sDNA: 3-d spatial network analysis for GIS, CAD, Command Line & Python

Crispin H.V. Cooper [a,*], Alain J.F. Chiaradia [b]

[a] *Sustainable Places Research Institute, Cardiff University, 33 Park Place, Cardiff, UK*
[b] *Faculty of Architecture, Department of Urban Planning and Design, 4/F, Knowles Building, The University of Hong Kong, Pokfulam Road, Hong Kong*

ARTICLE INFO

ABSTRACT

Spatial Design Network Analysis (sDNA) is a toolbox for 3-d spatial network analysis, especially street/path/urban network analysis, motivated by a need to use network links as the principal unit of analysis in order to analyse existing network data. sDNA is usable from QGIS & ArcGIS geographic information systems, AutoCAD, the command line, and via its own Python API. It computes measures of accessibility (reach, mean distance/closeness centrality, gravity), flows (bidirectional betweenness centrality) and efficiency (circuity) as well as convex hull properties, localised within lower- and upper-bounded radial bands. Weighting is flexible and can make use of geometric properties, data attached to links, zones, matrices or combinations of the above. Motivated by a desire to base network analysis on route choice and spatial cognition, the definition of distance can be network-Euclidean, angular, a mixture of both, custom, or specific to cyclists (avoiding slope and motorised traffic). In addition to statistics on network links, the following outputs can be computed: geodesics, network buffers, accessibility maps, convex hulls, flow bundles and skim matrices. Further tools assist with network preparation and calibration of network models to observed data.

To date, sDNA has been used mainly for urban network analysis both by academics and city planners/engineers, for tasks including prediction of pedestrian, cyclist, vehicle and metro flows and mode choice; also quantification of the built environment for epidemiology and urban planning & design.

Code metadata

Current code version	4.0.4
Permanent link to code/repository used for this code version	https://github.com/ElsevierSoftwareX/SOFTX_2019_336
Legal Code Licence	GNU General Public License version 3
Code versioning system used	git
Software code languages, tools, and services used	C++/OpenMP, Python, R, Visual Lisp, `shapefile.py` [1], GEOS [2], muparser [3], glmnet [4]
Compilation requirements, operating environments & dependencies	Windows XP, 7, 8, 10; Visual C++ 2008, Python 2.7, Boost C++ libraries [5], Advanced Installer, Sphinx.
If available Link to developer documentation/manual	https://sdna-open.readthedocs.io/en/latest/
Support email for questions	GitHub issue tracker (link above)

1. Motivation and significance

Spatial Network Analysis is the special case of network analysis in which nodes have positions in physical space and optionally, links between them have geometry. Within the literature on transportation and urban design, the term is used interchangeably with street/path/urban network analysis; in this paper we

* Corresponding author.
E-mail address: cooperch@cardiff.ac.uk (C.H.V. Cooper).

follow that tradition but do not exclude other applications. Current spatial network analysis makes extensive use of closeness, betweenness and reach measures taken from network analysis [6–12]. Despite widespread use, however, limited consideration has been given to how spatial network analysis varies from the parent discipline of network analysis. Aspatial measures are 'spatialised' in four ways: firstly, by using geometric properties of links to compute the network distances, howsoever defined, upon which these measures depend. Secondly, by restricting computation of measures to a local spatial neighbourhood surrounding each point (see [13] for further discussion of this technique). Thirdly, additional network measures can explicitly incorporate space, for example, straightness/circuity [14,15]. Finally, it is common to focus on the link rather than node as primary unit of analysis, as links (being 1- rather than 0-dimensional) in some sense occupy more space: this is physically true in the case of many real world networks e.g. road systems. This gives rise to a dual representation [16] in which links (or parts thereof) are encoded as nodes, and nodes as links between them.

Broadly speaking, the measures produced in spatial network analysis quantify either accessibility, flow or efficiency; many of these are also considered to be measures of *centrality*. Such measures can be used to predict multiple phenomena including but not limited to transport behaviours, land use change and health (see Section 4). The discipline suffers from a lack of common terminology to describe the same underlying concepts, so we define terms here:

- In the urban planning domain, a large volume of spatial network analysis is labelled as *space syntax* following [17]. The two terms are not quite congruent as the latter tradition focuses on analysis of isovists, axial and convex space; however these objects are often processed as spatial networks, hence the overlap of interests.
- In modern spatial network analysis, local spatial neighbourhoods are typically used to restrict the scope of analysis around each link. These can variously be interpreted as catchment areas, maximum trip lengths or network buffers; in sDNA we adopt the term *radius* to refer both to the distance defining the size of the neighbourhood, and the neighbourhood itself.
- The quantity of network within a given radius is an important measure of accessibility referred to as *reach*, or sometimes density (whether or not divided by radius length or area).
- Network distance is often referred to as weight, impedance, cost, or in the space syntax literature, depth. We prefer to reserve 'weight' for importance of origins and destinations, avoid 'impedance' and 'depth' as the physical analogies are not accurate, and avoid 'cost' except for financial costs e.g. of fuel or public transport. Instead, we describe *distance* between two points, noting that distance can be defined in many ways depending on the choice of *metric*. We distinguish between an *analytical metric*, defining the measures computed in analysis, with a *radial metric* which defines the locality of analysis.
- Common metrics include *crow-flight-Euclidean* (straight line distance between two points), *network-Euclidean* (Euclidean distance measured along the network), and *angular* (cumulative absolute directional change along the route). The literature is sometimes ambiguous on whether Euclidean refers to network- or crow-flight-Euclidean; unfortunately we have been guilty of this ourselves so to set the record straight, where sDNA literature refers to Euclidean metrics this means network-Euclidean unless otherwise stated. Space syntax literature also uses the term 'metric' to refer

exclusively to the network-Euclidean metric, so it bears emphasizing that sDNA differs in this respect, using 'metric' in the mathematical sense of *any* definition of distance.

- *Closeness* measures accessibility, usually the inverse mean distance of a given link or node on the network to all other links/nodes within a localised radius. Space syntax literature also uses the term *integration*, although this is not unambiguous, as integration and closeness have each been defined both as inverse sum of distances and as inverse mean. Although the original formulation of closeness [18] used the sum formulation, sum of distance does not allow valid accessibility comparisons between networks of varying size, or localised analysis on a single network with varying density, hence the use of inverse mean in more recent work. Inverse sums are still used in some definitions of integration but it is likely an error to interpret them as a measure of accessibility e.g. [19].
- *Gravity models* combine reach and mean distance into a single measure of accessibility. If reach measures quantity of access, and mean distance measures quality of access, gravity models measure both (with at least one parameter required to specify the relative importance of each).
- *Betweenness* is the sum total of shortest paths from everywhere to everywhere (possibly subject to a weighting and/or maximum trip length) which traverse a given link or node. In space syntax literature this is known as *choice*.
- Most analysis relies on shortest paths through the network, also known as *geodesics* which vary with the choice of metric.

Comparable open source tools for spatial network analysis include:

- MIT Urban Network Analysis (UNA) [20] plugin for ArcGIS and RhinoCAD. The ArcGIS version of UNA requires the proprietary ESRI Network Analyst, in addition to the base ESRI ArcGIS software [21], in order to run. UNA does not actually compute statistics for network links, but rather for a point or polygon buildings layer with a 2-d network used to deduce connections between these. It computes reach, gravity, betweenness, closeness, straightness, redundancy index and paths, and the wayfinding index; all of which can be based on a crow-flight-Euclidean as well as network-Euclidean radius.
- DepthmapX [22] a stand-alone space syntax tool, originally intended for analysis of axial lines in urban spaces and since updated to road centre lines by dividing each link into one or more straight line segments. It computes 2-d depth, integration and choice within a network-Euclidean or topological radius. To overcome the problem that choice of segments to use to represent a link is arbitrary, road network analysis including DepthmapX has often used link segments weighted by length [23], however given evidence that link density is also important in urban networks [24], there is a case to be made for software which treats links as primary both in weighting and analysis [25].
- the Place Syntax Tool (PST) [26] computes reach, closeness and betweenness using network-Euclidean, angular or other metrics, within a radius defined in the same way or alternatively using a crow-flight-Euclidean radius, for either straight link segments or optionally a point layer representing origins/destinations.
- the Spatialist Reach Analysis toolkit [27,28] computes various measures of reach quantified by length.

Both the network theory and software of sDNA differ from competing approaches by treating network links (or user chosen
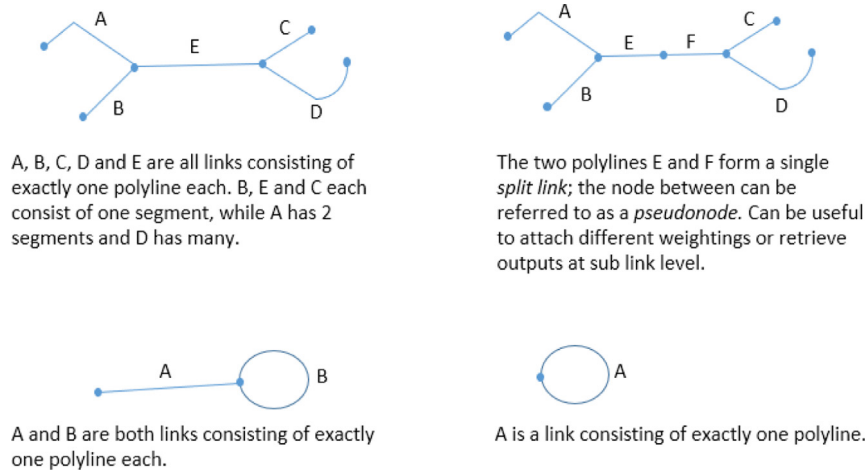
Fig. 1. Distinction between Links and Polylines in sDNA. Dots illustrate where polylines end for the purposes of this diagram. In practice sDNA does not need a point (node) layer but only requires polylines as inputs. No lines in this diagram overlap.
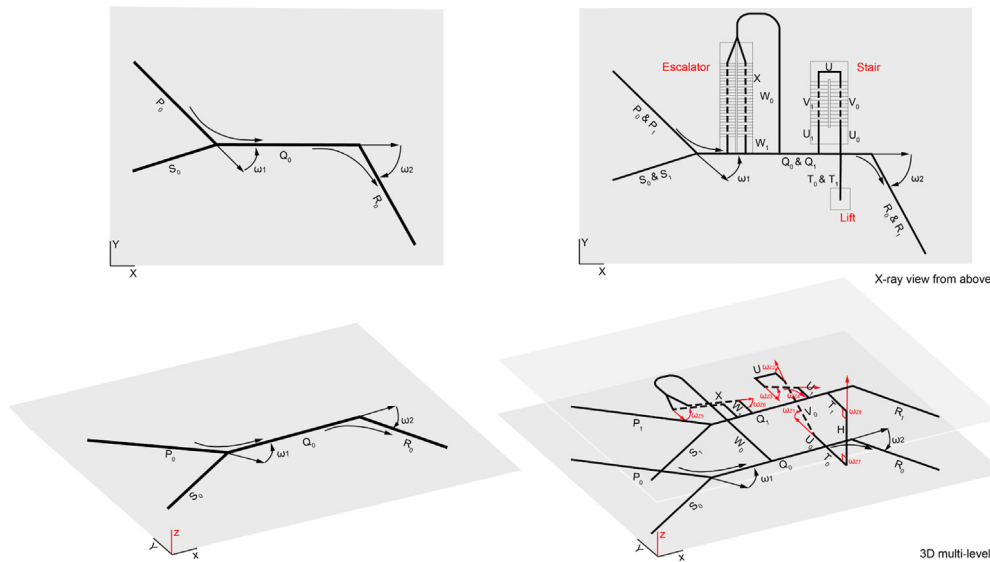

Fig. 2. Illustrating the difference between 2-d and 3-d turn angles. Left, links $P_0Q_0R_0S_0$ form turn angles ω_1, ω_2. Right, adding a second level with links $P_1Q_1R_1S_1$ and connecting lift HT_0T_1, stair $UU_0U_1V_0V_1$ and escalators W_0W_1X, these form 3-d turn angles $\omega_{z1}...\omega_{z8}$.

subdivisions thereof) – rather than straight link segments, or buildings – as the primary spatial unit of analysis. We define a network node as any point where it is possible to travel in 3 or more directions, and a link as the connection between two adjacent nodes, or between a single node and a dead end (Fig. 1). Contrasting other approaches, a curved link which would be treated as a large number of individual straight line segments in DepthmapX or PST, is handled as a single unit in sDNA. Reducing the number of units in this manner can lead to large increases in speed on road networks, as the time complexity of betweenness computation is proportional to the square of network density.

An additional motivation driving development of sDNA is the desire to base network analysis on route choice and spatial cognition grounded in human behaviour [29,30]. This is reflected in a wide variety of distance metrics available to use in sDNA (with even more available in its proprietary relative sDNA+ [31]).

2. Software description

sDNA is an ArcGIS/QGIS/AutoCAD plugin and Python/ command line tool for 2-d and 3-d network analysis. sDNA computes various measures of reach, including weight, junction and link count, and length; mean distance (the inverse of closeness), bidirectional betweenness (weighted either by product or a variant of the Huff [32] model without distance decay, as the latter can be handled using multiple radii), circuity [14,15] and geometric properties of the convex hull of the radius. Crucially, these can all be localised within lower- and upper-bound network-Euclidean radii and weighted by user defined expressions based on link geometry, zoning systems, origin–destination matrices or network attached data (the latter giving a means to import building data via GIS join). Once a localised network around each link has been defined, different distance metrics are available for its analysis: angular (least directional change), topological (fewest junctions), custom, cyclist-specific (based on slope and aversion to motorised traffic) and an angular-Euclidean mixture. (A greater variety of distance metrics and radius types, including 'hybrid' metrics based on user defined expressions incorporating angular change, height change, distance and custom data, and Monte Carlo randomization to handle individual preferences and analysis of regular grids, are available in the related proprietary software sDNA+ [31]. Networks can be read in 3-d, and turn angles are measured in 3-d so include turns 'uphill' and 'downhill' (Fig. 2); this is necessary in 3-d analysis otherwise it is possible
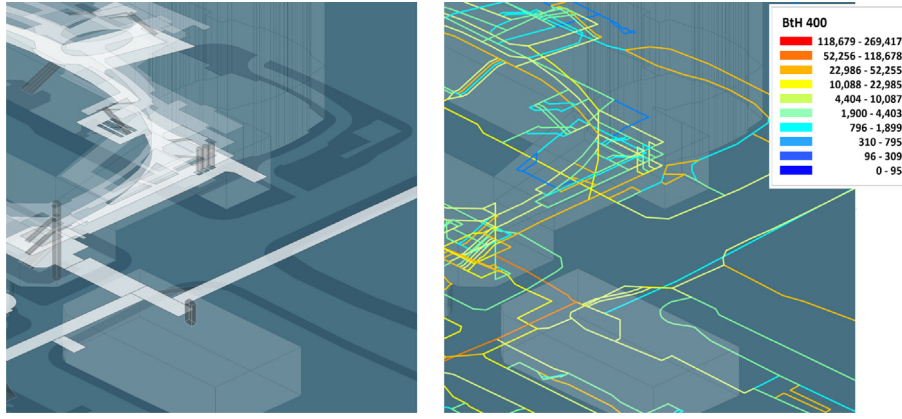
Fig. 3. sDNA used in 3-d pedestrian network analysis. Left, 3-d building model. Right, corresponding 3-d network model showing angular-Euclidean betweenness radius 400 m.

for a geodesic to 'cheat' angular analysis by turning in the 3rd dimension without accumulating angular distance. Where analytical geodesics exceed the size of the radius, sDNA can detect and correct 'problem routes' [13]. sDNA also includes tools for network preparation, provides a user-friendly means to use R for statistical inference and prediction from network data, and can output shapes of geodesics, accessibility maps, network buffers, convex hulls and flow bundles in addition to networks and skim matrices.

In practice, each link in a network is represented by one or more polyline objects in GIS or CAD (Fig. 1); our guiding principle is to compute statistics for the midpoint of each polyline and take them to be representative of the polyline (with midpoint defined by Euclidean or angular metric as appropriate). Where polylines form only a partial link, two approaches are available to the user: (1) to combine these into a single polyline using sDNA's Prepare tool; (2) to split the polyline into multiple parts as desired and compute statistics for each part separately, giving increased spatial accuracy albeit with more computational cost. If, in the latter case, the user is still interested in weighting the analysis by link, the "Link" weighting type will automatically apply partial weighting to partial links.

A further feature currently unique to sDNA is the extension of dual representation [16] by considering each *direction* on a link as a separate node, thereby allowing links with asymmetric distance metrics. At urban and wider scale this is useful for one way streets, at smaller pedestrian scale this feature can be used for escalators (Fig. 3) [33].

The issue arises of how to handle polylines which fall only partly within the radius of analysis. sDNA offers two modes [25]:

- Discrete space, in which the entire polyline is included/excluded depending on whether its midpoint falls within the radius;
- Continuous space, in which a partial polyline is included up to the point where the radius cuts the polyline, with metric distance along the polyline reduced as appropriate (partial polyline geometry is computed), and weighting scaled down as a function of partial length. This requires slightly more compute time, but is particularly useful in handling cases where long polylines would, in discrete space mode, create significant irregularities in analysis of small radii.

A further consequence of polylines occupying more than a single point in space, is that they will exhibit self-closeness and self-betweenness; properties which do not arise in aspatial networks [29]. The former arises because the average distance between all pairs of points on the polyline is nonzero; for a polyline

of length L the distance is in fact

$$\frac{1}{L^2} \int_{x=0}^{L} \int_{y=0}^{L} |x - y| \, dy \, dx = \frac{2}{L^2} \int_{x=0}^{L} \int_{y=x}^{L} y - x \, dy \, dx = \frac{1}{3}L \quad (1)$$

Self-betweenness of a polyline is derived by considering the mean betweenness of all points x on the line of length L, with origin weight W_o and destination weight W_d both distributed evenly over its length. Each point x is passed in one direction only by geodesics from origins $o : 0 < o < x$ with total weight $\frac{x}{L}W_o$, to destinations $d : x < d < L$ with total weight $\frac{L-x}{L}W_d$:

$$\frac{1}{L} \int_{x=0}^{L} (\frac{x}{L}W_o)(\frac{L-x}{L}W_d) \, dx = \frac{1}{6}W_oW_d \quad (2)$$

The self-betweenness contribution in both directions is therefore $\frac{1}{6}W_oW_d + \frac{1}{6}W_dW_o = \frac{1}{3}W_oW_d$. Additionally, the polylines which form the endpoints of a geodesic will themselves experience some degree of betweenness: on average, considering geodesics from all points on the polyline to one of its ends, an arbitrary point on the line will be passed by half of them.

sDNA includes a network preparation tool which can detect and/or fix the following, while preserving data attached to the network: (1) near miss connections, where link endpoints are almost but not quite coincident. This is essential for interface with some GIS which will display differing points as coincident if they fall below tolerance settings; (2) polylines which form only a partial link; (3) disconnected portions of the network; (4) duplicate links; (5) traffic islands (intended to correct features in some road data which encode spurious angular change).

Following feedback from urban design/transport partners, sDNA was updated to incorporate the Learn and Predict statistical modelling tools allowing for creation of predictive models based on network characteristics. As accessibility statistics are often strongly cross-correlated, a regression method robust to multicollinearity is required; for this reason ridge regression is employed with n-fold cross-validation to tune the ridge penalty, and bootstrapping to increase stability of results. A range of weighting schemes allow minimization of absolute or relative error terms, or a mixture of both. Model fit is reported both using weighted r^2 and the GEH statistic popular in transport planning ([34], section 3.2.7).

Table 1 shows the primary analysis outputs computed by sDNA version 4. Table 2 shows additional outputs of analysis, and Table 3 the outputs of the network preparation, learn and predict tools.

Table 1
Primary network analysis outputs (formatted as polylines) from sDNA software.

Field	Description
Length	Line length
Angular curvature	Distance along line under angular metric
Connectivity	Total number of other line ends connected
Hybrid metric	Line length under distance metric used (if not angular or Euclidean)
Sinuosity	Line length divided by crow-flight Euclidean distance between endpoints
Bearing	Compass bearing between endpoints
Link fraction	Fraction of a single link represented by line
[a]Mean distance	Mean distance under chosen metric (inverse of Closeness)
[a]Mean geodesic length	Mean network-Euclidean length of geodesics from this line
[a]NQPD	Network Quantity Penalised by Distance (Gravity model)
[a]Betweenness	Betweenness (product weighted)
[a]Two phase betweenness	Betweenness (Two phase/Huff variant weighted)
[a]Two phase destination	Total attractiveness under Two Phase/Huff variant weighting
[a]Weight	Total weight in radius (equivalent to attractiveness under product weighting)
[a]Links	Total links in radius
[a]Length	Total length in radius
[a]Angular distance	Total angular curvature in radius
[a]Junctions	Total junctions in radius
[a]Connectivity	Total number of link ends meeting junctions in radius
[a]Diversion ratio	Circuity for all destinations in radius
[a]Mean crow flight	Mean Crow Flight distance to all destinations in radius
[a]Convex hull area	Area of convex hull of radius
[a]Convex hull perimeter	Perimeter of convex hull of radius
[a]Convex hull radius	Maximum crow-flight-Euclidean radius of convex hull of radius
[a]Convex hull bearing	Compass bearing of maximum radius of convex hull
[a]Convex hull shape index	Shape index of convex hull of radius

[a]Denote output duplicated over multiple radii.

Table 2
Further analytical outputs from additional software in the sDNA toolbox.

Field	Description
Destination Map (polylines)	
[a]Origin	Origin from which destination stats apply
[a]Destination	Destination to which destination stats apply
[a]Radius	Radius under which destination stats apply
[a]Weight from origin	Geodesic weight received from origin (product weighting)
[a]Two phase weight from origin	Geodesic weight received from origin (two phase/Huff variant weighting)
[a]Analytic metric	Distance along geodesic under analytic metric
[a]Length	Network-Euclidean length of geodesic
Geodesics (polylines)	
[a]Origin	Origin of geodesic
[a]Destination	Destination of geodesic
[a]Radius	Radius of geodesic
[a]Weight from origin	Geodesic weight received from origin (product weighting)
[a]Two phase weight from origin	Geodesic weight received from origin (two phase/Huff variant weighting)
[a]Analytic metric	Distance along geodesic under analytic metric
[a]Length	Network-Euclidean length of geodesic
Convex Hulls (polygons)	
[a]Origin	Hull Origin
[a]Radius	Network radius which generated convex hull
[a]Weight	Weight of origin
Network Radius (multi-polylines)	
[a]Origin	Radius Origin
[a]Radius	Network radius (distance)
[a]Weight	Weight of origin
Skim Matrix (CSV file)	
Origin zone	–
Destination zone	–
Mean Distance	Mean distance between zones under chosen metric
Weight	Total weight of geodesics between zones
Number of geodesics	Number of geodesics between zones

[a]Denote output duplicated over multiple radii.

2.1. Architecture

Architecture is shown in Fig. 4. The network backend uses parallelization to process multiple origins simultaneously. Typical usage is via the QGIS or ArcGIS toolbox; a more limited interface is provided for Autocad due to its more limited data handling capabilities. As most sDNA users are not themselves programmers, an instance of R-portable [35] is included in the distribution to make use of relevant R libraries, and this is called from the front end tools without users needing to program in R. The sDNA source includes a suite of automated system level tests which are run from the Debug configuration of the Visual Studio project.
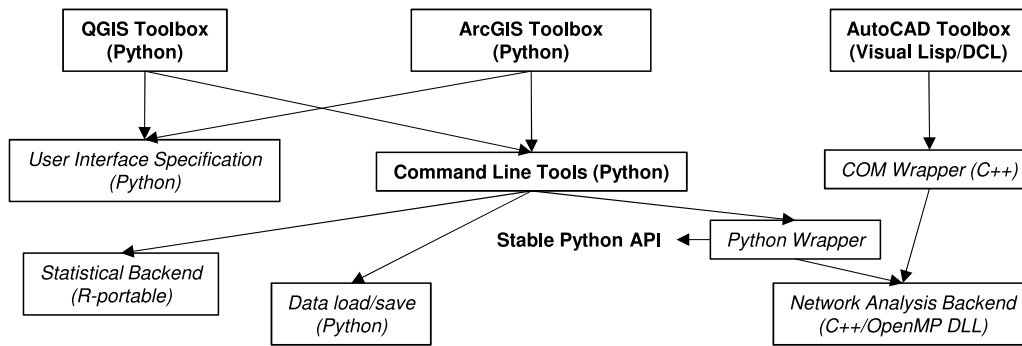
Fig. 4. sDNA software architecture. Components shown in bold represent front-ends/interfaces exposed to the user.

Table 3
Outputs of Prepare, Learn and Predict tools.

Output (data type)	Field	Description
Prepared network (polylines)	Original ID	ID of link before fixing errors
	Attached data	Data preserved by preparation process
Network preparation errors (polylines)	Original ID	ID of link
	Error type	Error type (text field)
Regression model (CSV file)	Parameters and diagnostics	
Residuals (polylines)	Prediction	Value predicted by regression model
	Name of variable	Actual value
	GEH	GEH Statistic
Predictions (polylines)	Values predicted by regression model (for non-training data)	

3. Illustrative examples

Betweenness can be used to approximate a transport model (see [30] for discussion); in the case of motor vehicles, angular betweenness can proxy a model based on travel time [36,37] as straight routes through cities tend to have priority and therefore offer quicker travel times to drivers. Fig. 5 shows angular betweenness with 8km radius (maximum geodesic length) for the road network of Cardiff, the capital city of Wales, with 479,000 inhabitants living in a 75 km^2 urban area. The network data is derived from Open Street Map (OSM) and contains approximately 23,000 links, displayed here using standard GIS tools. The following command was used for the network analysis:

```
d:\example> sdnaintegral.py --im "net=network"
             --om "net=outputs/output_ang"
             "metric=ANGULAR;radii=2000,8000"
```

Fig. 6 shows an example of using 3-d properties of links in analysis by displaying cyclist roundtrip metric per unit length (ignoring motorised traffic). This metric is derived from calibration in [38] and can be used in any of sDNA's accessibility and flow computations. The following command was used to compute metrics for individual links only (setting $t = 0$ to ignore traffic, and afterwards computing *HybridMetricForward/Length*):

```
d:\example> sdnaintegral.py --im "net=network"
             --om "net=outputs/output_crt"
             "metric=CYCLE_ROUNDTRIP;linkonly;t=0"
```

Fig. 7 shows the importance of correct use of one-way link data in vehicle transport models — without it, dual carriageways are not correctly modelled (with the further implication for cycling models based on these vehicle flows, that one side of a dual carriageway is always empty and hence forms an attractive cycle route). The following command was used to produce the corrected version (using the *sdnaoneway* and *wt* fields on the input network, the latter to restrict flows to both directions between a single origin/destination pair):

```
d:\example> sdnaintegral.py --im "net=oneway_prep"
             --om "net=outputs/output_oneway"
             "metric=ANGULAR;weight=wt;oneway=sdnaoneway"
```

Fig. 8 shows the difference between using continuous versus discrete space in a betweenness analysis weighted by length. Under discrete space mode, the long link generates high betweenness even though the radius of analysis is shorter than the link.

Further examples, along with data and batch file of command line calls to sDNA to generate output, are provided as supplementary material. These include a distance map from a single origin, different betweenness types at different radii, mean angular and mean euclidean distance, a 'flow bundle' of flows through a specific link, geodesics between specific origins/destinations, convex hull and network radius output.

Table 4 compares compute times for angular and network-Euclidean betweenness on the above network, between sDNA and the competing software described in Section 1. Note that each tool has a unique feature set and for many tasks these will take precedence over compute times as a basis for choice of tool. For OSM links, which may either be straight or curved, this comparison is a test not only of code efficiency but also of a theoretical approach which allows processing of curved links as a single unit. To provide a comparison based on processing speed alone, a simplified network is also tested in which all links are straightened (except where corners are required to preserve topology) and pre-broken into straight segments. For the straight line network tested, PST and DepthmapX are faster than sDNA while UNA is slower. For the OSM network in which many links are curved, the advantage of treating each link as a single entity becomes apparent as DepthmapX and PST, without this feature, must break the set of links into a set of straight segments 7–10 times the size before processing. DepthmapX in particular shows much longer compute times for curved links. The performance of PST benefits from highly optimised code (and like sDNA, parallel execution); its speed falls between those of sDNA Discrete and Continuous for network Euclidean analysis at a 3km radius, but is slower than either for angular 3km. For a 6km radius this
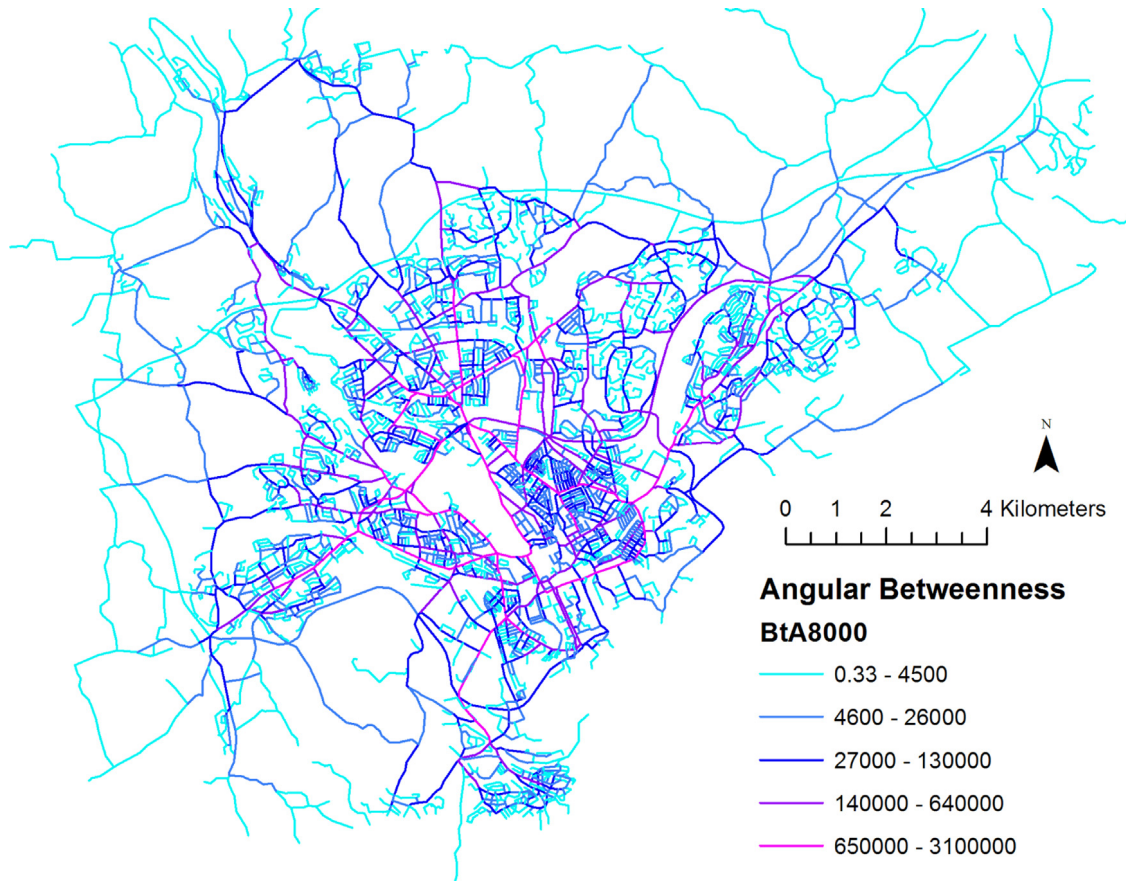
Angular Betweenness

BtA8000

— 0.33 - 4500

— 4600 - 26000

— 27000 - 130000

— 140000 - 640000

— 650000 - 3100000

Fig. 5. Angular Betweenness, 8km radius. Note classification for legend uses geometric intervals (recommended as Betweenness usually has exponential distribution).
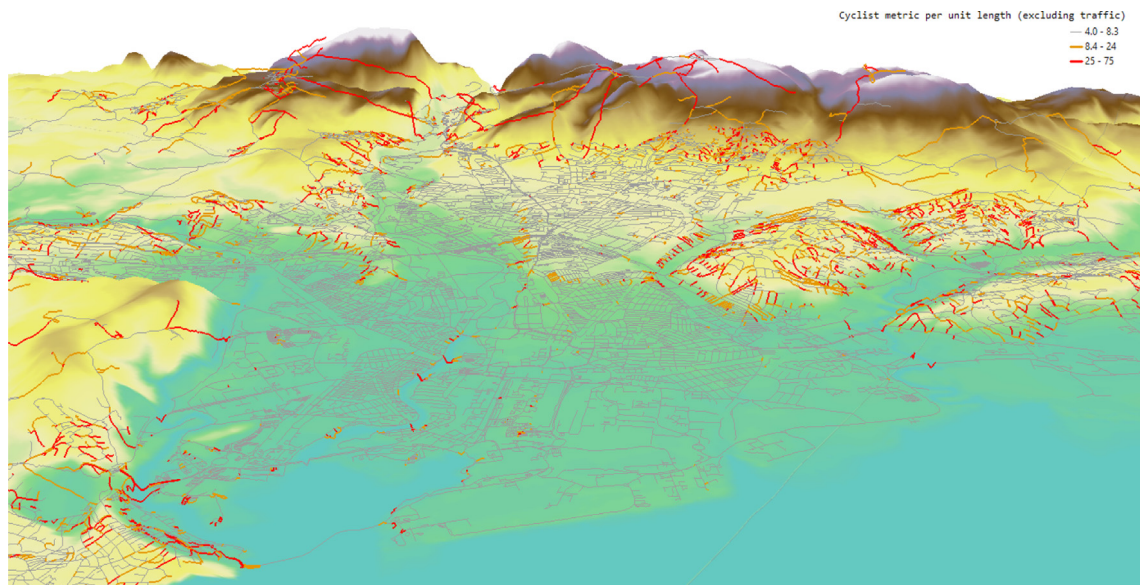

Fig. 6. Cyclist roundtrip metric per unit length (excluding deterrence from motorised traffic) in city of Cardiff. Links with high metric on slopes illustrate use of 3-d properties of links. Displayed in ArcScene with vertical exaggeration = 5.

trend continues further, and is likely to be more pronounced for larger radii, as the time complexity of betweenness scales with the square of the number of origins and destinations within the radius.

4. Impact

sDNA Open 4 is the first open source release of the sDNA software, which has been under development since 2011, funded
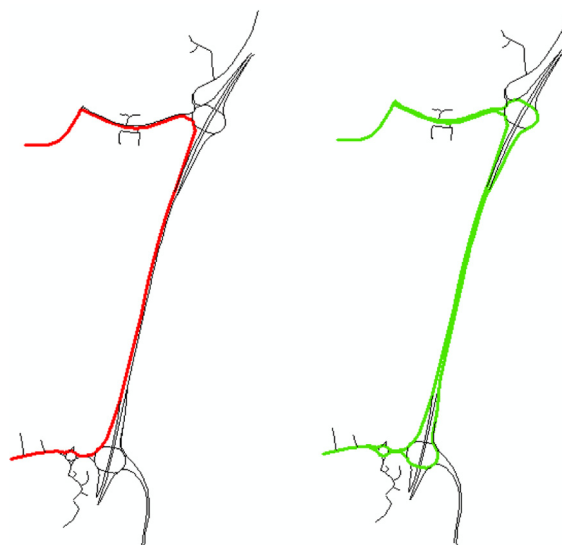
Fig. 7. Illustration of correct use of one-way links. Ignoring one-way data (left), angular betweenness flows used to approximate vehicle traffic will fail to correctly use dual carriageway. Correct flows shown on right.
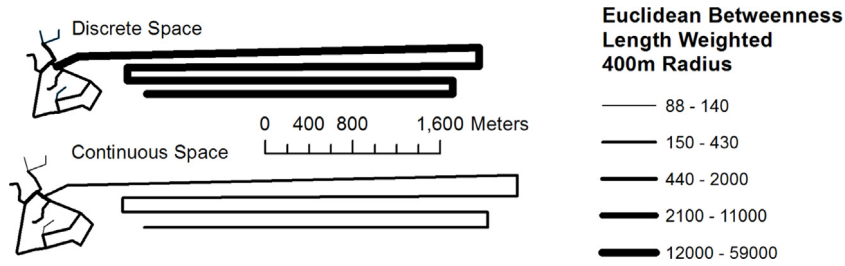

Fig. 8. Comparison of discrete and continuous space modes: in discrete space mode the long link appears to generate high levels of betweenness, as it exceeds the length of the radius. In continuous space a partial link is considered. In this case the long link has been drawn in unrealistic zigzag fashion in order to include a long link without the diagram exceeding the page width; however, similar ratios of link length are common in real spatial networks.

Table 4
Comparison of times to compute betweenness with different software, for Cardiff model on Intel i7-4810MQ, 2.8 GHz, 4 cores, 8 threads, 32GB RAM. net-Euc = network-Euclidean; n/a = not applicable (UNA does not offer this output); n/t = not tested.

Input Data # units	Simplified straight line 24,558		OSM Links 23,359		
Analysis metric	net-Euc	Angular	net-Euc	Angular	Angular
Radial metric	net-Euc	net-Euc	net-Euc	net-Euc	net-Euc
Radius	3000 m	3000 m	3000 m	3000 m	6000 m
sDNA discrete	3'54"	3'40"	3'24"	3'16"	9'25"
sDNA continuous	8'33"	7'32"	7'08"	7'00"	12'12"
DepthmapX	0'16"	1'06"[a]	45'00"[b]	72'00"[b]	n/t
UNA[c]	35'56"	n/a	34'52"	n/a	n/a
PST	0'07"	0'07"	6'00"[d]	8'52"[d]	16'01"[d]

[a] 1024-bin tulip analysis approximating angular.
[b] DepthmapX breaks link set into approx. 218,000 segments.
[c] Based on artificial buildings layer with a 'building' at the midpoint of each link (analysing all buildings expected to be orders of magnitude slower).
[d] PST breaks link set into approx. 180,000 segments.

from a mixture of sources including sponsorship of open source features and sales of the proprietary sister software sDNA+ [31]. A 2015 ESRC-funded project focused on removing barriers to use and uptake outside of academia. Prior to open source release, 1400 active installations were known to be in use; the sDNA QGIS plugin alone now records over 15,000 downloads with a rating of 4/5 stars. The user base can generally be considered experienced in using GIS or CAD software to load or input and display data, and perhaps experienced modellers and/or statisticians, but not necessarily having programming experience. sDNA has been used to deliver numerous sustainable transport projects

by global engineering consultancies, including Arup who used it in a project which won a Chartered Institution of Highways & Transportation's award (2016).

Academic applications of sDNA have included predicting flows and mode choice in transport networks: vehicle [39], metro [40], pedestrian [30,33,41–44] and bicycling [29,38]. Spatial network analysis shows especial promise as a technology for modelling slow, active and sustainable modes of transport, for which traditional zone-based transport models have been unable to model

details of trips within single zones, and feedbacks between transport and land use, in a cost-effective manner. This has allowed answering questions on the effect of infrastructure on active travel choices, and application of results in the public realm. Much of this has been through the customised forms of betweenness analysis offered by sDNA, however, the more novel convex hull based metrics we introduced have also been shown to have high correlation with flows of pedestrians [42]. Further applications include land use [45] and planning [46–48]: GIS arguably provides a *lingua franca* bridging the gap between planning and transport disciplines and improving spatial network analysis capability within this context has furthered study of the relationship between accessibility and land use, and research on quantification of urban plan quality. Another substantial field of application is health informatics, where sDNA has allowed quantification of built environment characteristics in epidemiological models, sometimes on a "big data" scale (sDNA outputs are used in UK Biobank [49]), enabling new research into the effects of built environment on health [50–54] ([54] won a Royal Town Planning Institute award for research excellence in 2019, and [55] was shortlisted for the same). [50] showed a link between convex hull statistics and social cohesion. A final, unexpected development is the use of sDNA in archaeology [56].

In all the above fields, we hope that by standardizing on the network link as a fundamental unit of analysis, in contrast to segments or axial lines used in previous research (which can be defined in multiple ways), more repeatable results might be obtained in the long term.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The authors are entitled to receive a small share of revenue from the proprietary sister project to sDNA, sDNA+. AC reinvests all revenue.

Acknowledgements

References

[1] Lawhead J. This library reads and writes ESRI shapefiles in pure Python. 2019, URL https://github.com/GeospatialPython/pyshp.

[2] Open Source Geospatial Foundation. GEOS. 2006–2019, URL https://trac.osgeo.org/geos.

[3] Berg I. Muparser fast math parser library. 2011–2015, URL https://beltoforion.de/article.php?a=muparser.

[4] Friedman J, Hastie T, Tibshirani R. Regularization paths for generalised linear models via coordinate descent. J Stat Softw 2010;33(1):1–22, URL http://www.jstatsoft.org/v33/i01/.

[5] Boost Community. Boost C++ libraries. 1999–2019, URL https://www.boost.org/.

[6] Bavelas A. A mathematical model for small group structures. Hum Organ 1948;7:16–30.

[7] Ford L, Fulkerson D. Flows in networks. 2nd ed. Princeton University Press; 1962.

[8] Freeman L. A set of measures of centrality based on betweenness. Sociometry 1977;40(1):35–41.

[9] Garrison W, Marble D. The structure of transportation networks. Armed Services Technical Information Agency; 1962.

[10] Haggett P, Chorley RJ. Network analysis in geography. Hodder and Stoughton Educational; 1969.

[11] Kansky K. Structure of transport networks: Relationships between network geometry and regional characteristics [Ph.D. thesis], Department of Geography, University of Chicago; 1963.

[12] Shimbel A. Structural parameters of communication networks. Bull Math Biophys 1953;15(4):501–7.

[13] Cooper CHV. Spatial localization of closeness and betweenness measures: a self-contradictory but useful form of network analysis. Int J Geogr Inf Sci 2015;29(8):1293–309. http://dx.doi.org/10.1080/13658816.2015.1018834.

[14] Huang J, Levinson D. Circuity in urban transit networks. Working Papers 201501. University of Minnesota: Nexus Research Group; 2015, URL https://ideas.repec.org/p/nex/wpaper/transitcircuity.html.

[15] New Zealand Transport Agency. Pedestrian planning and design guide. 2007, URL http://www.nzta.govt.nz/resources/pedestrian-planning-guide/docs/pedestrian-planning-guide.pdf.

[16] Añez J, De La Barra T, Pérez B. Dual graph representation of transport networks. Transp Res B 1996;30(3):209–16. http://dx.doi.org/10.1016/0191-2615(95)00024-0, URL http://www.sciencedirect.com/science/article/pii/0191261595000240.

[17] Hillier B, Hanson J. The social logic of space. Cambridge University Press; 1984.

[18] Bavelas A. Communication patterns in task-oriented groups. Acoust Soc Am J 1950;22:725. http://dx.doi.org/10.1121/1.1906679, URL http://adsabs.harvard.edu/abs/1950ASAJ...22..725B.

[19] Dhanani A, Vaughan L. Towards a walkability model for strategic evaluation of policy action and urban active transport interventions. In: (Proceedings). Universities' Transport Study Group (UTSG); 2016, URL http://discovery.ucl.ac.uk/1524124/.

[20] Sevtsuk A, Mekonnen M. Urban network analysis. MIT City Form Lab; 2012, URL http://cityform.mit.edu/projects.

[21] ESRI. ArcGIS. 1999–2019, URL https://www.esri.com/en-us/home.

[22] depthmapX development team. depthmapX. 2017, URL https://github.com/SpaceGroupUCL/depthmapX/.

[23] Turner A. From axial to road-centre lines: a new representation for space syntax and a new model of route choice for transport network analysis. Environ Plan B: Plann Des 2007;34(3):539–55, URL http://eprints.ucl.ac.uk/2092.

[24] Chiaradia AJ, Hillier B, Schwander C, Wedderburn M. Compositional and urban form effects in centres in greater London. Urban Des Plan - Proc ICE 2012;165:21–42.

[25] Cooper CHV, Chiaradia AJ. sDNA: How and why we reinvented spatial network analysis for health, economics and active modes of transport. In: Malleson N, editor. GISRUK 2015 proceedings. 2015, http://dx.doi.org/10.6084/m9.figshare.1491375.

[26] Ståhle A, Marcus L, Karlström A. Place syntax - geographic accessibility with axial lines in GIS. In: 5th International space syntax symposium; 2005, p. 131–44.

[27] Feng C. Spatialist reach analysis. 2019, URL https://www.food4rhino.com/resource/grasshopper-reach-analysis-toolkit.

[28] Feng C, Zhang W. Algorithms for the parametric analysis of metric, directional, and intersection reach. Environ Plan B: Urban Anal City Sci 2019;46(8):1422–38. http://dx.doi.org/10.1177/2399808319827299.

[29] Cooper CHV. Predictive spatial network analysis for high-resolution transport modeling, applied to cyclist flows, mode choice, and targeting investment. Int J Sustain Transp 2018;1–11. http://dx.doi.org/10.1080/15568318.2018.1432730.

[30] Cooper CHV, Harvey I, Orford S, Chiaradia AJF. Using multiple hybrid spatial design network analysis to predict longitudinal effect of a major city centre redevelopment on pedestrian flows. Transportation 2019. http://dx.doi.org/10.1007/s11116-019-10072-0.

[31] Cooper CHV, Chiaradia AJ, Webster C. sDNA+. 2011–2019, URL https://www.cardiff.ac.uk/sdna.

[32] Huff DL. Defining and estimating a trading area. J Mark 1964;28(3):34–8. http://dx.doi.org/10.2307/1249154, URL https://www.jstor.org/stable/1249154.

[33] Zhang L, Chiaradia A. Three-dimensional spatial network analysis and its application in a high density city area, central Hong Kong. Urban Plan Int 2019;34(1):46–53. http://dx.doi.org/10.22217/upi.2018.513, URL http://www.upi-planning.org/Magazine/Issue_Content.aspx?ID=47706.

[34] Department for Transport. Highway assignment modelling. In: Transport Analysis Guidance, (M3.1). 2014, URL https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/427124/webtag-tag-unit-m3-1-highway-assignment-modelling.pdf.

[35] Huber W. R portable. 2012–2019, URL https://sourceforge.net/projects/rportable/.

[36] Jayasinghe AB. A network centrality-based simulation approach to model traffic volume [Ph.D. thesis], Nagaoka University of Technology; 2017.

[37] Ciscal-Terry W, Dell'Amico M, Hadjidimitriou NS, Iori M. An analysis of drivers route choice behaviour using GPS data and optimal alternatives. J Transp Geogr 2016;51(Suppl. C):119–29. http://dx.doi.org/10.1016/j.jtrangeo.2015.12.003, URL http://www.sciencedirect.com/science/article/pii/S0966692315002239.

[38] Cooper CHV. Using spatial network analysis to model pedal cycle flows, risk and mode choice. J Transp Geogr 2017;58:157–65. http://dx.doi.org/10.1016/j.jtrangeo.2016.12.003, URL http://www.sciencedirect.com/science/article/pii/S0966692316307256.

[39] Collins A, Cooper CHV. Measuring and managing the environmental impact of festivals: the contribution of the ecological footprint. J Sustain Tourism 2016. URL http://orca.cf.ac.uk/90912/.

[40] Zhang L, Chiaradia A, Zhuang Y. Configurational accessibility study of road and metro network in Shanghai. In: Pan Q, Cao J, editors. Recent developments in Chinese urban planning. Springer; 2015, p. 219–45. http://dx.doi.org/10.1007/978-3-319-18470-8.

[41] Cubukcu E, Hepguzel B, Onder Z, Tumer B. Active living for sustainable future: A model to measure "walk scores" via geographic information systems. Procedia - Soc Behav Sci 2015;168:229–37. http://dx.doi.org/10.1016/j.sbspro.2014.10.228, URL http://www.sciencedirect.com/science/article/pii/S1877042814056869.

[42] Kang C-D. Measuring the effects of street network configurations on walking in Seoul, Korea. Cities 2017;71:30–40. http://dx.doi.org/10.1016/j.cities.2017.07.005, URL http://www.sciencedirect.com/science/article/pii/S0264275116305364.

[43] Sun G, Webster C, Chiaradia A. Ungating the city: a permeability perspective. Urban Stud 2018;55(12):2586–602. http://dx.doi.org/10.1177/0042098017733943.

[44] Sun G, Webster C, Chiaradia A. Objective assessment of station approach routes: Development and reliability of an audit for walking environments around metro stations in China. J Transp Health 2017;4:191–207. http://dx.doi.org/10.1016/j.jth.2017.01.010, URL http://www.sciencedirect.com/science/article/pii/S221414051630322X.

[45] He S, Yu S, Wei P, Fang C. A spatial design network analysis of street networks and the locations of leisure entertainment activities: a case study of Wuhan, China. Sustainable Cities Soc 2019;44:880–7. http://dx.doi.org/10.1016/j.scs.2018.11.007, URL http://www.sciencedirect.com/science/article/pii/S2210670718314914.

[46] Grimaldi D, Fernandez V, Carrasco C. Heuristic for the localization of new shops based on business and social criteria. Technol Forecast Soc Change 2019;142:249–57. http://dx.doi.org/10.1016/j.techfore.2018.07.034, URL http://www.sciencedirect.com/science/article/pii/S0040162517312271.

[47] Xiao Y, Sarkar C, Webster C, Chiaradia A, Lu Y. Street network accessibility-based methodology for appraisal of land use master plans: An empirical case study of Wuhan, China. Land Use Policy 2017;69:193–203. http://dx.doi.org/10.1016/j.landusepol.2017.09.013, URL http://www.sciencedirect.com/science/article/pii/S0264837717306336.

[48] Ye Y, Richards D, Lu Y, Song X, Zhuang Y, Zeng W, Zhong T. Measuring daily accessed street greenery: a human-scale approach for informing better urban planning practices. Landsc Urban Plan 2018. http://dx.doi.org/10.1016/j.landurbplan.2018.08.028, URL http://www.sciencedirect.com/science/article/pii/S0169204618309940.

[49] Sarkar C, Gallacher J, Webster C. Morphometric analysis of the built environment in UK Biobank: Data analyses and specification manual. Cardiff University; 2014, URL http://biobank.ctsu.ox.ac.uk/crystal/docs/ard-1195_StreetNtrkAccess.pdf.

[50] Cooper CHV, Fone DL, Chiaradia A. Measuring the impact of spatial network layout on community social cohesion: a cross-sectional study. Int J Health Geogr 2014;13(1):11. http://dx.doi.org/10.1186/1476-072X-13-11, URL http://www.ij-healthgeographics.com/content/13/1/11/abstract.

[51] Sarkar C, Gallacher J, Webster C. Urban built environment configuration and psychological distress in older men: Results from the Caerphilly study. BMC Public Health 2013;13(1):695. http://dx.doi.org/10.1186/1471-2458-13-695, URL http://www.biomedcentral.com/1471-2458/13/695/abstract.

[52] Sarkar C, Webster C, Pryor M, Tang D, Melbourne S, Zhang X, Jianzheng L. Exploring associations between urban green, street design and walking: results from the greater London boroughs. Landsc Urban Plan 2015;143(Suppl. C):112–25. http://dx.doi.org/10.1016/j.landurbplan.2015.06.013, URL http://www.sciencedirect.com/science/article/pii/S0169204615001383.

[53] Zhang L, Ye Y, Zeng W, Chiaradia A. A systematic measurement of street quality through multi-sourced urban data: A human-oriented analysis. Int J Environ Res Public Health 2019;16(10). http://dx.doi.org/10.3390/ijerph16101782.

[54] Sarkar C, Webster C, Gallacher J. Residential greenness and prevalence of major depressive disorders: a cross-sectional, observational, associational study of 94,879 adult UK Biobank participants. Lancet Planet Health 2018;2(4):e162–73. http://dx.doi.org/10.1016/S2542-5196(18)30051-2.

[55] Sarkar C, Webster C, Gallacher J. Neighbourhood walkability and incidence of hypertension: Findings from the study of 429,334 UK Biobank participants. Int J Hyg Environ Health 2018;221(3):458–68. http://dx.doi.org/10.1016/j.ijheh.2018.01.009.

[56] Orengo HA, Livarda A. The seeds of commerce: a network analysis-based approach to the Romano-British transport system. J Archaeol Sci 2016;66:21–35. http://dx.doi.org/10.1016/j.jas.2015.12.003, URL http://www.sciencedirect.com/science/article/pii/S0305440315003192.