42163	6626	6579	6392	6821	430	6154	8590
lw_closeness_500	41537	6.98e-07	4.14e-07	3.01e-07	6.13e-07	3.11e-07	1.60e-07	4.33e-04
lw_closeness_1000	42008	9.47e-08	5.32e-08	4.12e-08	7.94e-08	3.81e-08	2.31e-08	1.99e-04
lw_closeness_2000	42147	1.06e-08	6.91e-09	5.16e-09	9.61e-09	4.45e-09	3.89e-09	1.42e-05
lw_closeness_5000	42163	5.55e-10	4.18e-10	3.44e-10	5.60e-10	2.15e-10	2.60e-10	7.09e-08
lw_closeness_10000	42163	6.26e-11	5.50e-11	5.02e-11	6.39e-11	1.37e-11	4.37e-11	2.27e-09
lw_close_N1_500	41537	3.01e-03	3.00e-03	2.89e-03	3.11e-03	2.19e-04	2.00e-03	1.27e-02
lw_close_N1_1000	42008	1.52e-03	1.51e-03	1.46e-03	1.57e-03	1.13e-04	1.00e-03	2.90e-03
lw_close_N1_2000	42147	7.60e-04	7.55e-04	7.27e-04	7.85e-04	5.79e-05	5.03e-04	1.31e-03
lw_close_N1_5000	42163	2.99e-04	2.99e-04	2.91e-04	3.05e-04	1.40e-05	2.16e-04	4.13e-04
lw_close_N1_10000	42163	1.51e-04	1.52e-04	1.47e-04	1.56e-04	9.85e-06	1.16e-04	1.62e-04
lw_close_N1.2_500	41537	1.75e-02	1.79e-02	1.65e-02	1.91e-02	2.65e-03	2.99e-03	3.71e-02
lw_close_N1.2_1000	42008	1.16e-02	1.18e-02	1.09e-02	1.26e-02	1.70e-03	1.41e-03	1.67e-02
lw_close_N1.2_2000	42147	7.59e-03	7.74e-03	7.16e-03	8.19e-03	1.03e-03	1.03e-03	9.17e-03
lw_close_N1.2_5000	42163	4.34e-03	4.43e-03	4.14e-03	4.69e-03	5.49e-04	1.24e-03	4.97e-03
lw_close_N1.2_10000	42163	2.90e-03	2.94e-03	2.78e-03	3.10e-03	3.13e-04	1.14e-03	3.28e-03
lw_close_N2_500	41537	22	22	15	30	15	1.47e-02	55
lw_close_N2_1000	42008	43	43	29	56	27	5.19e-03	96
lw_close_N2_2000	42147	83	83	60	109	49	1.78e-02	155
lw_close_N2_5000	42163	207	212	157	265	108	1.360	340
lw_close_N2_10000	42163	405	420	346	479	133	7.206	569
lw_harmonic_500	42167	28	27	18	38	19	0.00e+00	74
lw_harmonic_1000	42167	56	55	38	73	35	0.00e+00	132
lw_harmonic_2000	42167	109	111	80	142	62	0.00e+00	219
lw_harmonic_5000	42167	275	278	211	355	144	0.00e+00	463
lw_harmonic_10000	42167	540	557	453	647	194	0.00e+00	799
lw_gravity_500	42167	818	803	532	1112	580	0.00e+00	2166
lw_gravity_1000	42167	3273	3231	2231	4358	2127	0.00e+00	7872
lw_gravity_2000	42167	12701	12858	9125	16376	7251	0.00e+00	26269
lw_gravity_5000	42167	77914	78117	59244	102658	43414	0.00e+00	131299
lw_gravity_10000	42167	309890	319353	251567	383456	131889	0.00e+00	465867

Table 7: Descriptive statistics for angular centrality measures.

	N	Mean	Median	Q1	Q3	IQR	Min	Max
far_500_ang	42167	260	209	112	375	263	0.00e+00	1592
far_1000_ang	42167	1300	1144	627	1836	1208	0.00e+00	7665
far_2000_ang	42167	6462	6209	3887	8985	5098	0.00e+00	27812
far_5000_ang	42167	51751	53623	30026	70395	40369	0.00e+00	165195
far_10000_ang	42167	355383	372069	282992	429857	146865	0.00e+00	802588
far_norm_500_ang	41537	3.054	2.935	2.393	3.584	1.192	9.03e-03	12
far_norm_1000_ang	42008	4.665	4.555	3.794	5.393	1.600	5.06e-02	14
far_norm_2000_ang	42147	7.248	7.069	6.121	8.148	2.027	0.384	21
far_norm_5000_ang	42163	12	12	11	14	2.625	2.412	37
far_norm_10000_ang	42163	19	19	17	21	3.574	9.871	48
closeness_500_ang	41537	3.18e-02	4.71e-03	2.64e-03	8.64e-03	6.00e-03	6.28e-04	111
closeness_1000_ang	42167	4.85e-03	8.67e-04	5.41e-04	1.58e-03	1.04e-03	0.00e+00	20
closeness_2000_ang	42147	7.18e-04	1.61e-04	1.11e-04	2.57e-04	1.46e-04	3.60e-05	2.603
closeness_5000_ang	42163	4.67e-05	1.86e-05	1.42e-05	3.33e-05	1.91e-05	6.05e-06	8.80e-02
closeness_10000_ang	42163	3.64e-06	2.69e-06	2.33e-06	3.53e-06	1.21e-06	1.25e-06	5.11e-04
close_N1_500_ang	41537	0.385	0.341	0.279	0.418	0.139	8.33e-02	111
close_N1_1000_ang	42167	0.235	0.219	0.185	0.263	7.84e-02	0.00e+00	20
close_N1_2000_ang	42147	0.147	0.141	0.123	0.163	4.06e-02	4.86e-02	2.603
close_N1_5000_ang	42163	8.36e-02	8.29e-02	7.40e-02	9.19e-02	1.79e-02	2.67e-02	0.415
close_N1_10000_ang	42163	5.33e-02	5.28e-02	4.81e-02	5.80e-02	9.97e-03	2.06e-02	0.101
close_N1_2_500_ang	41537	0.851	0.817	0.654	0.986	0.332	0.147	111
close_N1_2_1000_ang	42167	0.677	0.670	0.548	0.794	0.245	0.00e+00	20
close_N1_2_2000_ang	42147	0.545	0.545	0.458	0.629	0.171	8.87e-02	2.603
close_N1_2_5000_ang	42163	0.433	0.436	0.371	0.497	0.126	9.17e-02	0.789
close_N1_2_10000_ang	42163	0.379	0.380	0.338	0.424	8.69e-02	8.04e-02	0.665
close_N2_500_ang	41537	28	26	14	40	25	0.212	111
close_N2_1000_ang	42167	60	57	32	83	51	0.00e+00	231
close_N2_2000_ang	42147	127	116	72	177	105	9.12e-02	378
close_N2_5000_ang	42163	358	343	187	510	322	0.382	933
close_N2_10000_ang	42163	1024	1023	775	1317	542	3.257	2332
harmonic_500_ang	42167	24	22	12	33	21	0.00e+00	90
harmonic_1000_ang	42167	56	53	31	78	47	0.00e+00	208
harmonic_2000_ang	42167	128	120	75	177	102	0.00e+00	369
harmonic_5000_ang	42167	380	366	211	533	322	0.00e+00	951
harmonic_10000_ang	42167	1122	1123	835	1434	598	0.00e+00	2542

Table 8: Descriptive statistics for length-weighted angular centrality measures.

	N	Mean	Median	Q1	Q3	IQR	Min	Max
lw_far_500_ang	42167	18826	17213	10848	25695	14847	0.00e+00	84449
lw_far_1000_ang	42167	101824	97161	63036	136294	73258	0.00e+00	470177
lw_far_2000_ang	42167	555098	555054	384851	733383	348533	0.00e+00	1902248
lw_far_5000_ang	42167	4896063	5050740	3259989	6327224	3067235	0.00e+00	14512521
lw_far_10000_ang	42167	35232836	36230860	29151620	41564344	12412724	0.00e+00	76790216
lw_far_norm_500_ang	41537	3.031	2.922	2.358	3.584	1.226	9.03e-03	12
lw_far_norm_1000_ang	42008	4.631	4.545	3.761	5.384	1.623	5.06e-02	14
lw_far_norm_2000_ang	42147	7.198	7.079	6.072	8.146	2.074	0.384	20
lw_far_norm_5000_ang	42163	12	12	11	13	2.699	2.341	32
lw_far_norm_10000_ang	42163	19	18	17	20	3.602	10	47
lw_closeness_500_ang	41537	2.38e-04	5.73e-05	3.87e-05	8.98e-05	5.11e-05	1.18e-05	0.773
lw_closeness_1000_ang	42008	2.93e-05	1.03e-05	7.33e-06	1.58e-05	8.44e-06	2.13e-06	9.80e-02
lw_closeness_2000_ang	42147	3.33e-06	1.80e-06	1.36e-06	2.60e-06	1.23e-06	5.26e-07	5.32e-03
lw_closeness_5000_ang	42163	2.80e-07	1.98e-07	1.58e-07	3.07e-07	1.49e-07	6.89e-08	5.61e-05
lw_closeness_10000_ang	42163	3.20e-08	2.76e-08	2.41e-08	3.43e-08	1.02e-08	1.30e-08	1.69e-06
lw_close_N1_500_ang	41537	0.391	0.342	0.279	0.424	0.145	8.58e-02	111
lw_close_N1_1000_ang	42008	0.239	0.220	0.186	0.266	8.02e-02	7.20e-02	20
lw_close_N1_2000_ang	42147	0.148	0.141	0.123	0.165	4.19e-02	5.08e-02	2.603
lw_close_N1_5000_ang	42163	8.57e-02	8.40e-02	7.52e-02	9.44e-02	1.92e-02	3.08e-02	0.427
lw_close_N1_10000_ang	42163	5.54e-02	5.46e-02	4.96e-02	6.04e-02	1.08e-02	2.15e-02	9.56e-02
lw_close_N1.2_500_ang	41537	2.113	1.964	1.582	2.405	0.823	0.398	299
lw_close_N1.2_1000_ang	42008	1.694	1.629	1.355	1.943	0.588	0.294	58
lw_close_N1.2_2000_ang	42147	1.368	1.345	1.141	1.561	0.420	0.211	8.980
lw_close_N1.2_5000_ang	42163	1.118	1.116	0.958	1.266	0.308	0.292	3.208
lw_close_N1.2_10000_ang	42163	0.998	0.989	0.881	1.105	0.224	0.309	1.749
lw_close_N2_500_ang	41537	2185	2114	1398	2879	1481	2.476	33649
lw_close_N2_1000_ang	42008	4919	4738	3209	6382	3173	2.548	17205
lw_close_N2_2000_ang	42147	11114	10363	7302	14569	7267	3.655	31067
lw_close_N2_5000_ang	42163	35165	33939	22710	47038	24328	558	91400
lw_close_N2_10000_ang	42163	108127	107119	85866	133944	48078	2525	239304
lw_harmonic_500_ang	42167	1775	1756	1173	2359	1186	0.00e+00	5308
lw_harmonic_1000_ang	42167	4547	4446	3060	5897	2836	0.00e+00	14975
lw_harmonic_2000_ang	42167	11221	10653	7638	14521	6883	0.00e+00	31891
lw_harmonic_5000_ang	42167	37311	36086	24622	49110	24487	0.00e+00	98912
lw_harmonic_10000_ang	42167	119072	117615	93104	146571	53467	0.00e+00	272356

Table 9: Descriptive statistics for zone-level outcome variables (trip counts and trips per area).

	N	Mean	Median	Q1	Q3	IQR	Min	Max
origin_count	625	114	106	64	155	91	1.000	444
dest_count	625	114	106	64	156	92	1.000	458
origin_by_area	625	396	333	126	528	402	7.08e-02	5318
dest_by_area	625	400	331	131	520	389	3.54e-02	5383

Table 10: Descriptive statistics for zone-averaged centrality measures.

	N	Mean	Median	Q1	Q3	IQR	Min	Max
density_500	625	82	70	41	113	71	0.333	303
density_1000	625	316	294	179	437	258	1.667	1009
density_2000	625	1235	1223	799	1740	941	4.333	2547
density_5000	625	7441	7682	5257	9798	4541	35	12708
density_10000	625	26321	27463	22581	31338	8757	632	36024
far_500	625	27013	23059	13407	36681	23274	81	100420
far_1000	625	208657	195084	117325	285472	168147	1205	637461
far_2000	625	1637380	1608266	1032858	2330804	1297946	5381	3166530
far_5000	625	24622024	25368328	17765866	32018626	14252760	118084	42565468
far_10000	625	169181600	177521104	149557632	196353424	46795792	5506382	221925904
far_norm_500	625	333	332	324	340	16	228	467
far_norm_1000	625	668	670	644	695	51	488	846
far_norm_2000	625	1338	1347	1270	1409	139	895	1739
far_norm_5000	625	3327	3309	3197	3453	256	2488	4063
far_norm_10000	625	6495	6444	6195	6730	536	5903	8175
closeness_500	625	1.33e-04	5.29e-05	3.06e-05	1.21e-04	9.06e-05	1.01e-05	4.13e-03
closeness_1000	625	1.84e-05	5.96e-06	3.71e-06	1.28e-05	9.10e-06	1.57e-06	5.56e-04
closeness_2000	625	1.89e-06	6.52e-07	4.37e-07	1.09e-06	6.57e-07	3.16e-07	1.26e-04
closeness_5000	625	7.46e-08	3.96e-08	3.13e-08	5.71e-08	2.58e-08	2.35e-08	5.67e-06
closeness_10000	625	7.21e-09	5.64e-09	5.09e-09	6.72e-09	1.62e-09	4.51e-09	3.38e-07
close_N1_500	625	3.05e-03	3.03e-03	2.95e-03	3.11e-03	1.57e-04	2.14e-03	5.09e-03
close_N1_1000	625	1.51e-03	1.50e-03	1.44e-03	1.56e-03	1.15e-04	1.20e-03	2.13e-03
close_N1_2000	625	7.55e-04	7.45e-04	7.11e-04	7.90e-04	7.81e-05	5.76e-04	1.15e-03
close_N1_5000	625	3.02e-04	3.02e-04	2.90e-04	3.13e-04	2.30e-05	2.47e-04	4.13e-04
close_N1_10000	625	1.54e-04	1.55e-04	1.49e-04	1.61e-04	1.28e-05	1.23e-04	1.70e-04
close_N1.2_500	625	6.98e-03	7.03e-03	6.32e-03	7.76e-03	1.45e-03	2.14e-03	9.68e-03
close_N1.2_1000	625	4.58e-03	4.62e-03	4.13e-03	5.13e-03	9.98e-04	1.65e-03	6.65e-03
close_N1.2_2000	625	3.04e-03	3.08e-03	2.79e-03	3.38e-03	5.90e-04	1.18e-03	4.14e-03
close_N1.2_5000	625	1.77e-03	1.82e-03	1.66e-03	1.94e-03	2.75e-04	6.46e-04	2.11e-03
close_N1.2_10000	625	1.18e-03	1.20e-03	1.11e-03	1.27e-03	1.64e-04	4.45e-04	1.36e-03
close_N2_500	625	0.253	0.215	0.132	0.341	0.209	2.14e-03	0.917
close_N2_1000	625	0.482	0.435	0.266	0.656	0.390	3.52e-03	1.600
close_N2_2000	625	0.938	0.928	0.590	1.274	0.684	5.24e-03	2.194
close_N2_5000	625	2.256	2.313	1.576	3.029	1.453	1.53e-02	3.794
close_N2_10000	625	4.108	4.253	3.387	5.008	1.621	0.109	5.920
harmonic_500	625	0.328	0.281	0.173	0.453	0.280	8.18e-04	1.189
harmonic_1000	625	0.640	0.575	0.360	0.853	0.493	2.99e-03	2.153
harmonic_2000	625	1.253	1.232	0.791	1.693	0.903	4.75e-03	3.212
harmonic_5000	625	3.032	3.066	2.195	4.058	1.863	1.33e-02	5.840
harmonic_10000	625	5.591	5.771	4.549	6.953	2.404	8.04e-02	9.100
gravity_500	625	9.658	8.153	5.057	13	8.306	1.28e-02	35
gravity_1000	625	37	33	20	50	30	0.178	130
gravity_2000	625	143	138	88	193	105	0.594	410
gravity_5000	625	865	861	611	1187	575	3.634	1658
gravity_10000	625	3272	3372	2483	4183	1700	28	5179
cycles_500	625	89	72	38	125	87	0.00e+00	346
cycles_1000	625	374	333	192	523	331	0.333	1234
cycles_2000	625	1522	1439	890	2226	1335	2.333	3367
cycles_5000	625	9357	9469	6286	12628	6342	21	16617
cycles_10000	625	32896	34557	28164	39556	11392	705	45055
betw_500	625	168	106	48	239	191	0.00e+00	1096
betw_1000	625	1338	981	458	1862	1404	0.00e+00	7960
betw_2000	625	10448	9034	4600	14875	10275	0.00e+00	39782
betw_5000	625	160638	147249	69251	226304	157053	0.00e+00	825275
betw_10000	625	1020000	807722	431800	1398694	966894	0.00e+00	7732552
betw_wt_500	625	11	6.850	3.121	16	13	0.00e+00	74
betw_wt_1000	625	95	65	30	131	101	0.00e+00	595
betw_wt_2000	625	753	624	303	1071	768	0.00e+00	3666
betw_wt_5000	625	11554	10881	5387	16402	11016	0.00e+00	43901
betw_wt_10000	625	82753	71446	35153	119982	84829	0.00e+00	518176
NACH_500	625	0.364	0.397	0.306	0.457	0.150	0.00e+00	0.585
NACH_1000	625	0.462	0.495	0.421	0.545	0.124	0.00e+00	0.644
NACH_2000	625	0.529	0.551	0.497	0.599	0.102	0.00e+00	0.821
NACH_5000	625	0.585	0.594	0.53 ³⁴	0.636	8.31e-02	0.00e+00	1.182
NACH_10000	625	0.605	0.608	0.568	0.648	7.91e-02	0.00e+00	2.231

Table 11: Descriptive statistics for zone-averaged angular centrality measures.

	N	Mean	Median	Q1	Q3	IQR	Min	Max
far_500_ang	625	209	169	100	293	194	0.173	843
far_1000_ang	625	1129	1004	582	1513	931	5.767	3778
far_2000_ang	625	6186	6027	3766	8519	4753	26	15161
far_5000_ang	625	53368	55935	32839	71294	38455	528	105116
far_10000_ang	625	359098	376124	300312	423326	123014	21766	563114
far_norm_500_ang	625	2.909	2.898	2.546	3.255	0.709	0.410	4.621
far_norm_1000_ang	625	4.514	4.474	3.965	5.019	1.054	1.295	7.673
far_norm_2000_ang	625	7.089	6.992	6.265	7.696	1.432	3.550	13
far_norm_5000_ang	625	12	12	11	13	1.931	8.553	20
far_norm_10000_ang	625	19	18	17	20	2.986	13	34
closeness_500_ang	625	9.77e-02	8.46e-03	4.41e-03	2.22e-02	1.78e-02	1.25e-03	23
closeness_1000_ang	625	7.59e-03	1.22e-03	7.39e-04	3.28e-03	2.54e-03	2.74e-04	0.437
closeness_2000_ang	625	7.36e-04	1.85e-04	1.25e-04	3.83e-04	2.58e-04	6.81e-05	2.65e-02
closeness_5000_ang	625	3.55e-05	1.85e-05	1.43e-05	3.23e-05	1.80e-05	9.56e-06	1.59e-03
closeness_10000_ang	625	3.40e-06	2.67e-06	2.37e-06	3.38e-06	1.01e-06	1.78e-06	9.74e-05
close_N1_500_ang	625	0.470	0.374	0.329	0.448	0.119	0.226	23
close_N1_1000_ang	625	0.249	0.237	0.209	0.268	5.90e-02	0.130	1.431
close_N1_2000_ang	625	0.150	0.148	0.134	0.164	2.99e-02	8.24e-02	0.290
close_N1_5000_ang	625	8.55e-02	8.55e-02	7.89e-02	9.27e-02	1.38e-02	5.33e-02	0.126
close_N1_10000_ang	625	5.52e-02	5.48e-02	5.08e-02	6.00e-02	9.23e-03	3.09e-02	7.70e-02
close_N1.2_500_ang	625	0.918	0.855	0.748	0.956	0.207	0.365	23
close_N1.2_1000_ang	625	0.687	0.687	0.593	0.767	0.174	0.287	1.896
close_N1.2_2000_ang	625	0.555	0.564	0.489	0.623	0.134	0.246	0.852
close_N1.2_5000_ang	625	0.450	0.461	0.396	0.511	0.114	0.143	0.616
close_N1.2_10000_ang	625	0.396	0.398	0.358	0.440	8.16e-02	0.106	0.565
close_N2_500_ang	625	24	22	14	33	19	0.400	74
close_N2_1000_ang	625	55	52	31	76	45	0.652	153
close_N2_2000_ang	625	127	121	74	178	104	1.090	298
close_N2_5000_ang	625	389	383	229	544	315	3.348	797
close_N2_10000_ang	625	1099	1104	887	1388	501	21	1928
harmonic_500_ang	625	20	19	11	27	16	9.53e-02	60
harmonic_1000_ang	625	51	48	30	70	40	0.383	141
harmonic_2000_ang	625	128	121	76	176	100	0.718	291
harmonic_5000_ang	625	410	402	250	566	316	2.496	820
harmonic_10000_ang	625	1204	1222	969	1517	548	15	2067
betw_500_ang	625	134	88	39	187	148	0.00e+00	741
betw_1000_ang	625	990	736	319	1338	1018	0.00e+00	5380
betw_2000_ang	625	6974	5620	2608	9703	7095	0.00e+00	31342
betw_5000_ang	625	85320	57811	26618	124681	98063	0.00e+00	517862
betw_10000_ang	625	653772	413533	190716	823699	632983	0.00e+00	6203538
NACH_500_ang	625	0.656	0.726	0.589	0.812	0.223	0.00e+00	0.933
NACH_1000_ang	625	0.749	0.797	0.707	0.864	0.157	0.00e+00	1.059
NACH_2000_ang	625	0.789	0.817	0.749	0.876	0.127	0.00e+00	1.123
NACH_5000_ang	625	0.810	0.818	0.761	0.870	0.110	0.00e+00	1.381
NACH_10000_ang	625	0.826	0.824	0.774	0.877	0.104	0.00e+00	2.341

Table 12: Moran's I spatial autocorrelation for centrality measures. Higher values indicate stronger positive spatial clustering; k is the number of nearest neighbours used in the weights matrix.

Variable	Distance (m)	k (median density)	Moran's I	z-score	p (analytic)	p (permutation)
closeness_500	500	86	0.2095	294.77	0	0.001
close_N1_500	500	86	0.1800	253.26	0	0.001
close_N1.2_500	500	86	0.5277	742.31	0	0.001
close_N2_500	500	86	0.8398	1181.44	0	0.001
harmonic_500	500	86	0.7971	1121.34	0	0.001
gravity_500	500	86	0.7855	1105.01	0	0.001
betw_500	500	86	0.4641	652.90	0	0.001
betw_wt_500	500	86	0.4695	660.57	0	0.001
NACH_500	500	86	0.2848	400.63	0	0.001
closeness_1000	1000	330	0.1281	354.90	0	0.001
close_N1_1000	1000	330	0.1877	519.97	0	0.001
close_N1.2_1000	1000	330	0.4227	1170.93	0	0.001
close_N2_1000	1000	330	0.7933	2197.63	0	0.001
harmonic_1000	1000	330	0.7444	2062.24	0	0.001
gravity_1000	1000	330	0.7216	1998.93	0	0.001
betw_1000	1000	330	0.3373	934.31	0	0.001
betw_wt_1000	1000	330	0.3583	992.73	0	0.001
NACH_1000	1000	330	0.1367	378.82	0	0.001
closeness_2000	2000	1199	0.0641	342.97	0	0.001
close_N1_2000	2000	1199	0.1474	788.16	0	0.001
close_N1.2_2000	2000	1199	0.3487	1863.65	0	0.001
close_N2_2000	2000	1199	0.7303	3903.16	0	0.001
harmonic_2000	2000	1199	0.6657	3558.00	0	0.001
gravity_2000	2000	1199	0.6271	3351.90	0	0.001
betw_2000	2000	1199	0.1814	969.47	0	0.001
betw_wt_2000	2000	1199	0.2107	1126.41	0	0.001
NACH_2000	2000	1199	0.0610	325.95	0	0.001

Table 13: Effective sample size (N_{eff}) by centrality measure and distance threshold. Well-normalised measures with high spatial coherence have lower N_{eff} due to positive autocorrelation.

Variable	Distance (m)	Moran's I	N_{eff}	N_{eff}/N
closeness_500	500	0.2095	27,558	65.4%
close_N1_500	500	0.1800	29,301	69.5%
close_N1.2_500	500	0.5277	13,037	30.9%
close_N2_500	500	0.8398	3,671	8.7%
harmonic_500	500	0.7971	4,761	11.3%
gravity_500	500	0.7855	5,066	12.0%
betw_500	500	0.4641	15,434	36.6%
betw_wt_500	500	0.4695	15,220	36.1%
NACH_500	500	0.2848	23,474	55.7%
closeness_1000	1000	0.1281	32,591	77.3%
close_N1_1000	1000	0.1877	28,840	68.4%
close_N1.2_1000	1000	0.4227	17,111	40.6%
close_N2_1000	1000	0.7933	4,860	11.5%
harmonic_1000	1000	0.7444	6,178	14.7%
gravity_1000	1000	0.7216	6,819	16.2%
betw_1000	1000	0.3373	20,898	49.6%
betw_wt_1000	1000	0.3583	19,919	47.2%
NACH_1000	1000	0.1367	32,023	75.9%
closeness_2000	2000	0.0641	37,083	87.9%
close_N1_2000	2000	0.1474	31,330	74.3%
close_N1.2_2000	2000	0.3487	20,364	48.3%
close_N2_2000	2000	0.7303	6,573	15.6%
harmonic_2000	2000	0.6657	8,463	20.1%
gravity_2000	2000	0.6271	9,662	22.9%
betw_2000	2000	0.1814	29,219	69.3%
betw_wt_2000	2000	0.2107	27,488	65.2%
NACH_2000	2000	0.0610	37,321	88.5%

Table 14: Block bootstrap Spearman correlations (ρ) between centrality measures and PCA 1 (landuse intensity), with 95% confidence intervals. Negative correlations for unnormalized *Closeness* indicate pathological behaviour; near-zero correlations for *Normalised Closeness* indicate ineffective normalisation.

Variable	Distance (m)	ρ	CI Low	CI High	Moran's I	Neff
closeness_500	500	-0.704	-0.756	-0.641	0.2095	27,558
close_N1_500	500	-0.114	-0.158	-0.066	0.1800	29,301
close_N1.2_500	500	0.503	0.425	0.565	0.5277	13,037
close_N2_500	500	0.667	0.597	0.724	0.8398	3,671
harmonic_500	500	0.646	0.574	0.704	0.7971	4,761
gravity_500	500	0.625	0.553	0.685	0.7855	5,066
betw_500	500	0.576	0.524	0.621	0.4641	15,434
betw_wt_500	500	0.547	0.490	0.594	0.4695	15,220
NACH_500	500	0.516	0.467	0.560	0.2848	23,474
closeness_1000	1000	-0.759	-0.804	-0.704	0.1281	32,591
close_N1_1000	1000	-0.002	-0.072	0.085	0.1877	28,840
close_N1.2_1000	1000	0.613	0.537	0.675	0.4227	17,111
close_N2_1000	1000	0.771	0.712	0.821	0.7933	4,860
harmonic_1000	1000	0.744	0.682	0.797	0.7444	6,178
gravity_1000	1000	0.729	0.666	0.781	0.7216	6,819
betw_1000	1000	0.591	0.548	0.631	0.3373	20,898
betw_wt_1000	1000	0.590	0.543	0.633	0.3583	19,919
NACH_1000	1000	0.502	0.462	0.538	0.1367	32,023
closeness_2000	2000	-0.770	-0.820	-0.705	0.0641	37,083
close_N1_2000	2000	0.148	0.054	0.242	0.1474	31,330
close_N1.2_2000	2000	0.694	0.618	0.752	0.3487	20,364
close_N2_2000	2000	0.825	0.777	0.864	0.7303	6,573
harmonic_2000	2000	0.825	0.776	0.866	0.6657	8,463
gravity_2000	2000	0.816	0.764	0.858	0.6271	9,662
betw_2000	2000	0.536	0.492	0.578	0.1814	29,219
betw_wt_2000	2000	0.576	0.532	0.617	0.2107	27,488
NACH_2000	2000	0.438	0.402	0.472	0.0610	37,321