



GIS-Based Study on the Association Between Road Centrality and Socio-demographic Parameters: a Case Study

Cynthia Baby Daniel¹ · Samson Mathew¹ · Saravanan Subbarayan¹

Accepted: 19 November 2021 / Published online: 8 December 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

Realising the close relationship between transportation and socio-demographic parameters is a vital factor in designing a safe, structured and more dynamic transportation society. This paper investigates the relationship between road centrality and socio-demographic parameters (land-use intensity), namely population density, employment density and household density, for the bustling capital city of Thiruvananthapuram in southern India by adopting the Multiple Centrality Assessment (MCA) technique. Unlike previous studies, a weighted centrality approach is adopted in this research by weighing the nodes based on degree centrality. The results indicate that betweenness centrality has the highest correlation value followed by closeness centrality. This indicates that the location advantage of being traversed more frequently acts as a key factor that influences the three socio-demographic parameters considered in this exploratory study. Monocentricity is also another factor likely to exist in urban cities due to the substantial influence of the Central Business District (CBD). The relationship between the socio-demographic parameters and the location of the CBD is also analysed to confirm the needed dominance of network centrality over monocentricity. Thus, this study reflects the prominent factors that affect the land-use intensity and also affirms that road centrality plays a vital role in capturing the variation of land-use intensity. This research can act as a precursor for spatial allocation studies. Furthermore, this study also validates that the study area is not a monocentric city.

Keywords Weighted street centrality · Socio-demographic parameters · GIS · CBD · Kernel density estimation · Spatial demography

Introduction

Roads are urban entities that link places with each other and thereby improve the ease of access and the economic activities of an area. A well-configured road network remarkably improves the functional efficiency of a transport system (Xie and Levinson 2007). Topology refers to the arrangement and connectivity of the network elements (Xie and Levinson 2007). A precise comprehension of the topology of complex networks such as transport networks is, therefore, necessary to predict the overall system performance (Jamakovic et al. 2008). In graph theory, the road network can be examined as a directed graph $G = \{v, e\}$ of 'v' vertices or road intersections and 'e' edges that act as street segments connecting them. Urban artefacts like road networks are noticed to have

an unbalanced distribution in space. Road networks can be characterised in terms of connectivity (Chen et al. 2014; Liu and Zhao 2015; Sahitya and Prasad 2020), fractal dimension (Kim et al. 2003; Bai et al. 2010; Li et al. 2012), space syntax (Duan and Wang 2009; Telega 2016), etc. Land-use and economic activities are closely knit with the urban transportation network. The spatial development of an area can be quantified in terms of the different socio-economic activities in the region. The Lowry model, developed in 1964, defined three broad sectors of activities in the urban form, namely employment, population and household sector, and these areas remain the chief components in the subsequent models derived from it. Thus, land-use intensity can indisputably be quantified in terms of population density, employment density and household density. This research paper takes into account these three parameters for the study of socio-demography or land-use intensity. The two terms, land-use intensity and socio-demographic parameters, have been used synonymously in this paper. The population density and household density indicate the spatial variation of

✉ Cynthia Baby Daniel
cynthiababydaniel@gmail.com

¹ Department of Civil Engineering, National Institute of Technology, Tiruchirappalli 620 015, India

residential land-use whereas employment density is indicative of the commercial or business or related land-use.

A comprehensive knowledge of the spatial variation of centrality parameters will help urban planners to plan future constructions effectively. For instance, commercial buildings should be constructed in places with greater centrality advantage so as to improve the economic value. Places with greater betweenness centrality would accommodate more flow of people and hence these can be considered for allocation of commercial centres.

Land-use models, from the monocentric model (Burgess 1925) to the sector model (Hoyt 1939) and the multiple nuclei model (Harris and Ullman 1951), have evolved over a period of time depicting the influence of human activities in the spatial structure. Recent developments in network science have contributed immensely to the study of the properties of road networks.

The Concept of Space Syntax and Centrality

Space syntax (Hillier et al. 2007) is a set of techniques for representing, quantifying and interpreting the urban configuration. Spatial syntax studies incorporate the notions of visibility and integration into the urban research scenario. This facilitated a better understanding of the structural properties of urban space. However, due to inconsistencies in certain geometric configurations, its applicability has been questioned (Ratti 2004). Moreover, according to space syntax theory, a dual approach is practised for the representation of street networks which neglects the spatial topography. This paved the way for a new approach, the street centrality approach.

With the growing complexity of an area, new road networks evolve adjacent to the central place, thereby fuelling the growth of residential and economic activities (Sheikh et al. 2013). Centrality is a concept that denotes how pivotal or central a place is. It is a concept that has its roots in social sciences and is used to quantify the centrality of spatial networks. A node in a network with a greater centrality value is traversed more frequently. Centrality plays a vital role in sculpturing urban evolution. It has become one of the crucial tools for urban policymakers and city planners to understand how road networks of cities work and identify areas where redevelopment needs to be initiated (Porta et al. 2009). In recent years, Multi Centrality Assessment (MCA) has been gaining popularity over the space syntax model due to its increased efficiency. The MCA model computes the spatial distribution of centrality of entities such as road networks in terms of betweenness centrality, closeness centrality and straightness centrality. Furtherance in Geographic Information System (GIS) has made the computation of network centrality easier.

Related Studies

Few researchers have explored the relationship between street centrality and other points of interest such as retail stores (Lin et al. 2018; Han et al. 2019; Zhang et al. 2019) and health care facilities (Ni et al. 2016; Daniel et al. 2021). Porta et al. (2011) studied the relationship between economic activities and street centrality in Barcelona and suggested that central urban arterials should be considered the cores, not borders of neighbourhoods. Rui and Ban (2014) studied the association of street centralities and different land-use categories within Stockholm. Not many studies have used a weighted centrality approach for the computation of network centrality.

The concept of centrality has been applied in other transportation studies for identifying the critical locations in transport network (Demšar 2007; Henning et al. 2017) and for analysing the traffic flow patterns (Jayasinghe et al. 2015; Zhao et al. 2017) as well as in identifying areas with traffic congestion (Jayaweera 2017) etc.

Central Business District (CBD), which can be considered the focal point of a city, plays a significant role in the variation of land-use intensity of an area. However, there are cities that lack a single focal point. Most economic activities and services are found to cluster around the CBD area. The distribution of facilities in and around the CBD plays a significant role in CBD analysis (Yu et al. 2015). Starting from the Burgess model which highlighted the importance of a monocentric model, an evolution from monocentricity to polycentricity was noticed along with the changing urban forms and with the enhanced research in the field. Emphasising the functional areas within the CBD and rectifying the problems related to human activities have drawn the concerted attention of policymakers and academicians (Yu et al. 2015; Zhu and Sun 2017). The advancement of GIS techniques with its improved cartographic capabilities has simplified this task.

Objectives of the Study

The current paper aims to study the spatial variation of the socio-demographic characteristics in terms of population density, employment density and household density and evaluate their dependence with respect to the CBD and the street centrality for the city of Thiruvananthapuram, South India, the administrative capital of the state of Kerala. To the best of the authors' knowledge, no previous studies of this nature have been conducted in India. Node-weighted road centrality approach based on the nodal degree has been used in this study.

Study Area

Thiruvananthapuram recognised as Trivandrum in former times is the capital of the state of Kerala located in South India. This city is also the headquarters of the Thiruvananthapuram district. It has been described as an ancient temple city built on seven low forested hills (Abram and Edwards 2005). The picturesque coastal city secures an eternal place in the religious, socio-economic and cultural legacy of the Indian nation. The district is geographically located between the latitudes $8^{\circ} 17' N$ and $8^{\circ} 51' N$ and $76^{\circ} 41' E$ and $77^{\circ} 17' E$ longitude. The city located in the southwestern coastal area of the country is bounded by the Arabian Sea to its west and the **Western Ghats** to its east. The corporation area of the city covers over 215 km^2 with **hundred administrative units called wards**. Only 86 wards of the urban area are chosen for this study due to the unavailability of census data. East Fort located towards the south of Thiruvananthapuram is considered the Central Business District. Thiruvananthapuram is one of the chief IT and academic hubs of the nation and additionally the administrative capital of the state of Kerala. Thiruvananthapuram is a host to many state and central government agencies. A major bulk of the city's workforce comprises employees of the state and central government. **Technopark**, India's first IT park, which came into existence in the year 1990, has now emerged as the largest IT park in India in its spatial extent employing thousands of technocrats in

around 400 companies. As per the Indian Census 2011, Thiruvananthapuram city has a net population of 9,57,730.

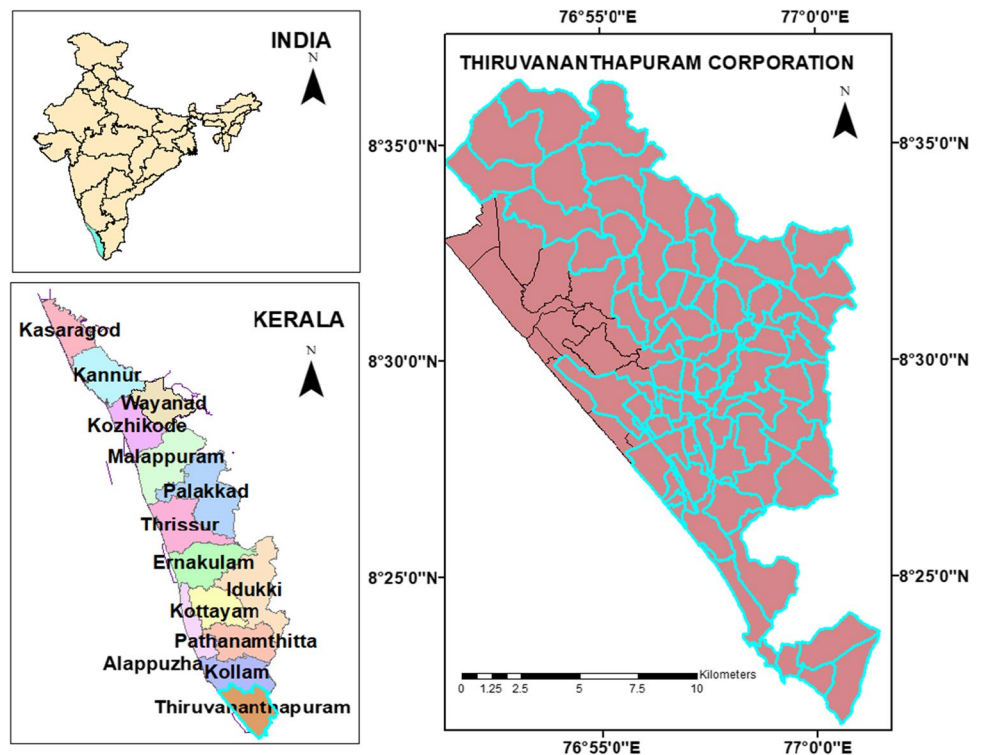
The district of Thiruvananthapuram has a notable transport connectivity, viz. roadways, railways, airways and waterways. The state-owned Kerala State Road Transport Corporation (KSRTC) provides a large percentage of the district's bus services, with private bus services being meagre in number. Thiruvananthapuram is home to the divisional headquarters of the Indian railways' Southern zone. Thiruvananthapuram International Airport, which is located in Chakai and is 6.7 km from East Fort, provides air travel. Figure 1 depicts the Thiruvananthapuram corporation's location map, highlighting the 86 wards chosen for the study out of a total of 100.

Methodology

Data Preparation

The road network for computing the centrality is obtained from the National Transportation Planning and Research Centre (NATPAC), Thiruvananthapuram. The network is first decomposed into its geometric features, nodes and links. A primal representation of the road network instead of the dual representation is chosen for this study. In primal representation, nodes are the intersections and road stretches are edges of a directed graph. Dual representation, on the contrary, considers nodes as streets and edges as intersections.

Fig. 1 Location map of the area of study



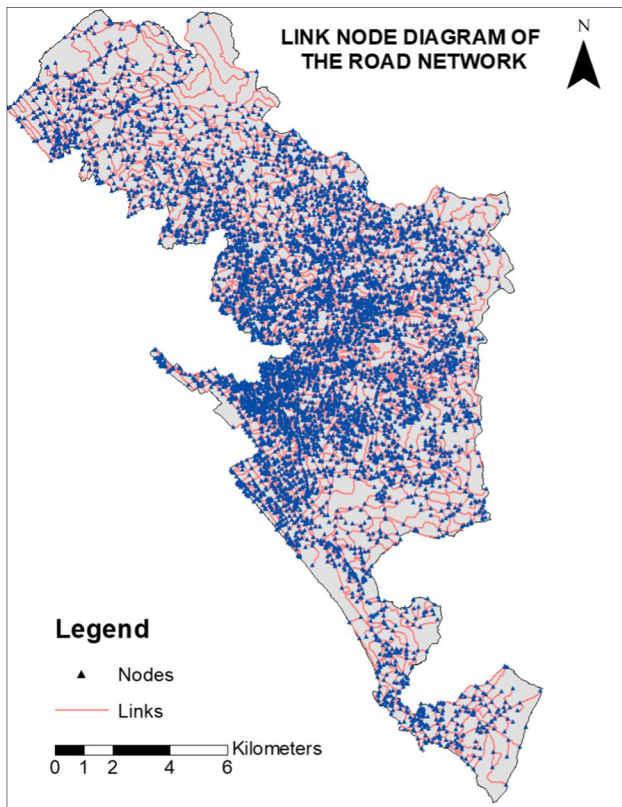


Fig. 2 Link node diagram of the road network for the study area

a better choice as far as the spatial layout and geographical properties are concerned. The study area had a total of 6885 nodes and 3556 links. The network is topologically corrected to remove the pseudo-nodes, dangling links and deviations such as overshoots and undershoots. The number of nodes and links are then reduced to 6507 and 3469 respectively as represented in Figure 2. All the roads consisting of arterial, sub-arterial and collector roads have been included in the analysis. The study involved the preparation of two datasets: the street network and the socio-demographic parameters. The socio-demographic parameters were extracted from the Census of India, 2011 database, and integrated into a GIS layer for each ward. The result was a polygon layer of wards with the population density, employment density and household density assigned as attributes to it. Figure 3 a–c show the population densities, employment densities and household densities in the 86 wards of Thiruvananthapuram city.

Centrality Measures of the Road Network

This study considers three main centrality measures that are indicative of accessibility, directness and intermediacy (Wang et al. 2014). The centrality measures adopted are betweenness, closeness and straightness (defined by Equations 2, 3 and 4) as they are identified as the most significant measures of centrality (Rui and Ban 2014). The street centrality method gives weightage to the nodes grounded on their ease of access or the repeatability with which it is traversed. They can be weighted based on the degree centrality of the node. It is based on the concept that nodes

The spatial topography is preserved in the former representation whereas the latter discards it. Primal representation is

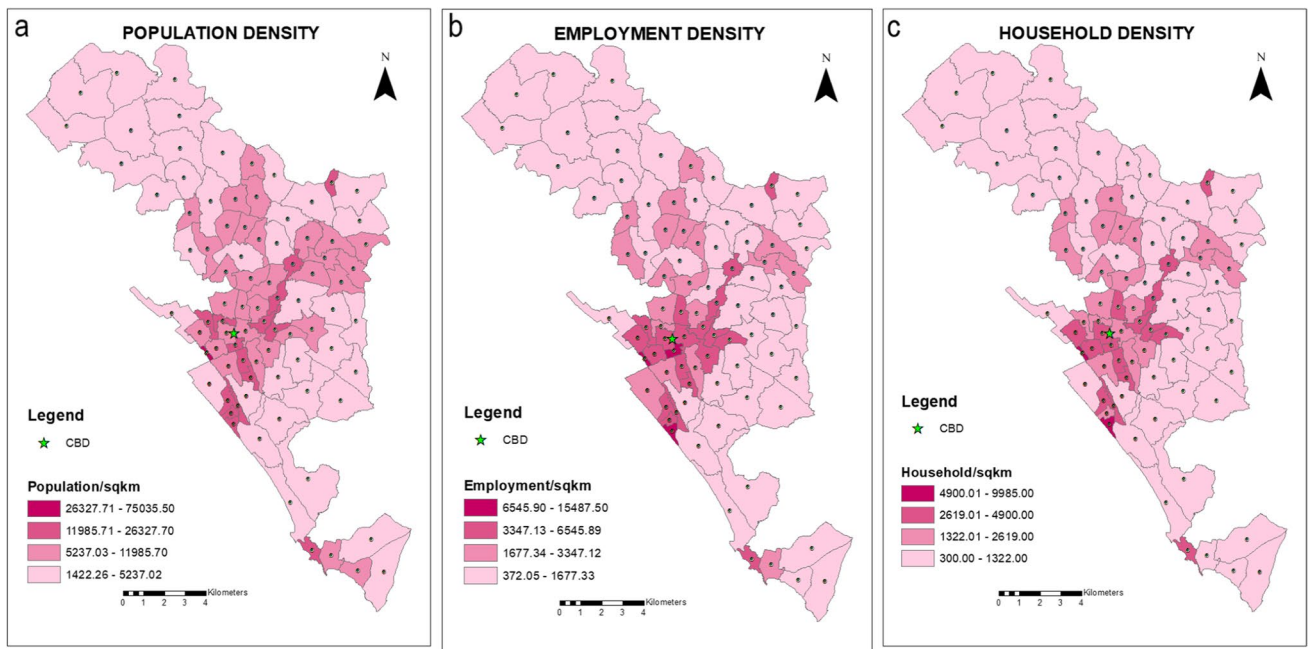


Fig. 3 Spatial variation of socio-demographic parameters: a population density; b employment density; c household density

with greater importance will have more edges attached to them. For a directed graph G , the degree centrality of a vertex is defined as the gross total of the in-degree (i_{deg}) and out-degree (o_{deg}) of the node considered. In-degree of the vertex is the number of incoming edges attached to it and out-degree of the vertex is the number of outgoing edges attached to it. For an undirected graph, total degree can be deduced as the number of edges connected to it. Thus, representing it mathematically by Eq. 1,

$$deg(v) = i_{deg} + o_{deg} \tag{1}$$

The degree of each node is calculated from the network dataset that is modelled from the network shapefile and it is fed as weights into the Urban Network Analysis (UNA) toolbox. The UNA toolbox is a freely available plugin developed by City Form Lab (Sevtsuk et al. 2013) that is compatible with ArcGIS. A weighted nodal approach of representing the transport network yields better outcomes as it portrays the real-world scenario in a better and efficient way (Liu et al. 2016).

Betweenness Centrality (C_B)

The betweenness centrality of a node quantifies the importance of a node in terms of the relative number of times it is traversed while considering the shortest paths from points in a network.

$$C_{B(i)r} = \sum_{j,k \in G - \{i\}, d[i,j,k] \leq r} \frac{nj_k(i)}{n_{jk}} \cdot W[j] \tag{2}$$

where $C_{B(i)r}$ is the betweenness centrality of junction i within a buffer of radius r and n_{jk} is the complete number of shortest paths from j to k . $n_{jk}(i)$ indicates the number of shortest paths from node j to node k that passes by node i and $W[j]$ denotes the weight of the destination node j . Betweenness centrality essentially calculates the number of shortest network paths that traverse through a particular vertex.

Closeness Centrality (C_C)

Closeness centrality is defined as the inverse of cumulative distance required to reach or travel from one node to all other nodes along the shortest paths. Closeness centrality describes the level of integration of a vertex into the network.

$$C_{C(i)r} = \frac{1}{\sum_{j \in G - \{i\}, d[i,j] \leq r} (d[i,j] \cdot W[j])} \tag{3}$$

where $C_{C(i)r}$ denotes the closeness centrality of node i within a search radius r of the buffer, $W[j]$ indicates the weight of the destination node j and $d[i,j]$ is the shortest distance path between node i and destination node j .

Straightness Centrality (C_S)

Straightness centrality emphasises on the directness of the network in terms of its deviation from or resemblance to straight-line distance between two points.

$$C_{S(i)} = \frac{1}{n - 1} \sum_{j \in V, j \neq i} \frac{d_{Euc_{ij}}}{d_{ij}} \tag{4}$$

where $d_{Euc_{ij}}$ is the Euclidean distance or the straight-line distance between nodes i and destination node j . The efficacy of communication between nodes in a network is represented by this metric. When the deviation from the straight-line distance is reduced, the efficiency improves.

Creation of Raster Dataset

Kernel density estimation (KDE) is a spatial interpolation approach that estimates the density of features at each location by computing the average value within a spatial search window or kernel by assigning weights based on the nearness of the objects (Fotheringham et al. 2000). KDE is used in this study to transform the centrality values and the socio-demographic parameters into a similar raster framework for analysing the relationship existing between them. By converting both the parameters into a raster dataset, a cell-to-cell correlation is facilitated. Numerous other smoothing and interpolation techniques exist in ArcGIS; however, KDE remains to be the most widely used technique due to its compatibility with the GIS platform. Many previous researchers have recognised that the selection of a certain method is not of key significance. Figure 4 shows a kernel function fitted over a line segment. Kernel density estimator is defined by Equation 5.

$$f(x_i) = \sum_{i=1}^n \frac{1}{h^2} K\left(\frac{x_i - x_j}{h}\right) \tag{5}$$

where h is the selected bandwidth, $\frac{1}{n}$ is the number of points or lines that lie inside the specified bandwidth and

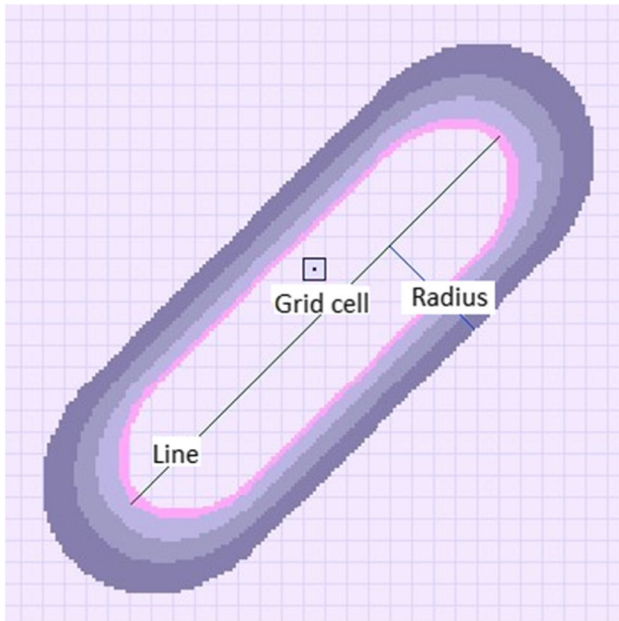


Fig. 4 Kernel density estimation of a linear feature

$d = x_i - x_j$ is the distance between the grid point x_i and event point x_j .

The spatial analyst toolbox of ESRI ArcGIS makes use of an Epanechnikov also known as a quadratic kernel function framed by Silverman (1986) and defined by Equation 6. It is given by the function;

$$K(x_i) = \begin{cases} \frac{1}{3\pi} (1 - s^2)^2 & \text{if } s^2 < 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{where } s = d_{ij}/h \quad (6)$$

There are plentiful earlier researches that have implemented KDE as the spatial interpolation method for converting the vector data points into a raster format (Porta et al. 2009; Wang et al. 2011; Rui and Ban 2014). KDE considers the decay of functions in terms of Tobler’s law, which states that ‘Everything is related to everything else, but near things are more related than distant things’ (Tobler 1979), and hence outperforms other commonly used statistical techniques. KDE employs a distance function to compute distances in Euclidean space and is used to obtain a continuous and smooth surface of a univariable or multivariable probability density (Cai et al. 2013). Moreover, KDE highlights the concept of the neighbouring area being equally significant rather than just the network alone.

A fixed bandwidth type is adopted in this study. The choice of bandwidth is a factor that influences the outcomes of the interpolation and hence needs more consideration. Porta et al. (2009) proposed bandwidths of 100m, 200m and 300m for his study as they were more extensively used in Italian urban planning. Since this idea is based on the nature

Table 1. Raster layers obtained after kernel density estimation

Raster layers	Bandwidth, h in metres		
Socio-demographic parameters			
Population density	6000	7000	8000
Employment density	6000	7000	8000
Household density	6000	7000	8000
Centrality parameters			
Betweenness centrality	6000	7000	8000
Closeness centrality	6000	7000	8000
Straightness centrality	6000	7000	8000

of the study area, the selection of similar bandwidths was not considered appropriate for this analysis. Contemplating the geometry of the area being a predominant factor, bandwidths are chosen based on the average distance between each of the 86 wards which is deduced as 7579.24m. Hence, bandwidths of $h = 6000m$, $h = 7000m$ and $h = 8000m$ are chosen for the current study. This method has been adapted from the study by Wang et al. (2011). For the three centrality layers and three land-use intensity layers, a total of 18 raster layers are created. A cell size of $100m \times 100m$ is chosen for this investigation and is consistently opted for all the KDE estimations. Table 1 depicts the 18 raster layers that are constructed.

Results and Discussion

The centrality indices are computed resulting in a polyline layer of the road network. Figure 5 a–c show the spatial variation of the three centrality indices within the 86 wards.

Density Distribution of Road Centrality and Socio-demographic Parameters

The statistical distribution of the centrality measures (Figure 6 a–c) and the socio-demographic parameters is plotted in the graphs (Figure 7 a–c). The statistical distribution with bandwidth $h = 7000m$ has been plotted here. Similar patterns are observed for all the bandwidths ($h = 6000m$, $7000m$ and $8000m$) for the single variable considered. The smoothing effect of KDE aids in outlining the spatial patterns more clearly. The ordinate axis represents the frequency of the cells or the number of raster cells. The abscissa values represent the kernel density of the variable considered. Both the axes are in linear scale. The number of cells is found to decline with the increasing value of the variable. This implies that the high-density areas are fewer in number. A two-peak graph is noticed for betweenness and straightness centrality and a single peak graph for closeness centrality denoting the concentration of centrality in a few

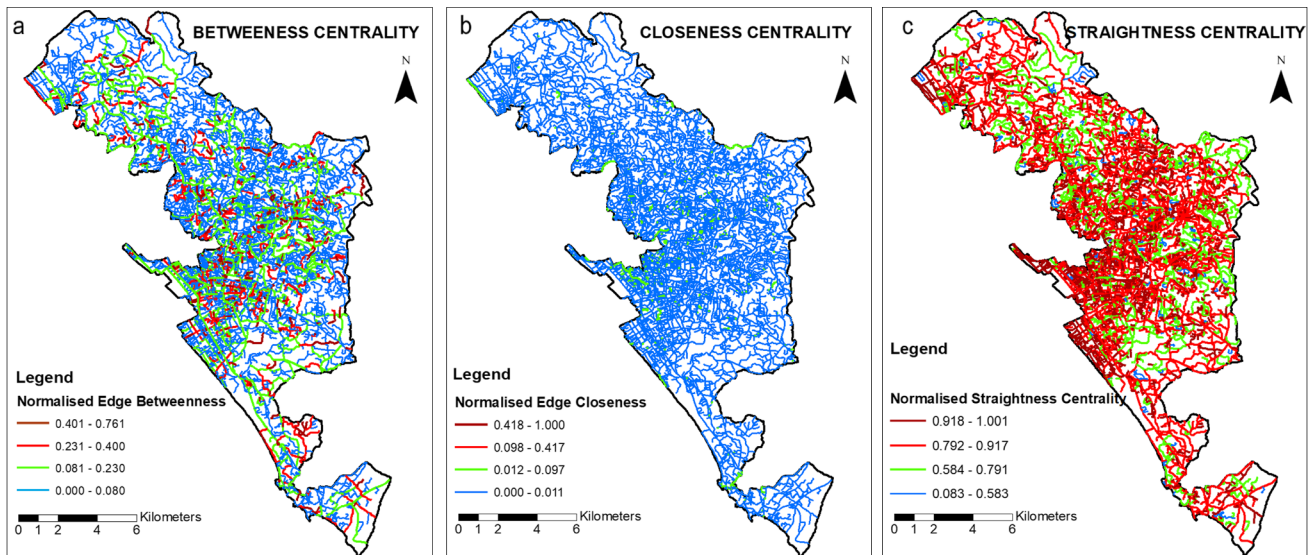


Fig. 5 Spatial variation of centrality indices: **a** betweenness centrality; **b** closeness centrality; **c** straightness centrality

places which in turn reflects the ease of accessibility being predominant in certain locations.

The distribution of density values of socio-demographic parameters shows a multi-peak graph followed by a decreasing trend denoting higher density values for a few places. It is followed by an exponentially decreasing tail. The results are found to be consistent with the study undertaken by Porta et al. (2011). The peaks in the graph denote the concentration of centrality values at certain locations in the study area implying that intermediacy, accessibility, etc. are concentrated at certain locations or its benefits are enjoyed by the population in a specific locality. This also indicates the presence of polycentric nature of the urban area of Thiruvananthapuram city.

Correlation Analysis

Scatterplots are used to display the relationship between the socio-demographic factors and the centrality indices and to visualise the correlation existing between them. Both the abscissa and ordinate values denote the average kernel density values of each ward. Betweenness centrality values give the best R square value in all the cases for all the bandwidths and are shown in Figure 8 a–c. The best fit is given by the power law. Due to space constraints, only the best fit cases for bandwidth ($h = 8000\text{m}$) have been depicted here.

The foremost proposition of this paper is that centrality is the key factor that affects land-use intensity. Correlation analysis is done to determine the relationship existing between the network centralities and socio-demographic parameters. Only variables with the same bandwidth can be correlated resulting in a total of 27 correlations. Table 2

shows the Pearson's correlation coefficient between the centrality values and land-use intensity values at 0.01 significance.

Two major findings can be formulated from the coefficient values. One, there is a strong correlation between the socio-demographic parameters and the centrality measures. Two, among the three measures considered in this study, betweenness centrality exhibits the highest correlation (correlation coefficient value is greater than 0.8 for all the three socio-demographic parameters considered). The spatial variation in the employment density values is specifically influenced by betweenness centrality values. It is followed by closeness centrality having a correlation coefficient value greater than 0.7 for all the three socio-demographic parameters considered. This affirms the fact that the location advantage of a place being traversed more frequently acts as a significant force that affects the socio-demographic diversity of Thiruvananthapuram city. Consequently, accessibility is the established criterion that determines socio-demographic diversity. Different bandwidths reflect the spatial variation of the parameters considered. Though not consistent, higher bandwidth is shown to provide a better correlation. In the centrality study conducted in Louisiana (Wang et al. 2011), a floating catchment area analysis was carried out after the kernel density estimation to smoothen the KDE values but it was found to yield lower correlation coefficients.

Socio-demographic Parameters and CBD

The spatial influence of the CBD on the socio-demographic parameters being a factor that cannot be neglected, and in order to cross-validate the results obtained from the

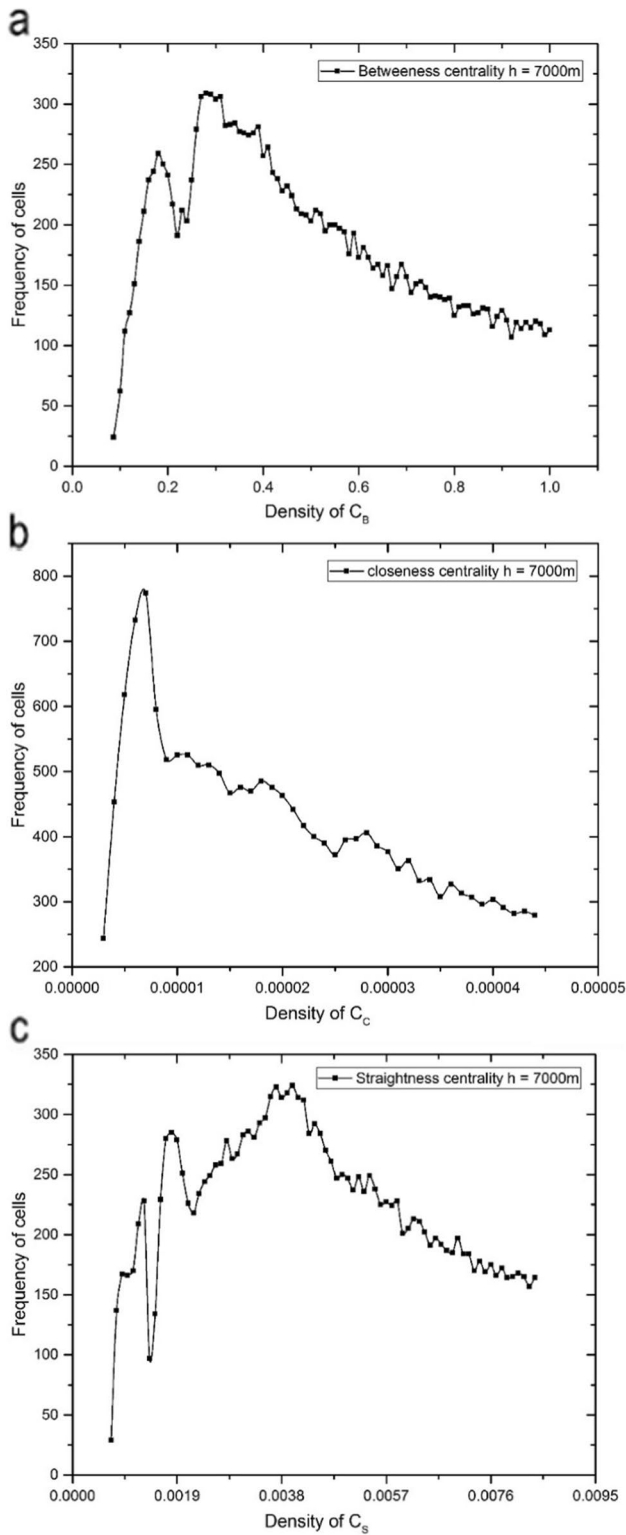


Fig. 6 Statistical distribution of centrality measures at $h = 7000m$: **a** betweenness centrality; **b** closeness centrality; **c** straightness centrality

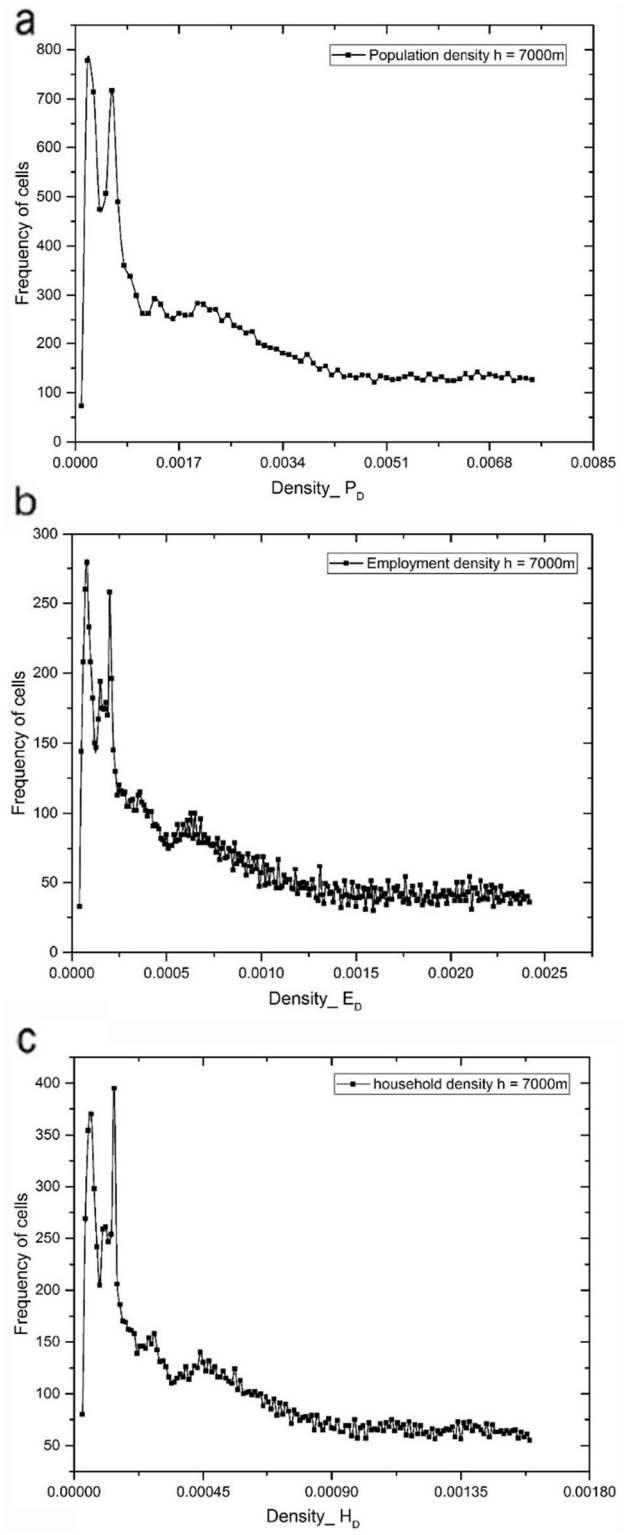


Fig. 7 Statistical distribution of socio-demographic parameters at $h = 7000m$: **a** population density; **b** employment density; **c** household density

Fig. 8 Socio-demographic parameters vs centrality measures at $h=8000m$: **a** population density; **b** employment density; **c** household density

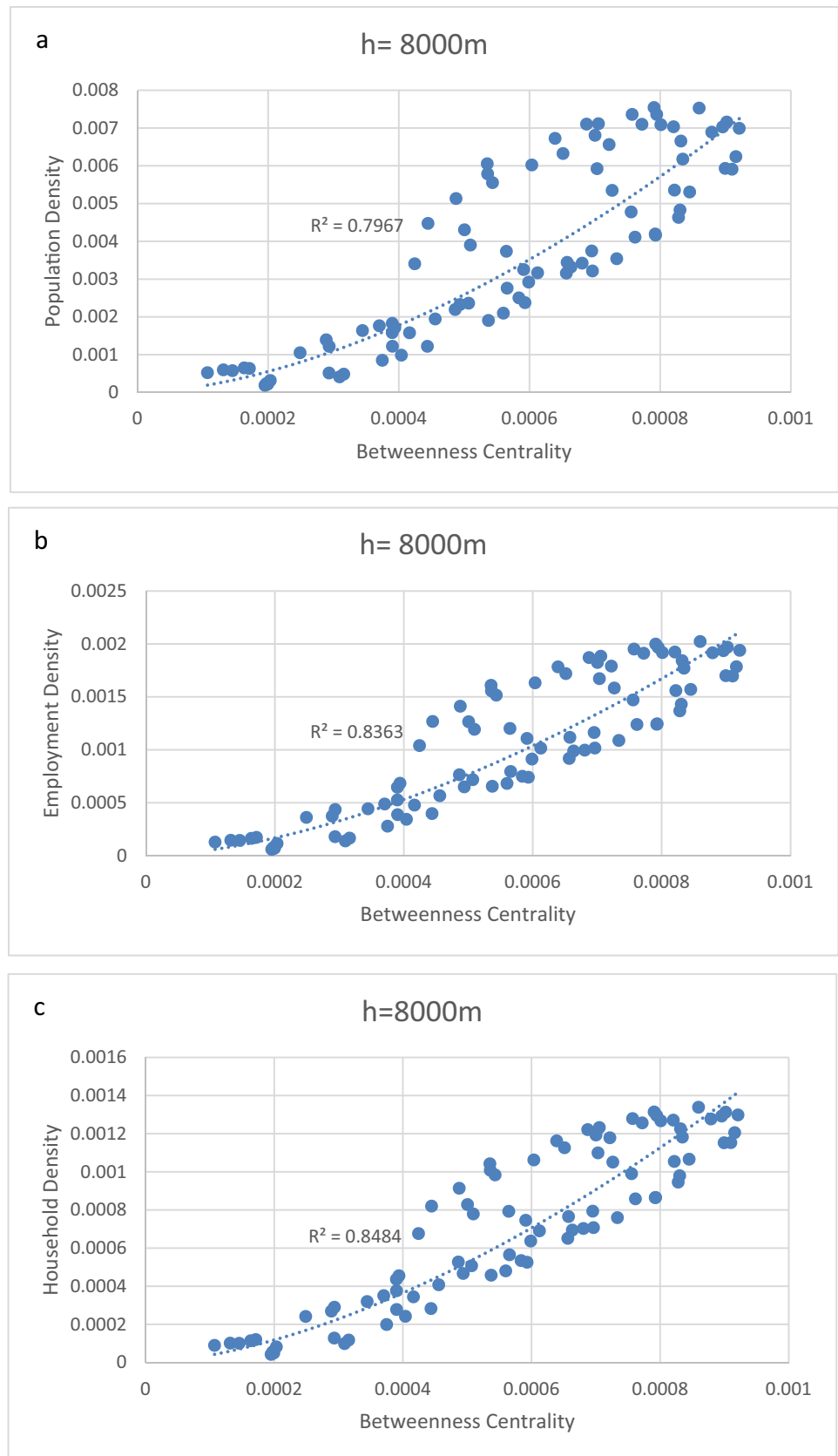


Table 2. Pearson’s coefficients between centrality and socio-demographic parameters for all bandwidths

Centrality measures	Bandwidth (m)	Population density	Employment density	Household density
Betweenness centrality	6000	0.842	0.845	0.868
	7000	0.750	0.746	0.767
	8000	0.857	0.887	0.905
Closeness centrality	6000	0.718	0.715	0.745
	7000	0.741	0.736	0.821
	8000	0.733	0.764	0.790
Straightness centrality	6000	0.683	0.684	0.715
	7000	0.714	0.711	0.740
	8000	0.705	0.739	0.765

*All correlations are significant at the 0.01 level (2 tailed)

correlation analysis, the location analysis with respect to the CBD is performed. This is done to check whether monocentricity (the centre being considered as the CBD) is the factor that dominates the area rather than the road centrality. This is also an attempt to explain the spatial variation of the variables considered in the study. East Fort is the commonly recognised Central Business District of Thiruvananthapuram. It has been declared as a heritage site by the state government as it is the abode of the world famous Sree Padmanabhaswamy temple. It can also be called a transport hub of Thiruvananthapuram as the Central railway station is just 1 km, and the International airport is 7 km away from East Fort. The bus depot is also situated at East Fort. The distance of the ward centroids from the CBD is computed using ArcGIS and is defined as the parameter *x*. Five bivariate functions, namely linear, logarithmic, quadratic, power and exponential functions, are used to test the relationship between density/centrality and location of CBD. The best fit functions for the relationship between centrality and distance from CBD and socio-demographic parameters and distance from CBD are presented in Table 3.

The *R* square values (all less than 0.5) clearly reveal that the socio-demographic parameters are not influenced by the

distance from the CBD. It is contradictory to the general notion that urban jobs and employment opportunities are largely concentrated in and around the CBD. Also, it can be confirmed from the analysis that a monocentric structure is not the dominant pattern that prevails in the city of Thiruvananthapuram. It also affirms the hypothesis that socio-demographic parameters are influenced by the centrality of the road network.

Due to the difficulty in obtaining real-time data, no traffic parameters were considered in this study and this study is limited to the corporation area of Thiruvananthapuram city neglecting the outskirts of the district considering the importance of the urban area.

Conclusions

This study documents the relationship between socio-demographic parameters and street centrality of 86 wards of the city of Thiruvananthapuram. Socio-demographic parameters considered for the study are population density, employment density and household density. Betweenness centrality, closeness centrality and straightness centrality are the three variables used to calculate centrality. The frequency with which a node is crossed by the shortest path considering all nodes in the network is referred to as node betweenness centrality. Node closeness centrality indicates how far away a certain node is from every other node or junction, meaning that nodes with low closeness values are closer to one another and hence more accessible (Mishra et al. 2012). Node straightness centrality, on the other hand, displays the ability of nodes in a network to communicate with one another. In essence, betweenness and closeness centrality refers to the network’s intermediacy and accessibility, respectively, while straightness centrality refers to the network’s directness.

A node-weighted centrality approach is used in the present study to give a better portrayal of the real road network. The nodes are weighted based on their degree centrality, i.e. the number of road segments connected to each of the nodes.

Table 3. Parameter estimates of best-fitting models for socio-demographic parameters

	Best fitting function $y = a + bx$	Constant, a	b	R squared
Socio-demographic parameters				
Population density	Exponential	12690	0.000	0.389
Employment density	Exponential	4374	0.000	0.496
Household density	Exponential	2873	0.000	0.452
Centrality measures				
Betweenness centrality	Power	3656	-0.945	0.475
Closeness centrality	Exponential	0.007	-3.043×10^{-5}	0.017
Straightness centrality	Power	24.122	-0.274	0.428

Kernel density estimation is used to convert the density and centrality measures into the same scale to permit the cell-by-cell correlation between them. The study also investigates the role played by the Central Business District in socio-demographic diversity.

The statistical distribution of both the centrality measures and the socio-demographic variables is seen to follow a geometrically decreasing trend indicating that with the increase in the variables considered, the frequency of cells is considerably reduced. Single/multi-peak graphs are obtained denoting the concentration of the variables in certain locations. Correlation analysis reveals that out of the three centrality indices considered in the study, betweenness centrality acts as the driving force behind the socio-demographic diversity. Pearson's correlation coefficient is above 0.8 for all the values with a maximum value of 0.905. Employment density, in particular, shows the highest correlation with betweenness centrality. Betweenness centrality is followed by closeness centrality in acting as the decisive criteria in the socio-demographic diversity. To warrant these results, a location analysis of the CBD is also carried out to find the effect of the distance from the CBD and both the centrality indices and socio-demographic parameters.

The study confirms that the socio-demographic diversity can be accounted to the centrality measures of the road network. It can be concluded that the highly complex urban space is shaped by the transport network. A weighted node approach augments the reliability of the results. A further addition to this research work would be to compute the centrality indices using a weighted edge approach which could be taken forward by future research endeavours.

Code Availability No codes were used in the study.

Declarations This research complies with ethical standards.

Ethics Approval No humans or animals were used for this research. Hence not applicable.

Human and Animal Rights No humans or animals were used for this research. Hence not applicable.

Informed Consent Not applicable

Conflict of Interest The authors declare no competing interests.

References

- Abram D, Edwards N (2005) *The Rough Guide to South India. Rough Guides*
- Bai C et al (2010) 'Fractal characteristics of the road network in Jiangsu Province', in the sixth advanced forum on transportation of China.
- Burgess EW (1925) *The growth of the city: An introduction to a research project, in the city.* University of Chicago Press, Chicago
- Cai X, Wu Z, Cheng J (2013) Using kernel density estimation to assess the spatial pattern of road density and its impact on landscape fragmentation. *Int J Geogr Inf Sci* 27(2):222–230. <https://doi.org/10.1080/13658816.2012.663918>
- Chen S, Claramunt C, Ray C (2014) A spatio-temporal modelling approach for the study of the connectivity and accessibility of the Guangzhou metropolitan network. *J Transp Geogr* 36:12–23. <https://doi.org/10.1016/j.jtrangeo.2014.02.006>
- Daniel CB, Mathew S, Saravanan S (2021) 'Network constrained and classified spatial pattern analysis of healthcare facilities and their relationship with the road structure: a case study of Thiruvananthapuram city', *Spat Inf Res* <https://doi.org/10.1007/s41324-021-00385-7>
- Demšar U (2007) 'Centrality measures and vulnerability of spatial networks', in *ISCRAM*, pp. 201–209
- Duan Z, Wang Q (2009) 'Road Network Analysis and Evaluation of Huizhou City Based on Space Syntax', 2009 International Conference on Measuring Technology and Mechatronics Automation, pp. 579–582. <https://doi.org/10.1109/ICMTMA.2009.505>
- Fotheringham S, Brunsdon C, Charlton M (2000) 'Quantitative Geography: Perspectives on Spatial Data Analysis', in. Sage, London. Available at: [https://risweb.st-andrews.ac.uk/portal/en/researchoutput/quantitative-geography-perspectives-on-spatial-data-analysis\(c7b6913a-d7b0-4660-9538-11a117a5d0bf\)/export.html](https://risweb.st-andrews.ac.uk/portal/en/researchoutput/quantitative-geography-perspectives-on-spatial-data-analysis(c7b6913a-d7b0-4660-9538-11a117a5d0bf)/export.html) Accessed: 21 July 2021.
- Han Z et al (2019) 'Identifying spatial patterns of retail stores in road network structure', *Sustainability (Switzerland)*, 11(17).<https://doi.org/10.3390/su11174539>.
- Harris CD, Ullman EL (1951) *The Nature of Cities.* Glencoe, Illinois: The Free Press
- Henning S et al (2017) Methodology for Determining Critical Locations in Road Networks based on Graph Theory. *IFAC-PapersOnLine* 50(1):7748–7753. <https://doi.org/10.1016/j.ifacol.2017.08.1065>
- Hillier B et al (2007) 'Metric and topo-geometric properties of urban street networks: some convergences, divergences and new results', in *Proceedings, 6th International Space Syntax Symposium, İstanbul, 2007*
- Hoyt H (1939) *The structure and growth of residential neighborhoods in American cities.* Federal Housing Administration, Washington, DC
- Jamakovic A, Uhlig S, Theisler I (2008) On the relationships between topological metrics in real-world networks. *Netw Heterog Media* 3(2):345–359. <https://doi.org/10.3934/nhm.2008.3.345>
- Jayasinghe A, Sano K, Nishiuchi H (2015) 'Explaining Traffic Flow Patterns Using Centrality', 5(2), 134–149
- Jayaweera IMLN (2017) Centrality Measures to Identify Traffic Congestion on Road Networks: A Case Study of Sri Lanka. *IOSR J Math* 13(02):13–19. <https://doi.org/10.9790/5728-1302011319>
- Kim KS, Benguigui L, Marinov M (2003) The fractal structure of Seoul's public transportation system. *Cities* 20(1):31–39. [https://doi.org/10.1016/S0264-2751\(02\)00094-X](https://doi.org/10.1016/S0264-2751(02)00094-X)
- Li B, Gu H, Ji Y (2012) 'Study on fractal features of transportation network in Xuzhou City', 2012 2nd International Conference on Remote Sensing, Environment and Transportation Engineering, RSETE 2012 - Proceedings 3(1):3–6. <https://doi.org/10.1109/RSETE.2012.6260745>
- Lin, G., Chen, X. and Liang, Y. (2018) 'The location of retail stores and street centrality in Guangzhou, China', *Appl Geogr*, 100(April 2017), pp. 12–20. doi: <https://doi.org/10.1016/j.apgeog.2018.08.007>.
- Liu Z, Zhao S (2015) Characteristics of road network forms in historic districts of Japan. *Front Archit Res* 4(4):296–307. <https://doi.org/10.1016/j.foar.2015.08.003>

- Liu Y et al (2016) Relationships between Street Centrality and Land Use Intensity in Wuhan , China. *J Urban Plan Dev* 142(1):1–14. [https://doi.org/10.1061/\(ASCE\)UP.1943-5444.0000274](https://doi.org/10.1061/(ASCE)UP.1943-5444.0000274)
- Mishra S, Welch TF, Jha MK (2012) Performance indicators for public transit connectivity in multi-modal transportation networks. *Transp Res A* 46(7):1066–1085. <https://doi.org/10.1016/j.tra.2012.04.006>
- Ni J et al (2016) Spatial Distribution Characteristics of Healthcare Facilities in Nanjing : Network Point Pattern Analysis and Correlation Analysis. *Int J Environ Res Public Health* 13(833):1–13. <https://doi.org/10.3390/ijerph13080833>
- Porta S et al (2009) ‘Street centrality and densities of retail and services in Bologna Italy. *Environ Plann B: Plann Des* 36:450–466. <https://doi.org/10.1068/b34098>
- Porta S et al (2011) ‘Street Centrality and the Location of Economic Activities in Barcelona’, *Urban Studies*, pp. 1–18. doi: <https://doi.org/10.1177/0042098011422570>.
- Ratti C (2004) Space Syntax: Some Inconsistencies. *Environ Plann Plann Des* 31(4):487–499. <https://doi.org/10.1068/b3019>
- Rui Y, Ban Y (2014) Exploring the relationship between street centrality and land use in Stockholm. *Int J Geogr Inf Sci* 28(7):1425–1438. <https://doi.org/10.1080/13658816.2014.893347>
- Sahitya KS, Prasad CSRK (2020) Modelling structural interdependent parameters of an urban road network using GIS. *Spat Inf Res* 28(3):327–334. <https://doi.org/10.1007/s41324-019-00295-9>
- Sevtuk A, Mekonnen M, Kalvo R (2013) ‘Urban Network Analysis’, p. 29. Available at: <http://cityform.mit.edu/projects/urban-network-analysis.html>. Accessed 20 May 2019.
- Sheikh A, Zadeh M, Rajabi MA (2013) Analyzing the effect of the street network configuration on the efficiency of an urban transportation system. *Cities* 31:285–297. <https://doi.org/10.1016/j.cities.2012.08.008>
- Silverman BW (1986) *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London
- Telega A (2016) ‘Urban Street Network Analysis Using Space Syntax in GIS -- Cracow Case Study’, 2016 Baltic Geodetic Congress (BGC Geomatics), pp. 282–287. <https://doi.org/10.1109/BGC.Geomatics.2016.57>.
- Tobler WR (1979) Smooth Pycnophylactic Interpolation for Geographical Regions. *J Am Stat Assoc* 74(367):519–530. <https://doi.org/10.1080/01621459.1979.10481647>
- Wang F, Antipova A, Porta S (2011) Street centrality and land use intensity in Baton Rouge , Louisiana. *J Transp Geogr* 19(2):285–293. <https://doi.org/10.1016/j.jtrangeo.2010.01.004>
- Wang F et al (2014) Location analysis of retail stores in Changchun, China: A street centrality perspective. *Cities* 41(PA):54–63. <https://doi.org/10.1016/j.cities.2014.05.005>
- Xie F, Levinson D (2007) Measuring the structure of road networks. *Geogr Anal* 39(3):336–356. <https://doi.org/10.1111/j.1538-4632.2007.00707.x>
- Yu W, Ai T, Shao S (2015) The analysis and delimitation of Central Business District using network kernel density estimation. *J Transp Geogr* 45:32–47. <https://doi.org/10.1016/j.jtrangeo.2015.04.008>
- Zhang Y et al (2019) Exploring Location Pattern of Commercial Stores in Shichahai, Beijing from a Street Centrality Perspective. *Chin Geogr Sci* 29(3):503–516. <https://doi.org/10.1007/s11769-019-1045-z>
- Zhao S, Zhao P, Cui Y (2017) ‘A network centrality measure framework for analyzing urban traffic flow : A case study of Wuhan, China. *Physica A* 478:143–157. <https://doi.org/10.1016/j.physa.2017.02.069>
- Zhu J, Sun Y (2017) Building an Urban Spatial Structure from Urban Land Use Data: An Example Using Automated Recognition of the City Centre. *ISPRS Int J Geo Inf* 6(4):122. <https://doi.org/10.3390/ijgi6040122>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.