

Exploring the dynamics of spatial inequality through the development of sub-city typologies in English Bazar Urban Agglomeration and its peri urban areas

Ipsita Dutta · Arijit Das

Published online: 6 June 2018
© Springer Nature B.V. 2018

Abstract The term peri-urban can be defined as the settlements beyond, about or around core cities. It has often observed that the areas congested with the core city are enjoying some basic facilities/amenities as urban entity whereas the areas beyond the city are lagging behind. Thus the uneven distribution of amenities increases the inequality from core to peripheries. These heterogeneous characteristics within a city and around its periphery increase the inequalities which are characterized by the formation of sub-city typologies within the city and its peripheral settlements. The goal of this study is to explore extent and magnitude of spatial inequality in levels of living and the resultant sub-city typologies of English Bazar city and its peri-urban settlements, one of the most dynamic urban agglomerations of West Bengal. Village level House listing and housing data from the 2011 census are used here to find out the magnitude of inequality and construct sub-city typologies. 19 variables from the census are selected to represent three broad class of attributes such as housing quality, access to amenities and availability of household assets through different standardize indices for developing sub-city typology. Hierarchical and non-

hierarchical cluster analysis methods are then used to identify empirical typologies considering relevant principal factors from PCA analysis. It identifies a five cluster solution corresponds to five spatial typological categories (High SE-Area, Average SE-Area, Good SE-Area, Average Area, Low socio-economic area). Identification of five spatial typological categories helps to explore the urban inequality at micro level for English Bazar Urban Agglomeration and its peri-urban settlements. The result also shows that settlements lying contiguous to the Old Malda city are deprived in asset holding and household quality whereas the peripheral settlements lying along English Bazar city have better standard of living conditions.

Keywords Peri-urban · HQLI · Spatial metrics · Sub-cities · Spatial inequality

Introduction

Spatial variations in socio-economic, political, environmental, infrastructural condition within and its surrounding settlements of a city are key elements of urban heterogeneity (Allen 2003; Balakrishnan and Anand 2015). Heterogeneity in this sense is relatively different than the terms ‘inequality’ and ‘inequity’, where ‘inequality’ refers to the uneven distribution of an observed attribute and ‘inequity’ refers to the unfair distribution of an attribute. But a particular kind of

I. Dutta (✉) · A. Das
Department of Geography, University of Gour Banga,
Malda, West Bengal 732103, India
e-mail: ipsitadutta25@gmail.com

A. Das
e-mail: arijit3333@gmail.com

heterogeneity can often leads to socio-spatial inequity. It affects mobility, asset values, infrastructure, and housing and determines the heterogeneous characteristics of the core city and its periphery (Bhagat 2011; Bhan and Jana 2015). From this perspective, study of urban heterogeneity holds significance in this context (Keil 2017) because this heterogeneous characteristic manifest in socio-spatial inequalities within and around the city. The heterogeneity within and around a city can be characterized by the formation of sub-city typologies (Bhan and Jana 2015; Balakrishnan and Anand 2015). Sub-city typologies refer to a character where the cities that are more similar to each other create a typology. This formation of sub-city typology easily helps to identify the cities with same character which will enhance the chance to identify the character of the cities more clearly and will help in implementing government policies (Samanta 2012, 2017) for development of those areas. These sub-city typologies could enable a better understanding to study the urban heterogeneity. It also provides a new empirical method for categorical classification of cities which is different from the classifications that often includes different indices to show the ward inequalities (Balakrishnan and Anand 2015).

This study focuses on the pattern of spatial inequality and formation of sub-city typologies in English Bazar UA and its peri-urban areas (Shaw and Das 2017). At the level of the spatial organization of the urban agglomeration and its peri urban area, the main objectives of the study is to identify the pattern of heterogeneity in terms of different spatial (Ramachandra et al. 2012, 2014; Sudhira et al. 2004) (various landscape matrices) and non-spatial aspects (provision of basic amenities and facilities) and to find out the sub-city typologies on the basis of these characteristics around English Bazar Urban Agglomeration.

The release of the ward level data in 2011 census expands the scope by allowing a set of insights into different patterns of spatial inequalities. This paper uses the data to trace three vectors of inequality-housing conditions, provision of basic amenities and asset holding capacity. The first is a Quality of Housing Index (QHI) which measures the material adequacy of housing conditions, second is basic amenity index (BAI) which shows access to basic services such as water, sanitation, kitchen etc. The third is Asset Index (AI) to measure the relative purchasing power of the households. Along with these

non-spatial indices, three spatial indices were obtained using the Landsat data such as, number of urban patches (NP), edge density (ED) of urban area and aggregation index (AI) to quantify urban spatial growth and the overall landscape characteristics. Combining these two different dimensions of indices the present study tries to address the following research questions, (1) how the distribution of different attributes increases the spatial inequality in and around the core city? (2) How this kind of heterogeneity could be studied through the development of sub-city typologies? (3) How the spatial and non-spatial metrics interact to each other to depict the linearity and non-linearity between them? In this present study, principal component analysis has been done followed by the process of hierarchical and non-hierarchical cluster analysis to identify the sub-city typologies.

Literature review

Urban heterogeneity has been studied by the researchers from a wide range of perspectives (Roy 2011, 2016). There is a large literature where urban heterogeneity has been studied in physical aspects such as landuse/landcover and overall urban and environmental structure (Forman 2008; Hoffman et al. 2008). Most of the other studies have focused on the intra-urban variation by means of socio-spatial differentiation and residential segregation (Dupont 2004; Jamil 2014).

In 2014, after the release of the ward level household and houselisting data systematically in census of India 2011 (Balakrishnan and Anand 2015), very few attempts are made to study the intra-urban heterogeneity in India (Bhan and Jana 2015; Das and Mistri 2013). Das and Mistri (2013) have used state level housing and houselisting data to show the regional variation at state level. Bhan and Jana (2013) made an attempt to show the variations of distribution of basic facilities in slums and non-slums households in Manipur and identified the urban poor. Again in 2015 they use the data at ward level to construct two composite indices to study intra-urban variation in Delhi and Bengaluru and argued that the kind of variation also exist within different scales of settlements. But this study fails to construct a method to study the pattern of urban heterogeneity in medium

and small towns. Balakrishnan and Anand (2015) examined the pattern of urban heterogeneity in Bengaluru through the development of sub-city typologies within the city by following the cluster analysis method using the non-spatial data available from the census of India 2011.

But the present study more emphasizes to deal with the intra-urban variation in different spatial and non-spatial parameters in and around the city. Shaw (2005), Ramachandra and Bharath (2013a, b) urged the need to study of peri-urban areas of small cities as the areas outlying the city boundary, have an unsatisfactory level of environmental infrastructure services such as water supply, sanitation, drainage, solid waste management etc. as most of these areas do not fall under the planning processes. The basic services that are provided to the urban core do not cover the peri urban areas. Thus the uneven distribution of amenities increases the inequality from core to periphery. The release of the census data enables a much wider range of explorations of intra-urban variation in small cities and its peripheral settlements than was previously possible. From this perspective, no studies are done yet to study the variation between the core and periphery using both spatial non-spatial data.

Justification for selection of the study area

English Bazar Urban Agglomeration is located between latitudes of 24°59'N to 25°04'N and 88°06'E to 88°10'E longitudes of Malda District. English Bazar Urban Agglomeration has got its recognition as an urban agglomeration in the recent 2011 census with more than 3 lacks population which shares 57.91% of total urban population and a total area of only 28 sq. km. English Bazar UA comprises of 2 urban statutory bodies and 3 census towns. A dramatic change of urbanization has been observed in Malda District where the urbanization rate jumped to 124.81% which is considered as the fastest growing town of West Bengal. It is also need to note that, in 1991 there was no census town in Malda District, but in 2011 census the number has grown to 27.

In this regard, different backgrounds play important roles for this huge urban growth in the district. The location of Malda district in the middle of West Bengal act as a nodal point for the connection of two Bengals and it is also emerging as a major urban centre of North

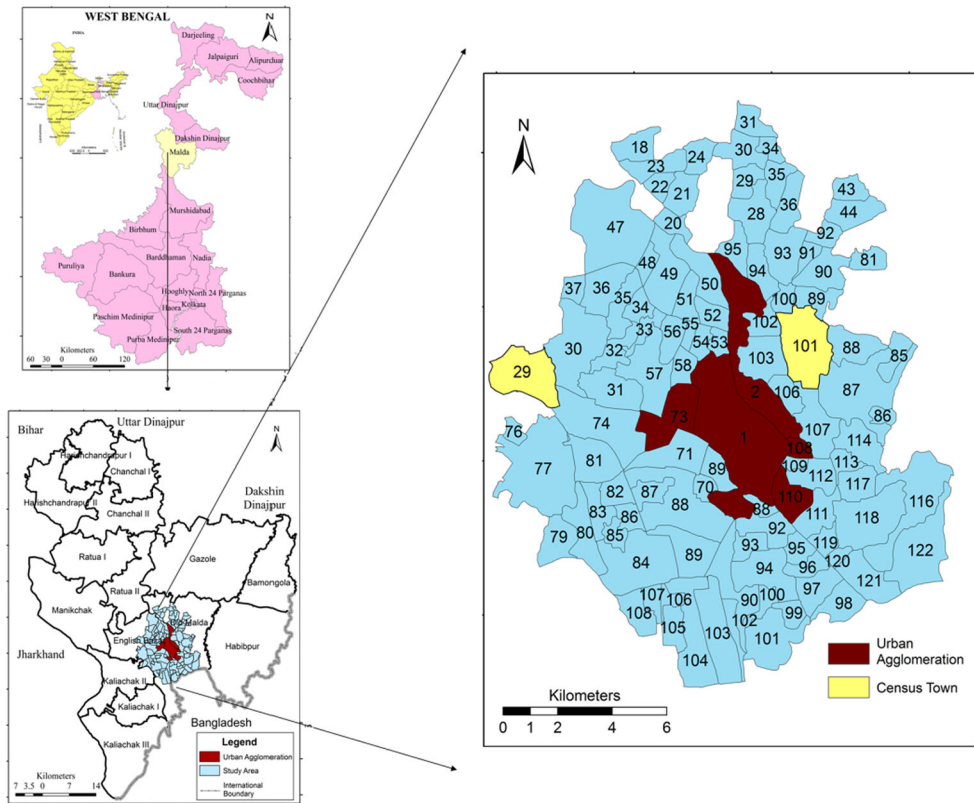
Bengal. The location of NH-34 and SH-10, and meeting point of N. F. Railway and Eastern railway across the main headquarter of Malda District and the longitudinal location of the river Mahananda made the transportation and communication easier. The figure of population even changed during the day time when most of the people not only from the surrounding villages but also from the neighboring districts migrated in the city core for their occupational purposes. All these account importance to the growth of the city core. Before 2011 census, no census town was identified around the headquarter of Malda District, i.e., English Bazar city. But in 2011 census, it got recognized as UA with constituting Old Malda Municipality and 3 census towns contiguous to the municipalities, namely Bagbari at the western boundary of English Bazar Municipality, Chhatianmore and Sahapur at the eastern boundary of Old Malda Municipality. English Bazar UA plays an important role in the urban as well as overall growth of the Malda District. However, the growth of the city core and its surroundings is not uniform across the peripheral areas in terms of different aspects. There is tremendous variation on distribution of amenities and facilities in the peripheral regions and urban core Intra-urban and inert-regional variation is high across different settlements. Some of the settlements within 103 peri-urban settlements which are considered for this study purpose may lack in some demographic characteristics, but most of them can be characterized by growing areas in terms of different social and demographic states (Fig. 1).

Materials and methods

This study uses the village level houselisting and housing data from 2011 Census of India for the purpose of the study. The full list of variables available from the 2011 Census which are selected at the village level is given in Fig. 2. The variables were selected from the Census data to represent three broad classes—housing conditions, access to basic amenities, and asset disparities. A total of 19 variables are selected for the purpose of the study (Fig. 2).

Methodology for computation of three indices

As part of the study, with this data in place, three indices were computed, (a) Quality of Housing Index



Name	J_L_NO	Name	J_L_NO	Name	J_L_NO	Name	J_L_NO
Patamari	18	Jot Gopal	49	Madhapur	86	Banshata	102
Kadamtai	20	Anandipur	50	Dilalpur	87	Krishna Prasad Nij	102
Badanpur	21	Uttar Gobindapur	51	Morgram	87	Maulpur	103
Chhotapara	22	Itakhola	52	Uttar Jadupur	88	Bara Phulbari	103
Lakshmi Kandar	23	Nima Sarai	53	Kamancha	88	Tiakati	104
Muralighat	24	Jatalpur	54	Dakshin Jadupur	89	Purana Guamatati	105
Jadupur	28	Jot	55	Maligram	89	Khaihata	106
Sonatala	29	Arapur	56	Gabgachhi	90	Arazi Phulbari	106
Kamaldighi	29	Sultanpur	57	Sukhandighi	90	Barachak	107
Shyampur	30	Nazirpur	58	Mehamanpur	91	Rasiladaha	107
Dhumadighi	30	Arazi Dilalpur	70	Maheshpur	92	Dharampur	108
Daulatpur	31	Uttar Ramchandrapur	71	Minapara	92	Chhatianmor	108
Gungaon	31	Bagbari	73	Jagadishpur	93	Nageshwarpur	109
Khalimpur	32	Lakshmiipur	74	Mandilpur	93	Sahapur	110
Uttar Nazirpur	33	Mahajampur	76	Nischintapur	94	Kadirpur	111
Sahazalapur	34	Madapur	77	Kaludewan	94	Nityanandapur	112
Gunshankrul	34	Karmmanigram	79	Gopalpur	95	Dilalpur	113
Jot Narasinha	35	Naoda Bazar	80	Sujapur	95	Dakshin Bhatra	114
Fatepur	35	Kamalpur	81	Raypur	96	Sadhail	116
Daibakipur	36	Ijjatpur	81	Gobindapur	97	Santipur	117
Kuriapara	36	Azimpur	82	Kumarpur	98	Sakamma	118
Khorda Jalalpur	37	Chandipur	83	Krishnapur	99	Jotgobinda	119
Jharpukuria	43	Sadullapur	84	Nrisimhapur	100	Azmatpur	120
Panchgaon	44	Mollapara	85	Aradpur	100	Adampur	121
Jot Basanta	47	Pathar Madhapur	85	Moslempur	101	Memua	122
Narhatta	48	Hazarat Nagar	86	Jhangra	101		

Fig. 1 Location of the study area

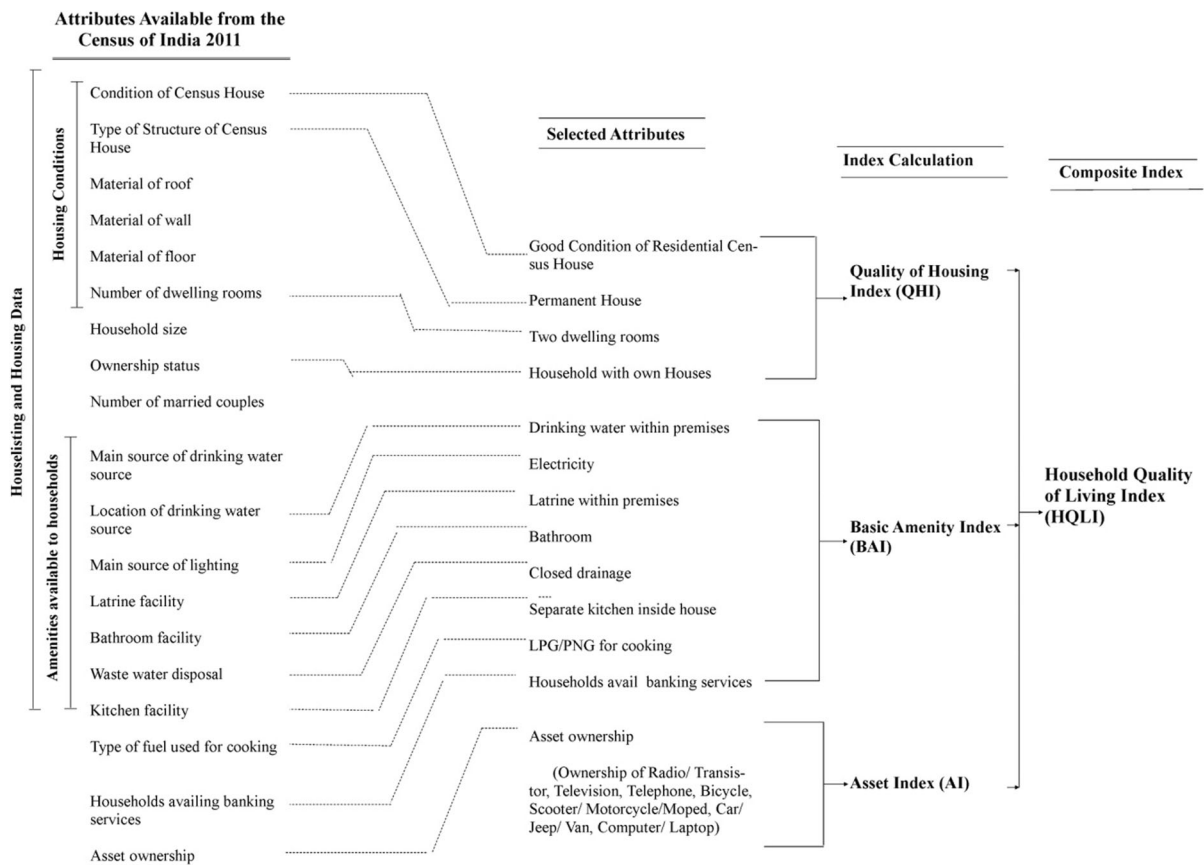


Fig. 2 List of selected variables for analysis

(QHI) to estimate the relative level of access to housing with certain characteristics, (b) Basic Amenity Index (BAI) to measure the access of basic amenities of households, and (c) Asset Index (AI) to estimate the relative purchasing power of households. The methodology used for building the indices was a modified form of composite index by applying z-score method. The index was calculated for each village to each characteristic. Apart from these, Household Quality of Living Index (HQLI) was calculated by the simple sum of all the three indices.

Methodology for computation of Household Quality of Living Index (HQLI)

To understand the present scenario of household quality of living for the periurban areas, Household Quality of Living Index (HQLI) is constructed on the basis of 19 variables opted from Census of 2011.

HQLI is computed with the three distinctive indices such as Quality of Housing Index (QHI), Basic Amenity Index (BAI) and Asset Index (AI). HQLI is then computed by the simple sum of these three indices.

$$HQLI = QHI + BAI + AI \tag{1}$$

The selected 19 variables are divided into these three indices as follow

Quality of Housing Index (QHI)

4 variables which are selected for computing QHI index are

1. HHs by good condition of residential census houses
2. HHs living in permanent houses
3. HHs with own houses
4. HHs having at least two or more dwelling rooms

Basic Amenity Index (BAI)

For computing BAI, 8 variables were taken into consideration.

1. Drinking water with in premises
2. Electricity
3. Latrine within premises
4. Bathroom
5. Closed drainage system for waste water outlet
6. Separate kitchen inside the house
7. LPG/PNG for cooking
8. Banking service

Asset Index (AI)

HHs having following 7 assets

1. Radio/Transistor
2. Television
3. Telephone facilities (mobile, landline or both)
4. Bicycle
5. Scooter/Motorcycle/Moped
6. Car/Jeep/Van
7. Computer/Laptop (with or without internet)

HQLI was generated by the combination of the above mentioned three indices.

Methodology to identify sub-city typologies

Combination of principal component analysis (PCA) and cluster analysis are the main methods used for better understanding and identifying typologies of cities. PCA and cluster analysis has been done in SPSS 16.0 and STATA 9.0.

Method for principal component analysis

Principle component analysis (PCA) is a mathematical process that transforms the data to a new projection system such that the greatest variance by some projection of the data lie on the first coordinate which is referred as the first principal component, the second greatest variance on the second coordinate, and so on (Chatterjee et al. 2015). PCA is used in this study as it helps to derive a small number of components on which cluster analysis is carried out. In this analysis,

household and housing data matrix x , with column-wise means where each of the n rows represents a different repetition of the experiment, and each of the p columns gives a particular kind of datum (say, the results from a particular sensor).

Mathematically, the transformation is defined by a set of p dimensional vectors of weights or loadings $w_k = (w_1 \dots w_p)_{(k)}$ that map each row vector x_i of x to a new vector of principal component scores $t_i = (t_1 \dots t_p)_{(i)}$, given by $t_{k(i)} = x_{(i)}w_{(k)}$ in such a way that the individual variables of t considered over the data set successively inherit the maximum possible variance from x , with each loading vector w constrained to be a unit vector (Chatterjee et al. 2015).

First component

The first loading vector w_1 thus has to satisfy,

$$\begin{aligned} w_{(1)} &= \arg \max_{\|w\|=1} \left\{ \sum_i (t_1)_{(i)}^2 \right\} \\ &= \arg \max_{\|w\|=1} \left\{ \sum_i (x_{(1)}w)^2 \right\} \end{aligned} \quad (2)$$

Equivalently, writing this in matrix form gives

$$w_{(1)} = \arg \max_{\|w\|=1} \left\{ \|xw\|^2 \right\} = \arg \max_{\|w\|=1} \left\{ w^T x^T x w \right\} \quad (3)$$

Since w_i has been defined to be a unit vector, it equivalently also satisfies

$$w_{(1)} = \arg \max \left\{ \frac{w^T x^T x w}{w^T w} \right\} \quad (4)$$

A standard result for a symmetric matrix such as $x^T x$ is that the quotient's maximum possible value is the largest Eigen values of the matrix, which occurs when w is the corresponding eigenvector (Chatterjee et al. 2015).

With $w_{(1)}$ found, the first component of a data vector x_i can then be given as a score $t_{1(i)} = x_{(i)}w_{(1)}$ in the transformed co-ordinates, or as the corresponding vector in the original variables, $\{x_{(i)}w_{(1)}\}_{w_{(1)}}$.

Further component

The k th component can be found by subtracting the first $k - 1$ principal components from x ,

$$\hat{x}_k = x - \sum_{s=1}^{k-1} xw_{(s)}w_{(s)}^T \tag{5}$$

and then, finding the loading vector which extracts the maximum variance from this new data matrix,

$$w_k = \arg \max_{\|w\|=1} \left\{ \|\hat{x}_k w\|^2 \right\} = \arg \max \left\{ \frac{w^T \hat{x}_k^T \hat{x}_k w}{w^T w} \right\} \tag{6}$$

This equation gives the remaining eigenvectors of $x^T x$, with the maximum values for the quantity in brackets given by their corresponding Eigen values. Thus the loading vectors are eigenvectors of $x^T x$ (Chatterjee et al. 2015).

The k th component of a data vector $x_{(i)}$ can therefore be given as a score $t_{k(i)} = x_{(i)}w_{(k)}$ in the transformed coordinates, or as the corresponding vector in the space of the original variables, $\{x_{(i)}w_{(k)}\}w_k$, where $w_{(k)}$ is the k th Eigenvector of $x^T x$.

The full principal components decomposition of x can therefore be given as

$$T = xw \dots \tag{7}$$

where w is a $p \times p$ matrix whose columns are the Eigenvectors of $x^T x$ of household quality data.

Method for cluster analysis

Cluster analysis is a technique normally used to classify cases into groups that are relatively homogeneous within themselves and heterogeneous between each other on the basis of a set of defined variables. This method can be grouped into hierarchical and non-hierarchical methods (Balakrishnan and Anand 2015). The study runs both hierarchical and non-hierarchical methods.

Hierarchical clustering

Hierarchical clustering is first used to explore the data set and identify the number of distinct clusters that may be present. Then this solution is used to specify the number of clusters in the non-hierarchical clustering procedure (Balakrishnan and Anand 2015). Hierarchical clustering algorithm is divided into two groups, i.e., agglomerative and divisive hierarchical algorithm. In this study agglomerative hierarchical

clustering was chosen for the analysis and to determine a number of clusters for the data set. The whole process includes the following stages,

Selecting linkage method

Linkage method calculates the distance from a cluster centre to a certain case. This study uses the Ward’s Linkage method. With this method, groups are formed so that the pooled within-group sum of squares is minimized. That is, at each step, the two clusters are fused which result in the least increase in the pooled within-group sum of squares.

$$\alpha_i = \frac{n_i + n_m}{n_k + n_m}, \quad \alpha_j = \frac{n_j + n_m}{n_k + n_m}, \quad \beta = \frac{-n_m}{n_k + n_m}, \quad \gamma = 0 \tag{8}$$

Representation of cluster results

Result of hierarchical clustering can be seen through dendrogram (Fig. 6) that represents each merge at the similarity between the two merged groups.

Distinctive breaking (step of elbow)

Hierarchical clustering procedures provide information that allows identifying the gaps that define logical clusters based on the output. From the agglomeration schedule table obtained from Ward’s Linkage algorithm, it is important to identify the step (step of elbow) where the “distance coefficients” makes a bigger jump. After then the number of cluster can be calculated as:

$$\begin{aligned} \text{Number of cluster} \\ = (\text{Number of cases} - \text{Step of elbow}) \end{aligned} \tag{9}$$

Stopping rules

The stopping rules are basically statistical tests that check for the presence of underlying clusters within the data and the number of such clusters that can be identified. Calinski–Harabasz rule and the Duda–Hart rule are used as statistical methods for identifying the optimum clustering solution (Balakrishnan and Anand 2015). For identifying the optimum number of clusters, the highest value for the Calinski–Harabasz

pseudo-F statistic is selected, while in the latter, the cluster solution with a combination of higher Duda–Hart index value and lower pseudo-T-squared value is selected (Balakrishnan and Anand 2015).

Calinski and Harabasz rule Variance ratio criterion (VRC) introduced by Calinski and Harabasz is a widely used criterion that computes ratio of between and within-cluster sums of squares for k clusters (Zenina and Borisov 2013). The expected optimal solution of this criterion is the number of clusters that maximizes the value of the variance criterion and minimizes ω_k value. It is expressed as,

$$VRC_k = \frac{B_k}{k-1} \cdot \frac{n-k}{W_k} \quad (10)$$

where VRC_k = Variance ratio criterion, k = the number of cluster, B_k = the overall between-cluster variation, W_k = the overall within-cluster variation with respect to all clustering variables, n = data objects.

ω_k value should be computed for each cluster solution to determine the suitable number of clusters It is expressed as,

$$\omega_k = (VRC_{k+1} - VRC_k) - (VRC_k - VRC_{k-1}) \quad (11)$$

The J-index The J-index proposed by Duda and Hart compares the within-cluster sum of squared distance with the sum of within cluster sum of squared distances and decides whether cluster should be partitioned into two clusters. The hypothesis that cluster could be subdivided is rejected if DH value is more than a standard normal quantile (Zenina and Borisov 2013).

$$DH = \left(-\frac{W_2}{W_1} + 1 - \frac{2}{\pi p} \right) \left(\frac{2(1 - 8/\pi^2 p)}{np} \right)^{-1/2} > Z_{1-\alpha} \quad (12)$$

where p = the number of variables, n = the number of objects in the studied cluster and $Z_{1-\alpha}$ = the standard normal quantile.

Non-hierarchical clustering For non-hierarchical clustering, k-means algorithm is used. In this algorithm, objects are moved to the group whose group means it is closet to. The k-means algorithm follows an entirely different concept than the hierarchical methods discussed earlier. This

algorithm is not based on distance measures such as Euclidean distance or Manhattan distance, but uses the within-cluster variation as a measure to form homogenous clusters. Specifically, the process aims at segmenting the data in such a way that the within-cluster variation is minimized.

While applying the k-means clustering process the researcher has to pre-specify the number of clusters to retain from the data. This makes k-means less attractive to some. However, the VRC discussed above can likewise be used for k-means clustering

Method for computing landscape metrics

Different spatial metrics such as, number of urban patches (NP), edge density (ED) of urban area and aggregation index (AI) were calculated for each settlement to quantify urban spatial growth and the overall landscape characteristics using FRAGSTATS 4.2.5 (McGarigal and Marks 1995) (Table 1).

Results and analysis

In this section, the result was carried out through regional and spatial distribution to look at the differentiation and disparities in case of distribution and development at the local level in different dimensions.

Regional and spatial pattern of housing facilities

Housing is the basic need for every person to live a better life. Regional level analysis from the census data shows that there is a huge disparity at different direction of housing condition. The result shows that areas outlying the EBM are ahead of the areas around OMM. Over 26.85% households are occupied as good condition residential house whereas it is 22.83% for the areas around OMM. Gungaon, Kuriapara, Nazirpur, Gunshankrul are the villages that occupies the highest proportion of house in good condition occupied by HHs and Shyampur, Khalimpur, Itakhola, Patamari etc. are some villages that are in lowest position. The spatial pattern of HHs by good condition of census houses depicts that more than 32% villages around EBM and 27% villages around OMM lives in good condition census houses (Fig. 3a).

Table 1 Selected landscape metrics calculated in this study

Indicator	Formula	Description	Significance
<i>Class level landscape metrics</i>			
Number of patches (built up) (NP)	n_i =number of patches in the landscape of patch type i Range, $NP \geq 1$	NP is the number of patches of a corresponding patch type	Increase in number of patches indicates the land fragmentation, while decline in patches indicates the formation of cluster
Edge density (ED)	E = total length (m) of edge in landscape A = total landscape area (m^2) Range, $ED \geq 0$	ED standardizes edge on a per unit area basis that facilitates comparison between landscapes of different sizes	It indicates the aggregation or dispersion of urban growth over an area, where increase in edge density indicates the formation of aggregated growth
Aggregation index (AI)	g_i = number of like adjacencies (joins) between pixels of patch i based on the single-count method $max-g_i$ =maximum number of like adjacencies (joins) between pixels of patch type i based on the single-count method	AI equals to 0 when the patch types have no like adjacencies. AI equals 100 when the landscape consists of a single patch	It shows the degree of patch compactness of an urban area

Nearly 41% of HHs live in permanent houses, 54% in semi-permanent and 3% in temporary houses in the study area. HHs living in permanent houses is more around EBM than the villages around OMM. Over 55% of HHs live in permanent houses around EBM which is nearly more than double of villages around OMM. The highest proportion of HHs lives in permanent houses in Dilalpur (91.7%), Jatalpur (89.2%), Krishnapur (87%) in English Bazar and Nageshwarpur (72.7%), Nityanandapur (71.5%), Chhatianmore (70.6%), Kadamtali (70.5%) in Old Malda according to Census 2011 data. The spatial distribution of HHs by permanent houses is shown in Fig. 3b.

Nearly 94% HHs in the study area live in their own houses and nearly 1% HHs in rented houses. The result shows that 97% HHs in the villages around OMM live in their own houses and it is nearly 90% for the HHs of the villages around EBM. In the villages around EBM, nearly 1.42% HHs lives in rented house. Some villages in English Bazar like Jatalpur, Maheshpur, Sultanpur and Aradpur, Sujapur, Jhangra in Old Malda are having a large number of HHs live in rented house. The reason may be the rapid urbanization of these villages.

Another important indicator to assess the household quality of living is congestion factor which is expressed as percentage of households with no

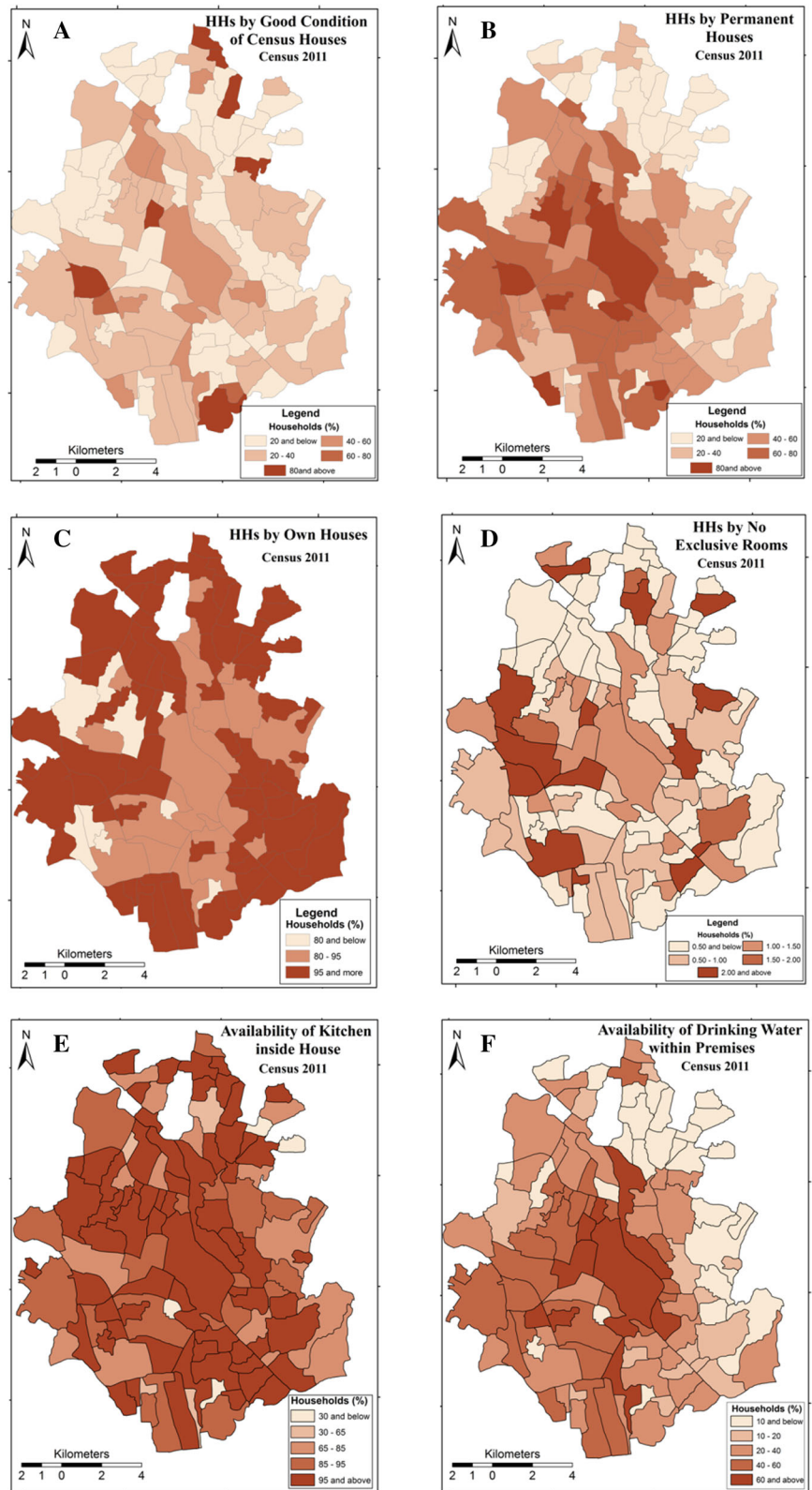
separate room for married couple (Das and Mistri, 2013). There could be HHs which have two or more rooms but may not have a separate room for couple's individual use. However, current census shows that total 0.96% HHs have no exclusive room for the couples to live in. The congestion factor for villages around EBM is 1.08% whereas it is 0.82% for villages around OMM. The highest proportion of HHs that has no exclusive room for the couples to live is Shyampur (15.4%) in English Bazar and Jadupur (5.1%), Azmatpur (4.3%) in Old Malda. This distribution is well shown in Fig. 3d.

Regional and spatial pattern of basic amenities

Provision of basic amenities is an important factor for measuring the household quality of living. There are various kinds of basic amenities. Among them, some of the most important elements are access to safe drinking water, sanitation, electricity etc.

Provision of drinking water source in HH premises designates more gender empowerment as women and girls spend a long time and energy in bringing water and that they will be relieved from that servitude (Das and Mistri 2013). On an average, about 39% HHs have drinking water within their premises in villages around EBM and 19% HHs in villages around OMM. This is

Fig. 3 Spatial distribution of housing (a–d) and amenity facilities (e, f)



clear that availability of drinking water source widely vary between areas around EBM and OMM. In Fig. 3f, it is found that there is a wide strip from north-west to south-west direction where more than 40% HHs collect water from premises.

The scenario for electricity facility also shows a wide regional and spatial variation. Nearly 50% HHs in areas around EBM get electricity facility where this is half for the areas around OMM. Census 2011 depicts wide regional and spatial variation of electricity usages by HHs. Villages from EBM like Arapur, Jatalpur, Jot, Nazirpur, Uttar Nazirpur, Gabgachi, Maheshpur, Azimpuretc and from OMM like Nageshwarpur, Aradpur, Jhangra etc. (Fig. 4c) are highly electrified (more than 60%).

Another important basic need is sanitation. While talking about the health and hygiene of rural HHs, the main two factors which are emerged are access to safe drinking water and sanitation (Das and Mistri 2013). Sanitary toilet within or near the premises, provides privacy and dignity to women (Das and Mistri 2013). Census 2011 portrays the clear picture of sanitation facilities of HHs. Nearly 40% HHs have their latrine within premises in villages of EMM and 24% in villages of OMM. Only in two villages in periurban areas around both EBM and OMM, more than 80% HHs enjoy latrine facilities within premises (Fig. 4a). It is clearly found that the villages which are congested with the two municipal boundaries have more number of HHs enjoying latrine facility within premises and the villages towards the outskirts have less number of HHs.

Availability of separate kitchen inside the house is another important factor of basic amenities. Apart from the drinking water facilities, latrine facilities, Census of India provides data on drainage facilities and households by availability of kitchen facilities. It is found that most of the HHs (90%) cook inside their house in the study area. For the areas around EBM, it is 93.26% and for the areas around OMM it is 89.84%. Shyampur, Daulatpur, Khalimpur, Kadamtali, Chhotapara are some villages which have maximum number of HHs having kitchen inside house (Fig. 3e). On contrary to this, the use of LPG/PNG is substantially low (6.03%). Only one village around EBM has household over 40% which use LPG/PNG for cooking.

Nowadays, banking service has become the most important public service for citizens (Das and Mistri 2013). Nearly 34% HHs have banking facilities in the

study area. However 45% HHs in villages around EBM and only 22% in villages around OMM enjoys banking facilities. Figure 4f shows the spatial variation of HHs having banking facilities.

Regional and spatial pattern of household assets

Household assets hold an important position for measuring poverty of households. Household asset is a broad concept which includes all financial, non-financial or human, social, natural, physical and financial assets (Das and Mistri 2013).

Census of India provides information on many assets. Possession of telephone (includes both mobile and landline) and bicycle is considerably high in the study area. Nearly 36% HHs from English Bazar area and 30% HHs from Old Malda area have telephone connection. After the telecom sector, another important sector is information technology. 6.84% HHs in the periurban areas of EBM and 9.48% HHs in the periurban areas of OMM use computer/laptops. Figure 4e, g show the spatial distribution of HHs using telephone and computer/laptop facilities. It is clearly highlighted from Fig. 4e that in most of the villages more than 30% HHs enjoys telephone facilities, whether Fig. 4g shows that there are very less villages where more than 10% HHs use computer/laptops. But this picture is different for four villages around OMM e.g. Maligram, Maulpur, Banshata, Jotgobinda where more than 40% HHs enjoy computer/laptops facility.

Number of HHs having television is also high for the periurban regions. HHs having television in the areas around EBM and OMM are 20% and 12% respectively. Dwellers of these periurban areas use bicycle for their daily work over car/jeep/van or scooter/motorcycle.

Spatial pattern of inequality and quality of household living

Table 2 shows the value of each index for the periurban areas around EBM and OMM. It is observed that the villages around EBM are having high index value than the villages around OMM. It may be because as English Bazar Municipality is called the heart of the District, therefore the villages surrounding the municipality enjoys some common facilities.

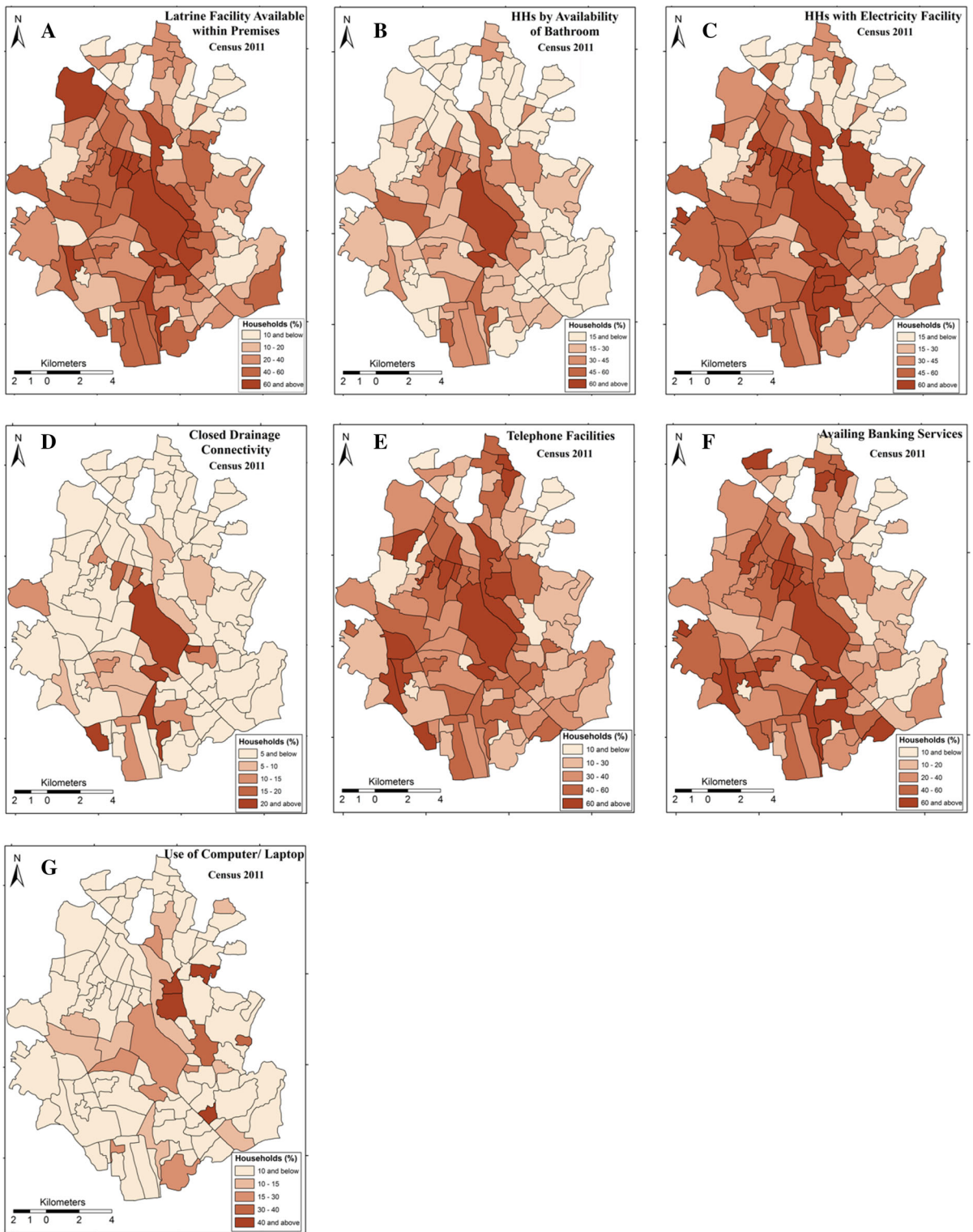


Fig. 4 Spatial distribution of amenities (a–d, f) and asset (e, g) facilities

Therefore the villages adjacent to the municipality ranked high for the indices. Table 2 also shows that there is a huge difference for the mean value HQLI between these two regions.

Figure 5a–d represent Quality of Housing Index (QHI), Basic Amenity Index (BAI), Asset Index (AI) and Household Quality of Living Index (HQLI) respectively. English Bazar Municipality holds the highest value for household quality of living followed by, Dharampur, Gabgachi, Jot, Jatalpur etc. Lowest value for household quality of living is observed among Mollapara, Shyampur, Minapara, Jot Narasinha, Ijjatpuretc (Fig. 5d). It has been observed that most of the villages having low quality of household living are from Old Malda area. This is because urbanization is more concentrated around English Bazar Municipality which reflects on the socio-economic standard of the entire region.

Table 3 shows the correlation among the computed indices. It highlights that the above mentioned indices are strongly and positively related to each other. The result shows that, there is very strong relationship among the indices. All the relations are significant at 1% level of significance. It clearly depicts the fact that, if one index is performing well then all the indices will tend to perform well.

Formation of sub-groups and identification of deprived regions

The analysis was carried out by following principal component analysis method and cluster analysis method respectively.

Principal component analysis

19 variables related to household quality are taken into consideration for PCA. The data were run through

PCA by applying varimax rotation method (Table 4). After passing the data through PCA, the PCs with Eigen values more than 1 have been considered. As such, five such PCs have been obtained.

The first Principal Component (PC-1) involves those variables which are more related to basic amenities of households, such as latrine facilities within premises, access to bathroom, having drinking water facilities within premises, access to electricity, use of LPG/PNG for cooking etc. These variables have highest value on PC-1. This first Principal Component explains 42.98% of the variation of data that are primarily included into household quality. The second Principal Component (PC-2) takes variables that represent the quality of house and asset holding capacity of HHs (access to scooter/motorcycle/moped, HHs having two or more dwelling rooms, telephone, television, good condition residential house etc.). It explains 9.95% of the variation of data. The third Principal Component (PC-3) includes two facilities as well (quality of house and asset holding capacity of HHs). PC-3 explains 7.08% of variation of data. The fourth Principal Component (PC-4) and fifth Principal Component (PC-5) involve the variables which are more related to asset facilities of HHs with an explanation of 6.37% and 5.94% variation of data respectively. These five PCs together explain about 72.31% variation in the data.

High score of PC-1 for a locality suggests that the locality enjoys good status with respect to the variables that under lie PC-1 i.e., latrine facilities within premises, access to bathroom, having drinking water facilities within premises, access to electricity, use of LPG/PNG for cooking etc. The localities with higher PC-2 and PC-3 score are more preferred by people for their living due to good quality of housing and high access to asset. Localities with high score of PC-4 and PC-5 are not found very suitable residential places as these localities only emphasizes on having access to assets.

Table 2 Regional pattern of household quality of living index. *Source:* Computed by author

	Indices	QHI	BAI	AI	HQLI
Peri urban areas around EBM	Mean	0.28	0.35	0.17	0.84
	SD	0.61	0.66	0.57	1.62
Peri urban areas around OMM	Mean	– 0.11	– 0.35	– 0.15	– 0.56
	SD	0.48	0.56	0.52	1.39
Range	Mean	0.39	0.70	0.32	1.40

Fig. 5 Spatial pattern of inequality

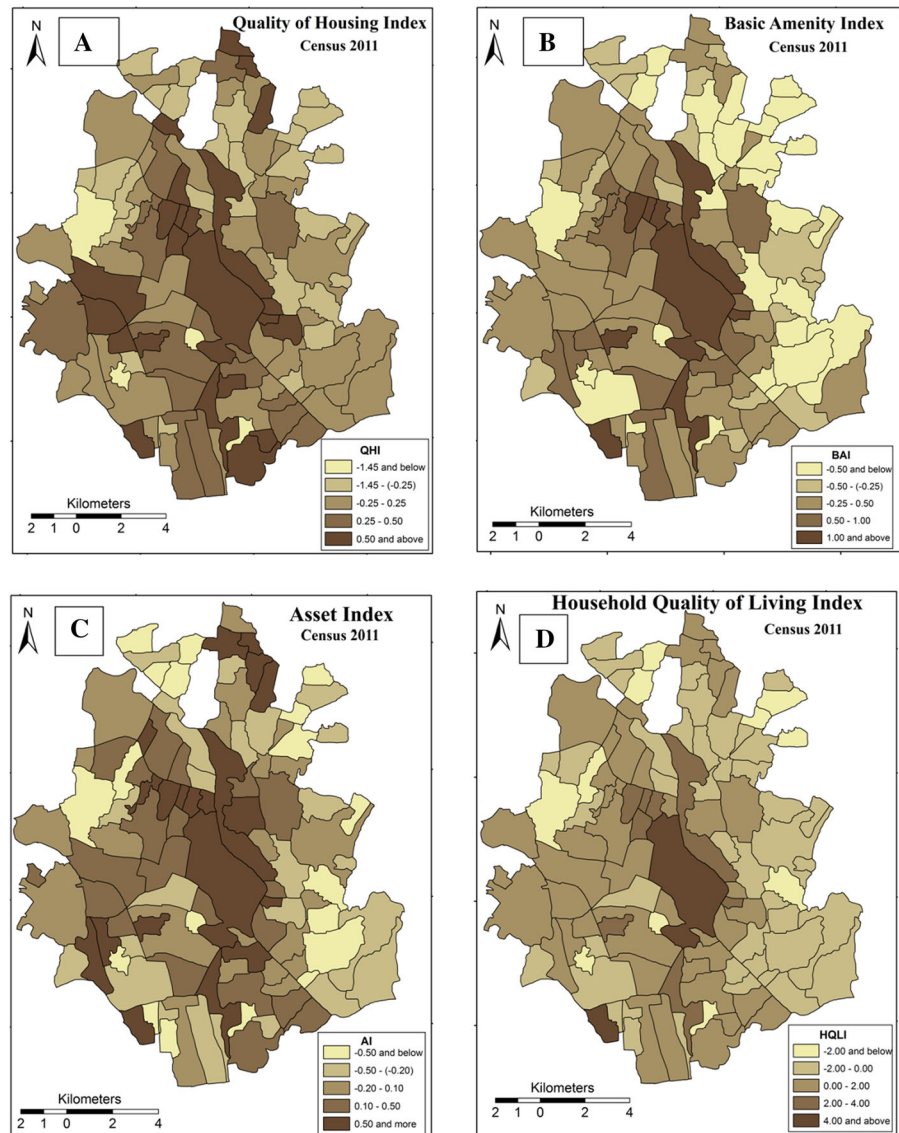


Table 3 Correlation among different indices. *Source:* Computed by author

	QHI	BAI	AI	HQLI
QHI	1	.666**	.649**	.873**
BAI	.666**	1	.745**	.912**
AI	.649**	.745**	1	.882**
HQLI	.873**	.912**	.882**	1

**Correlation is significant at the 0.01 level (2-tailed)

Cluster analysis method

In this section, cluster analysis was conducted initially on the principal components that are obtained by the principal component analysis method as it will better help to identify the form of cluster. Five principal components were obtained through principal component analysis method.

While analyzing the results, it is important to keep in mind that there are various decisions involved in arriving at a clustering solution which includes choice of variables, clustering method/s, algorithms, stopping rules, and starting partitions.

Table 4 Rotated component matrix of different variables

Variable	PC-1	PC-2	PC-3	PC-4	PC-5
Good condition of residential census house	0.283	0.313	0.273	− 0.247	0.415
Permanent house	0.719	0.301	0.281	0.114	− 0.238
Household with own houses	− 0.018	0.011	0.842	− 0.094	0.055
HHs having two or more dwelling rooms	0.321	0.428	0.504	0.309	0.096
Drinking water within premises	0.839	0.211	0.192	− 0.031	− 0.016
Electricity	0.808	0.141	0.215	0.207	− 0.188
Latrine within premises	0.856	0.213	0.195	− 0.077	0.01
Bathroom	0.841	0.111	0.068	− 0.118	0.009
Closed drainage	0.426	0.795	− 0.023	0.003	− 0.069
Separate kitchen inside house	0.187	0.032	0.725	0.021	0.025
LPG/PNG for cooking	0.784	0.199	− 0.063	− 0.062	0.031
Banking services	0.564	0.115	0.173	0.429	0.262
Radio/transistor	− 0.197	0.03	0.083	0.011	0.865
Television	0.856	0.353	0.142	0.136	− 0.01
Telephone (mobile, landline or both)	0.615	0.423	0.33	0.153	0.224
Bicycle	0.354	0.252	0.475	0.478	0.229
Scooter/motorcycle/moped	0.183	0.909	0.099	0.108	0.113
Car/jeep/van	0.585	− 0.058	− 0.078	0.196	0.397
Computer/laptop (with or without internet)	0.08	− 0.049	0.071	− 0.86	0.077

Dendrogram and distinctive breaking point (step of elbow)

Hierarchical cluster analysis using Ward's linkage on five PCs is shown in Fig. 6. This dendrogram suggests that a five-cluster solution would be optimal. This was again confirmed by plotting of coefficients taken from agglomeration schedule and using the stopping rules. The scree diagram is plotted in Fig. 7. It shows that in 102 steps the “distance coefficients” makes a bigger jump. Therefore the number of cluster obtained from this method will be 5 (as number of cases in this study is 107 and step of elbow is 102).

Calinski–Harabasz and Duda–Hart rules

Lastly a crosscheck was done using the Calinski–Harabasz and the Duda–Hart rule (Table 5). It is found that the five-cluster solutions have high values for the pseudo-F statistic as per the Calinski–Harabasz stopping rule. It also has the best combination of a high Duda–Hart Index value and low pseudo-T-squared value as per the Duda–Hart rule.

K-means cluster

After this, a five cluster solutions was generated using k-means non-hierarchical cluster analysis. The result is shown in Table 6.

Cluster 1 is very much related to the factors that explain Principal Component 1. It means, cluster 1 is involved with those variables which are involved for evaluating Principal Component 1 such as latrine facilities within premises, access to bathroom, having drinking water facilities within premises, access to electricity, use of LPG/PNG for cooking etc. This sub-group typology is referred as “High SE-Area” (High socio-economic area).

Cluster 2 is similar to Principal Component 5 which means this cluster involves the variables dominant in Principal Component 5. This typology can be called as “Average SE-Area” (Average socio-economic area).

Cluster 3 is characterized by the variables which are dominant in Principal Component 2, i.e., access to scooter/motorcycle/moped, HHs having two or more dwelling rooms, telephone, television, good condition

Fig. 6 Ward’s linkage dendrogram

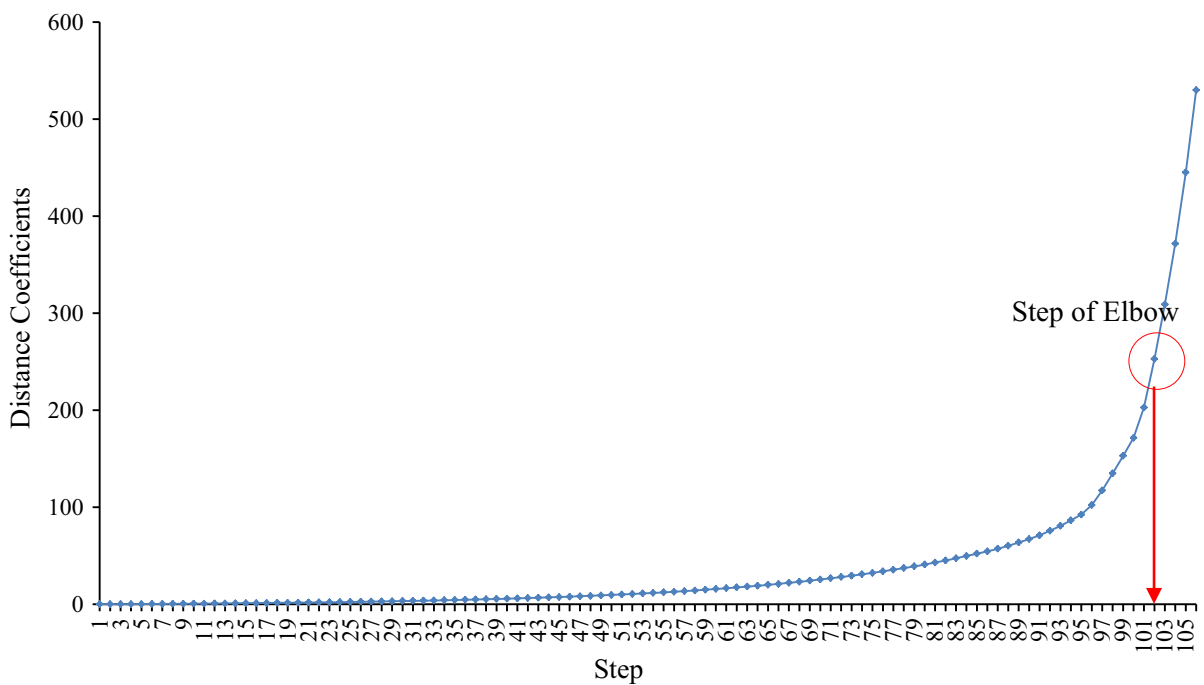
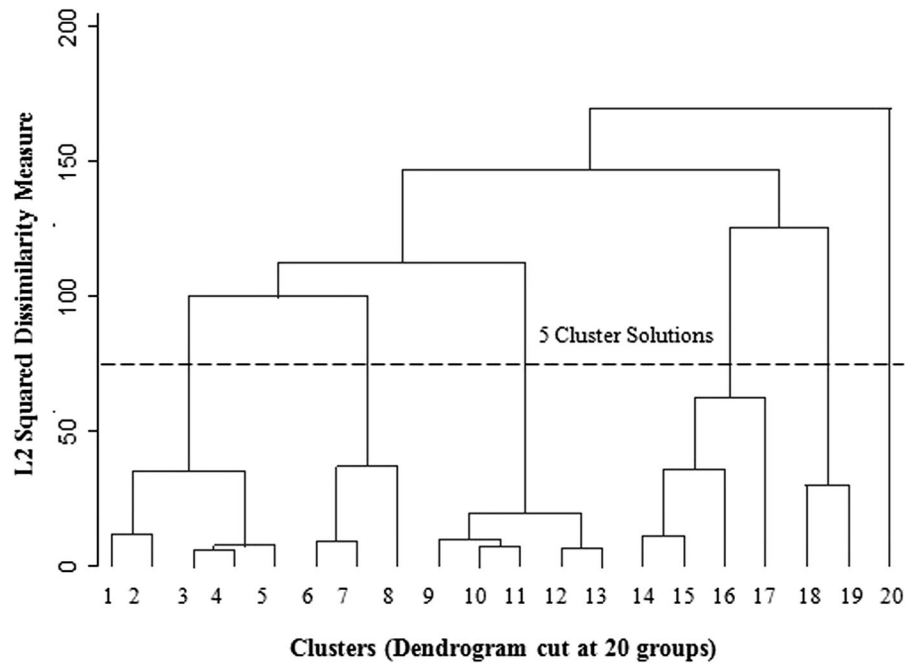


Fig. 7 Plotting of the coefficients of ward’s linkage

residential house etc. This sub-group typology is called “Good SE-Area” (Good socio-economic area).

Cluster 4 is very much similar to Principal Component 3 and Principal Component 5 which means it is

characterized by the dominant variables in these two components such as the variables under quality of house and asset holding capacity of HHs. This typology can be called as “Average Area”.

Table 5 Values for Calinski–Harabasz and Duda–Hart rules

Number of clusters	Calinski/Harabasz	Duda/Hart	
	Pseudo-F	Je(2)/Je(1)	Pseudo T-squared
2	20.04	0.84	20.55
3	22.16	0.58	13.81
4	24.55	0.75	28.04
5	27.98	0.61	24.14
6	32.61	0.53	11.39
7	34.85	0.43	13.21
8	34.88	0.48	12.12
9	35.91	0.60	16.54
10	37.93	0.22	14.39
11	40.15	0.75	14.53
12	40.93	0.45	6.04
13	40.22	0.58	5.08
14	39.82	0.75	9.12
15	39.45	0.66	4.55

Cluster 5 is far from Principal Component 3 and not particularly similar to any profile. This area is basically characterized by low development. Settlements that come under this typology are called as “Low SE-Area” (Low socio-economic area).

In summary, it can be said that, cluster analysis on major Principal Components provides us five sub-group typologies for the study area. These can be called High SE-Area, Average SE-Area, Good SE-Area, Average Area and Low SE-Area. This sub city typology is shown in Fig. 8.

The sub-group typology map in Fig. 8 shows that maximum numbers of settlements (64) are falling into the category of High SE-Area in terms of their household quality of living. 12 settlements are considered as average SE-area whereas only 1 settlement is classified as good SE-area. Average and low SE settlements are dispersedly distributed towards

different directions. It comprised those settlements which are located outwards from the core.

Relationship between spatial metrics and non-spatial data

Urbanization is seen as a process by which a region is grown not only by its built up extension but also it is expected to enjoy more quality in standard of living and housing condition. It has been seen from the study that, most of the settlements from English Bazar block performs well in different non-spatial indices, such as, QHI, BAI, AI, HQLI etc. But the distribution of high basic facilities is not linear with the built up growth. Therefore this study has attempted to visualize the growth of urban and the quality of living to show whether any linear relationship existed between them.

Table 6 Cluster classes along with principal components values

Components	Cluster				
	1	2	3	4	5
Principal component 1	0.40	– 0.06	– 1.59	– 0.39	– 0.95
Principal component 2	– 0.03	– 0.11	8.93	– 0.15	– 0.21
Principal component 3	0.29	0.06	0.37	0.26	– 1.06
Principal component 4	0.12	0.77	0.72	– 2.37	0.18
Principal component 5	– 0.31	2.06	– 1.09	0.28	– 0.30
No. of settlements	64	12	1	9	21

Housing Quality of Living Index (HQLI) is a strong indicator for non-spatial data. It shows the aggregated condition of quality of housing, basic amenities and asset holding capacities. Therefore HQLI is chosen as the non-spatial data for this purpose.

Different spatial metrics such as, number of urban patches (NP), edge density (ED) of urban area and aggregation index (AI) were calculated for each settlement. This analysis is helpful to identify the character of a settlement as spatial metrics can be effective to look into inequality perspectives.

Table 7 shows the characteristics of each spatial metrics that are used in this study. It is observed that, number of patches holds a negative relation with built up growth, which shows that increase number of urban patches indicates the fragmented growth whereas low number of urban patches indicates aggregated growth. Except this, all other metrics hold positive relation with built up growth.

Figure 9 shows the ideal relationships that are expected to come through the analysis. Figure 9a shows the ideal relationship between data when both the data are positively related with builtup growth and Fig. 9b shows the relationship between data where any

one of the data holds a negative relationship with builtup growth. In this case, number of patches has a negative relationship with builtup growth where increase in number of patches signifies fragmented growth of a region. By this kind of relationship, this study tries to identify the cases that mismatch the ideal relationship.

The actual presentation of this analysis has been shown in Fig. 10. It is observed from the analysis that, most of the cases follow a linear relationship between spatial and non-spatial data with some exceptions. There are some cases which are showing non-collinearity between them. As increase number of patches, low aggregation index or low edge density represent fragmented growth of a region, which indicates the formation of sprawl, it can be said that, the cases which are showing the combination of fragmented urban growth and high quality of living can be identified as the future urban growth areas.

Figure 10a shows a situation where there are some cases which are having high number of patches are also having high value for HQLI. As increase number of patches represents fragmented growth of a region, which indicates the formation of sprawl, it was expected that the cases with high number of patches will show a low value for HQLI. But the result is different for some cases.

The relation with other spatial metrics and HQLI also represents more or less the same condition for some cases. It clearly brings out the fact that, there are some cases where in spite of low builtup density, the areas are having high value for HQLI.

Conclusion

EBM and OMM are the main urban local bodies of Malda district towards which the urbanization and population growth is attracted for the easy accessibility of fast and superior civic amenities. The above analysis reflects the fact that, the growth of small urban and rural bodies surrounding the municipalities poses a wide variation in different parameters of quality of living. The release of the micro level data of house listing and housing data in census 2011 enables us to visualize the existing inequality and heterogeneity at micro level. Using the methodologies, the results show how English Bazar UA and its peripheral settlements can be characterized as consisting of five

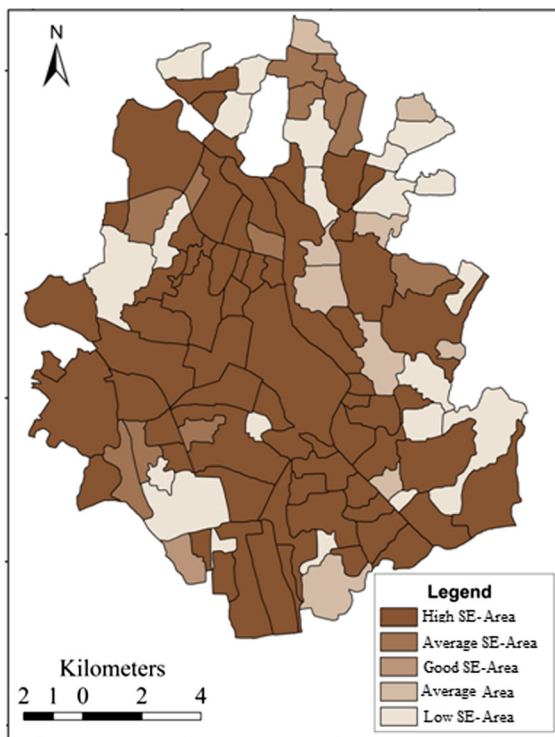


Fig. 8 Sub-group typologies map

Table 7 Description of spatial metrics

Spatial metrics	High value	Low value
Number of patches (NP)	Fragmented growth	Aggregated growth
Edge density (ED)	High builtup density	Low builtup density
Aggregation index (AI)	Aggregated growth	Fragmented growth

Fig. 9 Ideal relationship between spatial and non-spatial data, **a** positive relation with builtup, **b** negative relation with builtup

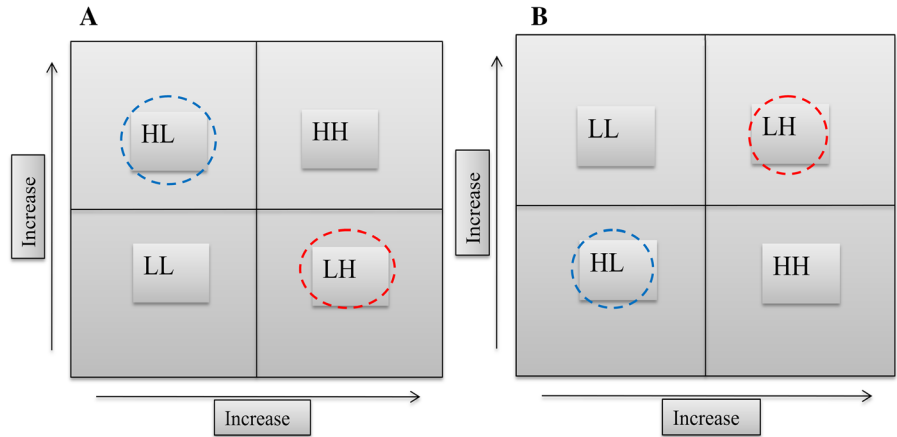
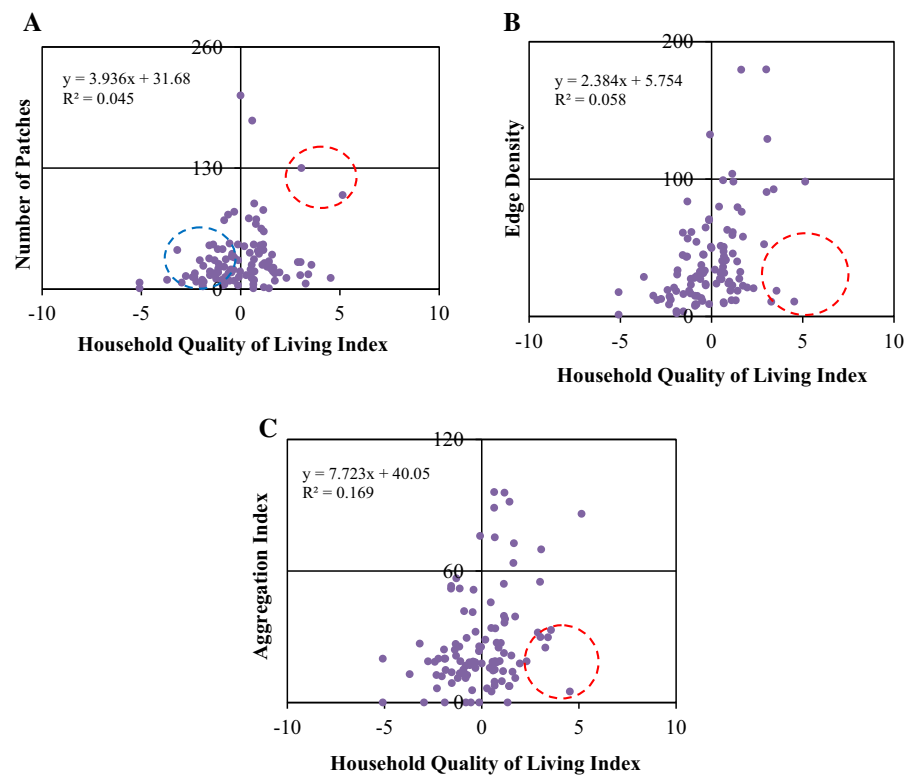


Fig. 10 Relationship between spatial metrics and housing quality of living index



sub-city typologies. Compared to the peripheral settlements, most of the settlements near the core city remain in good condition of sub-city typology. In

contrast to that, the settlements with low socio-economic status are seen to be very strongly clustered in three directions—to the north-east, south-east and

south-west. Though there are several limitations to be kept in mind while interpreting the cluster analysis results, the application of hierarchical and non-hierarchical cluster analysis in this study is useful to demonstrate a method for deriving empirical sub-city typologies for small and medium Indian cities (Balakrishnan and Anand 2015). The sub-city typologies explained in this paper provide a means of characterizing English Bazar city and understanding intra-urban variation across multiple dimensions. The correlation between different spatial and non-spatial indices also shows the linear relationship between the indices and the tendency to form co-linearity between them. Such an understanding could be helpful for city-scale vulnerability analysis. It could be also useful for the planner to adopt a suitable planning process for the settlements that are having rapid dynamic change in landuse pattern as well as social and economic aspects.

This study could be more enhanced if better resolution data is available for successive census years for Indian cities at ward level, the changes of the sub-city typologies that are identified here could also be examined for different temporal periods to understand the temporal stability.

References

- Allen, A. (2003). Environmental planning and management of the peri-urban interface. *Environment and Urbanization*, 15(1), 135–147.
- Balakrishnan, K., & Anand, S. (2015). Sub-cities of Bengaluru: Urban heterogeneity through empirical typologies. *Economic and Political Weekly*, 50(22), 63–73.
- Bhagat, R. B. (2011). Emerging pattern of urbanisation in India. *Economic and Political Weekly*, 46(34), 1–11.
- Bhan, G., & Jana, A. (2013). Of slums or poverty notes of caution from census 2011. *Economic and Political Weekly*, XLVIII(18), 13–16.
- Bhan, G., & Jana, A. (2015). Reading spatial inequality in Urban India. *Economic and Political Weekly*, 50(22), 5–10.
- Chatterjee, N., Das, A., Chatterjee, S., & Khan, A. (2015). Spatial modeling of urban sprawl around greater Bhubaneswar City, India. *Modeling Earth Systems and Environment*, 2(14), 1–21.
- Das, B., & Mistri, A. (2013). Household quality of living in Indian States: Analysis of 2011 census. *Environment and Urbanization Asia*, 4(1), 151–171.
- Dupont, V. (2004). Socio-spatial differentiation and residential segregation in Delhi: A question of scale? *Geoforum*, 35(2), 157–175.
- Forman, R. T. T. (2008). *Urban regions*. Cambridge: Cambridge University Press.
- Hoffman, P., Strobl, J., Blaschke, T., & Kux, H. (2008). Detecting informal settlements from QuickBird data in Rio de Janeiro using an object-based approach. In T. Blaschke, S. Lang, & G. Hay (Eds.), *Object-based image analysis: Spatial concepts for knowledge-driven remote sensing applications* (pp. 531–553). New York: Springer.
- Jamil, G. (2014). The capitalist logic of spatial segregation: A study of Muslims in Delhi. *Economic and Political Weekly*, 49(3), 52–58.
- Keil, R. (2017). Extended urbanization, “disjunct fragments” and global suburbanisms. *Environment and Planning D: Society and Space*, 36(3), 1–18.
- McGarigal, K., & Marks, B. J. (1995). FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure. USDA Forest Service General Technical Report PNW-351. www.umass.edu/landeco/research/fragstats/fragstats.html.
- Nkeki, F. N. (2016). Spatio-temporal analysis of land use transition and urban growth characterization in Benin metropolitan region, Nigeria. *Remote Sensing Applications: Society and Environment*. <https://doi.org/10.1016/j.rsase.2016.08.002>.
- Ramachandra T. V., & Bharath, H. A. (2013a). Spatio temporal patterns of urban growth in Bellary, Tier II City of Karnataka State, India. *International Journal of Emerging Technologies in Computational and Applied Sciences*, 3(2), 201–212.
- Ramachandra, T. V., & Bharath, H. A. (2013b). Understanding urban sprawl dynamics of Gulbarga-Tier II city in Karnataka through spatio-temporal data and spatial metrics. *International Journal of Geomatics and Geosciences*, 3(3), 1–12.
- Ramachandra, T. V., Bharath, H. A., & Sowmyashree, M. V. (2014). Urban structure in Kolkata: Metrics and modelling through geo-informatics. *Applied Geomatics*, 6(4), 229–244.
- Ramachandra, T. V., Setturu, B., & Aithal, B. H. (2012). Peri-urban to urban landscape patterns elucidation through spatial metrics. *International Journal of Engineering Research and Development*, 2(12), 58–81.
- Roy, A. (2011). Urbanisms, working practices and the theory of planning. *Planning Theory*, 10(1), 6–15.
- Roy, A. (2016). What is urban about critical urban theory? *Urban Geography*, 37(6), 1–14.
- Samanta, G. (2012). In between rural and urban: Challenges for governance of non-recognized urban territories. In N. Jana, et al. (Eds.), *West Bengal: Geo-spatial issues* (pp. 44–57). Burdwan: The University of Burdwan
- Samanta, G. (2017). New urban territories in West Bengal: Transition, transformation and governance. In E. Denis & M.-H. Zérah (Eds.), *Subaltern urbanisation in India: An introduction to the dynamics of ordinary towns, exploring urban change in South Asia* (pp. 421–441). India: Springer.
- Shaw, A. (2005). Peri-urban interface of Indian cities growth, governance and local initiatives. *Economic and Political Weekly*. <https://doi.org/10.2307/4416042>.
- Shaw, R., & Das, A. (2017). Identifying peri-urban growth in small and medium towns using GIS and remote sensing technique: A case study of English Bazar Urban

- Agglomeration, West Bengal, India. *The Egyptian Journal of Remote Sensing and Space Sciences*. <https://doi.org/10.1016/j.ejrs.2017.01.002>.
- Sudhira, H. S., Ramachandra, T. V., & Jagadish, K. S. (2004). Urban sprawl: metrics, dynamics and modelling using GIS. *International Journal of Applied Earth Observation and Geoinformation*, 5(1), 29–39.
- Zenina, N., & Borisov, A. (2013). Clustering algorithm for travel distance analysis. *Information Technology and Management Science*, 16, 85–88.