

# UoDSU



# disclaimer

## **Article:**

Wang F, Antipova A, Porta S, 2010, Street centrality and land use intensity in Baton Rouge, Louisiana, *Journal of Transport Geography*, In press.

## **Disclaimer:**

This paper not necessarily reflects the final definitive publication: it might be a pre-copy-editing or a post-print author-produced .pdf or in any case a different version of that. Therefore the reader is advised to refer to the publishing house's archive system for the original authenticated version of this paper.



Contents lists available at ScienceDirect

## Journal of Transport Geography

journal homepage: [www.elsevier.com/locate/jtrangeo](http://www.elsevier.com/locate/jtrangeo)

## Street centrality and land use intensity in Baton Rouge, Louisiana

Fahui Wang<sup>a,\*</sup>, Anzhelika Antipova<sup>a</sup>, Sergio Porta<sup>b</sup><sup>a</sup> Department of Geography & Anthropology, Louisiana State University, Baton Rouge, LA 70803, USA<sup>b</sup> Department of Architecture, University of Strathclyde, Glasgow, UK

## ARTICLE INFO

## Keywords:

Street centrality  
 Land use intensity  
 Kernel density estimation  
 Floating catchment area method

## ABSTRACT

This paper examines the relationship between street centrality and land use intensity in Baton Rouge, Louisiana. Street centrality is calibrated in terms of a node's closeness, betweenness and straightness on the road network. Land use intensity is measured by population (residential) and employment (business) densities in census tracts, respectively and combined. Two GIS-based methods are used to transform data sets of centrality (at network nodes) and densities (in census tracts) to one unit for correlation analysis. The kernel density estimation (KDE) converts both measures to raster pixels, and the floating catchment area (FCA) method computes average centrality values around census tracts. Results indicate that population and employment densities are highly correlated with street centrality values. Among the three centrality indices, closeness exhibits the highest correlation with land use densities, straightness the next and betweenness the last. This confirms that street centrality captures location advantage in a city and plays a crucial role in shaping the intraurban variation of land use intensity.

© 2010 Elsevier Ltd. All rights reserved.

## 1. Introduction

The interrelatedness of transportation and settlement has been a constant theme of theoretical and empirical inquiries, both in intra-urban and interurban settings. The focus of this paper is on their interdependence in an intraurban context. The classic *economic model* proposed by Mills (1972) and Muth (1969), often referred to as the monocentric model, assumes that all employment is concentrated in the city center. Intuitively, as everyone commutes to the city center for work, a household farther away from the center spends more on commuting and is compensated by living in a larger-lot house (also cheaper in terms of price per area unit). The resulting population density exhibits a declining pattern with distance from the city center. The monocentric model, with roots traced back to the famous “Chicago School” (Park et al., 1925), emphasizes the role of central business district (CBD) on the differentiation of citywide land use intensity. Since the 1970s, more and more researchers recognize the changing urban form from *monocentricity* to *polycentricity* (Ladd and Wheaton, 1991; Berry and Kim, 1993; Hoch and Waddell, 1993). In addition to the major center in the CBD, most large cities have secondary centers or subcenters, and thus are better characterized as polycentric cities. In a polycentric city, assumptions of whether residents need to access all centers or some of the centers lead to various population density functions (Heikkila et al., 1989; Wang, 2006, pp. 107–109).

The economic model is “simplification and abstraction that may prove too limiting and confining when it comes to understanding

and modifying complex realities” (Casetti, 1993, p. 527), but highlights the important role of transportation costs in shaping urban structure. Urban planners and geographers tend to be more cognizant of the complexity of urban structure, and have developed some models to capture the relationship between urban transportation networks and land use patterns. Lowry (1964) and Garin (1966) developed a model that emphasizes the interactions between population and employment distributions. The interactions between employment and population decline with distances, which are defined by a transportation network. The model has the flexibility of simulating population and employment distribution patterns corresponding to a given road network. Wang (1998) used the model to examine how the population and employment density patterns respond to changes in transportation networks. However, the *Garin–Lowry model* relies on the economic base theory that divides employment into basic vs. nonbasic employment, and requires the basic employment pattern as an input. Division of economy (employment) into basic (those independent of local economy) and nonbasic (e.g., service) sectors is not straightforward and often infeasible in practice. Wang and Guldmann (1996) proposed a gravity-based model to simulate urban densities (in general, perhaps a combination of population and employment densities) given a road network. The model assumes that density at a location is proportional to its accessibility to all other locations in a city, measured as a gravity potential. It emphasizes that location determined by the road network is the force that shapes the variation of land use intensity. Like any gravity models, the *Wang–Guldmann model* needs a value for the distance friction coefficient, which is not conveniently available and requires additional data and model calibration to derive.

\* Corresponding author. Fax: +1 225 578 4420.

E-mail address: [fwang@lsu.edu](mailto:fwang@lsu.edu) (F. Wang).

Most recently, urban planners have benefited from advancement of network science (Barabási, 2002; Batty, 2005, 2008). A street network can be characterized by a well-defined geometric structure consisting of nodes (points) and segments (lines). Miller (1999) developed a space–time accessibility measure based on a transportation network. Borruso (2003, 2005, 2008) used a simple road network density index to approximate land use intensity and thus delineate urban areas. Downs and Horner (2007a,b) used network-based measures in characterizing linear point patterns and estimating home ranges. Okabe et al. (2006a,b) developed a toolbox for network-based spatial analysis. With analogies to space syntax, Porta et al. (2006) implemented several indices to measure street systems and their connectivity. These network-based measures capture location advantage of various places in a city in terms of connectivity between opportunities or activities. This family of indices, grouped as the *multiple centrality assessment* (MCA) model, is used to explain densities of commercial and service activities in a city (Porta et al., 2009). The MCA model defines centrality of a place not only as being central in terms of closeness (proximity) to other places but also being “intermediary, straight ... and critical” to others. Therefore, the model is a more comprehensive assessment of location than the traditional gravity-based accessibility measures (Hansen, 1959), which were found to be good predictors of land development rate.

This paper will examine the relationship between street centrality measures and land use intensity. Street centrality is calibrated in terms of a node’s closeness, betweenness and straightness on the road network. Land use intensity is measured by population and employment densities in census tracts, respectively and combined. The population (evening-time) density pattern reflects the variation of residential land use. Employment (daytime) density captures business-related land uses including industrial, commercial and others. Population and employment distributions are more general and comprehensive indicators of land use intensity than service and commercial activities used in Porta et al. (2009). Such data are available and publicly accessible for any US metropolitan areas through the Census Transportation Planning Package (CTPP), and thus permit empirical studies in any other major cities in the US. Two GIS-based methods are used to transform data sets of centrality (at network nodes) and densities (in tracts) to one unit for correlation analysis. The kernel density estimation (KDE) converts both measures to raster pixels, and the floating catchment area (FCA) method computes average centrality values around census tracts. The study area for this research is Baton Rouge, Louisiana, with a total population of 430,770 and employment of 224,550 in 2000.

Major contributions made by this research may be summarized in three aspects:

- (1) The study supports the proposition that urban land use intensity, in terms of population and employment densities, is shaped by the street network instead of proximity to city center(s) as suggested by urban economic models.
- (2) Urban location is well captured by centrality metrics, some of which correlate with land use intensity better than others.
- (3) Either the kernel density estimation or the floating catchment area method enables the transformations of network centrality and land use intensity to the same data frame and permits the examination of their relationship.

## 2. Study area and data sources

Fig. 1 shows the study area, East Baton Rouge Parish of Louisiana and its surroundings (parish is a county unit in Louisiana). It comprises an area of 455.43 square miles including the City of Ba-

ton Rouge (i.e., the state capital) in the middle, a dozen satellite towns around it and surrounding unincorporated areas, with about 1809 miles of roadway. The Mississippi River separates the Parish from West Baton Rouge Parish on the west, and the Amite River runs between the Parish and the Livingston Parish on the east. Both rivers are the natural borders of the study area on the west and on the east. Unincorporated rural areas in the northwest, northeast and south serve as the buffer between the urbanized areas (occupying much of the central area) and neighboring parishes. Therefore, the area is fairly complete with minimal edge effects. Edge effects refer to instability or unreliability of conclusions drawn from a study area when bordering areas are included or excluded. Hereafter the study area is simply referred to as Baton Rouge.

The population and employment data were extracted from the Census Transportation Planning Package (CTPP) 2000 data sets. The CTPP data were downloaded from the Bureau of Transportation Statistics website ([www.bts.gov](http://www.bts.gov)). The CTPP data were compiled by local metropolitan planning agencies, in this case the Capital Region Planning Commission in Baton Rouge. Specifically, the 2000 CTPP Part 1 provides data similar to traditional census data by area of residence, and the 2000 CTPP Part 2 provides data by area of work (unique among all census products) such as the number of jobs (employment). We extracted the population data from the US Census Summary File 1 (SF1), and the employment data from the CTPP Part 2, both at the census tract level. The spatial GIS data of the study area (i.e., census tracts and road network) came from the Environmental Systems Research Institute web site at [www.esri.com/data/download/census2000\\_tigerline/index.html](http://www.esri.com/data/download/census2000_tigerline/index.html).

## 3. Research methods

### 3.1. Data preparation

This study uses a *primal* approach for the road network representation. In contrast to the dual representation that disregards the metric nature of a street network, the primal approach represents intersections as nodes and road segments as links (or edges) with lengths that connect nodes (see Jiang and Claramunt, 2004). The original road network file for the study area had 14,935 nodes and 19,892 links. In order to simplify the network structure and minimize computation time, we used some editing tools EditTools 3.6 for ArcView 3.x from the ET SpatialTechniques ([www.ian-ko.com](http://www.ian-ko.com)) to form a road network consisting of nodes that are truly the starting and end points of a street segment and links between these nodes. The resulting network has 12,235 nodes and 17,219 links with attributes such as updated link IDs, lengths, two end nodes’ IDs and their corresponding coordinates.

As explained earlier, population and employment densities are used as proxies for land use intensity. The study area has 89 census tracts. The CTPP data were processed and integrated into one GIS layer containing population and employment information for the census tracts. In addition to population density and employment density per square kilometer separately, this study developed a *combined density* to integrate employment (daytime) and population (evening-time) densities together for an overall measure of land use intensity. Considering only a fraction of residents participate in the labor force, one resident is discounted by the citywide *labor participation ratio* (LPR) to be comparable to one job. In other words,

$$\text{combined density} = \text{employment density} + \text{population density} \times \text{LPR}$$

where LPR in our study area is 0.5213 (i.e., ratio of total employment 224,550 vs. total population 430,770). Fig. 2a–c show the pop-

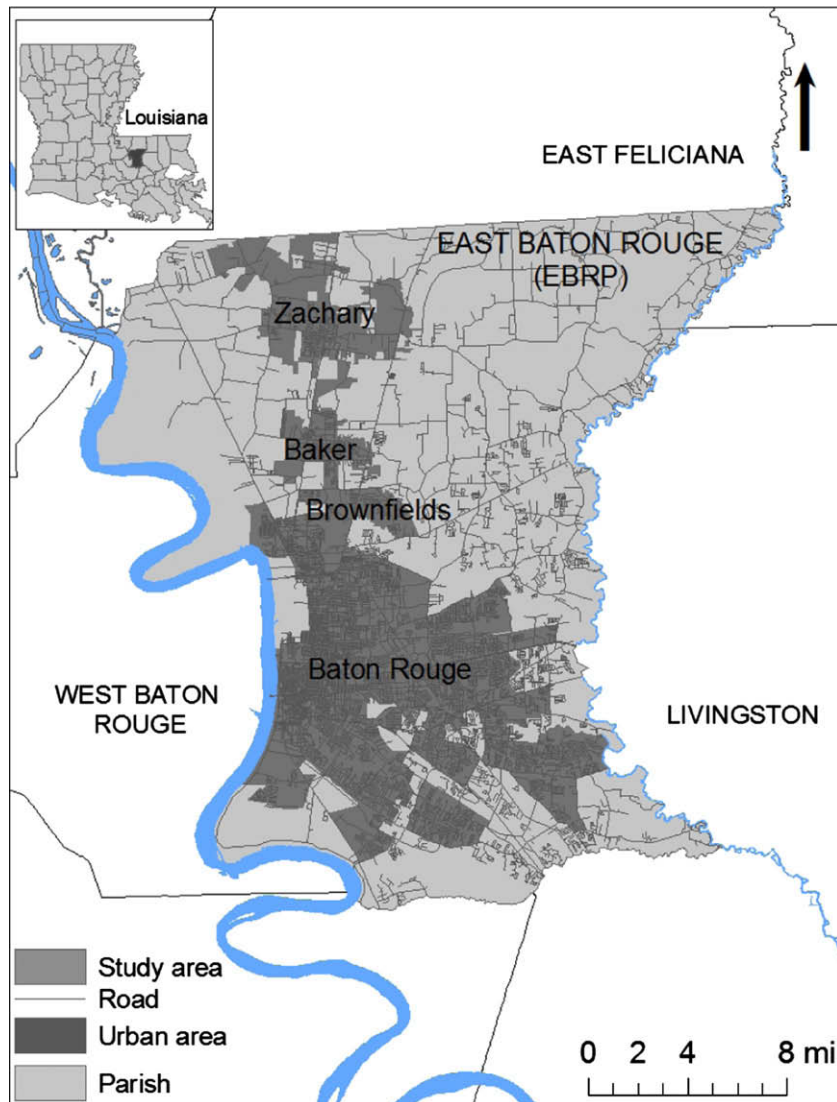


Fig. 1. Urbanized areas in Baton Rouge Region 2000.

ulation, employment and combined densities in Baton Rouge, respectively.

Results from network computation are various centrality values associated with nodes whereas land use densities are by tracts. The two sets of data need to be transformed to one unit so that their relationship may be examined. This is accomplished by two GIS-based methods: the kernel density estimation (KDE) converts both measures to raster pixels, and the floating catchment area (FCA) method computes average centrality values within a fixed distance from each census tract.

### 3.2. Street centrality measures

Among various centrality measures (Porta et al., 2006), three are critical and chosen here to measure a location being *close* to all others, being the *intermediary* between others, and being accessible via a *straight* route to all others. Namely, they are *closeness* ( $C^C$ ), *betweenness* ( $C^B$ ) and *straightness* ( $C^S$ ). Centrality-based metrics such as the connectivity metric and the control metric (Hillier and Hanson, 1984) and the PageRank-based metrics (Langville and Meyer, 2006; Jiang et al., 2008) do not consider link distance, and have limited value for transportation studies (Kuby et al., 2005,

p. 39). Other centrality metrics are not considered in this paper: e.g., efficiency centrality is similar to straightness, and information centrality captures a node's role in network resilience and is less central to the focus of this paper.

These centrality indices follow the tradition of *space syntax* in urban planning and design (Hillier and Hanson, 1984; Hillier, 1996) by utilizing a standard "primal" format for the street network representation without prior differentiation of nodal importance. More recent work uses the *complex weighted networks to scale nodes according to their capacities*, and thus more accurately depicts real-world transport systems (e.g., Barrat et al., 2004; Xu et al., 2006). However, the space syntax approach has its advantage mainly in avoiding the difficulty of modeling the endogeneity of node capacity. In other words, the relative importance of nodes is revealed in the centrality indices because of their locations on the network, and the centrality value at a node is not an input in computing the centrality values of other nodes. Access to all nodes is valued as each node represents an equal potential opportunity or activity.

*Closeness centrality*  $C^C$  measures how close a node is to all the other nodes along the shortest paths of the network.  $C^C$  for a node  $i$  is defined as:

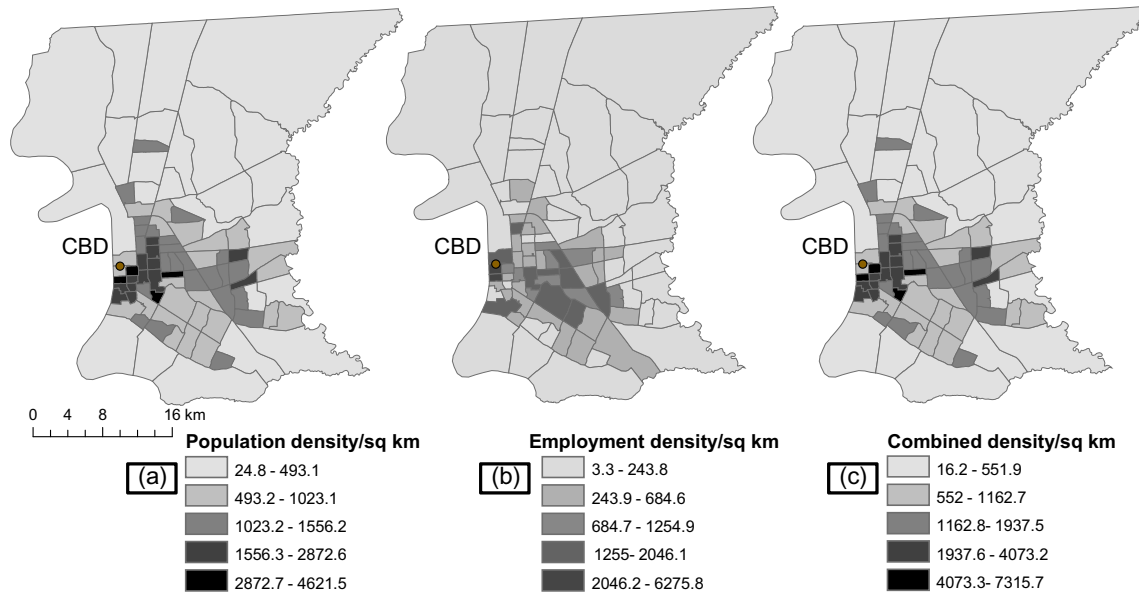


Fig. 2. Population, employment and combined densities in Baton Rouge 2000.

$$C_i^C = \frac{N-1}{\sum_{j=1, j \neq i}^N d_{ij}} \quad (1)$$

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,  $C^C$  is the inverse of average distance from this node to all other nodes. In essence, it reflects the travel cost of overcoming spatial separations between places with population and activities. All distances in this paper are in kilometers.

*Betweenness centrality*  $C^B$  measures how often a node is traversed by the shortest paths connecting all pairs of nodes in the network.  $C^B$  is defined as:

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

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  $i$ .  $C^B$  captures a special property for a place: 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.

*Straightness centrality*  $C^S$  measures how much the shortest paths from a node to all others deviate from the virtual straight lines (Euclidean distances) connecting them.  $C^S$  is defined as:

$$C_i^S = \frac{1}{N-1} \sum_{j=1, j \neq i}^N \frac{d_{ij}^{Eucl}}{d_{ij}} \quad (3)$$

where  $d_{ij}^{Eucl}$  is the Euclidean distance between nodes  $i$  and  $j$ .  $C^S$  measures the extent to which a place can be reached directly, like on a straight line, from all other places in a city. It is a quality that makes it prominent in terms of "legibility" and "presence" (Conroy-Dalton, 2003).

The above three global centrality indices were calculated in a C++ program, and the results were fed back into ArcGIS for mapping and other spatial analysis. Fig. 3a–c show the spatial distributions of closeness  $C^C$ , betweenness  $C^B$  and global straightness  $C^S$ . In the maps, a centrality value for one link is computed as the average of its two end nodes.

### 3.3. Using KDE to convert centrality and density values to one raster frame

As explained earlier, the kernel density estimation (KDE) method is used to transform both the street centrality and land use density values into a new framework (i.e., a raster system) so that the relationship between them can be assessed at the same scale. Data transformations from one scale or analysis unit to another utilize *spatial smoothing* or *spatial interpolation* techniques. There are rich choices for this task (Wang, 2006, pp. 37–47). While the choice of a particular smoothing or interpolation technique should not significantly affect the outcome of this research, this research uses the KDE.

The KDE estimates the density within a range (window) of each observation to represent the value at the center of the window. Within the window, the KDE weighs nearby objects more than far ones based on a *kernel function* (e.g., Fotheringham et al., 2000, pp. 146–149; O'Sullivan and Unwin, 2003, pp. 85–87). Among various kernel functions (Bailey and Gatrell, 1995, pp. 83–108; Gatrell, 1994; Gatrell et al., 1996), popular choices include the standard Gaussian (Levine, 2004, 2006; used in Crimestat) and quartic functions (Silverman, 1986, p. 76; used in ArcGIS). Epanechnikov (1969) finds that the choice among the various kernel functions does not affect significantly the outcomes of the process. The KDE generates a density of the events (discrete points) as a continuous field (e.g., raster). 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: it is not the place itself but rather its surroundings that make it special and explains its setting. Therefore, using the KDE here is not only convenient with a built-in tool available in ArcGIS but also a necessity of accurately capturing the true intention of analyzing the relationship between two neighborhood features.

A rectangular area  $R$  consisting of 214,368 grid cells (462 rows by 464 columns) was formed for the study area. Each cell was a 100 m  $\times$  100 m square. Various methods are available for selecting an appropriate bandwidth (e.g., Silverman, 1986; Sheather and Jones, 1991; Scott, 1992; Cao et al., 1994). Considering the average nearest distance between the census tract centroids of about 2200 m in the study area, we experimented with three search radii or bandwidths ( $h = 1000$  m, 3000 m and 5000 m) for every variable



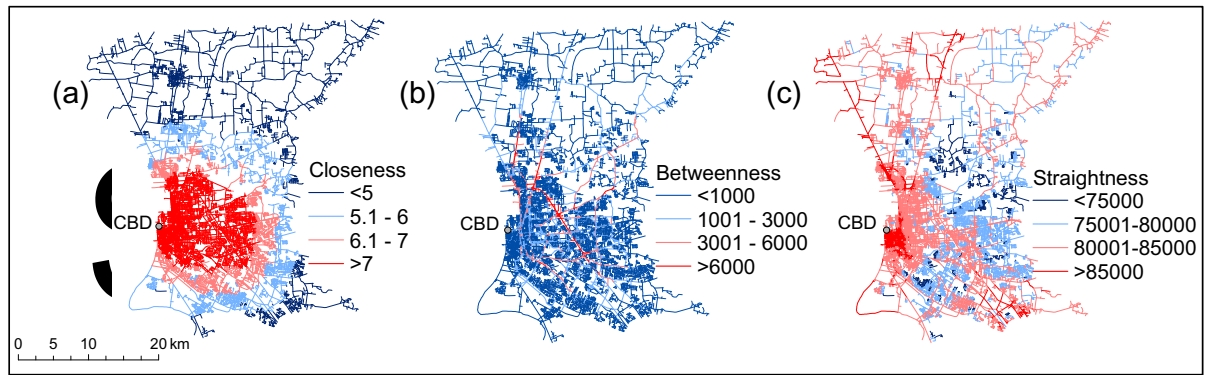


Fig. 3. Spatial distribution of global (a) closeness, (b) betweenness, and (c) straightness.

of interest. The choice of a fixed rather than adaptive bandwidth as suggested by Brunson (1995) pertains to the purpose of the study: we are interested in understanding the relationship between the street network and land use intensity in a medium-size city. With three centrality measures (closeness, betweenness and straightness) and three densities (population, employment and combined), this led to a total 18 raster layers (Table 1).

For illustrative purposes, Fig. 4a–d show the KDE results of closeness  $C^C$ , betweenness  $C^B$ , straightness  $C^S$ , and combined density using a bandwidth  $h = 1000$  m. The KDE results using different bandwidths ( $h = 3000$  m and  $5000$  m) show very similar patterns but with a stronger smoothing effect. The patterns of population and employment densities are similar to that of the combined density and thus not shown.

#### 3.4. Using FCA to compute average centrality values around census tracts

The KDE method converts both values of centrality and land use densities to a raster frame, and both sets of data in the raster are smoothed and thus secondary values. Analysis on the secondary data may be less intuitive. The floating catchment area method (FCA) can be used to compute average centrality values around each census tract's centroid and permits a direct analysis of relationships between densities in each tract and its location as measured by average centrality around it. Another alternative uses the FCA to compute average densities around each node and examine its relationship with the centrality values there. Considering a much larger sample of nodes than census tracts, it is more reliable to adopt the former. Furthermore, the first approach implies that land use intensity is affected by location in surrounding areas, not the other way around.

The FCA method is a common GIS operation used in various applications (e.g., Immergluck, 1998; Wang, 2000). A catchment area is defined as a circle with a fixed radius around a census tract

centroid. Nodes falling inside the catchment are identified and their average centrality values are calculated. The circle with the same radius floats from one tract to another, and average centrality values are obtained for nodes within each circle. The process continues until all census tracts are covered. The FCA method can be easily implemented by utilizing the Point Distance Tool in ArcGIS (Wang, 2006, pp. 38–41). Like in the KDE method, this study experimented with various radii ( $h$ ) such as 1000, 3000 and 5000 m. Fig. 5a–c show the average values of closeness, betweenness and straightness with a catchment of 1000 m.

## 4. Results

### 4.1. Statistical distributions of variables

Based on the KDE results, Fig. 6a–d show the statistical distributions of closeness centrality, betweenness centrality, straightness centrality and combined land use density, respectively. Again, as population, employment and combined densities are distributed similarly, only the pattern of combined density is shown (Fig. 6d). In each graph, the horizontal axis represents the variable of interest, and the vertical axis is the number of cells in the raster of KDE with the corresponding value. Both are in a logarithm scale.

The distribution of each variable generally conforms to the power law, consistent with previous studies (Goh et al., 2003; Lee et al., 2006; Porta et al., 2009). That is to say, the number of cells declines geometrically with the increasing value of a variable.

### 4.2. Spatial distributions of variables

In urban studies, there has been a long tradition of modeling the variation of population (and later employment) density with distance from the city center since Clark (1951). Our examination of spatial patterns with reference to the city center is by no means to endorse the monocentric model. As discussed earlier, polycentricity has increasingly been used to characterize urban structures. Its applicability goes beyond western cities (see Feng et al. (2009), for a recent example). Either monocentricity or polycentricity implies that location with reference to center(s) dictates the land use pattern. The very purpose of this research is to explain the land use density pattern by street network, which is not restricted to a monocentric structure but rather free from pre-selection of center(s). The reference to distance from the city center is more of an analytical way to describe the spatial variation of a particular variable, not a confirmation that the variable is dictated by its distance(s) from any center(s).

As commonly recognized in the study area, the city center is the Louisiana State Capitol Building, a landmark in the heart of Baton

Table 1

Eighteen raster layers from the KDE method.

Raster layer	Bandwidth, $h$ (m)		
<i>Land use density</i>			
Population density	1000	3000	5000
Employment density	1000	3000	5000
Combined density	1000	3000	5000
<i>Centrality</i>			
Closeness	1000	3000	5000
Betweenness	1000	3000	5000
Straightness	1000	3000	5000

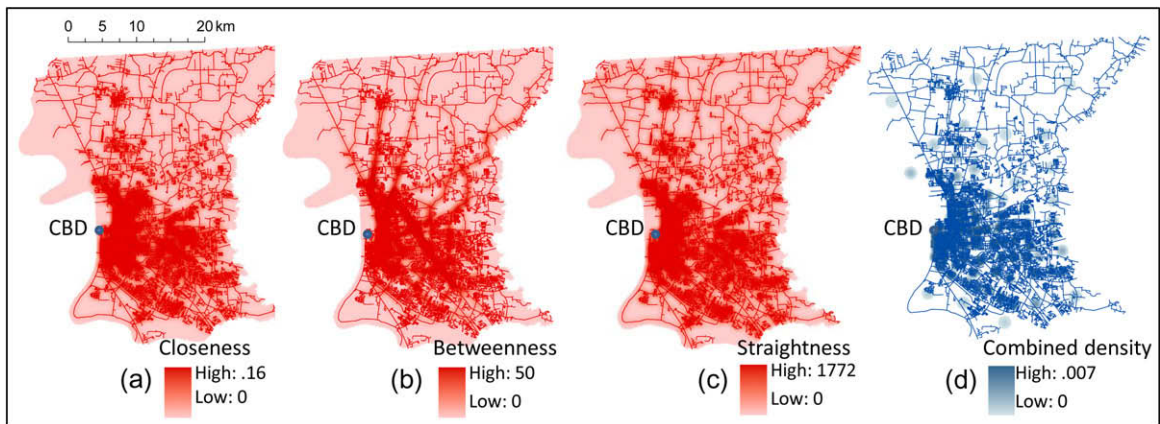


Fig. 4. KDEs of (a) closeness, (b) betweenness, (c) straightness, and (d) combined density ( $h = 1000$  m).

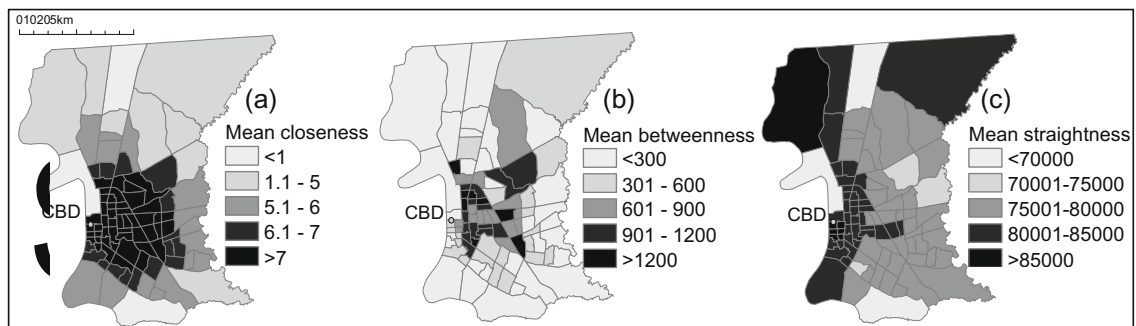


Fig. 5. Average (a) closeness, (b) betweenness, and (c) straightness around census tracts by FCA ( $h = 1000$  m).

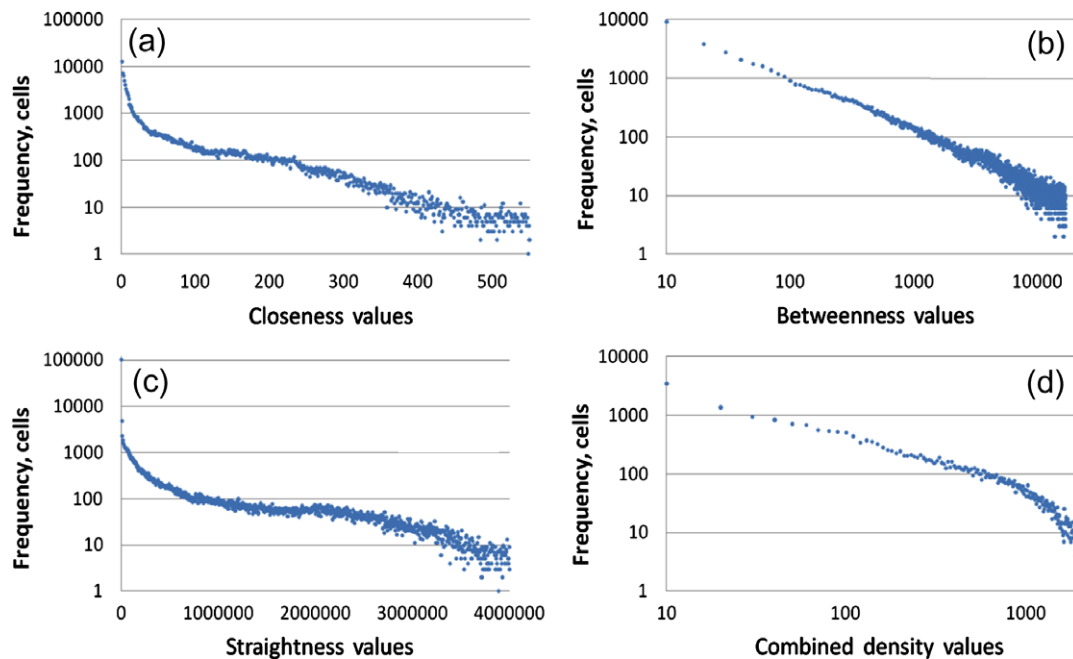


Fig. 6. Statistical distributions of KDE ( $h = 1000$  m) values of (a) closeness, (b) betweenness, (c) straightness, and (d) combined density.

Rouge. Among the four simple bivariate functions (linear, power, exponential and logarithmic) (Wang, 2006, p. 101), the exponential function best captures all three land use (population, employment and combined) densities, consistent with most of the literature (see Table 2). Note that the exponential function in the table is in its log transform.

Similarly, the four functions are used to model the spatial patterns of three centrality indices with increasing distance from the city center. The best fitting function for the closeness centrality index remains the exponential (see Table 2). Clearly, the closeness centrality exhibits a concentric pattern (also see Fig. 3a), captured by an exponential function with the highest  $R^2$  (0.80), like any of

**Table 2**

Best fitting models for land use densities ( $y$ ) and centrality indices ( $x$ ) vs. distance from city center ( $r$ ).

	Function	$a$	$b$	$R^2$	$n$
Population density	$\ln y = a + br$	7.8614	-0.1180	0.51	89
Employment density	$\ln y = a + br$	7.2178	-0.1665	0.47	89
Combined density	$\ln y = a + br$	8.0440	-0.1414	0.58	89
Closeness	$\ln x = a + br$	2.2201	-0.0321	0.80	12,235
Betweenness	$\ln x = a + br$	5.6933	-0.0440	0.02	12,235
Straightness	$x = a + b \ln r$	85872	-2168.11	0.27	12,235

the three land use densities. In other words, the commonly observed population density pattern (i.e., exponentially declining from the city center) is largely due to declining location advantage with distance from the city center (particularly accessibility to all other places) shaped by the road network. This is different from the reliance of monocentric assumption of a single job center in the central business district in the Mills–Muth economic model. The assumption has a significant gap from empirical evidence. In the study area, employment has increasingly been decentralized over time, and conforms to a negative exponential function with a slightly steeper gradient ( $-0.1665$ ) than that of the population density function ( $-0.1180$ ).

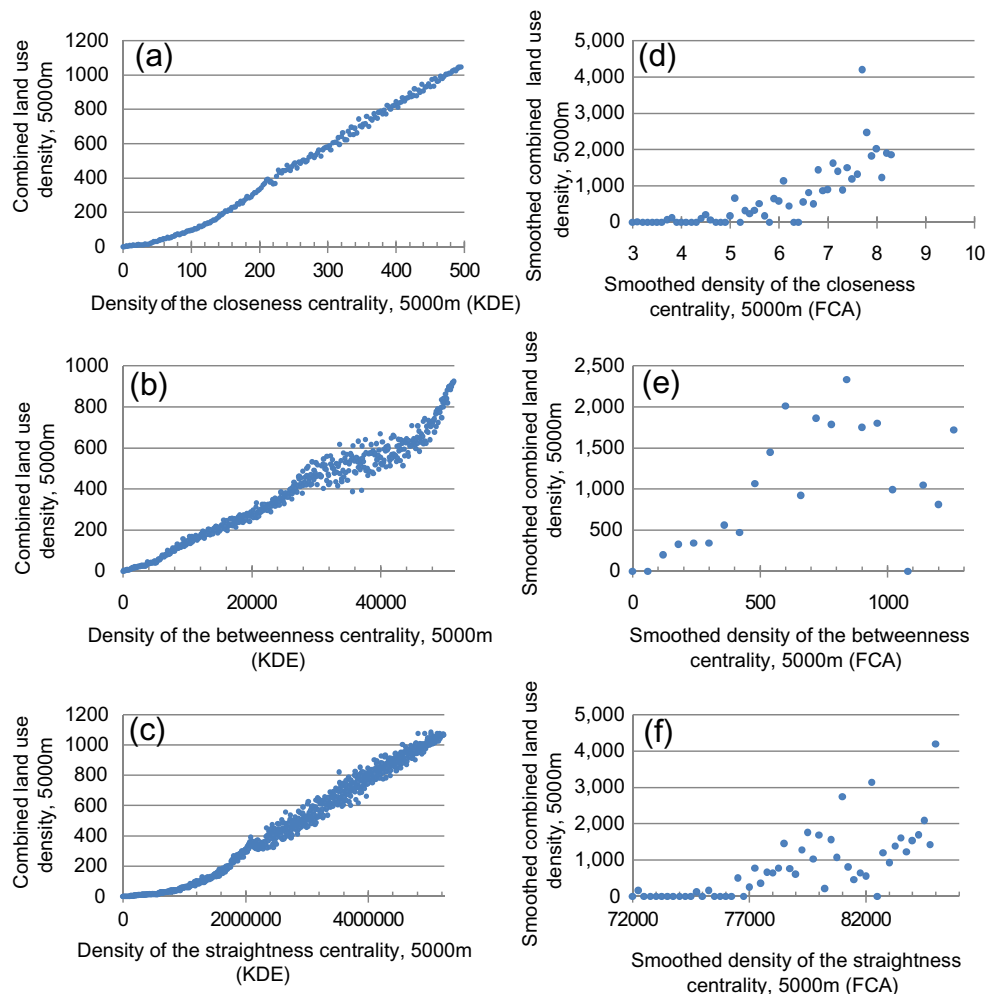
The spatial patterns of betweenness and straightness indices are more complex than concentric (also see Fig. 3b and c). Therefore, none of the bivariate functions generates  $R^2$  higher than 0.30. Nevertheless, the negative slopes in any regression models

remain statistically significant for both indices. The best fitting function for betweenness is still the exponential one with a low  $R^2 = 0.02$ , and for straightness is the logarithmic one with a moderate  $R^2 = 0.27$ . That is to say, both the indices decline with distance from the city center in general, more so for straightness than for betweenness; but there are local pockets of higher values outside of the city center.

#### 4.3. Correlation analysis

Fig. 7a–f show the relationships between centrality indices and land use densities (only combined density shown here as an example). The three graphs on the left (Fig. 7a–c) are based on the results from the kernel density estimation (KDE) results, where both sets of the centrality and combined density values are the KDE values around each node (using a bandwidth of 5000 m as an example). The three graphs on the right (Fig. 7d–f) are based on the results from the floating catchment area (FCA) results, where the centrality indices are average values within a catchment (also 5000 m here as an example) around each census tract and land use density values are the original ones in each tract. The correlations between them are visually evident; more so from the KDE results than the FCA ones.

Correlation analysis was used to further reveal the statistical relationship between centrality indices ( $x$ ) and land use densities ( $y$ ). In order to account for possible effect of each variable's measurement scale, we used log-transforms on either  $x$  or  $y$  or both



**Fig. 7.** Land use densities vs. centrality indices (KDE-based results on the left and FCA-based results on the right).



in addition to plain  $x$  vs.  $y$ , and thus yielded four scenarios ( $x$  vs.  $y$ ,  $\ln x$  vs.  $y$ ,  $x$  vs.  $\ln y$ ,  $\ln x$  vs.  $\ln y$ ). All four scenarios on the KDE results were very similar, and Table 3 reports the correlation coefficients between  $x$  and  $y$ . Table 4 reports the correlation coefficients based on the FCA results (highest coefficient values generated between  $\ln x$  and  $\ln y$ ).

Two important findings can be obtained from Tables 3 and 4. The first is the strong correlations between the centrality indices and land use densities in general. Secondly, among the three centrality indices, closeness exhibits the highest correlation with land use densities, straightness the next and betweenness the last. That is to say, the classic notion of location as being close to all places remains the strongest predictor for land use intensity in Baton Rouge. For a medium-size city like Baton Rouge, residents and businesses may value access to citywide opportunities, all reachable within a reasonable distance (time). For a large city with a wide territory, not all opportunities are relevant, and it is conceivable that local centrality indices (based on areas within a search range of a node) may be better predictors of land use patterns than their global counterparts (based on all areas). A prior study in an Italian city reports that betweenness has a higher correlation with densities of retail and service activities (one of many land use types) than closeness (Porta et al., 2009).

As explained earlier, the KDE values were smoothed and thus secondary values of centrality and land use densities. In contrast, the FCA method was only applied to obtain average centrality values around each census tract, and thus only the centrality values in the correlation analysis were smoothed. Therefore, understandably the correlation coefficients based on the KDE results (as high as 0.92 between combined density and closeness with a 5-km bandwidth) are noticeably higher than those based on the FCA results (e.g., 0.84 as the highest between combined density and closeness with a 5-km catchment radius). In Table 3 on the KDE results, a larger bandwidth leads to a stronger smoothing effect, and thus the correlation coefficients are higher. In Table 4, the effect of catch-

ment radius is less obvious. That is to say, land use densities are almost equally related to centralities within various ranges.

## 5. Conclusions

This paper investigates the relationship between street centrality and land use density in Baton Rouge, Louisiana. Based on the road network composed of links (roads) and nodes (intersections of the roads), street centrality is quantified by three indices. Closeness centrality measures how close a node is to all the other nodes along the shortest paths of the network. Betweenness centrality measures how often a node is traversed by the shortest paths connecting all pairs of nodes in the network. Straightness centrality measures how close the shortest paths from a node to all others resemble virtual straight lines connecting them. The set of centrality indices captures the very essence of location in terms of a place's accessibility, intermediacy and directness among others. Employment, population and their combined densities are used to measure land use intensity. The kernel density estimation (KDE) and floating catchment areas (FCA) are used to convert the data sets of centrality indices at nodes and land use densities in census tracts to the same unit in order to investigate the association between them.

The statistical distributions of the KDE values of both centrality and land use densities conform to the power law, indicating geometrically declining locations with increasing centrality or density values. The spatial patterns of population, employment and combined densities are concentric and are best captured by an exponential function. So is the closeness centrality index. The betweenness and straightness indices also decline with distance from the center (with negative slopes being statistically significant). However, the concentric pattern is far less evident for the betweenness index, and moderately so for the straightness. The correlation analysis shows a strong correlation between the street centrality indices and land use densities, supporting the notion of strong interdependence between the transport network and land use structure in an intraurban setting. At times development encourages the construction of transport systems, whereas at other times the development of transportation helps the growth of a city (Vance, 1986, p. xiii). Only by analyzing data of both systems over a significant period of time one may detect possible causality effects between them.

Among the three centrality indices, closeness exhibits the highest correlation with land use densities, straightness the next and betweenness the last. The highest correlation coefficient is 0.92 (i.e.,  $R^2 = 0.84$ ) between combined density and closeness with a 5-km bandwidth based on the KDE results, and 0.84 (i.e.,  $R^2 = 0.71$ ) between combined density and closeness with a 5-km catchment radius based on the FCA results. The less-than-perfect  $R^2$  is attributable to many factors. To list a few, the factors include the time lag between road development and land use patterns, various amenities valued by residents and business owners in addition to location captured by centrality indices, and other complex social issues (e.g., possible discrimination in housing and job markets, imperfect information in decision making). Given the wide availability of data (street network and population/employment densities), validity of the results can be tested by empirical studies in any cities in the US.

## Acknowledgment

The first author would like to acknowledge the support by the National Natural Science Foundation of China (No. 40928001). We thank Andrew Goetz and three anonymous reviewers for constructive comments on earlier versions of the paper.

**Table 3**  
Correlation coefficients between KDEs of land use densities and centrality indices.

Bandwidth (m)		Population density	Employment density	Combined density
1000	Closeness ( $C^C$ )	0.64	0.42	0.56
	Betweenness ( $C^B$ )	0.43	0.24	0.35
	Straightness ( $C^S$ )	0.62	0.41	0.54
3000	Closeness ( $C^C$ )	0.91	0.76	0.86
	Betweenness ( $C^B$ )	0.81	0.66	0.76
	Straightness ( $C^S$ )	0.89	0.73	0.83
5000	Closeness ( $C^C$ )	0.95	0.88	0.92
	Betweenness ( $C^B$ )	0.79	0.82	0.86
	Straightness ( $C^S$ )	0.92	0.85	0.90

**Table 4**  
Correlation coefficients between land use densities and FCA-smoothed centrality indices (both in log-scales).

Catchment radius (m)		Population density	Employment density	Combined density
1000	Closeness ( $C^C$ )	0.78	0.77	0.83
	Betweenness ( $C^B$ )	0.47	0.34	0.45
	Straightness ( $C^S$ )	0.56	0.51	0.55
3000	Closeness ( $C^C$ )	0.80	0.78	0.84
	Betweenness ( $C^B$ )	0.43	0.36	0.43
	Straightness ( $C^S$ )	0.42	0.56	0.51
5000	Closeness ( $C^C$ )	0.81	0.77	0.84
	Betweenness ( $C^B$ )	0.42	0.34	0.41
	Straightness ( $C^S$ )	0.50	0.59	0.57

## References

- Barabási, A.-L., 2002. *Linked: The New Science of Networks*. Plume, New York.
- Bailey, T.C., Gatrell, A.C., 1995. *Interactive Spatial Data Analysis*. Longman, Harlow, England.
- Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A., 2004. The architecture of complex weighted networks. In: *Proceedings of the National Academy of Sciences of the USA*, vol. 101, pp. 3747–3752.
- Batty, M., 2005. Network geography: relations, interactions, scaling and spatial processes in GIS. In: Unwin, D.J., Fisher, P. (Eds.), *Re-presenting Geographical Information Systems*. John Wiley and Sons, Chichester, pp. 149–170.
- Batty, M., 2008. Whither network science? *Environment and Planning B: Planning and Design* 35, 569–571.
- Berry, B.J.L., Kim, H., 1993. Challenges to the monocentric model. *Geographical Analysis* 25, 1–4.
- Borruso, G., 2003. Network density and the delimitation of urban areas. *Transactions in GIS* 72, 177–191.
- Borruso, G., 2005. Network density estimation: analysis of point patterns over a network. In: Gervasi, O., Gavrilova, M.L., Kumar, V., Laganà, A., Lee, H.P., Mun, Y., Taniar, D., Tan, C.J.K. (Eds.), *Computational Science and its Applications (ICCSA 2005)*. Berlin, Springer Lecture Notes in Computer Science No. 3482, pp. 126–132.
- Borruso, G., 2008. Network density estimation: a GIS approach for analysing point patterns in a network space. *Transactions in GIS* 12, 377–402.
- Brunsdon, C., 1995. Estimating probability surfaces for geographical points data: an adaptive Kernel algorithm. *Computers and Geosciences* 21, 877–984.
- 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.
- Casetti, E., 1993. Spatial analysis: perspectives and prospects. *Urban Geography* 14, 526–537.
- Clark, C., 1951. Urban population densities. *Journal of the Royal Statistical Society* 114, 490–494.
- Conroy-Dalton, R., 2003. The secrete is to follow your nose: route path selection and angularity. *Environment and Behavior* 35, 107–131.
- Downs, J.A., Horner, M.W., 2007a. Characterising linear point patterns. In: *Proceedings of the GIScience Research UK Conference (GISRUK)*, Maynooth, Ireland.
- Downs, J.A., Horner, M.W., 2007b. Network-based kernel density estimation for home range analysis. In: *Proceedings of the Ninth International Conference on Geocomputation*, Maynooth, Ireland.
- Epanechnikov, V., 1969. Nonparametric estimation of a multivariate probability density. *Theory of Probability and its Applications* 14, 153–158.
- Feng, J., Wang, F., Zhou, Y.-X., 2009. Spatial restructuring of population in Beijing metropolitan area: from monocentricity to polycentricity. *Urban Geography* 30, 779–802.
- Fotheringham, A.S., Brunsdon, C., Charlton, M., 2000. *Quantitative Geography: Perspectives on Spatial Data Analysis*. Sage, London.
- Garin, R.A., 1966. A matrix formulation of the Lowry model for intrametropolitan activity allocation. *Journal of the American Institute of Planners* 32, 361–364.
- Gatrell, A., 1994. Density estimation and the visualisation of point patterns. In: Hearnshaw, H.M., Unwin, D. (Eds.), *Visualisation in Geographical Information Systems*. John Wiley and Sons, Chichester, pp. 65–75.
- Gatrell, A., Bailey, T., Diggle, P., Rowlingson, B., 1996. Spatial point pattern analysis and its application in geographical epidemiology. *Transactions of the Institute of British Geographers* 21, 256–274.
- Goh, K.-I., Oh, E., Kahng, B., Kim, D., 2003. Betweenness centrality correlation in social networks. *Physical Review E* 67, 017101.
- Hansen, W., 1959. How accessibility shapes land use. *Journal of the American Institute of Planners* 25, 73–76.
- Heikkilä, E.P., Gordon, P., Kim, J., Peiser, R., Richardson, H., Dale-Johnson, D., 1989. What happened to the CBD-distance gradient? Land values in a polycentric city. *Environment and Planning A* 21, 221–232.
- 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.
- Hoch, I., Waddell, P., 1993. Apartment rents: another challenge to the monocentric model. *Geographical Analysis* 25, 20–34.
- Immergluck, D., 1998. Job Proximity and the Urban Employment Problem: Do Suitable Nearby Jobs Improve Neighbourhood Employment Rates? 35, 7–23
- Jiang, B., Claramunt, C., 2004. Topological analysis of urban street networks. *Environment and Planning B* 31, 151–162.
- Jiang, B., Zhao, S., Yin, J., 2008. Self-organized natural roads for predicting traffic flow: a sensitivity study. *Journal of Statistical Mechanics – Theory and Experiment (Article Number P07008)*.
- Kuby, M., Tierney, S., Roberts, T., Upchurch, C., 2005. A Comparison of Geographic Information Systems, Complex Networks, and Other Models for Analyzing Transportation Network Topologies. NASA/CR-2005-213522.
- Ladd, H.F., Wheaton, W., 1991. Causes and consequences of the changing urban form: introduction. *Regional Science and Urban Economics* 21, 157–162.
- Langville, A.N., Meyer, C.D., 2006. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, Princeton, NJ.
- Lee, S.H., Kim, P.-J., Jeong, H., 2006. Statistical properties of sampled networks. *Physical Review E* 73, 016102.
- Levine, N., 2004. *CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations (Version 3.0)*. National Institute of Justice, Washington, DC.
- Levine, N., 2006. Crime mapping and the CrimeStat program. *Geographical Analysis* 38, 41–56.
- Lowry, I.S., 1964. *A Model of Metropolis*. Rand Corporation, Santa Monica, CA.
- Miller, H.J., 1999. Measuring space–time accessibility benefits within transportation networks: basic theory and computational methods. *Geographical Analysis* 31, 187–212.
- Mills, E.S., 1972. *Studies in the Structure of the Urban Economy*. Johns Hopkins University, Baltimore.
- Muth, R., 1969. *Cities and Housing*. University of Chicago, Chicago.
- Okabe, A., Okunuki, K., Shiode, S., 2006a. SANET: a toolbox for spatial analysis on a network. *Geographical Analysis* 38, 57–66.
- Okabe, A., Okunuki, K., Shiode, S., 2006b. The SANET toolbox: new methods for network spatial analysis. *Transactions in GIS* 10, 535–550.
- O'Sullivan, D., Unwin, D.J., 2003. *Geographic Information Analysis*. John Wiley and Sons, Chichester.
- Park, R.E., Burgess, E.W., McKenzie, R.D., 1925. *The City: Suggestions for the Investigation of Human Behavior in the Urban Environment*. University of Chicago Press, Chicago.
- Porta, S., Crucitti, P., Latora, V., 2006. The network analysis of urban streets: a primal approach. *Environment and Planning B: Planning and Design* 33, 705–725.
- Porta, S., Latora, V., Wang, F., Strano, E., Cardillo, A., Iacoviello, V., Messora, R., Scellato, S., 2009. Street centrality and densities of retailers and services in Bologna, Italy. *Environment and Planning B* 36, 450–465.
- Scott, D., 1992. *Multivariate Density Estimation: Theory, Practice and Visualisation*. Wiley, New York.
- Sheather, S.J., Jones, M.C., 1991. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society Series B* 53, 683–690.
- Silverman, B.W., 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
- Vance Jr., J.E., 1986. *Capturing the Horizon: The Historical Geography of Transportation since the Transportation Revolution of the Sixteenth Century*. Harper & Row, New York.
- Wang, F., 1998. Urban population distribution with various road networks: a simulation approach. *Environment and Planning B: Planning and Design* 25, 265–278.
- Wang, F., 2000. Modeling commuting patterns in Chicago in a GIS environment: a job accessibility perspective. *Professional Geographer* 52, 120–133.
- Wang, F., 2006. *Quantitative Methods and Applications in GIS*. Taylor & Francis, Boca Raton, FL.
- Wang, F., Guldmann, J.M., 1996. Simulating urban population density with a gravity-based model. *Socio-Economic Planning Sciences* 30, 245–256.
- Xu, X.-J., Wu, Z.-X., Wang, Y.-H., 2006. Properties of weighted complex networks. *International Journal of Modern Physics C* 17, 521–529.