

Chapter 7

Key Driving Factors Influencing Urban Growth: Spatial-Statistical Modelling with CLUE-s

Sohel J. Ahmed, Glen Bramley, and Peter H. Verburg

Abstract In a rapidly urbanising megacity such as Dhaka, identifying the driving factors that influence urban growth at different spatio-temporal scales is of considerable importance. In this study, based on literature survey and data availability, a selection of drivers is chosen and then tested through logistic regression. Using the CLUE-s land use modelling framework, the ability of these drivers to simulate urbanisation for the periods of 1988–1999 and 1999–2005 was examined against observed data. The results indicated that the role of these driving factors, as contributors to explaining change dynamics of urban land in Dhaka, changes with time. The overall performance of the model, when validated against observed data, is similar to that reported for other urban growth models.

Keywords Calibration • CLUE-s • Dhaka • Drivers of change • LUCC • Spatial statistical models • Urban growth • Validation

7.1 Introduction

Research in land use change can be grouped into three broad categories (Rindfuss et al. 2004): (1) observation and monitoring of land use change, (2) identification of the drivers of land use change and (3) computer modelling of land use (change) that

S.J. Ahmed (✉)

The Bartlett Development Planning Unit, University College London (UCL), 34 Tavistock Square, WC1H 9EZ London, UK
e-mail: sohel.ahmed@ucl.ac.uk

G. Bramley

School of the Built Environment, Heriot-Watt University, Edinburgh EH14 4AS, UK
e-mail: G.Bramley@hw.ac.uk

P.H. Verburg

Institute of Environmental Studies, VU University, Amsterdam, The Netherlands
e-mail: peter.verburg@vu.nl

combines categories 1 and 2 in a dynamic and integrative manner. Given Dhaka's history of growth and rapid urbanisation, it is vital to know how the various different locational drivers influence its urban footprint, particularly in the years since independence from Pakistan. The key aim of this chapter is therefore to explore and analyse the key driving factors that influence changes in built-up areas of Dhaka. This research embraces all three concepts mentioned above; initially, it analyses land use patterns by identifying relations between these observed patterns and explanatory factors and then aims to link them with the processes that are accountable for those patterns in a GIS environment by dynamic land use modelling (Overmars et al. 2007; Verburg et al. 2006; Geist et al. 2006; Overmars 2006; Verburg et al. 2004a). To accomplish this, selected driving factors are used in multivariate statistical models to explore their degree of influence on the observed urban growth at two different periods. Calibrated results are subsequently used to generate simulations in the CLUE-s raster modelling framework, which are then compared with observed changes in urban land use.

7.2 Data and Methods

7.2.1 CLUE-s: Introduction and Modelling Framework

The CLUE-s (Conversion of Land Use and its Effects at small regional extent) modelling framework was developed at Wageningen University (Verburg et al. 2002). Initially, it was applied at regional scale (Veldkamp and Fresco 1996), but the model has now become one of the most widely applied Land Use Land Cover Change (LUCC) models in various settings, including agricultural intensification, deforestation and urbanisation (e.g. Batisani and Yarnal 2009; Verburg et al. 2002, 2008; Verburg and Overmars 2007). It was developed to simulate land use change using statistically quantified relationships between land use and its driving factors in combination with dynamic modelling of competition between land use types where each cell only contains one land use type. It is capable of identifying areas that have high probabilities for future land use changes and potential 'hot spots' of land use change (Batisani and Yarnal 2009; Verburg et al. 2002, 2008; Verburg and Overmars 2007). The model is subdivided into two distinct modules, a nonspatial demand module and a spatially explicit allocation module (Verburg et al. 2002).

The *nonspatial module*, also known as the land use demand module, can take different model specifications, ranging from simple trend extrapolations to complex economic models. Although the choice for a specific model varies, the results from this demand module provide an annual demand for land, for different land use types, which is then distributed by the allocation module. For the work reported here, a simple piecewise regression was plotted between two observed periods to calculate trend demand for different land use/land cover.

The *spatial module* translates these demands and allocates them spatially (at cell level) which results in a representation of land use changes at different locations within the study region. For every year in the specified time frame, this module creates a land use prediction map taking into account a number of location specific preferences or combinations of them, derived mainly from a multivariate statistical model (e.g. logistic regression) (see Verburg et al. 2002, 2004b; Verburg and Overmars 2007).

To evaluate the probability of a location becoming urban, different driving factors are included as independent variables in the multivariate regression procedure (Verburg et al. 2002, 2004b). The CLUE-s framework assumes that the relationships based on the current land use pattern remain valid during the simulation period (Verburg et al. 2002, 2004b). The dependent variable for the logistic regression is a presence or absence event where 1 denotes that a given cell is certain to convert to urban land, and where 0 shows no change (Lesschen et al. 2005). The logistic coefficients (known as unstandardised logistic regression coefficients or logit coefficients) or the β coefficients vary between plus and minus infinity, with 0 indicating that the given explanatory variable does not affect the logit; positive or negative β coefficients indicate the explanatory variable increases or decreases the logit of the dependent (Lesschen et al. 2005).

Figure 7.1 provides an overview of the CLUE-s model. As seen in the diagram, the allocation is based upon a combination of modelling blocks. In addition, a set of decision rules is specified to restrict the conversions that take place based on the actual land use pattern. These are termed within the model as *Land use specific conversion settings* and can be defined as conversion resistance within a conversion matrix.

Current land use, in many cases, is an important determinant of location preference. Land use conversions are often costly and some conversions are almost irreversible (e.g. residential area is not likely to be converted back into agricultural area). Therefore, a *conversion resistance* factor is assigned to each land use type so that differences in behaviour towards conversion are approximated by conversion costs. For each land use type, a value is specified that represents the relative resistance to change, ranging from 0 (easy conversion) to 1 (irreversible change). The factor can be based on either expert knowledge or observed behaviour in the recent past, or alternatively subjective judgemental parameters can be used to calibrate the model.

The final step before running the model is to set up the conversion settings within a *conversion matrix* where each land use category is allocated values to indicate which other land use categories it can convert to in the next time step. The matrix shows the probability of each current category being converted to another. Rows represent present land use while columns show potential future land use. A value of 1 indicates that the conversion is allowed, and a value of 0 means that the conversion is not allowed.

Thus, this process of spatial analysis and inductive dynamic modelling comprises four categories (i.e. driving factors to change, spatial planning policies and restrictions, neighbourhood effects and land use specific conversion settings) that

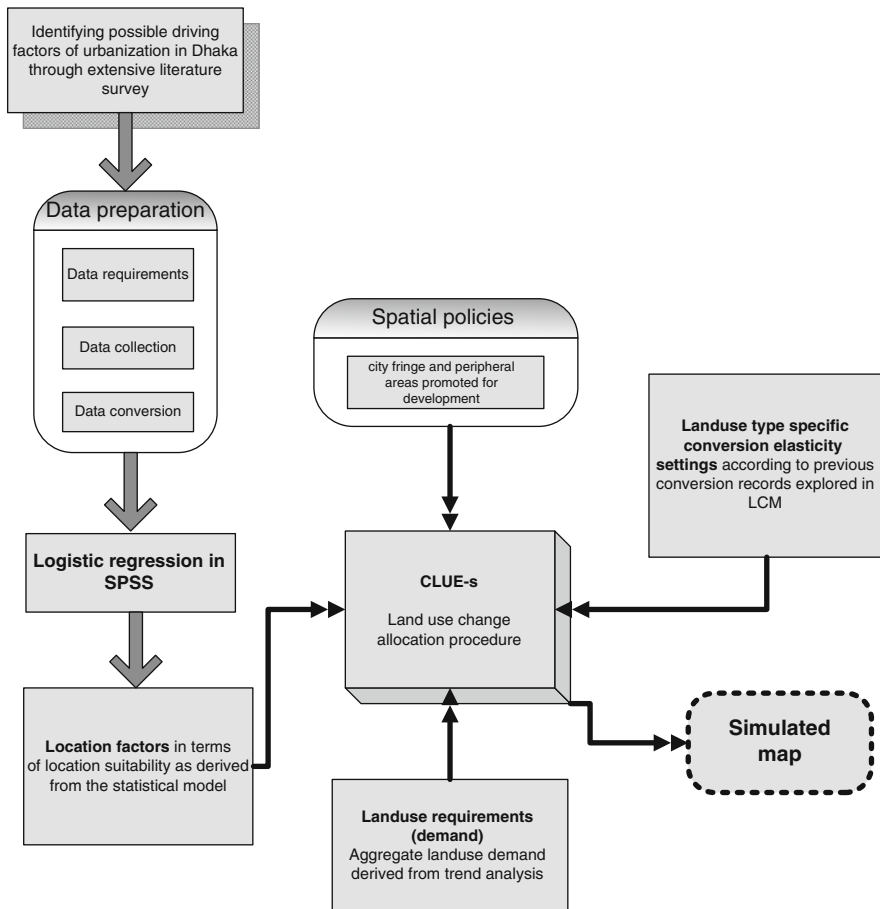


Fig. 7.1 Detailed modelling framework of the CLUE-s model (Modified after Verburg et al. 2002)

together create a set of conditions and possibilities for which the model calculates the best solution in an iterative procedure. Neighbourhood characteristics and spatial policies will be discussed in Sect. 7.3.

7.2.2 Type of Key Variables, Data Access and Management Issues

Two major types of variables are used within empirical land use change models: non-neighbourhood- and neighbourhood-based variables. While the former category

is included in all land use studies, only a limited number of studies consider neighbourhood effects and spatial autocorrelation in the modelling of land use change (Braumoha and Onishi 2007; Verburg et al. 2004a; Cheng and Masser 2003). A common set of variables encompassing biophysical and socioeconomic parameters are taken into account in most of the growth modelling studies such as in Clark's SLUETH or Metronomica from the Research Institute for Knowledge Systems (RIKS) or other spatial statistical models (Dietzel and Clarke 2006; White et al. 2004; Barredo et al. 2003, 2004; Engelen et al. 2002; White and Engelen 2000; Clarke and Gaydos 1998). Additionally, the selection of variables depends on both the adopted model and the characteristics of the area being modelled. For instance, Dhaka is surrounded by five major rivers which cause both fluvial and pluvial floods in the monsoon season, and therefore, urban land encroachment generally occurs onto relatively elevated agricultural lands and vegetated areas. In addition, some classical urban economic models like the Bid Rent model, Central Place theory and Tobler's first law of geography hold true in many contexts and provide sound justification in selecting the driving factors influencing urban expansion in Dhaka.

Research indicates that there are four primary factors that drive land use change in an area (Braumoha and Onishi 2007; Hu and Lo 2007; Verburg et al. 2004a, b). These are:

1. Socioeconomic factors
2. Biophysical constraints and potentials
3. Neighbourhood characteristics
4. Institutional factors in the form of spatial policies

The influence of these factors is often determined by an inductive approach, i.e. using multivariate statistical technique (e.g. linear or logistic regression) (Lesschen et al. 2005). One of the major pitfalls of this technique is that it does not account for spatial dependency; hence, autocorrelation may be present in the residuals. To overcome this problem, a stratified random or systematic spatial sampling scheme can be considered for spatial modelling with logistic regression (Cheng and Masser 2003). The major reason for doing this is to expand the distance interval between the sampling sites. Systematic sampling is effective in reducing spatial dependency but may cause loss of important information in isolated areas. In other words, its ability to accurately represent the population/study area may decrease with the increase of distance interval. Conversely, random sampling is efficient in representing population but has low efficiency in reducing spatial dependence, especially local spatial dependence.

In this study, none of the independent variables were found to be strongly correlated when tested using the Moran's *I* autocorrelation statistic (<0.4 on average). But they are moderately autocorrected, particularly the distance variables, which is to be expected (Braumoha and Onishi 2007; Hu and Lo 2007; Cheng and Masser 2003). Therefore, systematic sampling was initially performed which left one cell distance between sampled cells. This was later modified to spatial random

sampling in order to further minimise spatial autocorrelation. Hence, this study uses a randomly distributed sample (30 % of all observations) instead of the full dataset. This does not cause conflict between sample size and spatial autocorrelation as the sample size is still large. This approach is said to be effective at minimising spatial dependency to a level which will not impact the modelling outcome (Verburg et al. 2004a; Cheng and Masser 2003). In addition, the inclusion of neighbourhood characteristics in the model specification explicitly addresses the issue of autocorrelation in land use patterns. The sampling and consequent analysis is run at a 30×30 m resolution, which is the original resolution of the land use/cover maps (Dewan and Yamaguchi 2009a, b). Dependent and independent variables are, therefore, kept in or converted to raster format and are resampled where necessary to maintain this resolution. Consequently, datasets for all variables were in raster format with 30 m grid cell size on the Bangladesh Transverse Mercator projection system. In the CLUE-s model only one land use type is allowed for each grid cell. The raster data was then converted to ASCII format for input to the CLUE-s model. Regression modelling was performed in SPSS. Afterwards, the results from the statistical analysis were produced as probability maps using CLUE-s modelling framework and ArcGIS Desktop.

The study area includes Dhaka Metropolitan Area (DMA) (see Chap. 1) and beyond and covers an area of 420 km². A detailed description of the preparation of the land use/cover data can be found elsewhere (Dewan and Yamaguchi 2009a, b). This study uses land use/land cover maps for 1988 (t1), 1999 (t2) and 2005 (t2). The original land use/cover data contained six categories which are aggregated into urban and nonurban classes for the purposes of this study. Urban land use change, between two periods (1988–1999 referred to as t1–t2 and 1999–2005 as t2–t3), was considered as the dependent variable. The selection of dependent and independent variables is made between t1 and t2 as most of the independent variables are from t1 to t2, and also changes in urban surface started to be more pronounced during the late 1990s (Islam 1996, 2005). Besides, the lack of data in Dhaka precluded the use of more variables in this study. For instance, variables that are intrinsically dynamic in nature and that change over time play a crucial role in obtaining accurate result from the analysis and modelling. However, spatial data for roads and other facilities were only available for a single year despite the fact that urban growth was continuous. This can also be a crucial barrier to optimum results if the models are used to generate scenarios for the near future. Although it is true that demographic factors are important determinants of urbanisation in the study area, unavailability of finer resolution population data (at neighbourhood level) at the time of study prevented their inclusion in the simulation of urban growth. Similarly, other socioeconomic factors such as economic status, land value, type of water supply and sanitation could not be taken into account due to their unavailability.

7.2.3 Calibration and Validation Techniques

Until recently, efforts to calibrate and validate LUC models have received scant attention. This lack of attention can be attributed not only to the absence of universal techniques but also to lack of adequate data, particularly in developing countries (Veldkamp and Lambin 2001). However, with recent advances in spatial analysis and modelling techniques, significant progress has been made with calibrating and validating modelling results (Dietzel and Clarke 2006; Gardner and Urban 2005; Hagen-Zanker 2002a, b, 2003; Hagen-Zanker et al. 2005; Pontius and Schneider 2001; Pontius et al. 2004, 2008; Silva and Clarke 2005; Straatman et al. 2004; Visser 2004; Visser and de Nijs 2006).

There is no universally accepted procedure, nor is there an accepted set of guidelines for validating spatio-temporal models. Nevertheless, in principle, this procedure requires land use data at the start and end time of the period being modelled (Vliet 2009). It is also argued that model validation should use different time periods than those for calibration (Vliet 2009; Hagen-Zanker and Lajoie 2008). This is necessary since model results, which are consistent and accurate enough during calibration process, do not guarantee good capacity to predict at different time periods or during the validation process (Hagen-Zanker and Lajoie 2008).

Several map comparison tools were used for both calibration and validation, namely, the kappa statistic and its variants (particularly, fuzzy kappa) (Vliet 2009; Hagen-Zanker and Lajoie 2008; Pontius et al. 2004, 2008). Null or no-change models are also usually used for examining the model accuracy (Pontius et al. 2004, 2008). However, calibration and validation comparisons against null models can be misleading, since land uses with higher conversion cost possess high inertia to change and thereby results in higher false accuracy. This is why a variant of the neutral model was used. This model was created from land use maps that simulated the correct amount of land use change, randomly allocated over the map with minimal adjustments to the original land use map (Vliet 2009; Hagen-Zanker and Lajoie 2008). This neutral model provided additional benchmark maps for calibration and validation. Again as a rule of thumb, a simulation must perform better than the neutral model to indicate that the model possesses predictive power to make scenarios for the future.

The goodness of fit of logistic regression models may be assessed using the Relative/Receiver Operating Characteristic (ROC) value. The ROC method is used in LUC modelling to measure the relationship between simulated change and actual change (Pontius and Schneider 2001). In this study, the ROC method offers an answer to one important question: How well is urban growth concentrated at the locations of relatively high suitability for urban growth? The ROC statistic ranges from 0 to 1. An ROC value of 1 indicates that there is a perfect spatial agreement between the actual urban growth map and the predicted probability (perfect discrimination). An ROC value of 0.5 is the agreement that would be expected due to chance (Hu and Lo 2007; Lesschen et al. 2005; Verburg et al. 2002, 2004a; Pontius and Schneider 2001).

7.3 Description of Dataset

Based on extensive literature review, three categories of non-neighbourhood variables were used, apart from neighbourhood variables. These were biophysical, access to opportunities and planning variables. Access to opportunity variables were used as a proxy for socioeconomic variables. A brief description of each category is given below, followed by neighbourhood variables.

7.3.1 *Biophysical Variables*

Previous studies have demonstrated that in a highly flood prone study area such as Dhaka, land elevation plays a crucial role in determining urban expansion (Dewan 2013; Dewan and Yamaguchi 2008; Alam and Rabbani 2007; Dewan et al. 2007; Islam 2005; Huq 1999; RAJUK 1997). Because of recurrent floods and water logging in the monsoon season, flood-free land is usually the target of both public and private organisations for urban development. Therefore, elevation data has been considered here, and land with an elevation above five metres has been regarded as being relatively less vulnerable to flood (Dewan 2013; Dewan and Yamaguchi 2008; Dewan et al. 2007; Dewan et al. 2006a, b).

7.3.2 *Access to Opportunities as Proxies to Socioeconomic Variables*

Given the paucity of spatially encoded data at the local level in developing countries, information such as the distribution of employment opportunities at different spatial scales can be used as a socioeconomic indicator. The use of approximations like this can be further strengthened by relating this variable to classic utility maximising models like the Bid Rent or Lowry Models (Alonso 1964). A common variable that is used in economic models of land use change is the distance between residential location and employment which is in itself a surrogate for costs associated with travel. A number of studies, in various settings, have employed proxy variables in the absence of actual data (Jat et al. 2008; Braimoha and Onishi 2007; Hu and Lo 2007; Wu et al. 2006; Xiao et al. 2006; Verburg et al. 2006, 2004a, b; Henriquez et al. 2006; Liu et al. 2005; Sui and Zeng 2001; Veldkamp and Lambin 2001; Yeh and Li 1998). Therefore, accessibility representing attractiveness of any given land in terms of employment opportunities, services and facilities was included in this study. Although Euclidian distance has been used as an accessibility indicator in modelling urban land use in some studies (e.g. Cheng and Masser 2003), this may be unrealistic to use given that Euclidian distance is unable to account for the presence of infrastructure and natural features such as rivers and lakes. Apart from economic motives, travel time can be an important factor that influences the decision to live or work at a certain location.

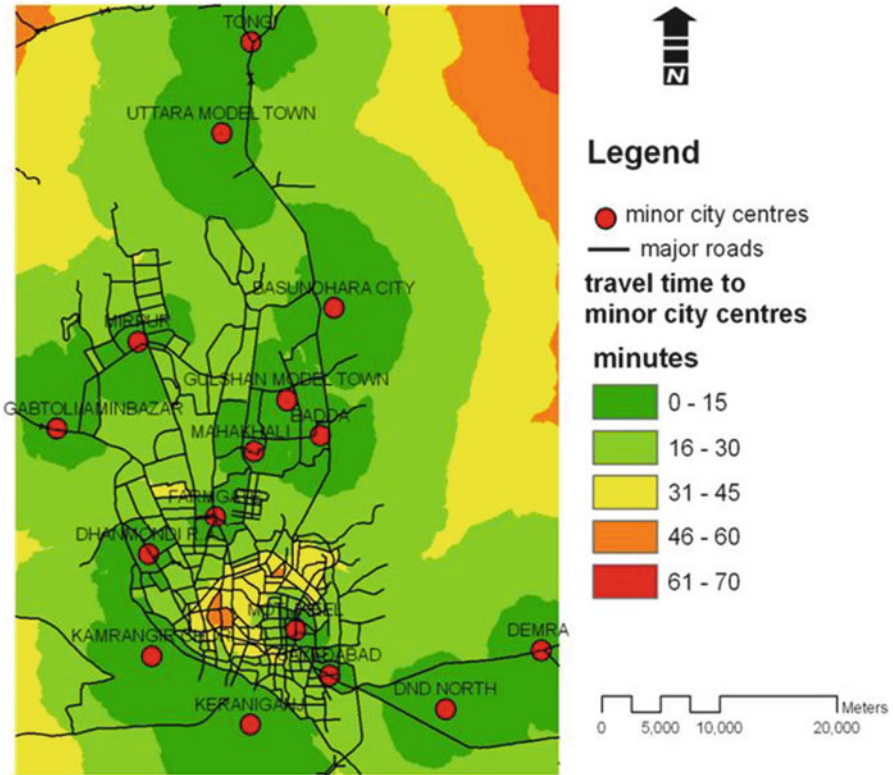


Fig. 7.2 An example of a network-based accessibility surface

Therefore, for this study, network-based accessibility measures were developed using average travel time along major road networks (NPS-ROMN 2008) (see Fig. 7.2 as an example).

7.3.3 Spatial Policy/Planning Variables

Spatial policies that provide guidance for development control or promotion can influence the patterns of land use change. Spatial policies and restrictions mostly indicate areas where land use changes are either restricted or allowed by the relevant authorities (RAJUK 1993). These can take the form of Boolean maps with cells set to 0 (areas not allowed to change) or 1 (areas allowed to change). In case of Dhaka, areas such as land of high agricultural value, areas designated as flood retention ponds and the city’s existing natural drainage systems have been selected as restricted planning variables, and the appropriate spatial data were extracted from the Dhaka Metropolitan Development Plan (1995–2015) (RAJUK

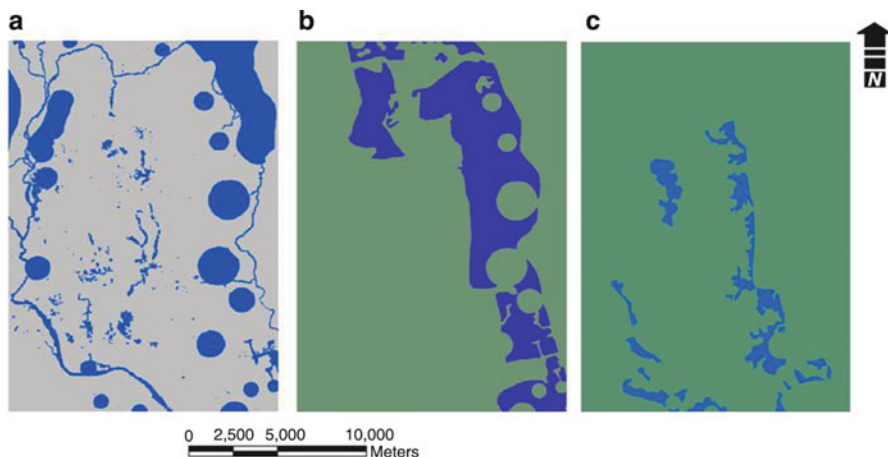


Fig. 7.3 Restricted planning variables: (a) areas restricted for development, (b) areas promoted for development, and (c) fringe areas for accelerated development (Adapted from DMDP 1997)

1997) (Fig. 7.3a). In addition, a category of permitted planning (or spatial policy) variables have been used in this study to delineate areas where urban development is permitted by the relevant authority. These areas include flood-free fringe and edge areas (e.g. northern edge areas around Uttora and Tongi) where new development is encouraged by the DMDP Plan (RAJUK 1997) (Fig. 7.3b, c).

7.3.4 Neighbourhood Variables

The incorporation of neighbourhood effects into a model acknowledges the importance of spatial interaction, proximity and spillover effects. These are known to have crucial roles in the spatial structuring of land use and can be expected in models in which relative location matters (Verburg et al. 2004c; White and Engelen 1993). The inclusion of this variable in empirical statistical models is believed to produce result similar to a constrained cellular automata model (Verburg and Overmars 2007). Spatial interactions can be represented in the model by incorporating a neighbourhood function as one of the drivers of the location preference. The relationship between location preference and neighbourhood composition may be determined either by (calibrated) decision rules, as in most cellular automata models, or can be based on statistical analysis of empirical results (Verburg et al. 2004c). Consequently, Verburg et al. (2004c) introduced the mean enrichment factor, a spatial metric of over- or under-representation of different land use types in the neighbourhood of a specific grid cell. The enrichment factor is presented at a logarithmic scale to obtain an equal scale for land use types that occur above the total

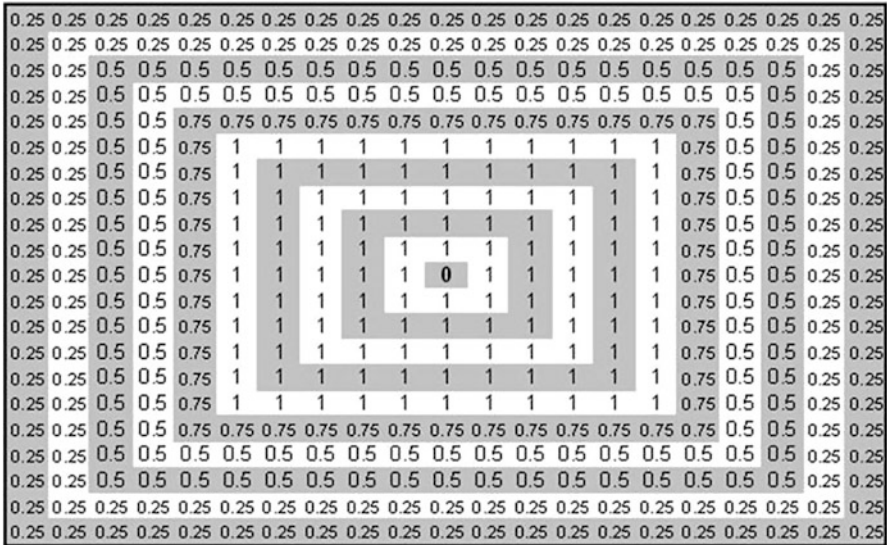


Fig. 7.4 Size and weights for 21 × 21 cell neighbourhood

study area average in the neighbourhood (enrichment factor >1) and land use types that occur below that average in the neighbourhood (enrichment factor <1).

Based on the analysis of enrichment factors for built-up areas on the 1988 and 1999 land use map, a 21 × 21 cell window was used (i.e. 10 cells from the central cell on all sides of the central cell) for the simulation (Fig. 7.4). Cells close to the central cells are given higher weight, and cells further away are given less weight to reflect the distance decay effect found for built-up areas by the mean enrichment factors.

7.4 Results

7.4.1 Statistical Models of Urban Land Change

Binary logistic regression was used to model the probability of conversion to the urban category in two periods (t1–t2 and t2–t3) for any given cell using a number of aforementioned driving factors (Sect. 7.3) as the independent variables.

From Table 7.1, based on the Wald statistics,¹ it can be deduced that the major determinators of urban development change over time. Distance from the CBD is important in the first period but is replaced by closeness to other minor urban centres in the second; proximity to major roads becomes less important with time,

¹The Wald statistic is the squared ratio of the unstandardised logistic coefficient to its standard error.

Table 7.1 Logistic regression results for driving factors

Driving factors	Urban t1–t2			Urban t2–t3		
	β^a	$\exp(\beta)$	Wald statistic	β	$\exp(\beta)$	Wald statistic
Travel time to commercial centres	-.03	.97	1,787.26	-.02	.98	444.62
Travel time to historical CBD	-.01	.99	2,256.08	-.01	.99	821.67
Travel time to industrial hubs	-.02	.98	1,117.15	-.02	.98	1,012.45
Travel time to minor city centres	-.02	.97	904.19	-.04	.96	2,586.42
Travel time to major roads elevation	-.17	.84	8,274.91	-.04	.96	1,411.04
Planning variable 1 (urban fringe acceleration)	.24	1.28	3,887.09	.13	1.13	1,650.64
Planning variable 2 (promoted areas for development)	* *	*	*	-.26	.77	104.52
Constant	-.10	1.07	4.31	-.10	.90	55.06
ROC	1.37	3.93	2,201.91	.96	5.77	1,300.03
	0.86			0.70		

*Not significant at 0.05 level, rest of the results are significant at 0.05 level

^a β = unstandardised coefficients and $\exp(\beta)$ = odds ratio

and the physical configuration of sites also becomes less significant. To some extent, this is an example of spatial pattern modelling at various temporal scales. It shows that various factors are changing their roles in the process of land development. It should be noted that this statistical analysis only takes the probability of development into account and does not consider detailed spatial patterns, such as changes in density or vertical expansions of urban growth.

A decentralised, polycentric suburbanizing trend in the metropolitan Dhaka area is also evident from the resultant coefficients. Probability (expressed as log-odds) of converting into urban areas decreases when cells are further away from the CBD, commercial centres, industrial hubs, minor city centres and major roads for both t1–t2 and t2–t3 periods. Nevertheless, the role of CBD becomes less important for the t2–t3 period as evidenced by Wald statistics. This is due to non-availability of land for development in the proximity of CBD since the late 1990s. Dhaka becomes more condensed and consolidated near and around CBD since the 1990s (see Chap. 3). Vibrant activity centres at the east and west fringe areas started to become more urban at or after the period t2–t3. This is reflected in the coefficients and Wald statistics for t2–t3. This can be further exemplified by travel time to minor centres in the later period, which shows the greatest effect in t2–t3. Travel time to industrial areas was found to be negatively related to the urban growth probability (i.e. increases in travel time to industrial areas decreased the logit of urban growth), and it seems to have kept the same importance between the observed periods considered.

Even though the probability of urban growth decreased with increasing travel time to commercial hubs/markets, this also becomes less significant at the period of t2–t3. This may be the result of fragmented growth away from these minor city centres. In other words, people are becoming more vehicle dependent.

Accessibility to major roads is a major determinant of urban land change, and this variable contributed significantly to the model in the period $t1-t2$; i.e. areas near to major road are 0.171 times more likely to convert to urban land than those further away. But this variable does not predict much growth during $t2-t3$ since updated data on roads was not available at the time of analysis.

As found in the logistic regression results (Table 7.1), most of the independent variables have significant influence in shaping urban growth in Dhaka despite variations in the degree of influence in different time periods. However, not all the results enable us to draw a solid conclusion. For example, planning variables, particularly urban fringe areas for accelerated development, are not significant during $t1-t2$. This is due to the fact that during that time, there was room for consolidation within the inner city areas, and this is reflected by the significant influence of factors like travel time to inner city facilities, namely, CBD, commercial and industrial areas.

According to the ROC values in Table 7.1, the driving factors explain the variations in urban growth in $t1-t2$ better than in $t2-t3$. While the lower level of fitting for the later period can be attributed to the absence of updated data for the later period, there may be other reasons for overall low performance of the statistical model. This shows that it may be important to include many other possible drivers, particularly those related to population, in the process of spatial allocation of new urban areas. Possibly, the location of urban land can be better explained by self-reinforcement of new urban locations through economies of scale and the emergence of new centres. When the observed pattern of urban locations cannot be explained by determinism alone, chance events become important determinators of city systems. Chance events, self-organisation, self-emergence and coincidences are all characteristics of complex systems (Batty and Xie 1994). Such elements can never be adequately included in an empirical analysis of LUCC. Reinforcement of patterns and path dependence can be studied over shorter time scales which require the application of complex theories (Cheng and Masser 2004; Batty et al. 1999). Previous case studies, therefore, have attempted to include neighbourhood variables in empirical analysis to mimic some of these complexity theories applicable to cities (Hansen 2008; Verburg et al. 2004c).

At this point, inclusion of neighbourhood variables, as enrichment factors, contributes significantly to the model. From Fig. 7.5, where the mean enrichment factor for urban land use 1999 is illustrated, it can be seen that neighbouring cells of existing urban ones are more likely to become urban than those further away. Outward expansion of central Dhaka and continuous infill and consolidation may be the major reason for this characteristic of urban land. Therefore, the enrichment factors developed with urban $t1$ and $t2$ as neighbourhood variables are included within the statistical models for $t1-t2$ and $t2-t3$, respectively. This has resulted in slightly better prediction overall by the models (with an ROC value of 87 and 73 % for $t1-t2$ and $t2-t3$ correspondingly). Both these ROC values are high enough to proceed with urban growth simulation.

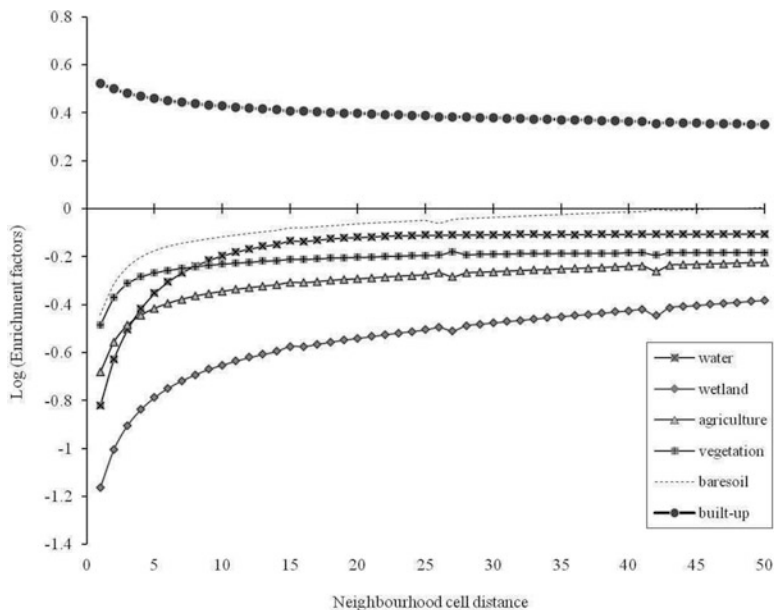


Fig. 7.5 Mean enrichment factors for built-up areas in 1999

Table 7.2 Conversion resistance factor values

	Water	Wetland	Agriculture	Vegetation	Built-up	Bare soil
Initial value	1	0.2	0.2	0.4	0.7	0.4
Final value	0.7	0.1	0.1	0.1	0.8	0.3

Table 7.3 Conversion matrix for CLUE-s simulation

Land use type	Water	Wetland	Agriculture	Vegetation	Built-up	Bare soil
Water	1	0	0	0	1	1
Wetland	0	1	0	0	1	1
Agriculture	0	0	1	0	1	1
Vegetation	0	0	0	1	1	1
Built-up	0	0	0	0	1	0
Bare soil	0	0	0	0	1	1

7.4.2 Conversion Resistance and Conversion Matrix

Based on recent land use change studies (Ahmed et al. 2012; Dewan and Yamaguchi 2009a, b), a tentative conversion resistance table and matrix were specified, and after calibration the values shown in Tables 7.2 and 7.3 were derived for each land use type. As seen in Table 7.2, the water and built-up categories were

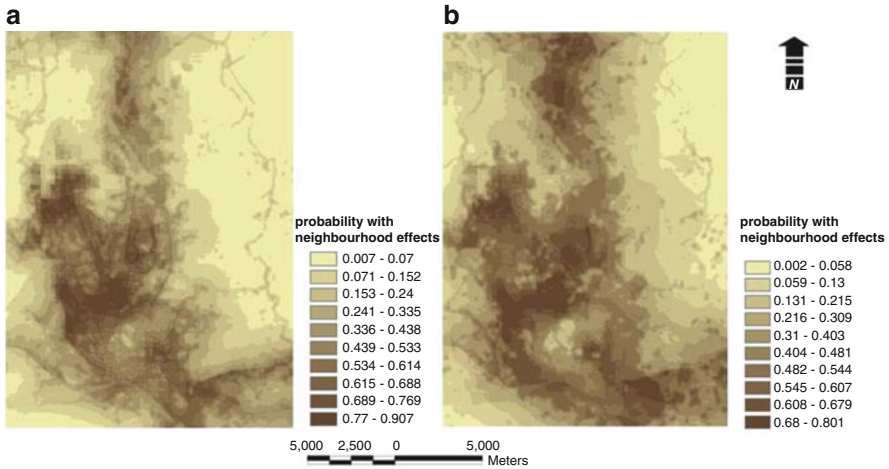


Fig. 7.6 Probability maps with neighbourhood effects: (a) t1–t2 and (b) t2–t3

given higher values as they are more difficult to convert and more likely to stay in the initial land use type. On the other hand, bare soil and agricultural land are easier to convert and less likely to resist change to built-up.

7.4.3 Simulation Runs in CLUE-s

Probability maps were generated in the CLUE-s model to examine the performance of the driving factors. Probability maps have been generated for t1–t2 and t2–t3, using both the driving factors and neighbourhood effects. For both time periods, areas at mid-north, west, central areas and areas at south-east have higher probabilities to change to urban land. The probability map for the period t2–t3 shows that the eastern part of the study area has the greatest probability of change to urban land. Figure 7.6 shows final simulation for both periods. This shows that the CLUE-S modelling framework produces better simulation maps when neighbourhood effects of the spatial pattern of urban development are included. Although the neighbourhood relationships only capture part of the processes of urban expansion, the empirical approach alone based on locational conditions is not able to capture the complexity of the processes involved. Although the model has been able to capture self-organisation process in urban growth, it lacks elements that can portray spontaneous development that the city is experiencing. Hence, there are many areas at the northern and eastern extensions of the city that are becoming urban but are not at all reflected within the simulation results (Fig. 7.7).

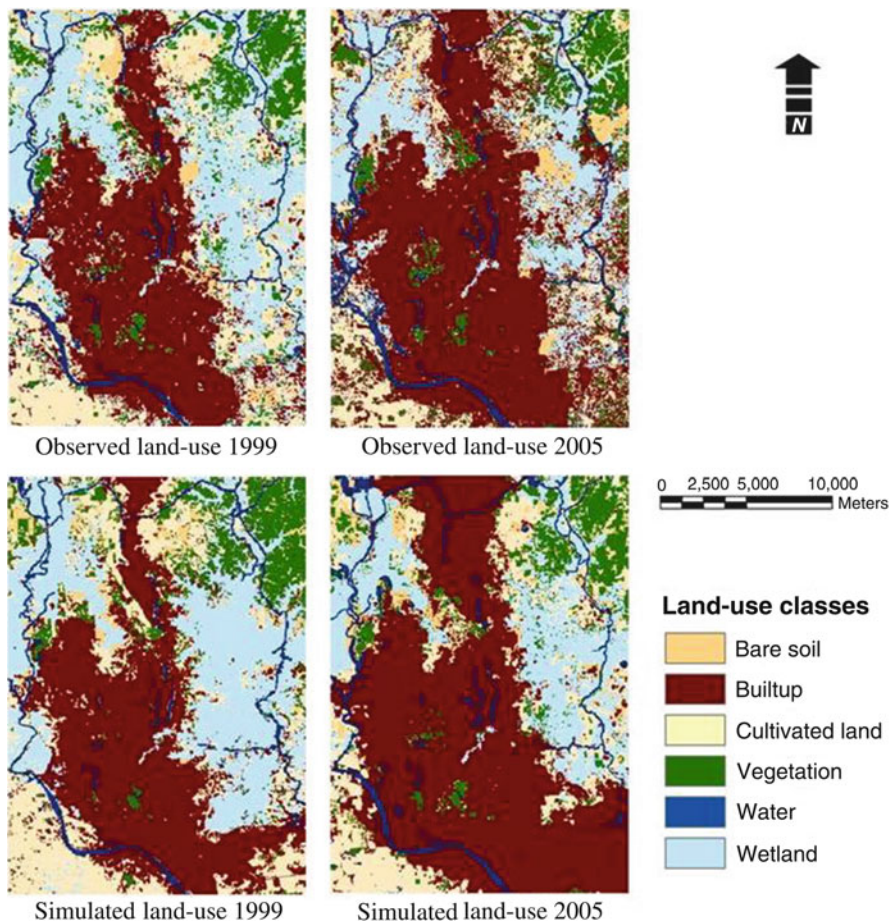


Fig. 7.7 Observed and simulated land-use for 1999 and 2005

7.5 Calibration and Validation

In this study, the model simulates observed change in urban land for t_2 using t_1 as the baseline during the calibration phase and simulates change in t_3 using t_2 as the baseline during the validation phase. The number of cells that are converted to urban land (between t_1 – t_2 and t_2 – t_3) is predefined, so errors in the calibration and validation are the result of misallocations in terms of changed cells. To check errors in terms of misallocations of cells, and assess the predictive power of the model, the neutral simulations for t_2 are compared to the observed urban land of t_2 . The comparison is conducted using variants of the kappa statistic (i.e. traditional and fuzzy kappa), and also the percentage accuracy at different scales is derived.

As shown in Table 7.4, the model simulation outperformed the neutral models in terms of kappa statistics, and the percentage accuracy for fuzzy kappa is better than

Table 7.4 Calibration and validation results compared to urban 1999 and urban 2005 (values in parenthesis)

		Methods									
	Overall kappa	KHisto	Kappa Loc	Fraction correct (model scale)	Overall fuzzy kappa	Fuzzy fraction correct (model scale)	Fraction correct at 05 cells aggregation	Fraction correct at 10 cells aggregation	Fraction correct at 20 cells aggregation	Fraction correct at 30 cells aggregation	
Urban 1999	1	1	1	1	1	1	1	1	1	1	1
Urban simulation 1999	0.650 (0.565)	0.993 (0.999)	0.655 (0.566)	0.841 (0.783)	0.167 (-.084)	0.881 (0.835)	0.857 (.808)	0.872 (.827)	0.888	0.900	
Random model1	0.476 (.418)	1	0.476 (.418)	0.762 (.709)	-.148 (-.044)	0.836 (0.841)	0.804 (.805)	0.829 (.843)	0.852	0.865	
Random model2	0.477 (.420)	1	0.477 (.420)	0.762 (.710)	-.149 (-.041)	0.836 (.841)	0.803 (.806)	0.829 (.847)	0.852	0.865	

that of the traditional kappa statistic. In addition, the model results did not improve dramatically when the percentage accuracy for a larger aggregation area is considered (Table 7.4). Nonetheless, the fuzzy kappa variants for the simulation results are lower than for the neutral. The simulation results here show a similar model accuracy to that of other LUC models reported in literature (Santé et al. 2010).

There may be several reasons why the model has not performed as expected at this stage. One possible reason could be other underlying factors not used here are equally important. These include water and sanitation networks, land value/price for different locations, population density etc. Additionally, according to Wu (2002), urban growth can be influenced by both spontaneous and self-organisational processes. Spontaneous growth results in a spatially homogeneous and sparse pattern, which contains more random components, whereas self-organisational growth results in spatial agglomeration, which is impacted by more self-organised socioeconomic activities. To understand spatial processes and patterns, we must take both types into account. Geographical phenomena can engage in a process of self-organisation in which locations with seemingly identical potential end up playing very different roles. As independent variables, the drivers in this study are able to capture the effect of self-organisation to a considerable extent. Nevertheless, while checking simulation results in the calibration and validation phases, the model results which include neighbourhood effects appear to cover the infill mechanism well but seem to fail to include the chance events of other distant areas becoming urban. This becomes particularly important during the later period, as there was much fragmented urban growth taking place especially in the eastern and north-eastern parts of the study area which are not well positioned in terms of land use preference (mostly low-lying areas that are not well linked to employment attractions of the city). Thus, as far this research is concerned, the results are successful in accommodating self-organisational process but have failed to simulate spontaneous random development.

7.6 Conclusions

Given the fact that significant urban changes have been observed between 1988 and 2005, the role of different spatial drivers, namely, distance to economic opportunities, elevation and planning, has been examined to simulate urbanisation for the periods of 1988–1999 and 1999–2005, using the CLUE-s land use modelling framework. As the city changed its growth pattern quite significantly and away from core areas, the contribution of drivers in explaining urban footprint dynamics also differed substantially. Nevertheless, local knowledge and observations in the study area lead to the deduction that the proposed eastern bypass may have significant influence on city's growth despite the fact that these areas are not free from monsoonal floods. For the same reason, the elevation variable has

considerable influence to urban growth at first period of modelling (1988–1999) but became less significant during 1999–2005. This might be due to artificial heightening of land levels from surrounding flood levels. The inclusion of neighbourhood effects improved model performance for both periods. Performance during calibration was better than that during validation, although both calibration and validation results are similar to some other reported urban growth models.

Overall, the methodology is deemed robust enough to be used for exploring and identifying factors responsible for urban growth dynamics that relate to observed patterns in a data poor country. Although the research tries to cover all the available and relevant drivers, the absence of information regarding infrastructural development (e.g. expansion, connection of existing major roads and newly built ones), land value, plans and expansions of utilities, services etc. obviously affects the ability to predict urban change accurately. It was found to be very difficult to simulate spontaneous processes, particularly during t_2 – t_3 as core and adjoining areas have become more consolidated and have left very little room for further development. Consequently, areas with a higher probability of change, as predicted by significant drivers during this period, do not coincide well with areas that actually experienced higher scatter development through leapfrogging. Such mismatches can be attributed partly to the absence of some of the variables mentioned earlier. In addition, the emerging land development pattern in Dhaka can arguably be an outcome of dominant market forces benefiting from largely weak and uncoordinated relevant institutions.

On the other hand, the land-based approach as adopted in CLUE-s model (like many other contemporary LUCC models) where the unit of analysis centres around an area of land, grid cell or census track cannot entirely mimic the decision-making or behavioural process made by individual plot owners or other agents that are involved for land use transition between different time periods. Many examples of agent-based models in land use modelling are being considered (Parker et al. 2003; Parker 2005) to mimic the exact features of urban expansion at local scale by modelling human-environment interactions and their emergent behaviour. However, these models are very data demanding which is a problem in an environment such as Dhaka. Nevertheless, prototype development of such models and their subsequent implementation may be able to complement results from simpler models like CLUE-s and lead to a better and more complete understanding of the behavioural complexity that exists in urban land development process in the megacity of Dhaka.

Acknowledgement This research was conducted as part of a Doctoral project in Urban Studies at the Heriot-Watt University, Edinburgh in the UK. The authors would like to thank the ORS and James Watt Scholarship for funding the study. Also, the authors like to thank Dr Ashraf M. Dewan for providing access to the land cover datasets. The authors are also grateful to Professor Mark Birkin (University of Leeds) and Professor Michael White (Nottingham Trent University) for constructive suggestions on the structuring of the simulation framework used in this study.

References

- Ahmed SJ, Bramley G, Dewan AM (2012) Exploratory growth analysis of a megacity through different spatial metrics: a case study on Dhaka, Bangladesh (1960–2005). *URISA Journal* 24(1):5–20
- Alam M, Rabbani MDG (2007) Vulnerabilities and responses to climate change for Dhaka. *Environ Urban* 19(1):81–97
- Alonso W (1964) Location and land use: toward a general theory of land rent. Harvard University Press, Cambridge, MA, 204 p
- Barredo JI, Kasanko M, McCormick N, Lavalle C (2003) Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. *Landsc Urban Plan* 64(3):145–160
- Barredo JI, Demicheli L, Lavalle C, Kasanko M, McCormick N (2004) Modelling future urban scenarios in developing countries: an application case study in Lagos, Nigeria. *Environ Plan B Plan Des* 32(1):65–84
- Batisani N, Yarnal B (2009) Urban expansion in Centre County, Pennsylvania: spatial dynamics and landscape transformations. *Appl Geogr* 29(2):235–249
- Batty M, Xie Y (1994) From cells to cities. *Environ Plan B* 21(7):41–48
- Batty M, Xie Y, Sun Z (1999) Modelling urban dynamics through GIS-based cellular automata. *Comput Environ Urban Syst* 23(3):205–233
- Braimoha AK, Onishi T (2007) Spatial determinants of urban land use change in Lagos, Nigeria. *Land Use Policy* 24(2):502–515
- Cheng J, Masser I (2003) Modelling urban growth patterns: a multiscale perspective. *Environ Plan A* 35(4):679–704
- Cheng J, Masser I (2004) Understanding spatial and temporal processes of urban growth: cellular automata modelling. *Environ Plan B Plan Des* 31(2):167–194
- Clarke KC, Gaydos LJ (1998) Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *Int J Geogr Inf Sci* 12(7):699–714
- Dewan AM (2013) Floods in a megacity: geospatial techniques in assessing hazards, risk and vulnerability. Springer, Dordrecht
- Dewan AM, Yamaguchi Y (2008) Effect of land cover changes on flooding: example from Greater Dhaka of Bangladesh. *Int J Geoinf* 4(1):11–20
- Dewan AM, Yamaguchi Y (2009a) Using remote sensing and GIS to detect and monitor land use and land cover change in Dhaka Metropolitan of Bangladesh during 1960–2005. *Environ Monit Assess* 150(1–4):237–249
- Dewan AM, Yamaguchi Y (2009b) Land use and land cover change in Greater Dhaka, Bangladesh: using remote sensing to promote sustainable urbanization. *Appl Geogr* 29(3):390–401
- Dewan AM, Yeboah KK, Nishigaki M (2006a) Flood hazard delineation in Greater Dhaka, Bangladesh using integrated GIS and remote sensing approach. *Geocarto Int* 21(2):33–38
- Dewan AM, Yeboah KK, Nishigaki M (2006b) Using synthetic aperture radar (SAR) data for mapping river water flooding in an urban landscape: a case study of Greater Dhaka, Bangladesh. *Jpn J Hydrol Water Resour* 19(1):44–54
- Dewan AM, Islam MM, Kumamoto T, Nishigaki M (2007) Evaluating Flood hazard for land-use planning in Greater Dhaka of Bangladesh using remote sensing and GIS techniques. *Water Resour Manag* 21(9):1601–1612
- Dietzel C, Clarke K (2006) The effect of disaggregating land use categories in cellular automata during model calibration and forecasting. *Comput Environ Urban Syst* 30(1):78–101
- Engelen G, White R, Uljee I (2002) The MURBANDY and MOLAND models for Dublin. Research Institute for Knowledge Systems (RIKS) BV, Maastricht, p 172

- Gardner RH, Urban DL (2005) Model validation and testing: past lessons, present concerns and future prospects. In: Canham CD, Cole JC, Lauenroth WK (eds) *The role of models in ecosystem science*. Princeton University Press, Princeton, pp 186–205
- Geist H, Mcconnell W, Lambin EF, Moran E, Alves D, Rudel T (2006) Causes and trajectories of land-use/cover change. In: Lambin EF, Geist H (eds) *Land-use and land-cover change: local processes and global impacts*. Springer, Berlin/Heidelberg, pp 41–70
- Hagen-Zanker A (2002a) Approaching human judgement in the automated comparison of categorical maps. In: 4th international conference on recent advances in soft computing. Nottingham Trent University, Nottingham
- Hagen-Zanker A (2002b) Multi-method assessment of map similarity. In: 5th AGILE conference on geographic information science, Palma
- Hagen-Zanker A (2003) Fuzzy set approach to assessing similarity of categorical maps. *Int J Geogr Inf Sci* 17(3):235–249
- Hagen-Zanker A, Lajoie G (2008) Neutral models of landscape change as benchmarks in the assessment of model performance. *Landsc Urban Plan* 86(3–4):284–296
- Hagen-Zanker A, Van Loon J, Maas A, Straatman B, De Nijs T, Engelen G (2005) Measuring performance of land use models: an evaluation framework for the calibration and validation of integrated land use models featuring cellular automata. In: 14th European colloquium on theoretical and quantitative geography, Tomar
- Hansen HS (2008) Quantifying and analysing neighbourhood characteristics supporting urban land-use modelling. In: Bernard L, Friis-Christensen A, Pundt H (eds) *The European information society: lecture notes in geoinformation and cartography*. Springer, Berlin/Heidelberg, pp 283–299
- Henriquez C, Azocar G, Romero H (2006) Monitoring and modeling the urban growth of two mid-sized Chilean cities. *Habitat Int* 30(4):945–964
- Hu Z, Lo CP (2007) Modeling urban growth in Atlanta using logistic regression. *Comput Environ Urban Syst* 31(6):667–688
- Huq S (1999) Environmental problems of Dhaka City. In Mitchell JK (eds) *Crucibles of Hazard: mega-cities and disasters in transition*. United Nations University Press, Tokyo/New York/Paris. <http://www.greenstone.org/greenstone3/nzdl?a=d&d=HASH9b6b117419a9f8e9a6e899.7&c=aedl&sib=1&dt=&ec=&et=&p.a=b&p.s=ClassifierBrowse&p.sa=>. Accessed 12 Dec 2008
- Islam N (1996) Dhaka: from city to megacity: perspectives on people, places, planning, and development issues. Urban Studies Programme, Dhaka
- Islam N (2005) Dhaka now: contemporary urban development. The Bangladesh Geographical Society, Dhaka
- Jat MK, Garg PK, Khare D (2008) Monitoring and modelling of urban sprawl using remote sensing and GIS techniques. *Int J Appl Earth Obs Geoinf* 10(1):26–43
- Lesschen JP, Verburg PH, Staal SJ (2005) Statistical methods for analysing the spatial dimension of changes in land use and farming systems. LUCR report series no.7. The International Livestock Research Institute & LUCR Focus 3 Office, Wageningen University, Wageningen, 81p
- Liu J, Zhan J, Deng X (2005) Spatio-temporal patterns and driving forces of urban land expansion in China during the economic reform era. *Ambio* 34(6):450–456
- NPS-ROMN (2008) Travel time cost surface model, version 1.7. Rocky Mountain Network-National Park Service, Fort Collins
- Overmars KP (2006) Linking process and pattern of land use change: illustrated with a case study in San Mariano, Isabela, Philippines. Institute of Environmental Sciences (CML), Leiden University, Leiden, 178 p
- Overmars KP, Groot WTD, Huigen MGA (2007) Comparing inductive and deductive modeling of land use decisions: principles, a model and an illustration from the Philippines. *Hum Ecol* 35:439–452

- Parker DC (2005) Integration of geographic information systems and agent-based models of land use: challenges and prospects. In: Maguire DJ, Goodchild MF, Batty M (eds) GIS, spatial analysis and modeling. ESRI Press, Redlands, pp 403–422
- Parker DC, Manson SM, Janssen MA, Hoffmann M, 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
- Pontius GR Jr, Schneider CL (2001) Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agric Ecosyst Environ* 85(1–3):239–248
- Pontius RG Jr, Huffaker D, Denman K (2004) Useful techniques of validation spatially explicit land-change models. *Ecol Model* 179(4):445–461
- Pontius RG Jr, Boersma W, Castella JS, Clarke K et al (2008) Comparing the input, output, and validation maps for several models of land change. *Ann Reg Sci* 42(1):11–37
- Rajdhani Unnayan Katripakko (RAJUK) (1993) Strategic growth options – Dhaka 2016. UNDP, UN-HABITAT, RAJUK, Dhaka
- Rajdhani Unnayan Katripakko (RAJUK) (1997) Dhaka Metropolitan Development Plan for Dhaka City (1995–2015), Vol I & II. RAJUK, Dhaka
- Rindfuss RR, Walsh SJ, Turner Ii BL, Fox J, Mishra V (2004) Developing a science of land change: challenges and methodological issues. *Proc Natl Acad Sci U S A* 101(39):13976–13981
- Santé I, García AM, Miranda D, Crecente R (2010) Cellular automata models for the simulation of real-world urban processes: a review and analysis. *Landsc Urban Plan* 96(2):108–122
- Silva EA, Clarke KC (2005) Complexity, emergence and cellular urban models: lessons learned from applying SLEUTH to two Portuguese metropolitan areas. *Eur Plan Stud* 13(1):93–115
- Straatman B, White R, Engelen G (2004) Towards an automatic calibration procedure for constrained cellular automata. *Comput Environ Urban Syst* 28(1–2):149–170
- Sui DZ, Zeng H (2001) Modeling the dynamics of landscape structure in Asia's emerging desakota regions: a case study in Shenzhen. *Landsc Urban Plan* 53(1–4):37–52
- Veldkamp A, Fresco LO (1996) CLUE-CR: an integrated multi-scale model to simulate land use change scenarios in Costa Rica. *Ecol Model* 91(1–3):231–248
- Veldkamp A, Lambin EF (2001) Predicting land-use change. *Editor Agric Ecos Environ* 85:1–6
- Verburg PH, Overmars KP (2007) Dynamic simulation of land-use change trajectories with the CLUE-s model. In: Koomen E, Stillwell J, Bakema A, Scholten HJ (eds) *Modelling land-use change: progress and applications*, 40. Springer, Berlin, pp 321–337
- Verburg PH, Soepboer W, Veldkamp A, Limpiada R, Espaldon V, Mastura SS (2002) Modeling the spatial dynamics of regional land use: the CLUE-S model. *Environ Manage* 30(3):391–405
- Verburg PH, Schot P, Dijst MJ, Veldkamp A (2004a) Land use change modelling: current practice and research priorities. *GeoJournal* 61(4):309–324
- Verburg PH, Ritsema Van Eck JR, Nijs TCMD, Dijst MJ, Schot P (2004b) Determinants of land-use change patterns in the Netherlands. *Environ Plan B Plan Des* 31(1):125–150
- Verburg PH, Nijs TCMD, Ritsema Van Eck JR, Visser HJ, de Jong K (2004c) A method to analyse neighbourhood characteristics of land use patterns. *Comput Environ Urban Syst* 28(6):667–690
- Verburg PH, Kok K, Pontius RG, Veldkamp A (2006) Modeling land-use and land-cover change. In: Lambin EF, Geist H (eds) *Land-use and land-cover change: local processes and global impacts*. Springer, Berlin/Heidelberg, pp 117–135
- Verburg PH, Eickhout B, Meul HV (2008) A multi-scale, multi-model approach for analyzing the future dynamics of European land use. *Ann Reg Sci* 42(1):57–77
- Visser H (2004) The map comparison kit: methods, software and applications. National Institute for Public Health and the Environment (RIVM), Bilthoven, 124 p
- Visser H, De Nijs T (2006) The map comparison kit. *Environ Model Software* 21(3):346–358
- Vliet JV (2009) Assessing the accuracy of changes in spatial explicit land use change models. In: 12th AGILE international conference on geographic information science Leibniz, Universität Hannover, Hanover

- White R, Engelen G (1993) Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. *Environ Plan A* 25(8):1175–1199
- White R, Engelen G (2000) High resolution integrated modelling of the spatial dynamics of urban and regional systems. *Comput Environ Urban Syst* 24(5):383–400
- White R, Straatman B, Engelen G (2004) Planning scenario visualization and assessment: a cellular automata based integrated spatial decision support system. In: Goodchild MF, Janelle D (eds) *Spatially integrated social science*. Oxford University Press, New York, pp 420–442
- Wu F (2002) Calibration of stochastic cellular automata: the application to rural-urban land conversions. *Int J Geogr Inf Sci* 16(8):795–818
- Wu Q, Li H-Q, Wang R-S, Paulussen J, Hec Y, Wang M, Wang BH, Wang Z (2006) Monitoring and predicting land use change in Beijing using remote sensing and GIS. *Landsc Urban Plan* 78 (4):322–333
- Xiao J, Shen Y, Ge J, Tateishi R, Tang C, Liang Y, Huang Z (2006) Evaluating urban expansion and land use change in Shijiazhuang, China, by using GIS and remote sensing. *Landsc Urban Plan* 75(1–2):69–80
- Yeh AGO, Li X (1998) Sustainable land development model for rapid growth areas using GIS. *Int J Geogr Inf Sci* 12(2):169–189