Cities 39 (2014) 68–86

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

Logistic regression and cellular automata-based modelling of retail, commercial and residential development in the city of Ahmedabad, India

Talat Munshi [a,*], Mark Zuidgeest [b], Mark Brussel [c], Martin van Maarseveen [c]

[a] School of Planning, CEPT University, Ahmedabad, India
[b] Centre for Transport Studies, Department of Civil Engineering, Faculty of Engineering and the Built Environment, University of Cape Town, Rondebosch 7701, South Africa
[c] Department of Urban and Regional Planning and Geo-Information Management, ITC, University of Twente, Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 14 May 2013
Received in revised form 12 February 2014
Accepted 16 February 2014
Available online 17 March 2014

Keywords:
Urban growth
Activity floor space
Logistic regression
Cellular automata
Ahmedabad

ABSTRACT

This study presents a hybrid simulation model that combines logistic regression and cellular automata-based modelling to simulate future urban growth and development for the city of Ahmedabad in India. The model enables to visualize the consequence of development projections in combination with present zoning and development control regulations. The growth in activities' floor space is projected at a zonal level using time series data. Then, a logistic regression model is used to calculate a probability surface of development transition, while a cellular automata-based spatial interaction model is used to simulate change in activity floor space per activity, and thus urban growth. The developed model has the capacity to simulate urban growth space and hence vertical growth. The structure of the model allows for a detailed urban growth simulation and is flexible enough to incorporate changes in development control regulations and settings for spatial interaction. Therefore, it carries scope of being used to visualize growth for other, similar, cities and help urban planners and decision makers to understand the consequences of their decisions on urban growth and development.

Introduction

Rapid urbanization and urban growth are recognized facts in India (MOUD., 2011). The planning of urban areas in India and, specifically, the methods used to prepare statutory urban development plans have been criticized as being ineffective (Adhvaryu, 2011). A likely reason for this poor practice of urban planning methods is mainly attributed to a limited understanding of the spatio-temporal processes and dynamics of urban growth (Adhvaryu, 2011) and the lack of tools that can help in making informed urban planning decisions (Geertman & Stillwell, 2004; Masser & Campbell, 1991). The dynamics of urban growth are driven by the availability of space in urban areas for urban areas to change or grow and the ageing process of development (referring to the ageing and deterioration of buildings, as a result of which buildings might have to be restored or reconstructed [Batty, Xie, & Sun, 1999a]). Nivola (1999) states that urban growth can progress in four directions: "in, up, down and out. Indicating that the objective of any urban growth model that is used to make informed urban planning decisions should be to identify what land use will develop, when and where, and what its repercussions are on other developments" (Cheng, 2003).

Over the years, urban growth models have proven to be effective in describing and estimating urban development (mostly outgrowth) and have consequently proven to be valuable for informed urban planning decisions (Clarke & Gaydos, 1998; Herold, Goldstein, & Clarke, 2003; Vaz, Nijkamp, Painho, & Caetano, 2012). Previous studies (e.g., Clarke, Hoppen, and Gaydos (1997), Dubovyk, Sliuzas, and Flacke (2011); Arsanjani, Helbich, Kainz, and Darvishi Boloorani (2013); Lo and Yang (2002), Thapa and Murayama (2012), Aljoufie, Zuidgeest, Brussel, and van Maarseveen (2013a); (Aljoufie, Zuidgeest, Brussel, van Vliet, & van Maarseveen, 2013b) also emphasized the need and importance of performing a spatio-temporal analysis of urban growth, particularly in assessing the impact of future (land use) scenarios in terms of locations, characteristics and consequences. The use of urban growth modelling and prediction dates back to the 1950s (Torrens, 2006). The early urban growth models were developed by Chapin and Weiss (1962), Tobler (1970) and Nakajima (1977). However, the interest in urban growth modelling faded in the following years and regained momentum only in the 1990s (Cheng, 2003) as a result of improvements in the availability of spatial data and computing ability (Wegener, 1994). Allen and Lu (2003) observed that in empirical studies on urban growth modelling, urban systems

* Corresponding author. Tel.: +91 9998171940; fax: +91 7926302470.
E-mail addresses: talat@cept.ac.in (T. Munshi), mark.zuidgeest@uct.ac.za (M. Zuidgeest), brussel@itc.nl (M. Brussel), maarseveen@itc.nl (M. van Maarseveen).

http://dx.doi.org/10.1016/j.cities.2014.02.007
0264-2751/

have been viewed in several contrasting ways, which have led to several subsequent theories and models (Southworth, 1995; Thapa & Murayama, 2012; Wegener, 1994) of land use change and urban growth, including rule-based models such as Cellular Automata (CA) models (Clarke, Hoppen, & Gaydos, 1996). As observed by Zhao (2011), Tobler's (1979)) work initiated several modifications to the then-existing CA models to make them suitable for simulating and predicting urban growth (Batty, Xie, & Sun, 1999b; Clarke & Gaydos, 1998; O'Sullivan, 2001; White & Engelen, 2000; Wu, 2002; Yeh & Li, 2002; Yen & Li, 2001). These modifications include the coupling of cellular automata with fuzzy logic (Liu, 2012), Markov Chain algorithms (Cheng & Cao, 2011), relative probability and/or regression models (Hu & Lo, 2007; Pinjanowski, Long, Gage, & Cooper, 1997), statistical models (Landis, 2001), and artificial neural network models (Almeida, Gleriani, Castejon, & Soares-Filho, 2008). The logistic regression-based Cellular Automata (CA) model was first proposed by Wu (2002). This hybrid approach, which was also implemented by Paulmans and Van Rompaey (2009), helps to overcome the main limitations of logistic regression, which is the inability to quantify spatial and temporal changes (Arsanjani et al., 2013), and of the CA approach that oversimplifies urban reality and does not provide enough evidence for informed urban planning (Allen & Lu, 2003; Sui, 1998) (Batty, 2000; Sipper, 1997).

CA-based models are simple and allow for dynamic spatial simulation (Torrens & O'Sullivan, 2001). CA, when used in combination with logistic regression models, as done by Wu (2002) and Arsanjani et al. (2013), has the ability to predict a complex urban reality and thus has a potential application in Indian cities. This hybrid framework can form the basic structure of the model; however, when such models are applied in the context of Indian cities, several due considerations will have to be made. The first important consideration deals with adapting the CA framework to suit the realities in the city in the sense that the outputs of the model should provide information at an appropriate scale for urban policy making. The second consideration deals with the input to the logistic regression models. Given the availability of data, it is important to identify the methods that can be used to quantify the inputs that go into the logistic regression model. The third consideration is to build a link with the policy framework, that is, the CA-based framework, and to operationalize the model by using inputs from the logistic regression model and urban policies in such a way that the model is able to establish forward and backward linkages with urban development and control policies. This consideration should ultimately be reflected in the simulation of urban development and growth. Therefore, this study furthers the work conducted on modelling the urban development and growth of Indian cities (Sudhira et al., 2004, Lata et al., 2001, Jat et al., 2008), which are mostly examples of modelling urban sprawl and are not able to model the development of different land uses. Thus, this study showcases how some recent applications, as found in the empirical literature on urban growth modelling, can be adopted for the Indian situation.

The paper is divided into six sections. The next section briefly introduces the study area, followed by a description of the data used in the study. Section 4 details the modelling approach and discusses how the models used in the study are built on the existing approaches and how they are adapted in the context of Indian cities. The results of the application of the modelling approach for the city of Ahmedabad and the validation of the results are presented in Section 5. The paper ends with a discussion of the conclusions in Section 6.

The study area

The study area for this research is the city of Ahmedabad in Gujarat, India. Ahmedabad is a typical rapidly urbanizing second-tier metropolitan city in India with an estimated population of 6.3 million (Census of India., 2011).

The present built-up agglomeration area of Ahmedabad, covering approximately 210 square kilometres, is used as the base-year area. For the future year 2030, a simulation area of approximately 650 square kilometres is considered, which is determined by using a 3-kilometre buffer around the existing built-up area and its surroundings. An analysis of past development trends suggests that the city would not extend beyond the three-kilometre buffer on either side, thus rendering the future area sufficient to accommodate all future urban development until 2031. This area is over three times larger than the present urban development area (part of which is sparsely populated and includes a green belt, which is now available for development). A uniform grid consisting of 100 m. $\times$ 100 m. cells is used in this study to represent the study area. These are marginally larger in size than a singular Town Planning Scheme, a typical urban development tool used for local area land use planning in Ahmedabad (Ballaney & Patel, 2008; Ilhamdaniah, Munshi and Sherif, 2005) and was therefore considered as an appropriate grid size for the simulation exercise (see Figs. 1 and 2).

Data quantification

The data used in this study are listed in Table 1, and the set of independent variables and their indicators quantified using these data sets are listed in Table 2. Each grid cell described above is populated with these three sets of indicators, which are later used to develop the logistics regression model in line with the work by Cheng (2003), Allen and Lu (2003) and Vaz and Nijkamp (2009). The urban structural variables listed in Table 1 are defined based on those by Adolphe (2001), that is, built form, land use intensity, land use heterogeneity and connectivity. In the past, the use of structural variables to predict activity locations has mainly been limited to the presence of land use and distances from urban features such as major nodes and roads, among others. In this study, urban morphology variables such as the entropy index (defining land use balance), dissimilarity index (defining land use mix), density of intersection and road space, and accessibility to activities, among others, are modelled in the logistic regression as independent variables that define the location of land uses. This was conducted because it is envisaged that these stated urban morphology variables affect the location of activities via their relationship with travel behaviour (Wagner and Wegener, 2007), and this relationship is likely to be even stronger in a developing country context such as India (Gakenheimer, 1999). In addition to these variables, the accessibility to similar and other activities is also used as a structural parameter that might influence the location of activities and thereby urban growth (Adolphe, 2001; Landis & Zhang, 1998). The property value is used as a proxy variable for other parameters that are not captured by the built form measure, such as locational characteristics, the age of the development, quality of construction, among others, which have not been included in this research.

Fig. 3 gives a schematic representation of the process that was followed to quantify the data required for this study. To quantify the area developed under each land use activity and related indicators, it is imperative to know the floor area per land use. To quantify these values, data on floor area and the age of buildings are available for each property tax zone, which is a larger unit than the grid cell. Moreover, data on the land use and built-up volume in each grid cell are also available from the land use map and the Digital Elevation Model prepared for the city by Bock (2008). Built-up volumes under each use were computed using spatial overlay methods in GIS. The volumes thus achieved were subse-