Modeling spatial variation of explanatory factors of urban expansion of Kolkata: a geographically weighted regression approach

Biswajit Mondal¹ · Dipendra Nath Das¹ · Gour Dolui²

Received: 20 July 2015 / Accepted: 14 August 2015 / Published online: 15 September 2015
© Springer International Publishing Switzerland 2015

Abstract Modern cities face a large number of issues related to boundary, land alteration and environmental degradation. Extensive urban expansion is an especially serious problem faced by most of the cities in developing countries. Urban management has extensive engagement with rapid urban expansion and their factors to achieve a sustainable form of development. A more inclusive understanding of the urban driver variables and their cross comparison greatly helps to investigate this rapid urban expansion. The major theme of this study will be to inquire into this factors of urban expansion. Both global and local regressions are performed to determine the spatial variability of the driving forces of urban expansion. The results indicate that geographically weighted regression model outperformed global regression. Urban expansion in the city of Kolkata can be observed mainly in two directions; northeast and south-east. The possible reasons include that of economic appraising, lack of environmental distress and good spatial settings along with investor interest, which together determine the biased bidirectional axial expansion. This study comes to the conclusion that local

variables are significant in driving the bi-directional urban expansion in the context of a biased urban management.

Keywords Urban expansion · Driving factor · Geographically weighted regression · Geographical variability test · Kolkata

Introduction

It has been noted by Chen et al. (2014) that the growth of cities can be seen as a result of the combined effects of agglomeration economies, migration, natural increase and policy initiatives. However, often the unplanned nature of such growth leads to problems. The possible detrimental effects of such expansion in cities and large metropolitan areas are not confined to its boundary. Besides directly impacting large segments of ancillary town areas and adjacent dependent rural spaces, they can have impacts on a national and international level. Polèse and Denis-Jacob (2010) stated that although, some mega cities in the developed countries have been facing stability in terms of population growth, mostly the situation is contrary in developing countries, for example, in the case of Mumbai, India (Shafizadeh-Moghadam and Helbich 2015). Hence, it is important to analyse the forces changing land use to develop a deeper understanding of the complex urban dynamics driving spatial growth in urban centres.

The process of urban expansion can be a complicated process involving multiple factors. The relationship between urban expansion and the environment, socio-economic and population factors has found focus in urban scientific research (Luo and Wei 2009; Li et al. 2013; Lu et al. 2013). These environmental and socio-economic factors or driving

✉ Biswajit Mondal
biswaj92_ssf@jnu.ac.in

Dipendra Nath Das
dipendra02@gmail.com; dipendas@mail.jnu.ac.in

Gour Dolui
gourdolui@gmail.com

¹ Centre for study of Regional Development, School of Social Science, Jawaharlal Nehru University, New Delhi, India

² Department of Geography and Environment Management, Vidyasagar University, Midnapore, West Bengal 721102, India

factors are always associated with urban expansion and environmental degradation (Du et al. 2014). Therefore, quantification of all these factors along with spatial modelling is the key to evaluating planning decisions. If we specifically focus on urban planning agendas, then housing has constantly been an essential part that has promoted spatial growth. Housing development has focused aspects of environmental protection, sustainable residential development, resource capitalization, and coherent regional infrastructure (Anthopoulos and Vakali 2012) which has fueled urban expansion beyond the existing spatial land allocation by local governments. By analysing factors like population and household growth, environmental quality, land price, per capita income, etc. one can find out the central influence of human-driven forces on urban change (e.g., Parker et al. 2003; Long et al. 2007; Luo and Wei 2009; Li et al. 2013; Lu et al. 2013; Chen et al. 2014).

Most studies of the driving forces of urban land transformation fall under one of the two following approaches—firstly, utilising local parameters that affect planning strategies and secondly, analysis relies on global parameters like literacy, employment status, and population growth etc. (Perz et al. 2005). These approaches enable us to identify the causes of LULC change and correlate them in terms of relationships between the considered set of independent variables and the observed phenomenon. Geographically weighted regression (GWR) and the global regression method are two statistical techniques which can be used to examine the spatial variability across the region (Fotheringham et al. 1998). However, the use of the global regression method over wide areas, such as Kolkata, could be inappropriate due to the application of stationary coefficient over the whole study area. This could mask local interactions that might have greater strength in explaining the phenomenon under study. The influence of local factors is more significant when there is a very high spatial diversity of bio-physical and social conditions. However, for global factors no such variability exists. Multiple regression analysis such as ordinary least squares (OLS) model is based on the assumption of independence of observations and results. As such, it fails to capture the spatial dependence of data, which is possible with the application of spatial data analyses. Thus, to overcome this limitation of the OLS method Fotheringham et al. (1998) developed a local regression technique, named as GWR.

Highly regarded and cited past studies like agent analysis, ‘travel cost demand’, multicriteria system emphasise the need to look broadly at socio-economic factors while keeping in mind individual factors that influence urban growth. Lu et al. (2014) develop a GWR model with a non-Euclidean distance metric to find out the relationship between London house price data set to study the housing market. By exploring the linkages between independent

and dependent factors in both global and local models, a general view of the whole area can arrive. In contrast, the local model would show spatial variability. Because of this advantage, the GWR is widely used in the area of urban planning (Lu et al. 2013; Huang et al. 2010; Kaligari and Ziberna 2014), ecological and environmental applications (Clement et al. 2009) and sometimes even in climate research (Brown et al. 2012). Although, GWR is not a problem-free statistical method, with major concerns focused on issues such as kernel and bandwidth selection, some researchers have applied this method even in diverse disciplines (e.g., Bagheri et al. 2009; Bitter et al. 2007; Du et al. 2014; Haitao et al. 2014; Lee and Schuett 2014; Megler et al. 2014). GWR is applied specifically to the urban expansion by Du et al. (2014) and Shafizadeh-Moghadam and Helbich (2015).

To analyse complex urban expansion process, this study obtained regression based findings of urban change by integrating remote sensing and GIS techniques, which were designed for simulated analytical models for widespread relevance (Mesev 1997; Batty et al. 1999) in the study of urban growth. The combined effects of the increasing availability of spatial data, advances in analytical methods, and the development of geographical information systems have made such studies possible. In fact, GIS has been instrumental in providing catalysts for the development of appropriate quantitative methods for spatial analysis (Thapa and Estoque 2012).

To fulfill the research gap on the association among local forces and bi-directional urban expansion in Kolkata, this study has accentuated two major objectives. The first is to question whether local models are a significantly better explanation for the urban expansion process in comparison to global models. In affirmation of this, we shall then address the spatial variability of local factors across space. The second is to explore the relationship between independent and dependent factors regarding the spatial non-stationary.

Study area

Kolkata faced short-term high population growth in the peripheral regions during 1991–2001 (Bhagat 2004) but this growth gradually declined during 2001–2011. Based on this empirically verified change in growth rate, Sivaramakrishnan et al. (2005) has categorized Kolkata as a ‘decline core decline periphery’ type city. This study considers a metropolitan core, and its immediate periphery contained by the Kolkata urban agglomeration consisting of both banks of the Hooghly River (Fig. 1). Two terms—core and periphery are frequently used in this study. KMC and Howrah (MC) have been considered as core city (according to the 1991 administrative boundary) and

Fig. 1 Location map of the study area

contiguous municipalities, towns and villages have been considered as peripheral (research boundary extent). The study area is located in the Eastern region of India with a latitude range of $22^{\circ}26'13''\text{N}$ to $22^{\circ}39'15''\text{N}$ and a longitude range of $88^{\circ}13'15''\text{E}$ to $88^{\circ}28'50''\text{E}$, with a total area of 644.72 km^2 . The study area consists of 2 municipal corporations, 7 municipalities, 48 towns and 87 villages according to the 2011 census. The total urban population share is 97.10 %, and rural population is 2.9 %. According to 2011 census, the study area contains a total population of 8.70 million.

Materials and methods

Materials

In this study, a series of Landsat TM/ETM⁺ images (1990, 2001, and 2011) were collected (Table 1), since its fine temporal and spatial resolution is ideal for urban growth analysis (Shafizadeh Moghadam and Helbich 2013) and converted to a projection system UTM-WGS 1984 Zone 45°N and a subset of the study area which contains 898 rows and 814 columns was used to generate land use land cover maps. A set of urban features (municipal and block boundary, primary (main) and secondary (minor) transport

networks, locational attributes, river and protected areas) were digitized from the census 2011 administrative atlas with the help of Google imagery. For the modelling purpose some other socio economic variables (population density, housing density, percentage of main workforce and employment rate) were calculated from the Census of India 2011 and housing price, land price, development cost were collected from the West Bengal Registration and field survey. Environmental variables (impervious surface) were developed by using the sup-pixel classifier and ISAT tool. Percentage of green space was derived by calculating zonal statistics by using classified images, and the temperature was derived from band 6 of ETM + 2011 Landsat imagery.

Methods

Image classification, accuracy assessment and change detection

LANDSAT ETM⁺/TM images were collected for LULC classification. Before implementing the maximum likelihood (ML) algorithm to classify the images, good quality training samples were collected and grouped into five major classes such as drainage, wetland and water bodies, built-up, cropland and open land, green space and plantation. After that, time series satellite images were classified using ML algorithm. Finally, all misclassified pixels were merged into more suitable classes. An accuracy assessment was performed to check the accuracy level of classified images using Kappa coefficient measure.

Driving forces and spatial analysis

Urban growth simulation is done in order to show attentively impact of policy measures on future scenarios. The selection of driving force and spatial measures are significant to increase the efficiency of the urban growth model. With respect to the study area three major forces which have the capability to drive the urban land use pattern are considered in this research. These are linear infrastructure, market force, and residential force. Before using GWR, Crammer's rule was applied to ensure the correlation of these forces to drive urban change.

Variable for linear infrastructure To fulfill the objective, influence of local transport infrastructures (like roads,

Table 1 Description of the collected geospatial data

Data type	Source	Year	Dataset	Resolution (m)
Satellite image	USGS	November, 1990	LANDSAT-TM	30
		November, 2001	LANDSAT-ETM	30
		November, 2011	LANDSAT-TM	30

railways, metrorails etc.) on urban expansion, spatial autocorrelation measure was implemented to define the spatial effects of transportation on urban expansion. A regression was also conducted to show the mutual effect of population, household and road upon spatial expansion. To find out the reciprocal relationship of transportation and spatial expansion, a spatial clustering measure was carried on. Spatial autocorrelation was widely used to define the correlation between the local and the average neighbour. The cluster maps (LISA maps) depict the significant locations and their associations. Random permutations have been used to assess the sensitivity of the findings (Anselin et al. 2006). Moran's Index is used to indicate the strength of the spatial similarity or dissimilarity of neighbouring districts. A positive Moran's I indicates the presence of degree of spatial autocorrelation (Aljoufie et al. 2013). The spatial arrangements of districts in the study area supported the rook weight matrix. Hence, a first-order connectivity weight matrix was constructed. Finally, on conducting the significance test according to 999 permutation combination procedure against the null hypothesis, there is no spatial autocorrelation. The location of significant positive or negative autocorrelation were identified and considered for LISA clusterity analysis. Local Moran's (LISA map) gives 5 output categories by producing 5 zones along with their significance. Where, Zone 1 = not significant, and Zone 2–Zone 5 present the positive and negative autocorrelation. H–H indicates a positive relation, meaning high local and higher neighbourhood average value. Similarly, L–L denotes the low local and low average neighbourhood value. H–L and L–H depict the negative autocorrelation among the local value and neighbour value.

Variables for housing environment The housing environment is helpful to find out the suitable living environment at a micro level within the city which is more preferable for urban housing, and can also possibly be a useful tool for making a decision to lead healthy lives. To analyse the quality of residence (QR), factor analysis has been performed. For QR assessment some indicators were used such as population density, housing density, impervious surface, percentage of green space, temperature, employment rate, percentage of main workers and median housing value. All indicators have been normalized between the ranges of 0–1 to reduce the scale bias (e.g., Zhou et al. 2005; Floridi et al. 2011).

Variables for investment potential Market force is another major component which has a considerable influence on urban change just like the residential force and which may be considered as an investment potential surface. The investment potential surface indicates the areas of the developer's choice that is where the developers are

likely to invest to get the highest returns by generating maximum profit. Three variables Land price, development cost and apartment price were used to compute investment potential. This study assumed that developers are the profit seekers and build property to achieve maximum profit. For this purpose in this model, another assumption has been taken that there is no other constraint on property development in the overall areas.

Geographical weighted regression (GWR) analysis

GWR is a type of regression model, it can be applied on geographically varying parameters. Due to its ability to deal with mixed fixed geographically changing variables, it can be said to have a semi-parametric formation. This study adopts a semi-parametric formation by applying a generalized linear modelling and incorporating the Gaussian regression, which allows the globally fixed and locally varying terms of explanatory variables concurrently. A semi-parametric GWR model can be expressed as (e.g., Nakaya 2009),

$$Y_a = \sum_{\kappa} \beta_{\kappa}(m_a, n_a) \chi_{\kappa,a} + \sum_{\rho} \gamma_{\rho} \vartheta_{\rho,a} + \varepsilon_a$$

where, γ_{ρ} is a fixed coefficient with $\vartheta_{\rho,a}$ independent variable $\beta_{\kappa}(m_a, n_a)$ are varying conditional and locational (m_a, n_a) is a coordinate of a^{th} location; Y_a is dependent variable; $\chi_{\kappa,a}$ is independent variable and ε_a is Gaussian error.

Non-stationary coefficient arrangements can be recognised easily through GWR although the results it generates are close approximations, rather than being exact determinants. The non-stationarity of the parameter estimates can be assessed by comparing twice the standard errors of the global ordinary least squares (OLS) estimates with the inter-quartile of local estimates from GWR, with larger values of the latter indicating significant spatial non-stationarity (Fotheringham et al. 2015).

The Spatial-Kernel function was used for weighting, where the weights were placed through the adaptive bi-square Euclidian distance method. The adaptive bi-square Kernel was kept constant. To identify the appropriate bandwidth, the Akaike-information criterion (AIC) was minimized (Fotheringham et al. 1998). Kernel type can be express as (e.g., Nakaya 2009);

$$\omega_{ab} = \begin{cases} \left(1 - d_{ab}^2 / \varphi_{a(k)}\right)^2 & d_{ab} < \varphi_{a(k)} \\ 0 & d_{ab} > \varphi_{a(k)} \end{cases}$$

where, a is index of the regression point; b is the index of location; ω_{ab} is the observations weight value of the location b for estimating the a coefficient at location. d_{ab} is the Euclidean distance between a and b ; φ is a fixed bandwidth size defined by a distance metric measure. $\varphi_{a(k)}$

is an adaptive bandwidth size defined as the k^{th} nearest neighbour distance. Geographical variability is also transformed to compare the varying coefficient. k^{th} varying coefficient was computed through GWR, and compared with a fixed k^{th} co-efficient of GWR model for the purpose of AIC. F test was used to find out the spatial variability.

Results and discussions

Urban spatial growth, population growth and household growth

Historically, Kolkata has shown a higher rate of built-up growth but this growth rate is not evenly distributed throughout Kolkata. Bhatta (2009) previously mentioned that the increasing trend of the built-up growth is concentrated towards the periphery. Also, Fig. 3 shows a decline in built-up growth during the period of 2001–2011 in comparison to 1990–2011 (e.g., Bhatta 2009). Therefore, growth is not evenly distributed over time either. This situation indicates the lack of land in the central part for residential development (Bagchi 1987; Shaw and Satish 2007; Sengupta 2006; Roy 2011; Bhatta 2009) and the availability of sufficient land in the periphery including suitable environment. During 2001–2011, Sonarpur and Rajarhat area have faced significant changes both in built-up areas and population. It is to be noted that during 1991–2001 the peripheral area had undergone changes without having any significant variation in the demography (Fig. 2). So, there are some other factors along with population that play a significant role in the spatial expansion of built-up areas.

Fig. 2 Classified image showing built-up area (1990, 2001 and 2011)

During 1991–2001 peripheral areas faced changes without having any significant demographic variation but during 2001–2011 two dominant areas of significant built-up and population change have occurred in the Sonarpur and the Rajarhat areas. It can be inferred that population has not necessarily influenced the built-up growth because the supply of land was very high, and demand was not necessarily high. So, there are some other factors along with population which play a major role in spatial expansion of built-up areas, as is evident from Fig. 3. Sonarpur has constantly been facing an exceptional high level of built-up growth without having significant population influx in 1991, with significant population increase happening only during 2001–2011. In contrast, Rajarhat has faced some level of vertical expansion (high rise buildings) because Rajarhat has high population growth, but stagnant built-up growth. Incorporating residential and investment potential data with built-up data can demonstrate three features. Firstly, it reveals the developers' interest to invest in residential development by analysing the past trend and the people's choice in property as investment yielding high returns. Secondly, allotted built-up area and its attractive prices attract not only the youth, but also middle-class migrants, who look for affordable housing in multi-storied buildings. Finally, large transportation and infrastructural projects in peripheral areas might also attract migrants and become responsible for the built-up growth.

Driving force analysis

During 1990–2011 urban expansion indicates a declining trend but growth is nonetheless present in the periphery. Townships planning with fast transportation, real-estate

Fig. 3 Dispersed association of exponential growth rate of population and built-up areas (left 1990–2001, right 2001–2011)

investments and new business hubs are induce growth in the periphery (Pal 2006; Roy 2011). Some inflexible government activities make it challenging to direct or manage ‘urban shrinkage’ (e.g., Haase et al. 2012). From different source, it has been identified that some group of forces like accessibility factors, an environmental factor of quality residence and market force of investment potentiality are the reasons behind urban expansion and shrinkage.

Linear infrastructure and urban growth

Linear infrastructure has always influenced new residential locational patterns; so, it may be said that during 1990–2001, Kolkata itself had a competitively high built-up growth (4.35 %) in its frontiers due to better transportation and proximity to central locations. Whereas, total road length and built-up area was relatively lower than KMC (Fig. 4). Thus, Fig. 4 indirectly indicate linear built-up expansion in periphery. A Moran’s Index analysis was performed using the log of urban growth and transportation variables. Global Moran’s values for the variables are 0.47 for population density, 0.72 for major road density, 0.45 for household density and 0.63 for built-up growth respectively. The Moran’s Index is positive and statistically significant ($p < 0.001$) for spatial expansion, population density, household density and major road density. This result indicates that the nearby districts tend to have similar attributes. So, it is also a fact that higher accessibility makes the land value higher, and density decreases due to the unfavourable land value for the middle class citizen.

LISA results identify the local spatial clustering of spatial expansion and road density variables at the block/ward level (Fig. 5). Blocks with a considerable LISA, was classified by the type of spatial correlation. The high–high and low–low

Fig. 4 Imbalance of built-up area as a function of road length

locations recommend clustering of related values of one variable, whereas the high–low and low–high locations designate spatial outliers of the similar variable (Anselin et al. 2006). By comparing LISA outputs, it accounts for the significant spatial clustering of spatial expansion (2001–2011) and the road density, population density and household density of 2011. The road density, population growth and household density had a resilient association with built-up growth ($R^2 = 0.44$) in 2001–2011. It might be possible that the incorporation of residential force increased the proficiency of the model because all these forces combine to drive the urban expansion.

Housing environment and urban growth

In this study, three factors have been extracted without having the Eigen value greater than 1.0 for the third factor. The variance criteria was utilised for factor extraction, where three factors have been combined to explain more than 75 % variance. These entire set of indicators are somehow interrelated and associated with each other, which have been combined into a connotation within

Fig. 5 LISA cluster map of urban spatial expansion variable

housing environment (e.g., Weng 2010). These three factors, namely, spacing quality (Population density and household density), environmental quality (percentage of green area, impervious surface and temperature) and economic quality (apartment price, employment rate and percentage of main worker) can be seen to play a residential force (Fig. 6) for urban expansion. Factor analysis concurs that the urban expansion correlates with the areas that are economically viable, environmentally attractive and less compact (Mondal 2014). On analysing the LISA cluster map (Fig. 5), and accounting for the residential force, one can find the clear association between spatial expansion and present housing environment. So, the density of population and residential quality has strong negative association whereas built-up growth and residential quality have a strong association. Therefore, quality of residence can be a healthier criterion for the spatial expansion process.

Investment potentiality and urban growth

In case of Kolkata, the areas of higher profitability are mainly concentrated in central Kolkata and in the Rajarhat planning area. Although, high accessibility, quality of the environment, commercial and business-friendly market

makes the land profitable for developers, it nonetheless remains within the financial reach of lower middle-class or middle-class families (Mondal 2014). There is an insignificant relation, possibly because of sample inaccuracy, between built-up growth and developer profitability because developers are interested in investing only in large townships and in specific potential pockets of investment.

Local modelling result

GWR model is primarily based on the weighted least square function considering an adaptive Kernal function to shows the local relationship. GWR model is more specific and reliable than OLS in terms of information (Lee and Schuett 2014). The GWR (Gaussian) model with combining sample data can be expressed as;

$$\begin{aligned} \text{Builtupgrowth}_a = & \beta_0(m_a, n_a) + \beta_1(m_a, n_a)\text{Population growth} \\ & + \beta_2(m_a, n_a)\text{Investment Profit} \\ & + \beta_3(m_a, n_a)\text{Spacing Quality} \\ & + \beta_4(m_a, n_a)\text{Environment Quality} \\ & + \beta_5(m_a, n_a)\text{Economic Quality} \\ & + \gamma_1\text{Population density} + \gamma_2\text{Road density} \\ & + \gamma_3\text{Median Land Price} + \varepsilon_i \end{aligned}$$

Fig. 6 Driving factors map used for the evaluation of urban expansion (POPD = population density, HHD = household density, TEMP = temperature, IMS = impervious surface, PGS = percentage of green space, ER = employment rate, AP = apartment price, PMW = percentage of main worker)

Adjusted R^2 values showed significant improvement in GWR model (adjusted $R^2 = 0.81$). AIC values using the GWR model (AIC = -640.51) is much lesser than that in the global regression model (AIC = -591.85 , adjusted $R^2 = 0.69$) (Table 2). Global Moran I test showing random distribution of the residuals of GWR result indicates absence of spatial autocorrelation in the residual values.

Geographical variability test (GVT) was done to find out the actual global and local term (Table 3). GVT shows that land price and spacing quality have a high positive DIFF of criteria that indicates this variable is better assumed as global. As compared to global models, GWR model considers five local terms due to their variability and three global terms that is population density, road density and median land price due to their non-variability. Out of the three fixed factors, two variables, that is population density, road density and median land price showed negative influence on urban expansion across the study area.

Table 2 Comparison of local and global regressions result

Geographically weighted regression result		2001–2011
AICc		-640.513
R square		0.850
Adjusted R		0.807
Moran I		0.102
Global regression result		2001–2011
AICc		-591.854
R square		0.712
Adjusted R		0.692
Moran I		0.143

Global regression model indicates that all three variables have some influence on the study area as the Table 3 shows. This result also revealed that road density and population density are more or less invariable across space. High

Table 3 Geographical variability tests of local coefficients

2001–2011	F	DOF	For F test	DIFF of criterion
Intercept	126.030	4.799	115.685	−260.060
Population growth	61.674	5.201	115.685	−183.063
Investment profit	0.603	3.308	115.685	8.683
Spacing quality	−9.780	3.635	115.685	67.788
Environment quality	4.999	4.452	115.685	−11.552
Economic quality	10.479	3.240	115.685	−27.815

Positive value of DIFF-Criterion (AICc, AIC, BIC/MDL or CV) suggests no spatial variability in terms of model selection criteria. F test: in case of no spatial variability, the F statistics follows the F distribution of DOF for F test

population density associated with congestion and low growth and negative influence of road density indicate linear expansion or growth along the linear infrastructure. Median land price is negatively influenced due to investor interest and vertical expansion not induces spatial expansion.

Estimation of local coefficient

GWR divided the local co-efficients for each location and each variables for 2001–2011 (Fig. 7). The main three criteria that this coefficient sufficiently explains are-

1. Variation of coefficient for the five independent variables to find out the influence of the variable.
2. Positive and negative influence of the independent variable for the particular location can be driving factor by the sign of the coefficient.
3. Finally, the significance level was derived to find out the influence of the variable on urban expansion.

Population growth and urban expansion The estimated value of population growth co-efficient for the global model was −0.048 and the standard error of 0.011. Local co-efficient show that the influence of the variable varied across the study area with very short positive influences in the southern peripheries and strong negative influence in central and northern zone. This finding reveals that population growth in southern periphery is positively affect the urban expansion process.

Investor profit and urban expansion Land market positively influences the urban expansion process as the global model showed a co-efficient 0.0001 with standard error 0.002. Local model shows that the variable substantially varied across the study area with co-efficient range of −0.007 to 0.015. High positive influences occurred further in the periphery where the investor profit is medium, but the north-eastern portion (Rajarhat), where the investor profit is very high was negatively associated with urban expansion. These results revealed that highly profitable areas do not have urban expansion.

Spacing quality and urban expansion Local coefficient of variable spacing quality was 0.033 with a standard error of 0.006. Local coefficient reveals a high spatial variability ranging from −0.026 to 0.005. Strong positive influences occurred in south eastern peripheries, and the northern area shows negative influence on urban expansion process due to high congestion and crowding (highly dense area). So, the result revealed that the quality of spacing affects the urban expansion.

Environmental quality and urban expansion Coefficient value of the environmental quality variable was 0.01 with a standard error of 0.001 as the global model shows. The local model, however, shows considerable variability of its influences across the study area. Local coefficient of the environment quality was in the range of −0.01 to 0.008. Environmental quality positively influenced the north eastern zone, whereas moderately influenced the south-eastern area. Strong negative influence in the old core zone was observed in this case.

Economic quality and urban expansion Economic quality positively influences urban expansion as revealed by the global model. Coefficient value of the economic quality variable was 0.002, with a standard error of 0.004. The local model revealed that the influence of the variable considerably varied across space, with a range of −0.04 to 0.02. The periphery of the study area except south western zone shows positive influence. Therefore, it can be concluded that economic quality positively influences urban expansion in the peripheral area of study, except in the south western zone.

Local R^2 value

The GWR results represent the noted variability of the relationship between all explanatory variable and predicted variables.

Local and global R^2 value range from 0 to 1, 1 denotes that the variable is explained perfectly by the model. In terms of the goodness-of-fit statistics, Global model

Fig. 7 Forces and their geographically varying local coefficient

($R^2 = 0.71$) explains that the relationship reliability is lesser than that of the local model ($R^2 = 0.85$). Figure 8 illustrates, that the local R^2 surface shows the maximum relationship area. This means that in the area, the dependent variable (urban expansion) is more or less perfectly explained by the independent variables. AIC score was minimized to generate an adaptive bandwidth. The GWR model is based on a 152 sample data point. Socio-economic driver variables considerably explain urban expansion during 2001–2011. Although, the model shows the significant results there is some possibility of uncertainty

due to the lack of temporal socio-economic driver variables. The results of GWR model are summarized in Table 4. Figure 8 depicts the relationship between observation and all driver variables. It is noticeable that the northern portion of the study area faces quite a low expansion and the negative impact of driver variables. The strongest adverse impact area is the central high-density residential area (old Kolkata).

So, the Sonarpur and Rajarhat localities are the present growth centre. Figure 8 shows the positive influence areas of the driver variables. Areas situated distant from linear

Fig. 8 Interpolated local R^2 surface

infrastructure (major roads) alter (non-built-up to built-up) more quickly than areas which are more accessible as they are of a higher price. Another emerging fact is that township developments consider the requirements of higher income groups (Sengupta 2007). Thus all townships are facilitated by efficient transportation systems (for the travel of low-end labour), institutions for education (for the children of the high-income groups), and recreation and shopping malls. These facilities allow the area to develop as they imbue a high potential for growth (Mondal 2014). Nonetheless, the adjacent areas, which remain underdeveloped, remain as catchment for the habitation of those providing low-end labour in the townships. For example, Saltlake city produced the efficiency of the nearest blocks belongs to N 24 Pargana, East Kolkata township and Baishnabghata-Patuli township which increase the effectiveness of Sonarpur area and transport link between Sonarpur to Saltlake (up to Airport and efficient linked connection with Newtown) by EM Bypass which have a major role in sub-urban expansion in 2001–2011.

Middle-class people prefer the peripheral areas for the quality of environment along with transport for diurnal migration, because township areas are made unpreferable, as high price makes them unaffordable. Therefore, it is

preferred to find housing land nearest to the township, which allows one to get some infrastructural benefit despite lower cost of land. Overall, New Town displays certain potentiality to entice the population and built-up growth, along with excellent possibilities of economic growth, Foreign Direct Investments, industrial development and sufficient opportunities for employment generation shortly.

Conclusion

This study explored the association between urban expansion and its driving factors by developing a local regression model. The global model is useful for regional planning or suitable for the purpose of general large area planning. To improve the local understanding and make it more specialized, the introduction of local models was useful. On the methodological ground, this study adopts an existing approach of GWR to show the spatial heterogeneity.

To improve the relationship of urban expansion and driving factors, this study incorporates environmental, economic, and population component as local factors, including fixed term spatially. The global model, unable to compare local variation can deliver an inappropriate urban land demand. Further, GWR has the capability to explain the linkages between factors and urban expansion. Till now there is no such detailed study to evaluate the relationship between biased and unidirectional urban expansion and driving factors in Kolkata. So, this study provides a valuable contribution to organized zonal planning agendas. In the analysis, it has been found that the GWR model is significantly better than the global regression model (e.g., Huang and Leung 2002). The local variables have great significance in the urban expansion process, and these variables can be used for decision-making purposes and in improving our perception of urban expansion. It is expected that the local model can considerably improve the urban development plan and be helpful in reducing the biased planning initiative (Shafizadeh-Moghadam and Helbich 2015), which occurred in Kolkata. The GWR

Table 4 Summary statistics for varying (local) coefficients

2001–2011	Lower quartile	Median	Upper quartile	Min	Max	Range	Test for non-stationarity ^o
Intercept	−75.967	−69.706	−27.722	−90.964	35.238	126.201	*
Population growth	−924.417	−848.247	−337.536	−1106.922	428.385	1535.307	*
Investment profit	0.000	0.002	0.004	−0.007	0.021	0.028	*
Spacing quality	−0.023	−0.018	−0.011	−0.026	0.008	0.034	*
Environment quality	−0.006	−0.004	−0.001	−0.011	0.008	0.019	*
Economic quality	−0.008	−0.002	0.000	−0.046	0.035	0.081	*

Lwr lower, *Upr* upper, *Min* minimum, *Max* maximum, ^o = the significance of parameter. If the local coefficient's interquartile range is larger than two standard errors of the global coefficient, this might be an indication of parameter variation. * non-stationary parameter

model fitted exceptionally well into this study area. This model has the proficiency to grapple the complex urban expansion process.

The driving force analysis also reveals the ‘economically-good’ with infrastructural investment and the market force drive of the built-up growth in the Rajarhat area. However, another fact is that the ‘environmental-bed’ with infrastructural facility along with the absence of market force can also drive the built-up as well as the population growth. Therefore, it is clear that the infrastructural investment is one of the major determinants to urban spatial expansions. Planning for the environment, economy and local factor reflects an urgent initiative for efficient urban growth management. Precisely summed up sustainable urban management principles for the Kolkata have listed below:

- Environmental planning must contemplate the space conservation (wetlands and green spaces) and utilization of land to optimize the existing resource (ground water and open spaces), living space allocation.
- Planning for the local factor must include efficient transportation through the corridor development, pulling down the travel time and cost that can moderate disjointed development and induced economic agglomeration.
- Kolkata must have to consider comprehensive land use planning rather than just a township plan to determine the land allocation and intensity of different uses for individual land parcels.
- More emphasis on land market, population spacing, and active local political regime (e.g., Todes 2012) to protect the habitat obliteration.

However, this GWR model has some limitations because of its difficulties in reading the results, where the tool fluctuates or delivers different results depending on bandwidth. The GWR techniques are still under much debate (Anselin et al. 2006). Further, the datasets in this study not adequate to catch a potent inference. Till now no commonly accepted methods exist to find out the complex relationship between urban growth and driving factor in the city level. However, further recurrence of the urban growth analysis can help to identify the growth externalities, direction and pattern of change. The insertion of supplementary potential time series driving factors in a future study would also be valuable. Also, while we investigated the temporal dynamics and effects of the driving factors on urban expansion, we must also have to examine the spatial heterogeneity of the driving factors.

References

- Aljoufie M, Brussel M, Zuidgeest M, Van Maarseveen M (2013) Urban growth and transport infrastructure interaction in Jeddah between 1980 and 2007. *Int J Appl Earth Obs Geoinf* 21:493–505
- Anselin L, Syabri I, Kho Y (2006) GeoDa: an introduction to spatial data analysis. *Geogr Anal* 38(1):5–22
- Anthopoulos LG, Vakali A (2012) Urban planning and smart cities: interrelations and reciprocities urban planning: principles and dimensions. *The future internet: lecture notes in computer science*, pp 178–189
- Bagchi A (1987) Planning for metropolitan development: Calcutta's basic development plan, 1966–1986: a post-mortem. *Econ Polit Weekly* 22(14):597–601. Retrieved from <http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/pdfplus/4376875.pdf?acceptTC=true>
- Bagheri N, Holt A, Benwell GL (2009) Using geographically weighted regression to validate approaches for modelling accessibility to primary health care. *Appl Spat Anal Policy* 2(3):177–194
- Batty M, Xie Y, Sun Z (1999) Modelling urban dynamics through GIS-based cellular automata. *Comput Environ Urban Syst* 23(3):205–233
- Bhagat RB (2004) Dynamics of urban population growth by size class of towns and cities in India. *Demogr India* 33(1):47
- Bhatta B (2009) Analysis of urban growth pattern using remote sensing and GIS: a case study of Kolkata, India. *Int J Remote Sens* (May 2014), 37–41. doi:10.1080/01431160802651967
- Bitter C, Mulligan GF, Dall'erna S (2007) Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. *J Geogr Syst* 9:7–27. doi:10.1007/s10109-006-0028-7
- Brown S, Versace VL, Laurenson L, Ierodiaconou D, Fawcett J, Salzman S (2012) Assessment of spatiotemporal varying relationships between rainfall, land cover and surface water area using geographically weighted regression. *Environ Model Assess* 17:241–254. doi:10.1007/s10666-011-9289-8
- Chen J, Chang K, Karacsonyi D, Zhang X (2014) Comparing urban land expansion and its driving factors in Shenzhen and Dongguan, China. *Habitat Int* 43:61–71. doi:10.1016/j.habitat.2014.01.004
- Clement F, Orange D, Williams M, Mulley C, Epprecht M (2009) Drivers of afforestation in Northern Vietnam: assessing local variations using geographically weighted regression. *Appl Geogr* 29:561–576. doi:10.1016/j.apgeog.2009.01.003
- Du S, Wang Q, Guo L (2014) Spatially varying relationships between land-cover change and driving factors at multiple sampling scales. *J Environ Manag* 137:101–110
- Floridi M, Pagni S, Falorni S, Luzzati T (2011) An exercise in composite indicators construction: assessing the sustainability of Italian regions. *Ecol Econ* 70(8):1440–1447
- Fotheringham AS, Charlton M, Brunsdon C (1998) Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Plan Environ C* 30:1905–1927
- Fotheringham AS, Crespo R, Yao J (2015) Geographical and temporal weighted regression (GTWR). *Geogr Anal*. doi:10.1111/gean.12071
- Haase D, Haase A, Kabisch N, Kabisch S, Rink D (2012) Actors and factors in land-use simulation: the challenge of urban shrinkage. *Environ Model Softw* 35(7):92–103. doi:10.1016/j.envsoft.2012.02.012

- Haitao Z, Long GUO, Jiaying C, Peihong FU, Jianli GU, Guangyu L (2014) Modeling of spatial distributions of farmland density and its temporal change using geographically weighted regression model. *Chin Geogr Sci* 24(2):191–204. doi:[10.1007/s11769-013-0631-8](https://doi.org/10.1007/s11769-013-0631-8)
- Huang Y, Leung Y (2002) Analysing regional industrialisation in Jiangsu province using geographically weighted regression. *J Geogr Syst* 4(2):233–249
- Huang B, Wu B, Barry M (2010) Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *Int J Geogr Inf Sci* 24(3):383–401. doi:[10.1080/13658810802672469](https://doi.org/10.1080/13658810802672469)
- Kaligari M, Zibera I (2014) Geographically weighted regression of the urban heat island of a small city. *Appl Geogr* 53:341–353
- Lee KH, Schuett MA (2014) Exploring spatial variations in the relationships between residents' recreation demand and associated factors: a case study in Texas. *Appl Geogr* 53:213–222
- Li X, Zhou W, Ouyang Z (2013) Forty years of urban expansion in Beijing: what is the relative importance of physical, socioeconomic, and neighborhood factors? *Appl Geogr* 38:1–10. doi:[10.1016/j.apgeog.2012.11.004](https://doi.org/10.1016/j.apgeog.2012.11.004)
- Long H, Tang G, Li X, Heilig GK (2007) Socio-economic driving forces of land-use change in Kunshan, the Yangtze River Delta economic area of China. *J Environ Manag* 83(3):351–364
- Lu C, Wu Y, Shen Q, Wang H (2013) Driving force of urban growth and regional planning : a case study of China' s Guangdong Province. *Habitat Int* 40:35–41. doi:[10.1016/j.habitatint.2013.01.006](https://doi.org/10.1016/j.habitatint.2013.01.006)
- Lu B, Charlton M, Harris P, Stewart A (2014) Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. *Int J Geogr Inf Sci* 28(4):660–681. doi:[10.1080/13658816.2013.865739](https://doi.org/10.1080/13658816.2013.865739)
- Luo J, Wei YHD (2009) Modeling spatial variations of urban growth patterns in Chinese cities: the case of Nanjing. *Landsc Urban Plan* 91(2):51–64. doi:[10.1016/j.landurbplan.2008.11.010](https://doi.org/10.1016/j.landurbplan.2008.11.010)
- Megler V, Banis D, Chang H (2014) Spatial analysis of graffiti in San Francisco. *Appl Geogr* 54:63–73
- Mondal B (2014) Modeling urban development potential surface by integrating cellular automata and Markov chain: a study on Kolkata and its surroundings. Jawaharlal Nehru University, New Delhi, India. Dissertation
- Mesev V (1997) Remote sensing of urban systems: hierarchical integration with GIS. *Comput Environ Urban Syst* 21(3):175–187
- Nakaya T (2009) GWR4 User Manual
- Pal A (2006) Scope for bottom-up planning in Kolkata: rhetoric vs reality. *Environ Urban* 18(2):501–521. doi:[10.1177/0956247806069628](https://doi.org/10.1177/0956247806069628)
- Parker DC, Manson SM, Janssen MA, Hoffmann MJ, Deadman P (2003) Multi-agent systems for the simulation of land-use and land-cover change: a review. *Ann Assoc Am Geogr* 93(2):314–337. doi:[10.1111/1467-8306.9302004](https://doi.org/10.1111/1467-8306.9302004)
- Perz SG, Aramburú C, Bremner J (2005) Population, land use and deforestation in the Pan Amazon Basin: a comparison of Brazil, Bolivia, Colombia, Ecuador, Perú and Venezuela. *Environ Dev Sustain* 7(1):23–49
- Polèse M, Denis-Jacob J (2010) Changes at the top: a cross-country examination over the 20th century of the rise (and fall) in rank of the top cities in national urban hierarchies. *Urban Stud* 47(9):1843–1860
- Roy AUK (2005) Development of new townships: a catalyst in the growth of rural fringes of Kolkata Metropolitan Area (KMA), 1–7. Retrieved from <http://www.atiwb.gov.in/U2.pdf>
- Roy A (2011) Re-forming the megacity: calcutta and the rural–urban interface. *Megacities: library for sustainable urban regeneration* 10(1):93–109. doi:[10.1007/978-4-431-99267-7_5](https://doi.org/10.1007/978-4-431-99267-7_5)
- Sengupta U (2006) Government intervention and public–private partnerships in housing delivery in Kolkata. *Habitat Int* 30(3):448–461. doi:[10.1016/j.habitatint.2004.12.002](https://doi.org/10.1016/j.habitatint.2004.12.002)
- Sengupta U (2007) Housing reform in Kolkata: changes and challenges. *Hous Stud* 22(6):965–979. doi:[10.1080/02673030701608217](https://doi.org/10.1080/02673030701608217)
- Shafizadeh Moghadam H, Helbich M (2013) Spatiotemporal urbanization processes in the megacity of Mumbai, India: a Markov chains-cellular automata urban growth model. *Appl Geogr* 40(6):140–149. doi:[10.1016/j.apgeog.2013.01.009](https://doi.org/10.1016/j.apgeog.2013.01.009)
- Shafizadeh-Moghadam H, Helbich H (2015) Spatiotemporal variability of urban growth factors: a global and local perspective on the megacity of Mumbai. *Int J Appl Earth Obs Geoinf* 35:187–198
- Shaw A, Satish MK (2007) Metropolitan restructuring in post-liberalized India: separating the global and the. *Cities* 24(2):148–163. doi:[10.1016/j.cities.2006.02.001](https://doi.org/10.1016/j.cities.2006.02.001)
- Sivaramakrishnan KC, Kundu A, Singh BN (2005) Handbook of urbanization in India: an analysis of trends and processes. Oxford University Press, Oxford
- Thapa RB, Estoque RC (2012) Geographically weighted regression in geospatial analysis. In: Murayama Y (ed) *Progress in geospatial analysis*. Springer, Japan, pp 85–96
- Todes A (2012) Urban growth and strategic spatial planning in Johannesburg, South Africa. *Cities* 29(3):158–165. doi:[10.1016/j.cities.2011.08.004](https://doi.org/10.1016/j.cities.2011.08.004)
- Weng Q (2010) Remote sensing and GIS integration: theories, methods, and applications. McGraw-Hill, New York, p 416
- Zhou P, Ang BW, Poh KL (2005) Comparing aggregating methods for constructing the composite environmental index: an objective measure. *Ecol Econ* 9:305–311