Environment and Planning B: Planning and Design 2009, volume 36, pages 450 – 465

doi:10.1068/b34098

Street centrality and densities of retail and services in Bologna, Italy

Sergio Porta, Emanuele Strano, Valentino Iacoviello, Roberto Messora
Human Space Lab, DIAP, Polytechnic of Milan, Via Bonardi 3, 20133 Milano, Italy;
e-mail: s.porta@humanspacelab.com, e.strano@humanspacelab.com,
v.iacoviello@humanspacelab.com, r.messora@humanspacelab.com

Vito Latora, Alessio Cardillo
Dipartimento di Fisica e Astronomia, Università di Catania, and INFN Sezione di Catania,
Via S. Sofia 64, 95123 Catania, Italy; e-mail: latora@ct.infn.it, alessio.cardillo@ct.infn.it

Fahui Wang
Department of Geography and Anthropology, Louisiana State University, Baton Rouge,
LA 70803, USA; e-mail:fwang@lsu.edu

Salvatore Scellato
Scuola Superiore di Catania, Via San Nullo, 5/i, 95123 Catania, Italy;
e-mail: sascellato@ssc.uni.ct.it
Received 6 August 2007; in revised form 8 October 2007; published online 14 July 2008

Abstract. This paper examines the relationship between street centrality and densities of commercial and service activities in the city of Bologna, northern Italy. Street centrality is calibrated in a multiple centrality assessment model composed of multiple measures such as closeness, betweenness, and straightness. Kernel density estimation is used to transform datasets of centrality and activities to one scale unit for analysis of correlation between them. Results indicate that retail and service activities in Bologna tend to concentrate in areas with better centralities. The distribution of these activities correlates highly with the global betweenness of the street network, and also, to a slightly lesser extent, with the global closeness. This confirms the hypothesis that street centrality plays a crucial role in shaping the formation of urban structure and land uses.

1 Location and centrality in a city

> "No matter how good its offering, merchandising, or customer service, every retail company still has to contend with three critical elements of success: location, location, and location"
>
> Taneja (1999, page 136).

What is *location*? Why does it matter? A simple and intuitive answer is: *centrality*.

A central place has one special feature to offer to those who live or work in a city: easy accessibility from immediate surroundings and more distant places. Accessibility may be transformed to visibility and popularity. Therefore, a central place tends to attract more customers and has a greater potential to develop into a social catalyst. Important landmarks such as museums, theatres, or office headquarters favour the central locations. A more central location commands a higher real estate value and is occupied by a more intensive land use. Central locations in an urban area have the potential to sustain higher densities of retail and services, and are a key factor for supporting the formation and vitality of urban 'nodes' (Newman and Kenworthy, 1999). Centrality emerges as one of the most powerful determinants for urban planners and designers to understand how a city works and to decide where renovation and redevelopment need to be placed.

Centrality not only affects how a city works today, but also plays an important role in shaping its growth. If one looks at where a city centre is located, it is most likely to sprout from the intersection of main routes, where some special configuration of the terrain or some particular layout of the river system (or water bodies in general) makes the place compulsory to pass through. That is one of the dominant theories that explain where a city begins. Then, departing from such central locations, the city grows up over time with gradual additions of dwellings, residents, and activities: first along the main routes, then filling the in-between areas, and then developing streets that realize loops and points of return. As the structure becomes more complex, new central streets and places are formed and stimulate a growth in the number of residents and activities around them. This evolutionary process has been driving the formation of urban fabrics and the advancement of human civilization throughout most of the seven millenniums of city history.

Centrality appears to be somehow at the heart of that marvellous hidden order that supports the formation of 'spontaneous' and organic cities (Jacobs, 1961; 1993). It is also a crucial issue in the contemporary debate on searching for more bottom-up and 'natural' strategies of urban planning beyond the modernistic heritage. Centrality has been studied in many branches of urban research, especially in economic geography and regional analysis (Wilson, 2000) and in transportation planning (Goulias, 2002; Meyer and Miller, 2000). In most cases centrality has been dealt with as a means of measuring the relationship between activities among places. In essence, this has led to an interpretation of centrality through an intuitive notion that a more central location is a place 'closer' to all others. In urban planning and design, centrality is the core issue addressed by *space syntax*, a methodology of spatial analysis, through the notions of 'visibility' and 'integration' (Hillier, 1996; Hillier and Hanson, 1984). Space syntax has opened a whole new range of opportunities for urban designers to develop a deeper understanding of several structural properties of city spaces. The model has achieved significant successes in the practice of countless urban regeneration programmes in the UK and elsewhere, and has helped urban planners and designers in making good decisions and reframing the debate on pivotal issues such as crime, self-surveillance, community building, and renovation of large housing estates in the last two decades or so. Despite these successes, urban designers often perceive space syntax as a quantitative threat to the creativity embedded in the art of city design, while on the other side researchers in spatial analysis and geocomputation often find it lacks rigorous expression and clear disciplinary references.

The *multiple centrality assessment* (MCA) model follows broader traditions in centrality assessment, which draw on structural sociology since the early 1950s (Bavelas, 1948; 1950; Freeman, 1977; 1979; see also an overview by Wasserman and Faust, 1994), and the more recent new physics of complex networks (Boccaletti et al, 2006). Implementing these traditions in a spatial environment, MCA works at the forefront of a growing wave of interest for geographic information research (Batty, 2005). Therefore, the MCA model shares with space syntax the fundamental *values* that refer to the structural interpretation of urban spaces for urban planning and design, while offering a new and deeply alternative *technical* perspective (Cardillo et al, 2006; Crucitti et al, 2006a; 2006b; Porta et al, 2006a; 2006b; Scellato et al, 2006; Scheurer and Porta, 2006; Scheurer et al, 2007).

The hypothesis for this study is that centrality captures the essence of location advantage in an urban area, and its values should be reflected in the intensity of land uses—in this case, densities of retail and service activities. We test the hypothesis in Bologna, a northern Italian city. The remainder of this paper is organized as follows. Section 2 describes the study area and data preparation. Section 3 reviews the MCA model for measuring urban network centralities. Section 4 shows how kernel density estimation (KDE) is used to transform both the MCA measures and densities of retails and services to one scale unit, which permits the correlation analysis between them.

Section 5 presents the analysis results with some discussion. The paper is concluded in section 6 with a brief summary.

2 Study area and data preparation

Bologna is a regional capital in northern Italy with some half a million residents. It is an important urban centre located in the middle of the River Po Plain. Historical major routes connecting Florence and southern Italy with northern Italy converge here (see figure 1). Bologna is a relatively wealthy city and a national transportation hub with a rich history and culture: it is where the 'Alma Mater', the most ancient university in the world, originated. The aim of this study is to examine how the variation of street centrality (measured by the MCA model) correlates with the distributions of retail shops and community services in Bologna.

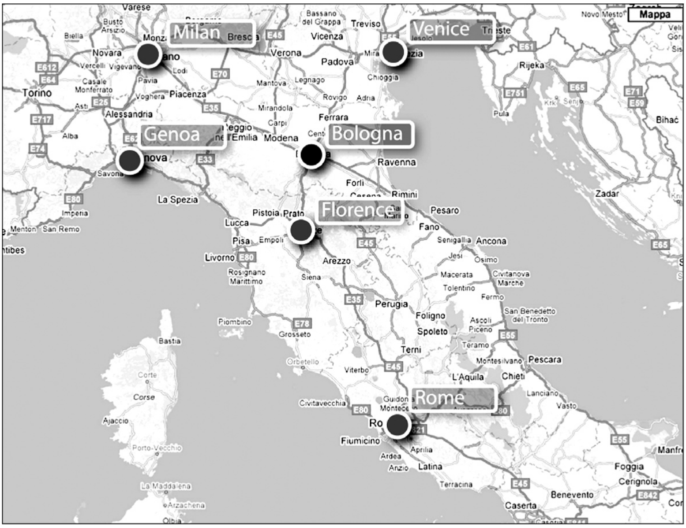

Figure 1. Location of Bologna in northern Italy.

Data for this study were provided by the municipality of Bologna in two separate datasets. The first dataset is an ArcGIS (ESRI Inc., Redlands, CA) point shapefile including all ground-floor retail and service activities in Bologna in 2003. Each of the n activities is indicated in the following as X_i, where $i = 1, 2, ..., n$. In Bologna these retail and service entities are in general comparable in floor size. Ideally, we would like to have the data of floor space or even the rental value for each activity to capture more effectively the intensity of land uses. We are in the process of collecting such data and will report the results when available. For the purpose of this study, we then extracted two separated layers from this dataset: one for retail shops alone ($n_{\mathrm{comm}} = 7257$ points), and one for retail and service activities altogether ($n_{\mathrm{comm+serv}} = 9676$ points). The second dataset is an ArcGIS line shapefile containing all streets in Bologna in 2004. The street network has 7191 street segments (edges) and 5448 intersections (nodes). Figure 2(a) shows the distribution of retail (commercial) and service activities on the street network in Bologna.

3 Measuring urban network centralities with multiple centrality assessment

The MCA model focuses on measuring *centrality* in urban networks constituted by streets as *links* (or *edges*), and by intersections as *nodes*. The main characteristics of the MCA are: (1) utilizing a standard 'primal' format for the street network representation;

(2) anchoring all measures on a metric computation of distances along the real street network; and (3) defining centrality by a set of multiple peer indices (Porta et al, 2006b). While the first two characteristics differentiate the MCA from space syntax, the third one makes it distinctive from traditional regional analysis, transportation planning, and urban geography. The MCA model in fact measures a location as 'being central' not only in terms of being *close* to all others, but also in terms of being the *intermediary* between others, being accessible via a *straight* route to all others, and being *critical* for the efficiency of the system as a whole.

Results from the MCA analysis include several maps of a street network, each of which shows one set of centrality values for the edges of the network. The following is a brief review of the MCA model. For more detailed discussions of MCA, see the references cited in section 1.

Streets (links or edges) are represented in a GIS system as linear features with two end nodes and, possibly, one or more intermediate vertices. The MCA model assigns a set of centrality values to each street segment (Crucitti et al, 2006a; 2006b; Porta et al, 2006a; 2006b). Here we discuss three of them: *closeness* (C^{C}), *betweenness* (C^{B}), and *straightness* (C^{S}).

Closeness centrality C^{C} measures to what extent a node is close to all the other nodes along the shortest paths of the network. C^{C} for a node i is defined as:

$$C_i^{\mathrm{C}} = \frac{N-1}{\sum\limits_{j=1;\ j \neq i}^{N} d_{ij}}, \tag{1}$$

where N is the total number of nodes in the network, and d_{ij} is the shortest distance between nodes i and j. In other words, the closeness centrality for a node is the inverse of the average distance from this node to all other nodes.

After calibrating the shortest path between any two nodes, it is straightforward to compute C^{C} for all the nodes in the network. C^{C} may be interpreted as proximity, and also captures the notion of *accessibility* of a place. The closer a place is to other places, the more accessible it is. The family of closeness measures has been widely used in urban and regional analysis. In essence, it reflects the *cost* of overcoming spatial separations between places with population and activities.

Betweenness centrality C^{B} is based on the idea that a node is more central when it is traversed by a larger number of the shortest paths connecting all couples of nodes in the network. C^{B} is defined as:

$$C_i^{\mathrm{B}} = \frac{1}{(N-1)(N-2)} \sum_{j=1;\ k=1;\ j \neq k \neq i}^{N} \frac{n_{jk}(i)}{n_{jk}}, \tag{2}$$

where n_{jk} is the number of shortest paths between nodes j and k, and $n_{jk}(i)$ is the number of these shortest paths that contain node i. Using a social network analogue, C^{B} is like the kind of prominence of a person who acts as intermediary among a large number of other persons. In MCA terms, C^{B} captures a special property for a place in a city: it does not act as an origin or a destination for trips, but as a pass-through point. C^{B} represents a node's volume of through traffic. A place with better betweenness may benefit from this important property.

Straightness centrality C^{S} originates from the idea that efficiency of communication between two nodes increases when there is less deviation of their shortest path from the virtual straight line connecting them—that is, a greater 'straightness' of the shortest path. C^{S} is defined as:

$$C_i^{\mathrm{S}} = \frac{1}{N-1} \sum_{j=1;\ j \neq i}^{N} \frac{d_{ij}^{\mathrm{Eucl}}}{d_{ij}}, \tag{3}$$

where d_{ij}^{Eucl} is the Euclidean distance between nodes i and j, or the length of the virtual straight connection.

C^{S} was originally proposed in nonspatial networks as a normalization procedure (Vragovìc et al, 2004). In spatial networks C^{S} reveals a totally different meaning related to human cognitive processes in navigating complex spatial structures. C^{S} measures the extent to which a place can be reached directly, on a straight line, from all other places in a city. This is a quality that makes it prominent in terms of 'legibility' and 'presence' (Conroy-Dalton, 2003).

In this study all three *global* centrality indices were calculated first, as all nodes and edges in the network participated in the computation: namely, global closeness $C_{\mathrm{glob}}^{\mathrm{C}}$, global betweenness $C_{\mathrm{glob}}^{\mathrm{B}}$, and global straightness $C_{\mathrm{glob}}^{\mathrm{S}}$. As an example, figure 2(c) shows the variation of global betweenness $C_{\mathrm{glob}}^{\mathrm{B}}$ across the street network in Bologna. In addition, two *local* centrality indices were calculated for the nodes located within a certain distance d from each node i. As shown in a previous study (Porta et al, 2006b), local measures are useful to overcome the *edge effect*—that is, the distortion that lowers the centrality values near the edge of a network. Such a distortion turns out to be very significant for the closeness index when calculated on highly fragmented networks. Moreover, the global centrality measures do not reveal network properties on a local scale, whereas local measures nicely capture properties of space at the neighbourhood or district scale that have to do with a nonmotorized experience of cities. In Bologna the search range for local centrality measures was set at $d = 800$ m, and, therefore, the local closeness and local straightness are denoted as C_{800}^{C} and C_{800}^{S}, respectively.

We have developed an ArcGIS extension to prepare the street network data for MCA computation. The module first cleans up the street network in an ArcGIS shapefile format for most common errors, then generates nodes at intersections and links the node identification to the polyline attribute table, and finally generates a 'connectivity table' that stores, for each street, its length, identification of the two end nodes, and their coordinates. The connectivity table is then processed outside of GIS by a C++ script that computes centralities of each node and each street (ie the average centrality values of the street's two end nodes). The result from the C++ program is fed back to ArcGIS for mapping and other spatial analysis such as the KDE in the next phase.

4 Using kernel density estimation to compute densities of street centralities and retail and service activities

4.1 Transforming data to one analysis unit

The objective of this research is to examine whether the variation of street centralities is indeed reflected in the intensity of land uses—in our case densities of retail and service activities in Bologna. In accordance with the MCA model, three centralities (C^{B}, C^{C}, and C^{S}) for each of the 5448 nodes were computed, on the basis of which, centralities for each of the 7191 edges were calculated as the average of two end nodes. Correspondingly, there were 7257 retail shops (or 9676 retail shops and service facilities altogether) in the study area. The street network and points of retail and service activities were two distinct spatial features. In order to analyze the relationship between them our first task was to transform the two datasets to one scale (analysis unit), so that such a comparison could be made.

Several approaches may be employed to facilitate such a comparison study. One is to use the distribution of retail and service activities as the base framework, to compute the street network centralities at (or around) each retail and service facility, and then to examine the relationship between them. The second approach is the other way around—that is, computing the density of retail and service activities along each street and matching with the centralities for that street. The third approach is to transform both activities to a new framework (eg a raster system), and to examine the relationship between the street centralities and land-use intensity at the same scale. In all cases, data transformations from one scale or analysis unit to another utlize *spatial smoothing* and/or *spatial interpolation* techniques (Wang, 2006, page 47). The current research uses the third approach, mainly for the convenience of taking advantage of some built-in tools available in ArcGIS. Our future work will explore the first and second approaches.

There are also rich choices of spatial smoothing (eg floating catchment area, KDE, and empirical Bayes estimation) and spatial interpolation methods (eg trend surface analysis, inverse distance weighted, thin-plate splines, and kriging) (Wang, 2006, pages 35–53). Here, the KDE method was used. Basically, the KDE uses the density within a range (window) of each observation to represent the value at the centre of the window. Within the window, the KDE weighs nearby objects more than distant objects, on the basis of a *kernel function* (Bailey and Gatrell, 1995; Fotheringham et al, 2000, pages 146–149; Silverman, 1986). By doing so, the KDE generates a density of the events (discrete points) as a continuous field (eg raster), and therefore converts the two datasets to the same raster framework and permits the analysis of relationships between them. The roots of KDE may be traced back to several pioneering studies in mathematical statistics (Akaike, 1954; Parzen, 1962; Rosenblatt, 1956), and it has since been used in many fields such as geography, epidemiology, criminology, demography, ethology, hydrology and others, including two recent examples (Anselin et al, 2000; Borruso, 2003).

While the choice of a particular smoothing or interpolation technique should not affect the outcome of this research, our choice of KDE was made for at least three reasons.

- First, and most importantly, by using the density (or average attributes) of nearby objects to represent the property at the middle location, the KDE captures the very essence of location measured by centralities and reflected by densities of retail and service facilities. In contrast to studies of a direct correlation between street integration and socioeconomic and environmental indicators (eg Penn and Turner, 2003), this approach emphasizes that it is not the place itself but rather its surroundings that make it special and explain its setting. Using the KDE here is not only a requirement for converting the data scale but also a necessity for accurately capturing the true intention of analyzing the relationship between two neighbourhood features.
- Secondly, the KDE uses a kernel function to value the contribution of a nearby object to the density estimate more than a remote one, as stated in Tobler's (1970) first law of geography—that is, "everything is related to everything else, but near things are more related than distant things." This property of distance decay for spatial interaction is widely recognized by urban researchers. The family of gravity models follow the same notion with strong theoretical foundations and have many successful applications in urban and regional studies (Fotheringham et al, 2000, pages 213–235).
- Finally, the KDE is a standard tool in the ArcGIS spatial analyst module, and the results can be easily integrated in ArcGIS for mapping.

4.2 Kernel density estimation

A kernel function looks like a bump centred at each point x_i and tapering off to zero over a bandwidth or window. The kernel density at point x at the centre of a grid cell is estimated to be the sum of bumps within the bandwidth:

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} \mathrm{K}\left(\frac{x - x_i}{h}\right), \tag{4}$$

where K is the kernel function, h is the bandwidth, n is the number of points within the bandwidth, and n is the total number of events. All events x_i within the bandwidth of x generate some bumps reaching the point x, and contribute to the estimated kernel density there.

The kernel function $\mathrm{K}(\boldsymbol{y})$ is a function satisfying the normalization for a two-dimensional vector $\boldsymbol{y}$ such as:

$$\int_{R^2} \mathrm{K}(\boldsymbol{y})\mathrm{d}\boldsymbol{y} = 1 .$$

A regularly adopted kernel is the standard normal curve:

$$\mathrm{K}(\boldsymbol{y}) = (2\pi)^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{y}^2\right) .$$

For convenience, our computation in ArcGIS used the following kernel function, as described in Silverman (1986, page 76):

$$\mathrm{K}(\boldsymbol{y}) = \begin{cases} (3\pi)^{-1}(1 - \boldsymbol{y}^2)^2, & \text{if } \boldsymbol{y}^2 < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

One advantage of equation (5) is its faster calculation than the regular kernel. As the formula indicates, any activity beyond the bandwidth h from the centroid of the considered cell does not contribute to the summation.

For our study area, the raster framework is a rectangular region R with 2 771 956 grid cells, and each cell is a 10 m $\times$ 10 m square. Experimenting with other cell sizes yielded similar results. The following explains how the distributions of retail shops and services and the centrality indices of street networks are both transformed to this raster system by the KDE method.

4.3 Estimating kernel densities of retail and service activities and street centralities

As discussed in section 2, retail and service facilities are represented as points in a GIS system. Such points were also provided with information about the size of the shop or service, but we decided in this first application to correlate just with the *presence* of shops and services—that is, their broad numer and location—and to leave more detailed analysis to further investigations: therefore facilities were not weighted here. ArcGIS has a built-in tool for kernel estimation. To access the tool in ArcGIS, click the Spatial Analyst dropdown arrow > Density > choose Kernel for Density Type in the dialogue. Applying the tool to the dataset of retail and service activities yielded the kernel densities. For this study we computed the densities of retail shops alone, and also the densities of retail (commercial) and service activities together in Bologna.

To compute the kernel densities of the street network we used centrality values for each street segment (edge) to weigh the contribution of each edge to the kernel 'bump' within a grid cell. In other words, a kernel function is applied to each street, yielding values which are greatest on the line, diminish with distance from the line, and reach 0 at the distance h from the line. Unlike the densities of retail or service activities, the kernel density of street centrality at each grid cell in region R is the sum of all the kernel surfaces within the bandwidth multiplied by the value of centrality at each surface. In ArcGIS, this is implemented by selecting one of the centrality indices as the 'population' (weight) field. By doing so, we are not computing just the density of streets, but the *density of street centrality.*

One problem in using the KDE is the choice of a particular kernel function and bandwidth h. Several methods have been proposed to pick up the best kernel function (Fotheringham et al, 2000, pages 155–157) or to optimize h (Cao et al, 1994) according to the global structure of the dataset. However, Epanechnikov (1969) finds that the choice among the various kernel functions does not affect significantly the outcomes of the process. Williamson et al (1998) and Levine (2004) point out that the choice of bandwidth is an important issue in any KDE application. Recent advancement has suggested using an *adaptive*, rather than *fixed*, bandwidth h—that is to say, h is larger in areas where events are sparser, and smaller where they are denser (Brunsdon, 1995; Fotheringham et al, 2002).

As explained earlier, the KDE is not the methodological focus of this research, and is used here to transform the two data features to the same analysis unit. Our study used a fixed bandwidth, but we did experiment with different h values to show the robustness of the results. The choice of a fixed bandwidth pertains to the purpose of the study: we are interested in understanding the relationship between the street network and basic services in an ordinary city. In Bologna we chose h = 300 m, 200 m

Table 1. Twenty-one raster layers from kernel density estimation (KDE) in Bologna. MCA denotes multiple centrality assessment.

Layer	Centralities			KDE
	index	description	MCA distance factor $d(m)$	bandwidth distance factor $h(m)$
1	C^{B}_{glob}	global betweenness centrality	all	300
2	C^{C}_{glob}	global closeness centrality	all	300
3	C^{S}_{glob}	global straightness centrality	all	300
4	C^{B}_{glob}	global betweenness centrality	all	200
5	C^{C}_{glob}	global closeness centrality	all	200
6	C^{S}_{glob}	global straightness centrality	all	200
7	C^{B}_{glob}	global betweenness centrality	all	100
8	C^{C}_{glob}	global closeness centrality	all	100
9	C^{S}_{glob}	global straightness centrality	all	100
10	C^{C}_{800}	local closeness centrality	800	300
11	C^{S}_{800}	local straightness centrality	800	300
12	C^{C}_{800}	local closeness centrality	800	200
13	C^{S}_{800}	local straightness centrality	800	200
14	C^{C}_{800}	local closeness centrality	800	100
15	C^{S}_{800}	local straightness centrality	800	100
16	comm + serv	retail commerce and commercial service		300
17	comm	retail commerce		300
18	comm + serv	retail commerce and commercial service		200
19	comm	retail commerce		200
20	comm + serv	retail commerce and commercial service		100
21	comm	retail commerce		100

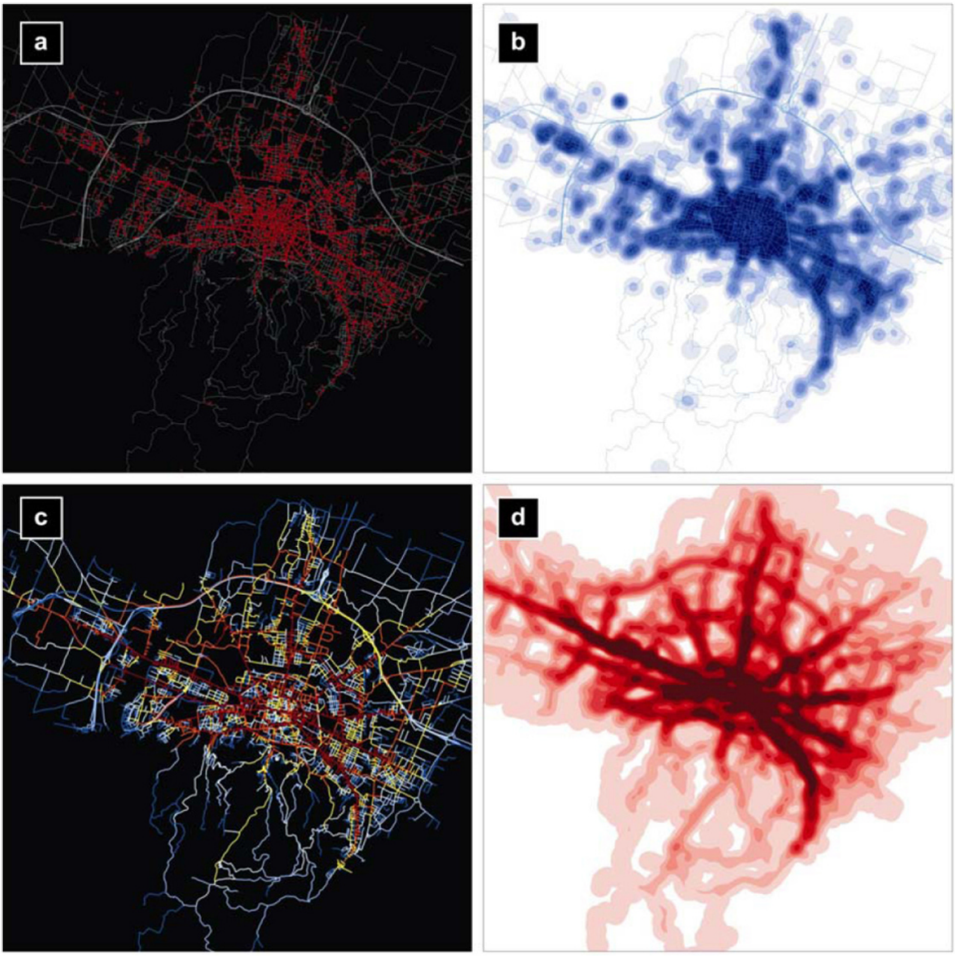

Figure 2. [In colour online, see http://dx.doi.org/10.1068/b34098] Density of activity and street centrality: (a) location of commercial and service activities (red dots); (b) kernel density estimation (KDE) (h = 300 m) of commercial and service activities; (c) street global betweenness C^{B}_{glob} (blue for lower values and red for higher); (d) KDE (h = 300 m) of C^{B}_{glob}.

and 100 m, which are widely used in urban planning and design to model the pedestrian catchment area at the scale of *neighbourhood*, *block*, and *street*, respectively (Calthorpe and Fulton, 2001; Cervero, 1998; 2004; Frey, 1999; Urban Task Force, 1999).

In summary, the KDE analysis was performed on: (1) two point layers (one is retail shops and the other is retail and service activities altogether) using no weights, and (2) the street layer using five different weights (C^{C}_{glob}, C^{B}_{glob}, C^{S}_{glob}, C^{C}_{800}, and C^{S}_{800}). For each of these seven measures, three bandwidths (h = 100 m, 200 m, and 300 m) were tested, yielding a total of twenty-one kernel density raster layers (see table 1). For example, figure 2(b) shows the KDE of retail and service activities, and figure 2(d) shows the KDE of global betweenness (C^{B}_{glob}) of the street network in Bologna. Both used a bandwidth of 300 m.

5 Analyzing the relationship between street centralities and retail and service activities

The above KDE analysis converted the measures of street centralities and the densities of retail and service activities to the same raster framework, where each cell contains attributes in multiple raster layers (see figure 3) for an illustration). In this section we

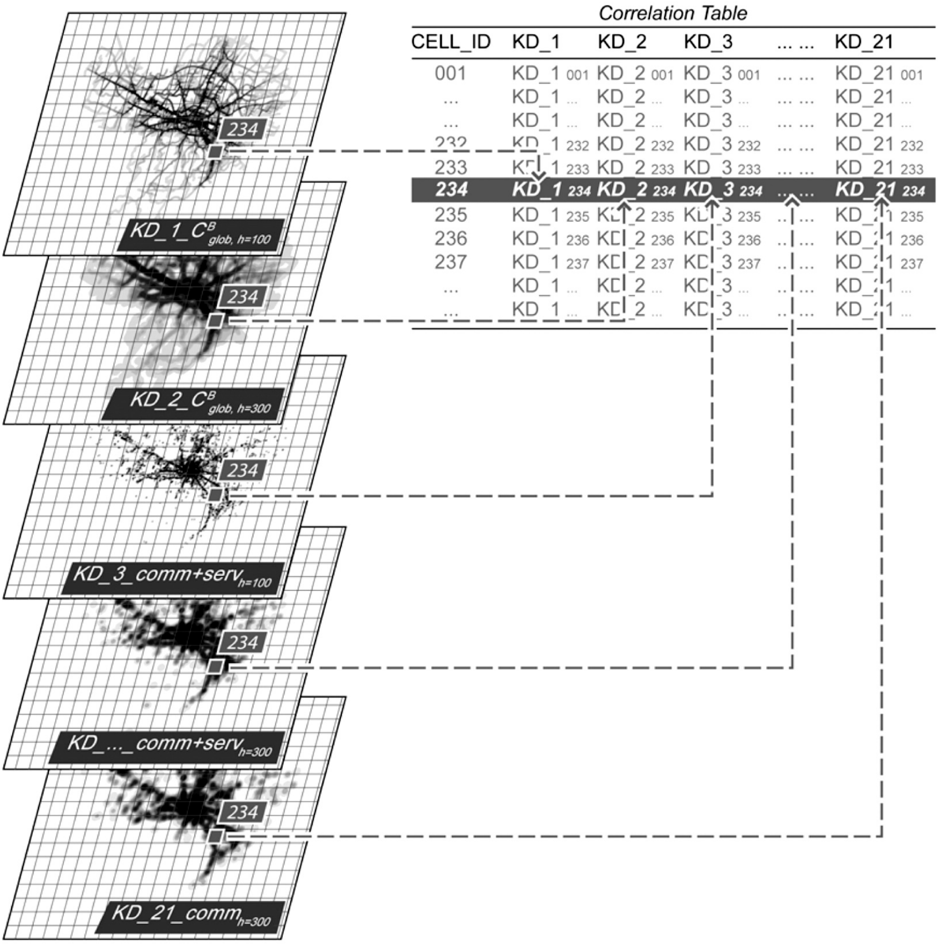

Figure 3. Illustration of a grid cell with attributes in various raster layers. KD denotes kernel density.

evaluate the statistical distribution of each of the density measures in the raster layers, and analyze the relationship between the densities of street centralities and the retail and service activities. For the latter, one of the two activity variables is paired with each of the five street centrality variables to examine their relationship—that is, $2 \times 5 = 10$ pairs of relationships. Only variables calculated at the same bandwidth can be matched—for example, the density of $C^{\mathrm{C}}_{\mathrm{glob}}$ with $h = 300$ m matches with that of retail shops with $h = 300$ m; with three bandwidths tested, this led us to examine thirty correlations.

5.1 Density distributions of street centrality and commercial and service activities

Figure 4 shows the distribution of density of activities for the case $h = 300$ m. Namely, we report the number of cells with a density of activities (in this case, commercial and service activities) in the range $[a, a + \Delta a]$. The increment was set to be equal to $\Delta a = 3.396 \times 10^{-5}$. The distribution follows closely the power law, which is a straight line in the log–log scale used in the plot. That is to say, as the density declines, the number of cells increases geometrically. Higher density areas are few, and lower density areas are plentiful. Results for the density of retail shops and other bandwidths are very similar to figure 4, and are thus not shown here.

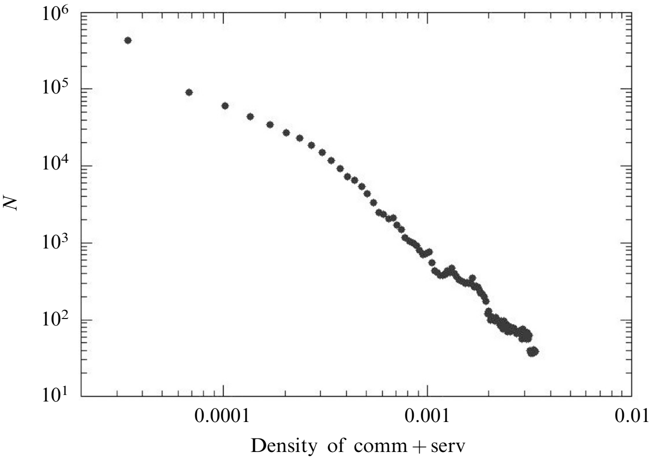

Figure 4. Distribution of kernel density of retail commerce (comm) and commercial service (serv) activities ($h = 300$ m) in Bologna.

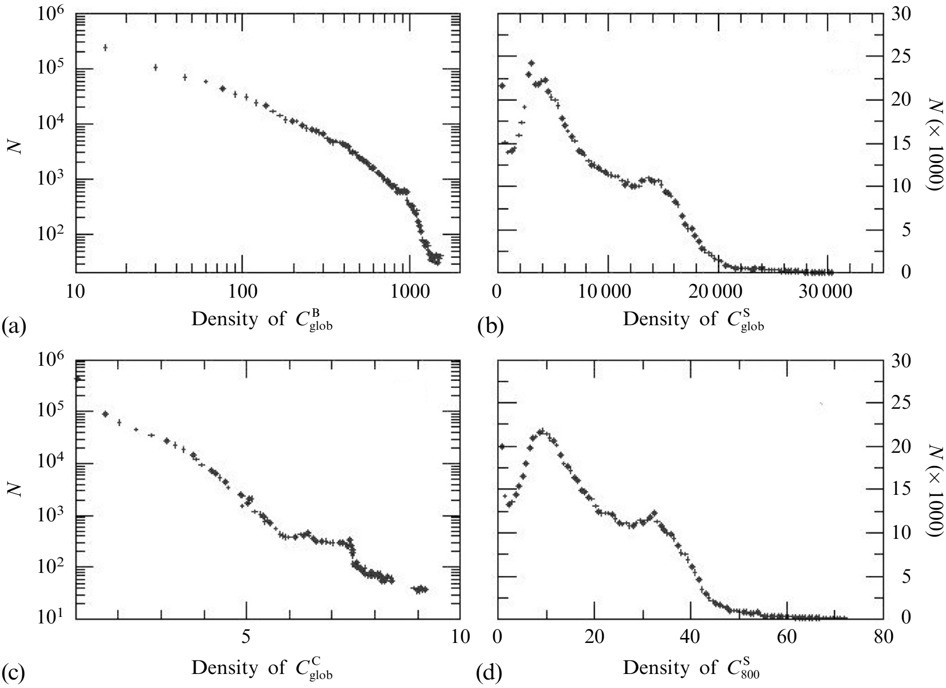

Figure 5. Distribution of kernel densities of street centrality ($h = 300$ m) in Bologna.

We now examine how many cells have a density range of a centrality index c. Figures 5(a) – 5(d) show only four of the fifteen centrality indices calculated by MCA, namely global betweenness [$C^{\mathrm{B}}_{\mathrm{glob}}$, in figure 5(a)], global straightness [$C^{\mathrm{S}}_{\mathrm{glob}}$ in figure 5(b)], global closeness [$C^{\mathrm{C}}_{\mathrm{glob}}$ in figure 5(c)], and local closeness evaluated within 800 m [C^{C}_{800} in figure 5(d)]. In all four cases the KDE bandwidth was $h = 300$ m. Defining the density range as $[c, c + \Delta c]$, the increment $\Delta c = 45\,310$ for $C^{\mathrm{B}}_{\mathrm{glob}}$, $\Delta c = 476\,902$ for $C^{\mathrm{S}}_{\mathrm{glob}}$, and $\Delta c = 1767$ for both $C^{\mathrm{C}}_{\mathrm{glob}}$ and C^{C}_{800}. All centrality measures exhibit significant variations across Bologna. Such a wide range of difference in centrality reflects the typical concentric structure of the medieval street network.

The concentration of best centrality in a very few places indicates that the best accessibility to the functional backbone of community life is enjoyed only by a small segment of the population or businesses. This appears to be the natural outcome of a historical evolution of Bologna's street network. Among the four centrality densities, the distribution of $C^{\mathrm{B}}_{\mathrm{glob}}$ [shown in figure 5(a)] is the only graph in a log–log scale. It resembles most that of the commercial and service activities, and can be fitted by a power law. This indicates that the distribution of betweenness is the most heterogeneous one: in fact, the density values of $C^{\mathrm{B}}_{\mathrm{glob}}$ span from 10 to over 1000, with over 100 000 cells having a density equal to 10 and only a few hundred cells having a density larger than 1000. The distributions of the other three centralities ($C^{\mathrm{S}}_{\mathrm{glob}}$, $C^{\mathrm{C}}_{\mathrm{glob}}$, and C^{C}_{800} exhibit more uniform patterns with rapidly declining tails. Note that for $C^{\mathrm{C}}_{\mathrm{glob}}$ [shown in figure 5(c)], the vertical axis is a log scale and the horizontal axis is a linear scale. Therefore, the distribution of $C^{\mathrm{C}}_{\mathrm{glob}}$ is well approximated by a decreasing exponential curve. For either $C^{\mathrm{S}}_{\mathrm{glob}}$ or C^{C}_{800} both the axes are plotted in a linear scale, and distributions are characterized by two peaks. A similar pattern with the presence of two peaks has also been observed for straightness and closeness indices evaluated locally but on different scales.

5.2 Correlating street centrality and commercial or service activities

The main hypothesis of this study is that centrality acts as a driving force in the formation and constitution of urban structure in terms of land uses such as commercial and service activities. Here the relationship is examined by analyzing the correlations between the densities of street centrality and retail or service activities in the same raster framework. As explained earlier, the study involves examining thirty paired correlations of kernel densities based on the same bandwidth. The methodology shares some common ground with an earlier study by Thurstain-Goodwin and Unwin (2000) that combined estimated density values on a cell-by-cell basis in a correlation of the same factors.

As our focus is on the developed areas along the street network, we first excluded cells of zero values in either of the coupled raster layers. This reduced the number of cells for analysis from 2 771 956 to 1 500 000 – 1 800 000. For each pair of variables, a simple linear correlation coefficient (or Pearson's r) was computed. Pearson's r, ranging from -1 to 1, determines the extent to which values of the two variables are 'proportional' to each other. In general, the value of Pearson's r decreases as the sample size increases, owing to statistical fluctuations (Taylor, 1982).

The emphasis of this research is on examining how the street network centrality and retail and service density are related to each other. It is not our objective to explain, one factor by another, which would require rigorous regression analysis. Spatial data (including those in this study) often exhibit spatial autocorrelation, in which values of a variable are systematically related to geographic location (eg similar or dissimilar values are near each other). In that case spatial regression models such as the spatial lag model or the spatial error model (Fotheringham et al, 2000, pages 167 – 169; Wang, 2006, pages 181 – 185) ought to be used, and more advanced techniques such as the maximum likelihood method (Anselin and Bera, 1998) are needed to calibrate the models. This will be explored in our future work, which may consider more factors in addition to the centrality measures.

The analysis in Bologna shows that all thirty paired correlations are positive. A very strong positive association emerges, especially considering the large sample size. Among the fifteen pairs with the highest r (see table 2), all the r values are higher than 0.5. In table 2 six of the correlations between the global betweenness centrality ($C^{\mathrm{B}}_{\mathrm{glob}}$) and retail activities are among the top fifteen; so are five of the correlations between the

Table 2. Top fifteen Pearson correlations between kernel densities of street centralities and commercial and service activities in Bologna. KDE denotes kernel density estimations. comm and serv denote retail commerce and commercial services, respectively.

Rank	Correlated variables		KDE bandwidth (m)	Linear correlation (Pearson's r)
	centralities	activities		
1	$C^{\mathrm{B}}_{\mathrm{glob}}$	comm + serv	300	0.727
2	$C^{\mathrm{B}}_{\mathrm{glob}}$	comm	300	0.704
3	$C^{\mathrm{B}}_{\mathrm{glob}}$	comm + serv	200	0.673
4	$C^{\mathrm{B}}_{\mathrm{glob}}$	comm	200	0.653
5	$C^{\mathrm{C}}_{\mathrm{glob}}$	comm	300	0.641
6	C^{S}_{800}	comm + serv	300	0.620
7	$C^{\mathrm{S}}_{\mathrm{glob}}$	comm + serv	300	0.615
8	$C^{\mathrm{C}}_{\mathrm{glob}}$	comm + serv	300	0.608
9	$C^{\mathrm{C}}_{\mathrm{glob}}$	comm + serv	200	0.583
10	$C^{\mathrm{B}}_{\mathrm{glob}}$	comm + serv	100	0.567
11	C^{C}_{800}	comm + serv	300	0.565
12	$C^{\mathrm{B}}_{\mathrm{glob}}$	comm	100	0.555
13	$C^{\mathrm{C}}_{\mathrm{glob}}$	comm	200	0.547
14	$C^{\mathrm{S}}_{\mathrm{glob}}$	comm + serv	200	0.546
15	$C^{\mathrm{C}}_{\mathrm{glob}}$	comm	300	0.533

global closeness centrality ($C^{\mathrm{C}}_{\mathrm{glob}}$) and retail activities. However, the global straightness centrality ($C^{\mathrm{S}}_{\mathrm{glob}}$) is only present twice in the table. Global betweenness $C^{\mathrm{B}}_{\mathrm{glob}}$ emerges to have by far the strongest correlation with retail activities at all scales, with r values above 0.7. Interestingly, betweenness centrality captures the potential of a place to be traversed by passenger or freight trips between other places in the system, or 'through-traffic', even if it does not serve as an origin or destination. In other words, a place itself may not attract people or cargo as a major trip destination, but it may take advantage of its unique location as merely a pass-through nexus to generate great business opportunities. Hence a high value of betweenness centrality often implies a high concentration of commercial or service activities.

Figure 6 plots the densities of retail shops and services versus centrality. Like the histograms in figure 5, each graph in figure 6 is drawn on the basis of the average densities of retail and service activities within each range of the street centrality. The graphs indicate that higher values of centrality generally correspond to higher average densities of retail and service activities. The pattern is clear for global betweenness [figure 6(a)] and global closeness [figure 6(c)], particularly for $C^{\mathrm{B}}_{\mathrm{glob}}$, which has an almost linear relationship. Note the inset of figure 6(a) (using a bandwidth $h = 100$ m for comparison) showing greater fluctuations. With regard to the global straightness [figure 6(b)] and the local closeness [figure 6(d)], the curves are flat in the lower range of centrality values and then climb up a steep slope when the centrality values are higher. In other words, in terms of global straightness or local closeness, concentrations of activities are slow in responding to the improving centrality *initially*, but get more intense after the centrality passes a certain threshold.

In summary, the results from the correlation analysis in the city of Bologna support the hypothesis that street network centrality, measured by the MCA model, acts as a driving force in the formation and constitution of urban structure, as reflected in the distribution of commercial and service activities at the neighbourhood level.

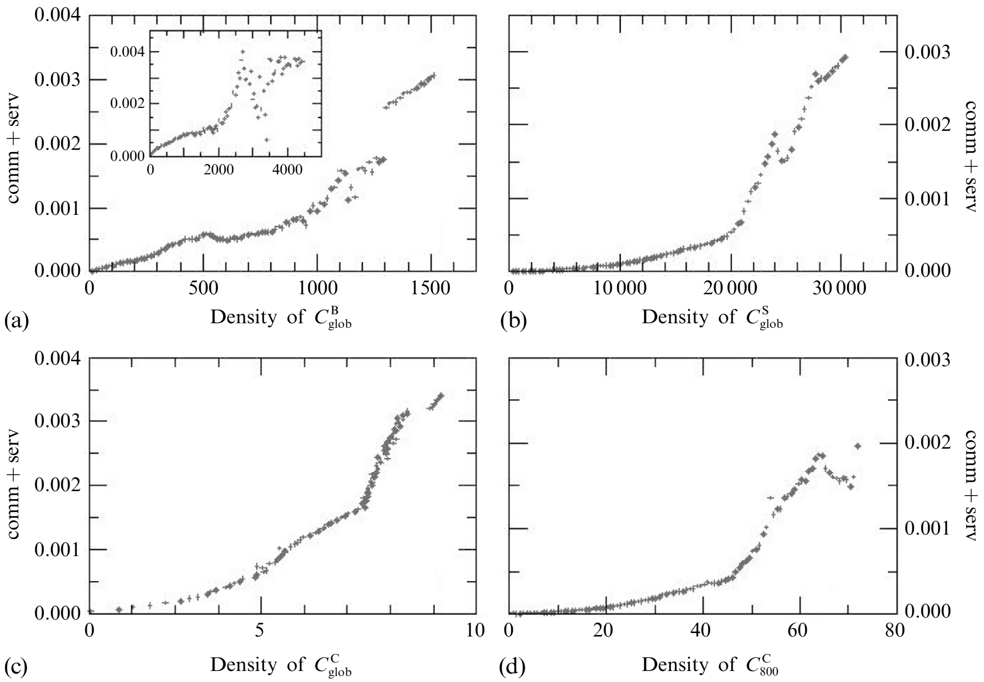

Figure 6. Correlations between kernel densities of street centrality and retail commerce and commercial service activities (comm and serv, respectively) ($h = 100$ m for inset and $h = 300$ m in all other cases).

6 Conclusions

This paper examines the relationship between street centrality and land-use intensity in the city of Bologna, northern Italy. Street centrality is calibrated in a multiple centrality assessment model composed of multiple measures such as closeness, betweenness, and straightness. In this study land-use intensity is measured in terms of ground floor commercial and service activities. Kernel density estimation is used to transform both the centrality measures and the densities of retail and services to one scale unit, which permits the correlation analysis between them.

Results indicate that retail and service activities in Bologna tend to concentrate in the areas that enjoy better centralities. The distribution of these activities correlates highly with the global betweenness of the street network, and also with the global closeness, but to a slightly lesser extent. This confirms the hypothesis that street centrality plays a crucial role in shaping the formation of urban structure and land uses. Betweenness centrality captures the potential of a place to be traversed or to attract 'through-traffic'. That is to say, a place itself inside a city may not need to be a major trip destination but merely a pass-through nexus in order to generate great business opportunities.

The study also suggests that the multiple centrality assessment model be an effective tool for mapping street centrality as a fundamental property in a city—that is, 'location advantage'—and thus can serve as a very useful guide for urban planning and design. By analyzing the statistical distributions both of densities of centrality and of activities, it is observed that values of betweenness centrality and the density of activities follow a power function, which indicates that very few places enjoy the best location and thus command a high concentration of business in contrast to the vast majority of urban space. Such a distribution pattern is found in most self-organized complex systems in nature, technology, and society, implying that cities may be a self-organized system following some 'organic' order in their evolution over time.

References

Akaike H, 1954, "An approximation to the density function" *Annals of the Institute of Statistical Mathematics* **6** 127 – 132

Anselin L, Bera A, 1998, "Spatial dependence in linear regression models with an introduction to spatial econometrics", in *Handbook of Applied Economic Statistics* Eds A Ullah, D E Giles (Marcel Dekker, New York) pp 237 – 289

Anselin L, Cohen J, Cook D, Gorr W, Tita G, 2000, "Spatial analysis of crime", in *Criminal Justice 2000 Volume 4: Measurement and Analysis of Crime and Justice* Eds R Kaminski, N La Vigne (US Department of Justice, Washington, DC) pp 213 – 262

Bailey T, Gatrell T, 1995 *Interactive Spatial Data Analysis* (Addison-Wesley, Edinburgh)

Batty M, 2005, "Network geography: relations, interactions, scaling and spatial processes in GIS", in *Re-presenting GIS* Eds P Fisher, D Unwin (John Wiley, New York) pp 149 – 170

Bavelas A, 1948, "A mathematical model for group structures" *Human Organization* **7** 16 – 30

Bavelas A, 1950, "Communication patterns in task oriented groups" *Journal of the Acoustical Society of America* **22** 725 – 730

Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D-U, 2006, "Complex networks: structure and dynamics" *Physics Report* **424** 175 – 308

Borruso G, 2003, "Network density and the delimitation of urban areas" *Transactions in GIS* **7** 177 – 191

Brunsdon C, 1995, "Estimating probability surfaces for geographical point data: an adaptive kernel algorithm" *Computers and Geosciences* **21** 877 – 894

Calthorpe P, Fulton W, 2001 *The Regional City: Planning for the End of Sprawl* (Island Press, Washington, DC)

Cao R, Cuevas A, González-Manteiga W, 1994, "A comparative study of several smoothing methods in density estimation" *Computational Statistics and Data Analysis* **17** 153 – 176

Cardillo A, Scellato S, Latora V, Porta S, 2006, "Structural properties of planar graphs of urban street patterns" *Physical Review E* **73** 0661071-8

Cervero R, 1998 *The Transit Metropolis* (Island Press, Washington, DC)

Cervero R, 2004 *Developing Around Transit: Strategies and Solutions That Work* (Urban Land Institute, Washington, DC)

Conroy-Dalton R, 2003, "The secrete is to follow your nose: route path selection and angularity" *Environment and Behavior* **35** 107 – 131

Crucitti P, Latora V, Porta S, 2006a, "Centrality measures in spatial networks of urban streets" *Physical Review E* **73** 0361251-5

Crucitti P, Latora V, Porta S, 2006b, "Centrality in networks of urban streets" *Chaos* **16** 0151131-9

Epanechnikov V, 1969, "Nonparametric estimation of a multivariate probability density" *Theory of Probability and its Applications* **14** 153 – 158

Fotheringham A S, Brunsdon C, Charlton M, 2000 *Quantitative Geography: Perspectives on Spatial Data Analysis* (Sage, London)

Fotheringham A S, Brunsdon C, Charlton M, 2002 *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships* (John Wiley, New York)

Freeman L, 1977, "A set of measures of centrality based on betweenness" *Sociometry* **40** 35 – 41

Freeman L, 1979, "Centrality in social networks: conceptual clarification" *Social Networks* **1** 215 – 239

Frey H, 1999 *Designing the City: Towards a More Sustainable Urban Form* (Taylor and Francis, London)

Goulias K, 2002 *Transportation Systems Planning: Methods and Applications* (CRC Press, Boca Raton, FL)

Hillier B, 1996 *Space is the Machine: A Configurational Theory of Architecture* (Cambridge University Press, Cambridge)

Hillier B, Hanson J, 1984 *The Social Logic of Space* (Cambridge University Press, Cambridge)

Jacobs A, 1993 *Great Streets* (MIT Press, Cambridge, MA)

Jacobs J, 1961 *The Death and Life of Great American Cities* (Random House, New York)

Levine M, 2004 *CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations (Version 3.0)* (National Institute of Justice, Washington, DC)

Meyer M, Miller E, 2000 *Urban Transportation Planning* (McGraw-Hill, Columbus, OH)

Newman P, Kenworthy J, 1999 *Sustainability and Cities: Overcoming Automobile Dependence* (Island Press, Washington, DC)
Parzen E, 1962, "On estimation of a probability density function and mode" *Annals of Mathematical Statistics* **33** 1065 – 1076
Penn A, Turner A, 2003, "Space layout affects search efficiency for agents with vision" *Environment and Behaviour* **35** 30 – 65
Porta S, Crucitti P, Latora V, 2006a, "The network analysis of urban streets: a dual approach" *Physica A* **369** 853 – 866
Porta S, Crucitti P, Latora V, 2006b, "The network analysis of urban streets: a primal approach" *Environment and Planning B: Planning and Design* **33** 705 – 725
Rosenblatt F, 1956, "Remarks on some nonparametric estimates of a density function" *Annals of Mathematical Statistics* **27** 832 – 837
Scellato S, Cardillo A, Latora V, Porta S, 2006, "The backbone of a city" *The European Physical Journal B* **50** 221 – 225
Scheurer J, Porta S, 2006, "Centrality and connectivity in public transport networks and their significance for transport sustainability in cities", paper presented at the World Planning Schools Congress, Mexico City, 13 – 16 July; copy available from S Porta
Scheurer J, Curtis C, Porta S, 2007, "Spatial network analysis of public transport systems: developing a strategic planning tool to assess the congruence of movement and urban structure in Australian cities", paper presented at the Australasian Transport Research Forum, Melbourne, 25 – 27 September; copy available from S Porta
Silverman B W, 1986 *Density Estimation for Statistics and Data Analysis* (Chapman and Hall, London)
Taneja S, 1999, "Technology moves in" *Chain Store Age* **75** (May) 136 – 138
Taylor J, 1982 *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements* (University Science Books, Mill Valley, CA)
Thurstain-Goodwin M, Unwin D, 2000, "Defining and delimiting the central areas of towns for statistical modelling using continuous surface representations" *Transactions in GIS* **4** 305 – 317
Tobler W R, 1970, "A computer movie simulating urban growth in the Detroit region" *Economic Geography* **46** 234 – 240
Urban Task Force, 1999 *Towards an Urban Renaissance* (E&FN Spon, London)
Vragovìc I, Louis E, Diaz-Guilera A, 2004, "Efficiency of informational transfers in regular and complex networks", http://orxiv.org/abs/cond-mat/0410174
Wang F, 2006 *Quantitative Methods and Applications in GIS* (CRC Press, Boca Raton, FL)
Wasserman S, Faust K, 1994 *Social Networks Analysis* (Cambridge University Press, Cambridge)
Williamson D, McLafferty S, Goldsmith V, Mollenkopf J, McGuire P, 1998, "Smoothing crime incident data: new methods for determining the bandwidth in Kernel estimation", paper presented at the 18th ESRI International User Conference, San Diego, 27 – 31 July; copy available at http://gis.esri.com/library/userconf/proc98/PROCEED.HTM
Wilson G, 2000 *Complex Spatial Systems: The Modelling Foundations of Urban and Regional Analysis* (Prentice-Hall, Upper Saddle River, NJ)

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 38 2009

www.perceptionweb.com