Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 00 (2018) 000–000

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research - WCTR 2019 Mumbai 26-31 May 2019

Calibration of a simulation model of land-use dynamics and its transport implications. Case study: Bogotá and the region

Luis A. Guzman[a]*, Javier Rodrigo Peña Bastidas[a], Rafael Cardona[a]

[a]*Department of Civil and Environmental Engeenering, Faculty of Engeenering, Universidad de los Andes, Colombia*

Abstract

Bogotá and its 17 neighboring municipalities make up one of the biggest metropolitan areas in South America. However, in spite of strong functional interactions and influences between cities, there is no official government body at this level in charge of coordinating all authorities to give insights into a wide variety of urban phenomena over the regional and urban land system. This paper presents a cellular automata-based model that allows decision makers to test different growth scenarios based on expected land-use policies, environmental suitability and transport infrastructure.

Bogotá has never had an urban planning tool that integrates its region. CA-based urban simulations can represent a useful approach to understanding the consequences of current planning and independent policies in the region. This tool can help to better inter-territorial and inter-institutional articulation, which through planning and management policies seek a spatial integrated development, through public, private and social actions, with a long-term perspective. This integration is key in the sustainable development of the region.

The model was calibrated to reproduce historic land-use changes from 2007 to 2016 using different methods and indicators. The model was developed in three stages, starting with 1) assembling of the transport infrastructure and land-use data base; 2) the calibration process; and 3) the simulation and analysis of different territory occupation scenarios. The results of the simulations reflect the dynamics of territorial occupation. The experiment indicated that the CA Bogotá model has a strong accuracy, with higher kappa indices, proving the effectiveness of the model.

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY.

Keywords: Cellular automata; policy support; urban dynamics; urban growth; land-use modelling; Bogotá

* Corresponding author.
E-mail address: la.guzman@uniandes.edu.co

Peer-review under responsibility of WORLD CONFERENCE ON TRANSPORT RESEARCH SOCIETY

1. Introduction

The process of urbanization in developing countries has drawn very little attention, primarily due to a lack of detailed spatial data (Sridhar 2007). Latin American and the Caribbean (LAC), is the most urbanized region on the planet with 80% of its population living in cities (UN-Habitat 2012). The economic and demographic growth along with high migration rates from rural areas is resulting in low quality urban expansion of these regions (Guzman 2018). These trends emphasizes the importance of territorial occupation planning, a process which seeks to estimate the impact of policies aimed at controlling the negative effects of urban development including congestion, pollution, and settlement on risky land. Still, urban planning is a highly complex process, especially in large cities for which neighboring municipalities become hubs of employment, business, industry, recreation, and residence.

A quick walk through the cities of the region, reveals a worrying reality: hours lost in traffic, public services with little coverage and low quality, social inequalities and high levels of pollution. Despite being the main drivers of economic growth LAC cities have not generally been characterized by adequate urban planning. In short, with the negative effects on quality of life of the people and the productive potential of the countries, lack of planning has slowed the social and economic development of the region. In this context, urban planners must foresee how travel demand is affected by territorial changes. In the same way, it is necessary to estimate how new developments in transport and land regulations may modify human activities and their associated land-uses. In LAC cities, the processes and conventional practices of urban planning are usually quite complex and a better understanding about urban dynamics, including its drivers and long terms impacts, is essential.

The evolution of cities has attracted interest from urban planners as it is believed to affect economic development, population, and in general, sustainable development (Meijers, Burger 2010). However, in the cities of the Global South, there is a widespread lack of useful planning tools. Currently, there is no proper tools or process that appropriately accounts for important factors involved in urban development: transport infrastructure, land regulation, and suitability. The development and implementation of a support tool is essential, and it is time to apply it in our LAC context to try to understand and assess the influence of the interaction of the location of transport infrastructures and land regulations through land-use models.

Quantitative methods have ben been used for many years in understanding spatial processes of urban growth. Land-use models based on urban cellular automata (CA) have been used due to their simplicity, flexibility, intuition, and principally, their capacity to incorporate the spatial and temporal dimensions of land occupation processes (Santé et al. 2010). These models are popular for simulating urban growth and complex dynamics (White, Engelen 1997, Lau, Kam 2005, Aljoufie et al. 2013, van Vliet et al. 2013, Barredo et al. 2004) similar to those found in real cities using simple rules (García et al. 2013). CA models provide a dynamic modeling environment that can simulate complex changes in land-use, transport, and their interactions and have been widely used to study the spatial process of changes in land-uses over time (Aljoufie et al. 2013, van Vliet et al. 2012, Liu, Phinn 2003). Because CA models operate on a space divided into an array of cells, it is possible to incorporate raster data obtained from aerial photographs and satellite images. In cities of the Global South, this is a key issue because land-use related information is very scarce, of low quality or simply does not exist.

In recent years, CA models have been used to simulate a great variety of urban phenomena, like urban growth, urban sprawl and development of urban form, as well as to evaluate the distribution of population and services, analyze traffic flow and services, and to model spatial competition for location between firms (Lau, Kam 2005). Unfortunately, most of these systems are tailor-made products that require software developers to configure the CA model. In addition to being expensive and requiring a high level team of coders, these products offer little flexibility to the planner.

The calibration of CA land-use models is challenging and most traditional methods for the calibration of urban CA models are based on trial and error (García et al. 2013). However, the calibration process is a crucial step of the model design and validation process. Good calibration parameters ensure that the model can make reasonable predictions about current and future land-use changes in real cities (Al-Ahmadi et al. 2009, Wang et al. 2011, He et al. 2008). Especially challenging, is the determination of optimal parameters for several factors of a CA land-use model (Stevens, Dragicevic & Rothley 2007). Because the CA models available in the market are generic, parameters well-adjusted to a particular geographic area, do not necessarily function with the same quality in other locations (Silva, Clarke 2002). This indicates the need for a parameter adjustment process to improve results.

Although it has slowed in recent years, Bogotá has experienced rapid urban growth that has caused complex planning challenges for local authorities. The neighboring municipalities of Bogotá have experienced rapid population growth, with great space variability and a complex urban dynamic. The population of the region went from 7,942,843 inhabitants in 2005 to 9,396,697 in 2016 and the urban area has also expanded dramatically, mainly in neighboring municipalities. This expansion has coincided with the use of conventional planning practices. Additionally, the lack of an appropriate and coordinated policy between the local, regional and national authorities, has led to an almost accidental regional development.

This paper presents the development, calibration and application of a generic integrated spatial decision support system, called Metronamica for Bogotá and 17 neighboring municipalities in a spatial two-dimensional grid, where urban growth is high and fast. The model was calibrated to reproduce historic land-use changes from 2007 to 2016 using different methodologies and indicators. Previous experiences with this type of models in Bogotá have shown some limitations, mainly due to lack of information and calibration process adjustment (Escobar, Páez 2018). The new model and calibration process considers the complete Bogotá region area, and the socioeconomic behaviors unique to each urban area, therefore directly associating the results with the site of application.

2. The urbanization in Bogotá

In Bogotá, as in most Colombian cities, there has historically been a high unsatisfied demand for housing which, in response, has encouraged the low-income population to occupy mostly un-planned and informal urban settlements. As a consequence, several informal neighborhoods with poor urban living conditions have emerged on the outskirts of the city and in neighboring municipalities, particularly in Soacha (Oviedo, Dávila 2016).

As of 2017, Bogotá has approximately 8.1 million people and an urban area of around 400 km^2. Additionally, Bogotá is surrounded by cities that have a close relationship with the capital making it, in practice, a large, although not officially constituted, metropolitan area (Guzman, Oviedo & Bocarejo 2017). These municipalities, which include Zipaquirá, Gachancipá, Tabio, Tocancipá, Cajicá, Chía, Sopó, Tenjo, Cota, La Calera, Funza, Madrid, Mosquera, Facatativá. Bojacá, Soacha, and Sibaté, (see Fig. 1), consist of about 1.4 million inhabitants in 2,272 km^2, and make around 2.4 million trips per day.

These municipalities make up a metropolitan region due to their functional interactions, however there is no formal government body at this level and each of the contiguous municipalities still makes autonomous decisions (Guzman, Oviedo & Bocarejo 2017). Historically this has led to a disorganized growth process, where, as population increases, territory develops according to available land and proximity to centers of employment, recreation, and cultural venues.

Fig. 1. Study area

Current transport and land-use planning practices in Bogotá region cannot keep up with the rapid urban growth, particularly in the surrounding municipalities. This region has grown not only through a process of economic development and industrialization, but also due to internal violence. Terrorist violence has given way to massive waves of migration from rural areas, where migrants have settled in urban areas that were not planned or conducive to urbanization. This has caused serious problems with congestion, exclusion and uncontrolled territory occupation. Between 1938 and 1992, two thirds of the country's population was rural and a third was urban. Today, it is the exact opposite and Bogotá, is no stranger to this phenomenon. In the last 30 years, the population of Bogotá has grown 76%, going from 4.6 to 8.1 million inhabitants, and population in the city's surrounding areas has grown around 84%.

The residential land of the region is divided into socioeconomic strata, SES (in Spanish, *estratos*) from one (the poorest) to six (the wealthiest). In the lower, SES one to three, people receive subsidies for essential utilities (water, gas, electricity). In the highest, SES five and six, residents pay for those subsidies with bills higher than their consumption. The classification of housing by strata was done based on dwelling characteristics and the urban environment. This unique model was devised in the mid-nineties, in a country that, at the time had poverty rates close to 40%. Now, the spirit of the model has been perverted, as the classification has encouraged spatial segregation.

As SES is a good proxy with the average household income, for the purposes of this study, residential areas were divided in 3 categories based on SES: low, medium and high. Economic land-uses were divided into 4 categories: commercial, services, industrial and mixed. For more details about this classification, see section 3.1. Fig. 2 shows the location of this different land-uses categories over the study area.

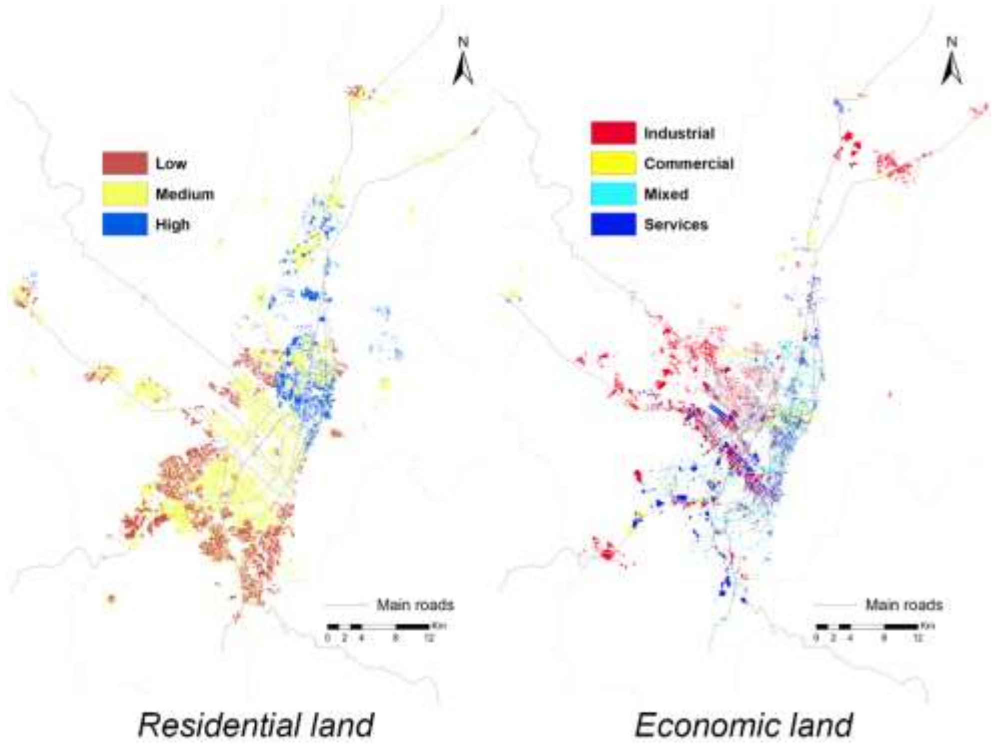

Fig. 2. Location of land-use by category

As these maps show, land occupation trends in Bogotá indicate a concentration of low-income settlements in the south and west of the city, with high-income areas located in the north and center. Lack of available land for expansion and high land prices have forced the low-income population out to peripheral areas. As available land within Bogotá's boundaries runs out and cost increases for the little that remains, urban trends suggest that most of the region's urban growth occurs outside the municipality of Bogotá and that most of inhabitants of these outer municipalities are either low or middle-income.

During the ten-year period 2007-2016, the measured built-up area in Bogotá region grew by 89.1 km^2 (+17.3%). It is noticeable that the high-income residential land-use classes showed most increase in this period, with a growth of 8.3 km^2 (+44.5%), while the low-income residential land-use, was the one that less grew (+3.5%). This may be because the high-income low-density residential land-use growing in the north (Fig. 2) and the re-densification of low-income population is concentrated in the south and west of the region. Those are one of the main characteristics of Bogotá region in the last few years.

It seems that the locations of residential areas, through this informal growth process, have segregated the residents of the region. However, Bogotá's social segregation has always been clearly visible, with the north of the city occupied primarily by wealthy people and the south by the poor (Guzman, Oviedo & Bocarejo 2017). Regarding job location, the most valuable commercial and service uses are located in a large central core along major road corridors as well as the northern part of the city center (the richest zones).

3. The integrated spatial decision support system Metronamica

This study applied and calibrate a CA-based model utilized in Metronamica (Aljoufie 2014, van Vliet et al. 2012). The model is composed of a grid of cells, where each cell represents a land-use category. The state of adjacent cells, and hence the land-use in the neighborhood of a location, is input to the transition rules (van Vliet et al. 2012). The state change potential of each cell is calculated in discrete steps while following a set of transition rules that depend on geographic features, which include environmental suitability, zoning, and accessibility. Here, environmental suitability is defined as the effect of the physical characteristics of the land for the generation of a future use and contemplates, the presence of flood or landslide risk. Zoning represents external restrictions due to political decisions. Accessibility measures the effect of proximity to transport infrastructure with certain land-uses activities.

The Metronamica model has three categories of land-use: Function, Vacant, and Featured. Active (Function) uses are defined as those to which location is assigned depending on the potential derived from the transition rules. Passive (Vacant) uses, only change as changes occur in active uses. Finally, featured land-uses are the classes not programmed to change in the simulation period. For each moment of time, active land-uses are located in the places with the greatest potential for development. The potential of each cell is given by the following formula:

$$P_{k,i} = r_{k,i} * A_{k,i} * S_{k,i} * Z_{k,i} * N_{k,i}$$

Where, $P_{k,i}$ is the potential for development of a land-use k in the cell i, $r_{k,i}$ is a random disturbance t $A_{k,i}$ term, is the accessibility for each of land-use category, $S_{k,i}$ is the physical suitability for the development of a certain land-use, and $N_{k,i}$ represents the neighborhood interactions. A detailed explanation of the Metronamica model can be found in van Vliet et al., (2012).

These variables are used to calculate “transition potential”, a measure of the probability of a land-use change, which indicate that land-use change is influenced not only by current use, but also by surrounding land-uses and transport infrastructure (Hagoort, Geertman & Ottens 2008). In other words, the development of a future land-use is determined by the attraction/repulsion effects exerted by neighboring uses at a specific point (van Vliet et al. 2013). The way chosen to represent the relationship between groups of neighboring cells is through the transition rules (Verburg, de Nijs et al. 2004). These rules, known as enrichment factor, characterizes the occurrence (enrichment) of a land-use category in the neighborhood of a location relative to the occurrence of this land use category in the study area as a whole. These can be obtained from over or under representation curves that demonstrate the density of a land-use in the surroundings of another use or transport infrastructure. In this study, we used a modified version of the enrichment factor (*F*) as was presented in van Vliet et al., (2013).

As seen in Fig. 3, the CA simulation process is divided into three stages. First it is necessary to collect information for a study area at two different points of time, usually separated by periods of several years. These points are later referred to as the beginning and end of the calibration period. In the second stage, transition rules must be established between land-uses and transport infrastructure factors, in order to calibrate the model and achieve simulation patterns similar to reality.

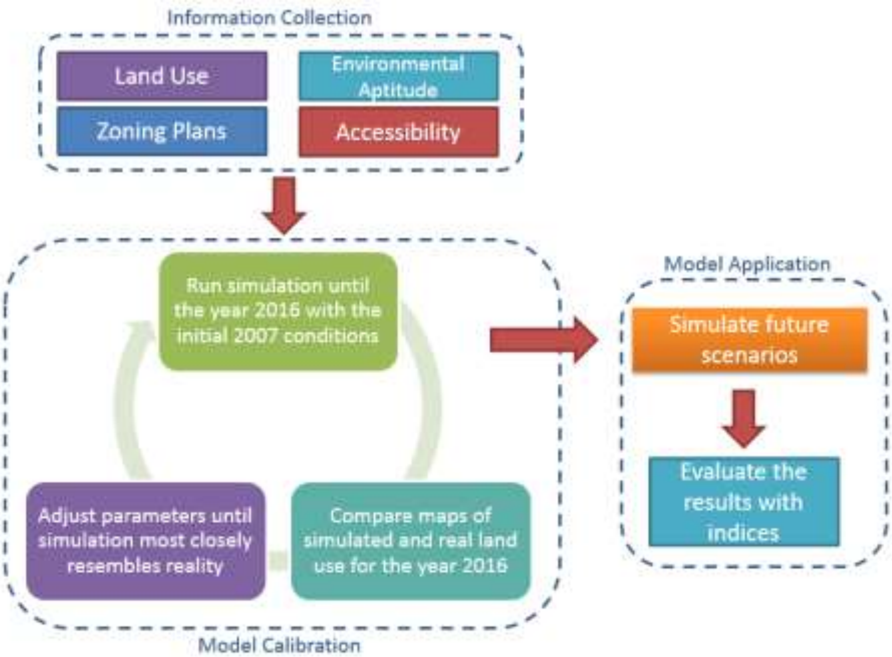

Fig. 3. Application of the Bogotá CA model

The third stage is the generation of results for each scenario using the calibrated model. At this stage there is no real map for comparison, but the model uses the parameters calibrated in stage two to ensure the accuracy of results. In this step, a simulation end year is determined, and a set of possible scenarios are designed to incorporate the introduction of new transport infrastructure, land demand changes by land-use category, and zoning regulations.

3.1. Data source and model preparation

The cell size for the study area was 60x60m, and the calibration process considered 2007 (143,099 cells) and 2016 (167,858 cells) as the base and final years, respectively. The year 2050 was used as the evaluation year for the different scenarios. To meet the data requirements, multiple sets of data were needed. Complete land-use information for different years was only available for the city of Bogotá, so in municipalities where there was no enough information, satellite image analysis and official documents was used.

The satellite image analysis procedure begins with the obtainment of Landsat 7 etm+slc-off satellite images (USGS 2017) for the year of interest. Images from the years 2007 and 2016 with the least amount of cloud cover and the best spatial resolution were chosen and processed to correct the parallel lines of error generated by the slc-off sensor. These two images are then classified through an automatic real image interpretation process. The methodology used is known as “supervised classification” and utilizes the “Image classification” and “Maximum likelihood classification” tools of ArcGIS. Significant samples of each class are collected, and the software executed on the satellite photographs to look for pixels with the same statistical criteria as the selected samples. The final result of the image processing and supervised classification is a raster for each period, classified by water bodies, available land, visible industries and populated centers. The land identified as occupied is further assigned a land-use compatible with Metronamica in accordance with available information.

To input the accessibility parameters, additional geographic information had to be collected, such as the road and public transport network, Transmilenio stations (BRT system), and the bicycle network. This information was taken from various official sources. Also was used information on landslide, forest fires, and flooding phenomena for the environmental suitability layers. Finally, zoning information came from the Territorial Ordinance Plans of the municipalities.

Land-use categories were classified into the following classes: 0) Available, 1) Low-income residential, 2) Medium-income residential 3) High-income residential, 4) Industrial, 5) Commercial, 6) Mixed (residential land-use between 40 and 60%), 7) Services, 8) Parks and country clubs, 9) Superficial water, 10) Institutional, 11) Airport, 12) Main roads, 13) Landfill, and 14) Mining. Of these, uses 1 through 7 were considered active uses, use 0 was considered passive, and uses 8 through 14 were classified as fixed use (feature).

3.2. Calibration procedure

The most crucial issue of the CA models is the calibrating process. On this depends its successful implementation because the quality of urban CA models contain a number of parameters that must be calibrated (Straatman, White & Engelen 2004). Estimation of the level of concordance between maps was realized with Kappa and Fuzzy Kappa indices.

The model was calibrated by running a simulation for the period 2007 to 2016. The simulation was initiated using the historical datasets for the year 2007, both in Bogotá and municipalities, in order to test the simulation results using the reference datasets. It is necessary to calibrate the CA model with the help of reliable historic data before using the model to predict future land-use patterns. The calibration process is centered on the adjustment of model parameters to obtain a simulated map as similar as possible to the 2016 real map. This is an iterative use trial and error process (Fig. 3) in which after each parameter adjustment the differences.

As one of the cell-to-cell evaluation methods, the Kappa index (*KI*) for both quantity and location are calculated to estimate the actual degree of agreement between two maps (Congalton 1991). The *KI* is based on a direct cell by cell comparison of the maps and vary between -1 and +1. *KI* is more reliable than simple percent agreement calculation because it considers chance agreement (Zheng et al. 2015). The equation for *KI* is:

$$KI = \frac{P - E}{1 - E}$$

Where P is the relative observed agreement between two maps and E is the expected agreement due to chance. Positive values indicate greater convergence than can be expected when comparing a map generated with random parameters, and negative values demonstrate that a completely random map would have better agreement than the simulation of the model. If two maps are in total agreement, *KI*=1. This indicator has been utilized before to measure the agreement between the results of a simulation and a real map for the end of the simulation period (Berberoğlu, Akın & Clarke 2016, He et al. 2008). A negative aspect of using this indicator to evaluate the model is that even small displacements in location are classified as incorrect. For example, a residential area located in the simulation just one cell away from the actual location would be, in the perspective of the human modeler, almost correct.

The Fuzzy Kappa index (*FKI*) was developed to improve the *KI* limitations (Hagen-Zanker 2009). Although also based on a cell by cell evaluation of maps, the *FKI* compares the set of surrounding cells to express similarity with a value between 0 (completely different) and 1 (completely identical). This method was developed with consideration of the main purpose of these indices, the calibration of high resolution territorial occupation models (Hagen-Zanker 2009, Hagen 2003). For the *FKI* statistic the equation is the same than for *KI*, but the mean agreement P considers fuzziness of categories and fuzziness of location. Fuzziness of categories is expressed in a categorical similarity matrix M expressing the similarity between category i in the map A and category j in the map B. Fuzziness of location is introduced through a distance decay function that specifies the degree to which a cell belongs to the categories found in its proximity (Hagen-Zanker 2009).

The Kappa indices were calculated with the Map Comparison Kit 3 software developed by RIKS. Calculation of the *FKI* requires the configuration of parameters such as the radius of the cell group to evaluate and the decay function that expresses the loss of importance as you move away from it. This study used a 4 cell radius and a linear decay function with a parameter of 0.5.

Bearing it in mind, different methods were utilized in calibration of the model to improve the quality of both the input data and results of the model. The first correction method was visual interpretation, in which changes that occurred during the years of study were visually identified, giving a clear idea of how land-uses were distributed throughout the period (Engelen et al. 1997). Some studies utilize this methodology as a first step to identify the most significant differences between the simulated and real data (Aljoufie et al. 2013). Visual calibration requires modelers familiar with the territory because it depends on the human eye to detect possible errors in the input data.

Subsequently, contingency tables were prepared to compare the overall performance of the model. These tables calculate the total number of cells that changed from a certain category (current use) to another in the data set as determined by the simulation. The contingency tables are considered a useful complement to visual interpretation as they identify errors that escape the perception of the expert eye by highlighting the volume and type of changes that occurred during the simulation period.

The values obtained from Kappa indices can vary enormously depending on the quality of the information collected for the land-uses observed. If Kappa indices values are above 0.8 represents strong agreement or accuracy between two maps. If the values are between 0.6 and 0.8 represents substantial agreement. Values ranging from 0.4 to 0.6 means moderate agreement; and below 0.4, the agreement is poor (Landis, Koch 1977).

3.3. Scenarios

Eight (8) development scenarios were created with different levels of complexity. In scenarios 1 and 2, Bogotá and the municipalities maintain their current expansion perimeter. Additionally, restrictions are applied to the land occupation with potential for agricultural development. The main difference between these two scenarios is the introduction of new public transport infrastructures and the construction of regional highways.

In scenario 3, areas in the north of Bogotá are enabled for residential uses and the municipalities implement their proposals of urban and suburban expansion plans. Then, in scenario 4, Bogotá regulates the occupation of the southern transition edge and eliminates the restriction on residential and commercial uses to occupy land with agricultural potential. For scenarios 5 and 6, a simulation of regional rail transport is included. Accordingly, Transport Oriented Development (TOD) strategies are implemented for the developable land with a buffer around the new stations. Finally, in scenarios 7 and 8, the second stage of the El Dorado Airport is enabled, residential land-uses in the peripheries of western municipalities are permitted, and commercial occupation strategies in the area of influence of the new airport are generated.

The purpose of each scenario is to evaluate the impact of several combinations of territorial planning policies and transport infrastructure projects on the territorial occupation of Bogotá region for the year 2050. The scenarios were developed from the formal and informal regional plans and projections discussed in recent years. The results serve as support to the decision making and policy formulation.

For space limitations, the results for two scenarios with the greatest differences in zoning and transport infrastructure are discussed in the following sections and a summary of their definitions are presented below.

- Scenario 1 – BAU: Bogotá and the municipalities maintain their current expansion perimeter, in addition restrictions are applied to prevent the occupation of land allocated to agricultural capacity. In terms of infrastructure, there is no introduction of new transport systems or the construction of regional roads or any important infrastructure.

- Scenario 2 – Infrastructure: Bogotá and the municipalities maintain the same land expansion restrictions. This scenario includes new BRT lines (Transmilenio phases 4 and 5), a proposal of regional road network, the north city road network and public transport interchange hubs.
- Scenario 3 – Development areas: enables residential area development in the north and west of Bogotá, while most of the municipalities enable both urban and suburban land expansion. There is no new infrastructure projects.
- Scenario 4 – Southern transition edge: Bogotá regulates the occupation of the southern transition edge and includes a new transport mode (cable car phases 1 and 2). Additionally, the restrictions on the occupation of agricultural land are eliminated.
- Scenario 5 – Regional rail transport system: The first Metro line is included in Bogotá, and in the municipalities the regional train is developed. The allocation of land for residential and commercial development is implemented under TOD strategies with a buffer around the stations.
- Scenario 6 – El Dorado Airport II: The region enables the second stage of the El Dorado Airport as a new infrastructure facility. This doesn't include a land occupation pattern around the airport, so it is considered an isolated infrastructure area.
- Scenario 7 – Mosquera development: new residential and economic land-use are enabled for the city of Mosquera and transport infrastructure such as the road network and the expansion of the Metro to this city is included.
- Scenario 8 – All changes: Bogotá enables residential area development in the north of the city while the municipalities enable both urban and suburban land expansion. Additionally the restrictions on the occupation of agricultural capacity are eliminated. In terms of infrastructure, this scenario includes public transport systems (new BRT lines and the first Metro line) as well as a regional train and the new airport. The allocation of land for development is implemented under TOD strategies with a buffer around the stations of new modes of transport in the region.

Figure 3 shows all the transport infrastructure improvements include private and public network that were taken into account for this model as the general infrastructure plan for the study area. All these projects are contemplated within the new Land-Use Master Plan. To complement these infrastructure scenarios in which either improved roads or a public transport were constructed, land regulation options were also modeled.

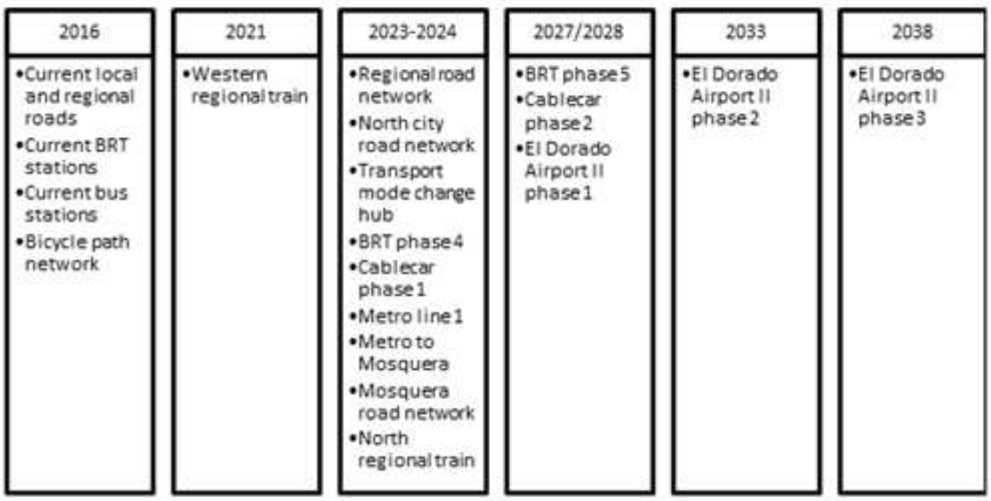

Fig. 4. New infrastructure project for the Bogotá region

Finally, as an external variable, the Urban Planning Office of Bogotá projected the land-use demands for the year 2050, as is shown in Table 1. These projections considered population growth, residential density, and GDP to estimate the demand area by land-use category.

Table 1. Demand by land use type (Km^2)

Land Use	**Current 2016**	**Demand 2050**
Low Residential	66.8	207.7
Medium Residential	118.4	235.3
High Residential	26.9	93.4
Industrial	51.9	104.0
Commercial	11.2	22.4
Mixed	21.0	42.1
Services	27.2	54.5
Total	**323.4**	**759.4**

In order to present the results clearer, the intermediate scenarios has been omitted in the results section. Thus, only scenarios 1 and 8 are shown and analyzed.

4. Results and discussion

The Metronamica CA model was calibrate the land-use patterns between 2007 and 2016 in Bogotá region and to simulate the spatial pattern of urban growth to 2050. The current land-use map in 2007 was derived from the land-use database provided by Bogotá city (cadastral data) and other sources. These empirical data were applied to develop and calibrate the transition rule of the Bogotá-CA model. Given the importance of the data quality and model calibration, results for these processes are presented before the discussion of the tool application in the simulation of scenarios.

Firstly, errors found during the visual analysis stage were mostly due to cloudiness and other issues with the satellite images used for the construction of the maps. For example, water bodies that did not exist in 2007 appeared in 2016 because the images were taken during different seasons. In order to reduce further the errors related to the quality of the input data, contingency tables were implemented. These tables calculate the total number of cells that changed from a certain land-use category (current use) to another in the data set as determined by the simulation. The contingency tables are considered a great complement to visual interpretation as they identify errors that escape the perception of the expert eye by highlighting the volume and type of changes that occurred during the simulation period. Table 2 shows the last contingency table constructed which was the basis for calculating the kappa indice.

Table 2. Final contingency table

Land-Use 2007 (Units: 60x60m cells)	**Land-Use 2016** (Units: 60x60m cells)														
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	**615,495**	3,852	11,104	2,690	6,673	1,121	864	8		1,526	858		1,071		
2	950	**13,034**	2,245	14	460	238	316	20		300	260		84		
3		536	**16,698**	271	937	317	589	33		512	187		386		
4	515	12	145	**3,978**	107	104	124	11		71	43		53		
5	1,871	507	828	69	**4,958**	258	238	36		1,603	379		229		
6	189	69	193	53	148	**713**	44			148	26		38		
7	196	195	519	138	259	99	**3,227**	7		214	181		98		
8								**9,486**							
9	259				11				**19,321**				115		

10	552	83	659	143	708	206	160			**2,834**	314		144		
11	482	265	505	104	157	50	279	213		342	**4,155**		100		
12												**1,550**			
13													**30,709**		
14														**1,213**	
15													38		**6,657**

Note: Number represents: 1. Available, 2. Low Residential, 3. Medium Residential, 4. High Residential, 5. Commercial, 6. Industrial, 7. Mixed, 8. Parks, 9. Water bodies, 10. Services, 11. Institutional, 12. Airport, 13. Highways, 14. Landfill, 15. Mining.

It is possible to see and correct changes in land-uses that should not occur. For example, Table 2 shows that 11,104 cells of available land (40 km^2) were converted into middle-income residential land. It also shows that 2,245 cells (8.1 km^2) of low-income residential land were converted to mid-income residential land. Of course, some errors still exist, but stabilize at a low level.

4.1. Neighborhood rules

One of the most important factors to be calibrated is the neighborhood attraction and repulsion effects between land-uses. The factors were calibrated in order to minimize the differences between the simulated land-use map for 2007 and the 2016 land-use map. The calibration of these parameters for any pair of land-use categories is based on a rational evaluation of the actual land-use patterns in the study area and their historical evolution. The process used was a manual calibration. The enrichment factors show a distance decay effect, as the values decrease with increasing distance. Commercial land-use does not appear, but it is similar to that of services. The curves confirm that urban land-uses have the strongest neighborhood effect, mainly the residential ones. Fig. 5 shows the enrichment factor of land-uses in the neighborhood of residential and economic land between 2007 and 2016 as a function of the distance (in cells) to the locations of land-use changes. The results show that, at a distance zero there is an attraction effect from one land-use to another. The values at distance greater than zero indicate the over (positive values) or under (negative values) representation of land-uses in the neighborhood of locations of land-use changes.

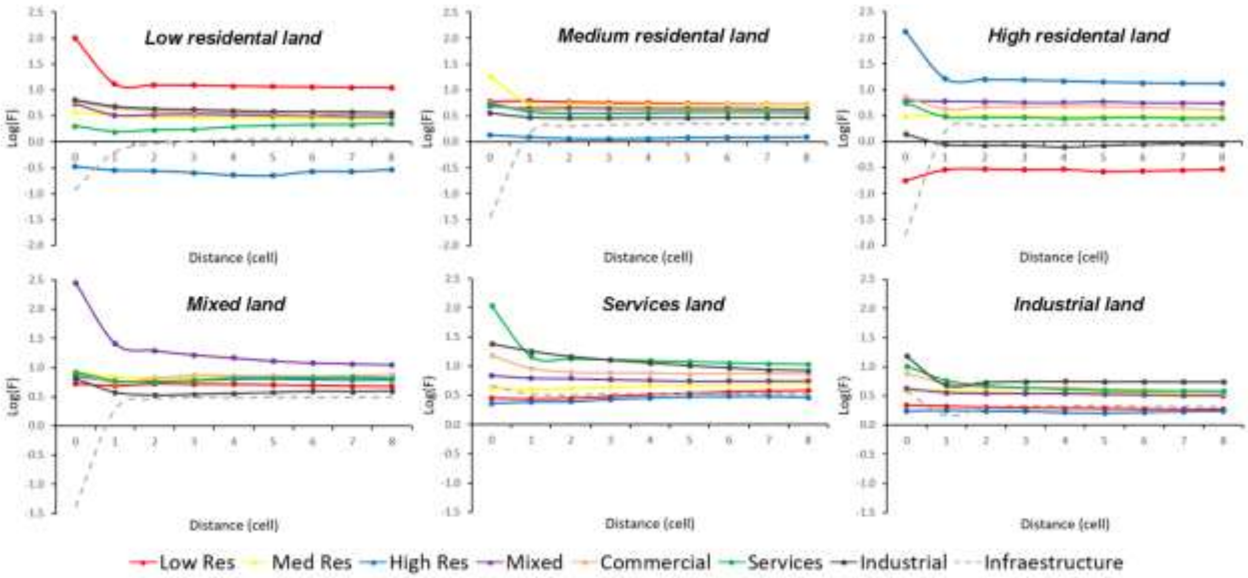

Fig. 5. Neighborhood rules between 2007 and 2016

Due to the great mix of land-uses in Bogotá, the curves show that new urban land-uses tend to be located next to other urban land-uses. For example, high-residential land-uses are underrepresented in the neighborhood of new low-residential land-uses, and low-residential land-uses are also underrepresented in the neighborhood of new high-residential areas. This implies that wealthy people do not want to live near the poor. It could also be said that the poor cannot live near the rich, because of the high land value and housing costs. It is also observed that no inhabitant of Bogotá wants to live in front of a main road (first cell), being the wealthy ones who have a bigger repulsion.

The calibration did not yield such sign changes between land-uses as there was no indication from data that these were appropriate. Therefore, the manual calibration for neighborhood rules is probably more realistic in developing cities, where urban regulations are not fully met. In addition, the spatial segregation includes informal settlements, with a large land property division, a high mix of uses and very high densities (Guzman, Bocarejo 2017). In those zones, the available data related with the land-uses may not be very reliable. Finally, the results confirm the existence of a neighborhood effect, mainly for the residential land-uses. This suggests that it is appropriate to include calibrated neighborhood rules to simulate urban land-use changes in the Bogotá region.

4.2. Model validation

A land-use changes were validated by comparing model results of an observed period with the actual changes of land use: simulated 2016 vs observed 2016 (Verburg, Schot et al. 2004). As mentioned before, the simulated land-use pattern of 2016 was achieved on the basis of land utilization map in 2007. Fig. 6 shows the simulated land utilization map of 2016 with the Kappa indices results. To achieve this, several iterations were made. Each iteration resulted in a better level of agreement.

The overall accuracy for the *KI* is 79.6%. This shows the percent of the number of cells simulated correctly to the total number of cells. For the *FKI* the accuracy is 85.3%. This value representing a substantial agreement, which indicates the effectiveness of the model's setting to conduct the future simulation. Although the *KI* never rose above 0.80, the value 0.796 was considered sufficient to be adopted. With these results it was concluded the model was valid to continue with the next stage in the simulation process.

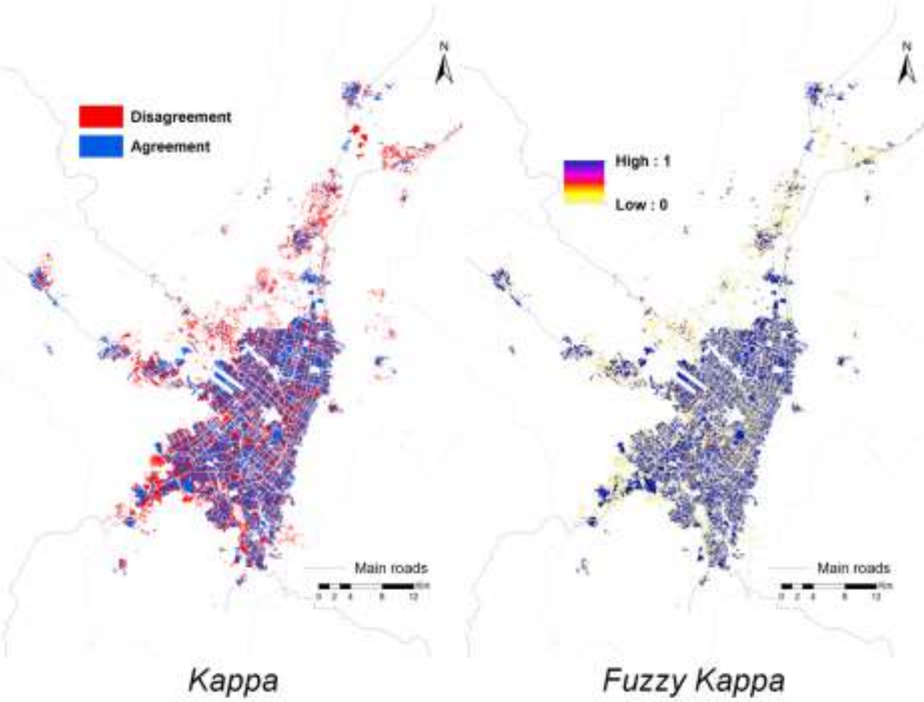

Fig. 6. Maps of Kappa indices result

4.3. Scenario simulation

Based on the above calibration and validation, results were obtained for each of the proposed scenarios through the modification of inputs entered into the calibrated model, and a set of numerical indicators where calculated in order to quantify the changes and differences between scenarios. The scenarios simulated to 2050 are shown in Fig. 7. Notice that in order to make the maps clearer, the large number of fifteen land-use categories have been grouped into residential (low, medium and high), economic urban built-up areas (in grey) and other categories (feature uses).

The resulting maps demonstrate the different patterns of development around the northwestern border of study area between with and without land regulations. This is due to zoning policies. For example, in scenario 8, urban expansion of the northwestern area is permitted. Industries evidently choose to locate as close as possible to the urban center, generating an industrial cluster and broader a conflict area just on the periphery, around north-west main road.

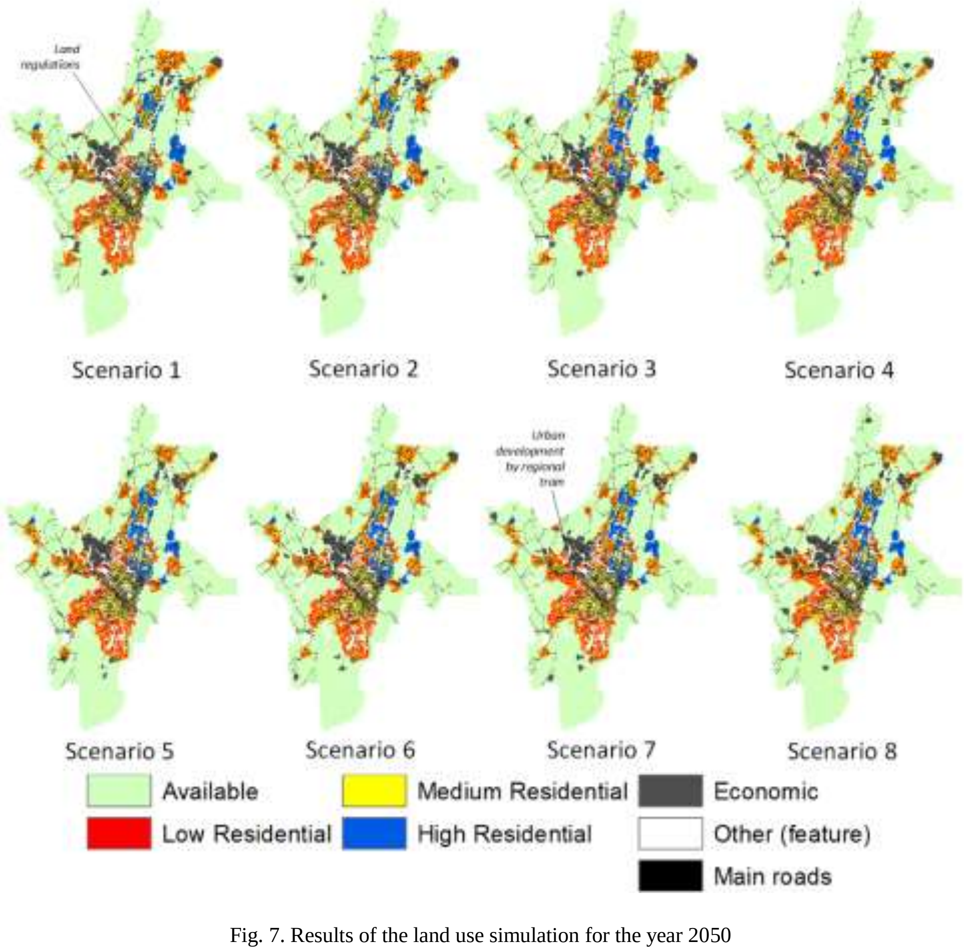

Fig. 7. Results of the land use simulation for the year 2050

A second analysis of this maps, indicates that with this enabling of expansion into the north zone, there is a relocation of the lower and middle residential land-uses, mainly populating the road that connects Bogotá with the municipality of Cota (west). This phenomenon, in turn, affects the location of the high residential uses. In scenario 1 the high SES residential areas are mostly grouped on the western side of the northern highway but for scenario 8 they tend to spread to other areas outside of Bogotá. Potentially, as the amount of available land increases, the upper strata prefer not to compete for exclusive areas north of the capital and instead looks for quieter sites in the surrounding municipalities -

both groups then taking advantage of greater accessibility supplied by the introduction of new transport infrastructure in the region and the relocation of industries.

It must be noted that the worth and validity of the simulation results are still uncertain, as land-use change is inherently linked to human decisions, which hardly have a deterministic behavior (Brown et al. 2005). Additionally, the changes are caused by the combination of physical and socioeconomic factors which influence each other, implying that many probabilities of the occurrence of an event or development of a certain type of land-use are subject to past events of multiple actors. In economics this is known as "path dependence", and indicates that this evaluation is not simple or direct (Batty, Torrens 2005).

5. Concluding remarks

This research proposed a CA model to explore land-use change mechanism in Bogotá region. Never before has a tool to help decision-making in urban planning been developed that covers Bogotá and 17 its surrounding municipalities. The cellular automata based model was successfully applied to the Bogotá region, contributing to the development of territorial planning tools in the Global South cities context. A comprehensive understanding of land-use change mechanisms and directions is crucial for land-use planning decisions and the sustainability of the region. This tool will be very helpful in bringing the differences of local authorities closer together and evaluating joint projects. The results of this research can help to better inter-territorial and inter-institutional articulation, which through planning and management policies seek a spatial integrated development, through public, private and social actions, with a long-term perspective.

It has been shown that CA urban models can reasonably represent the future of developing cities (Barredo et al. 2004). This study has introduced a new dynamic land-use planning approach for Bogotá region using a Metronamica model. Considering both spatial and temporal dimensions, a land-use change model was developed and validated to simulate four land-use scenarios in 2050 in the Bogotá region. The consequences of different policy interventions on land-use and transport were depicted by the model over space and time through different scenarios.

The calibration of CA models is a complex process. There is still discussion about which tools and results should be used to estimate the accuracy of land-use pattern represented in these models. In the model developed for Bogotá region, an iterative method of manual calibration was used and model validity determined by Kappa and Fuzzy Kappa indices. According to the values reported in literature, both indices indicated a strong accuracy model.

The model was used to simulate a set of 8 scenarios of interest to the Bogotá decision makers. These scenarios range from a BAU scenario with highly restricted land development and low investment in transport infrastructure, to a final scenario that enables large areas in the region for new urban development and introduces new transport infrastructures, which are not currently available, like a cable car and rail systems.

The simulations of the 8 scenarios are in accordance with the expectations of city authorities, in that they reflect the dynamics of territorial occupation, particularly when comparing policies that effect the zoning of the north zone of Bogotá. With the enabling of expansion into the north zone, both economic and the middle and lower residential uses start to occupy, since they offer proximity to employment centers and connectivity to public transport. In addition, the results also show that available land has a higher probability of being utilized when a new development is required, particularly in the region. To improve sustainability and diminish the negative effects of urban expansion, it is necessary to consider urban renewal policies. This is one of the elements to improve in the model. Due to data limitations and quality and the complexity of the Bogotá region land-use patterns, the analysis results may not be enough, and more factors and scenarios should be included into the model to better facilitate the decision-making process. This is a first step to try to better understand the urban land system of this region so important for the country, which can serve as an example for other urban systems in the region.

Bogotá CA modelling appears to be a useful tool for urban planners. It can be used to study different planning policies, to measure the spatial consequences of local and regional decisions, and, in particular, future land-use patterns. The results of this investigation provide several directions for future urban dynamics in Bogotá region. First, given the promising results of calibration and validation, a solid and complete basis is left for the exploration of future urban dynamics in the study area. Second, taking into account the rapid growth of Bogotá's neighboring municipalities and the rapid growth of the vehicle fleet (automobiles and motorcycles), this tool is both useful and necessary. However, further research is needed.

Acknowledgements

The authors express their gratitude to Urban Planning Office of Bogotá city (Secretaria Distrital de Planeacion de Bogotá) for funding this research.

References

Al-Ahmadi, K., See, L., Heppenstall, A. and Hogg, J. (2009) 'Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia', *Ecological Complexity*, 6(2), pp. 80-101.

Aljoufie, M. (2014) 'Toward integrated land use and transport planning in fast-growing cities: The case of Jeddah, Saudi Arabia', *Habitat International*, 41, pp. 205-215.

Aljoufie, M., Zuidgeest, M., Brussel, M., van Vliet, J. and van Maarseveen, M. (2013) 'A cellular automata-based land use and transport interaction model applied to Jeddah, Saudi Arabia', *Landscape and Urban Planning*, 112, pp. 89-99.

Barredo, J.I., Demicheli, L., Lavalle, C., Kasanko, M. and McCormick, N. (2004) 'Modelling Future Urban Scenarios in Developing Countries: An Application Case Study in Lagos, Nigeria', *Environ Plann B Plann Des*, 31(1), pp. 65-84.

Batty, M. and Torrens, P.M. (2005) 'Modelling and prediction in a complex world', *Futures*, 37(7), pp. 745-766.

Berberoğlu, S., Akın, A. and Clarke, K.C. (2016) 'Cellular automata modeling approaches to forecast urban growth for adana, Turkey: A comparative approach', *Landscape and Urban Planning*, 153, pp. 11-27.

Brown, D.G., Page, S., Riolo, R., Zellner, M. and Rand, W. (2005) 'Path dependence and the validation of agent-based spatial models of land use', *International Journal of Geographical Information Science*, 19(2), pp. 153-174.

Congalton, R.G. (1991) 'A review of assessing the accuracy of classifications of remotely sensed data', *Remote Sensing of Environment*, 37(1), pp. 35-46.

Engelen, G., White, R., Uljee, I. and Timmermans, H.P.J. (1997) *Decision Support Systems in Urban Planning*, .

Escobar, F. and Páez, D. (2018) 'Urban Transportation Scenarios in a LUCC Model: A Case Study in Bogota, Colombia', in Camacho Olmedo, M.T., Paegelow, M., Mas, J. and Escobar, F. (eds) *Geomatic Approaches for Modeling Land Change Scenarios*, Springer International Publishing, Cham.

García, A.M., Santé, I., Boullón, M. and Crecente, R. (2013) 'Calibration of an urban cellular automaton model by using statistical techniques and a genetic algorithm. Application to a small urban settlement of NW Spain', *International Journal of Geographical Information Science*, 27(8), pp. 1593-1611.

Guzman, L.A. (2018) 'A strategic and dynamic land-use transport interaction model for Bogotá and its region', *Transportmetrica B: Transport Dynamics*, pp. 1-19.

Guzman, L.A. and Bocarejo, J.P. (2017) 'Urban form and spatial urban equity in Bogota, Colombia', *Transportation Research Procedia*, 25C, pp. 4495-4510.

Guzman, L.A., Oviedo, D. and Bocarejo, J.P. (2017) 'City profile: The Bogotá Metropolitan Area that never was', *Cities*, 60, Part A, pp. 202-215.

Hagen, A. (2003) 'Fuzzy set approach to assessing similarity of categorical maps', *International Journal of Geographical Information Science*, 17(3), pp. 235-249.

Hagen-Zanker, A. (2009) 'An improved Fuzzy Kappa statistic that accounts for spatial autocorrelation', *International Journal of Geographical Information Science*, 23(1), pp. 61-73.

Hagoort, M., Geertman, S. and Ottens, H. (2008) 'Spatial externalities, neighbourhood rules and CA land-use modelling', *The Annals of Regional Science*, 42(1), pp. 39-56.

He, C., Okada, N., Zhang, Q., Shi, P. and Li, J. (2008) 'Modelling dynamic urban expansion processes incorporating a potential model with cellular automata', *Landscape and Urban Planning*, 86(1), pp. 79-91.

Landis, J.R. and Koch, G.G. (1977) 'The Measurement of Observer Agreement for Categorical Data', *Biometrics*, 33(1), pp. 159-174.

Lau, K.H. and Kam, B.H. (2005) 'A Cellular Automata Model for Urban Land-Use Simulation', *Environ Plann B Plann Des*, 32(2), pp. 247-263.

Liu, Y. and Phinn, S.R. (2003) 'Modelling urban development with cellular automata incorporating fuzzy-set approaches', *Computers, Environment and Urban Systems*, 27(6), pp. 637-658.

Meijers, E. and Burger, M. (2010) 'Spatial structure and productivity in US metropolitan areas', *Environment and Planning A*, 42(6), pp. 1383-1402.

Oviedo, D. and Dávila, J.D. (2016) 'Transport, urban development and the peripheral poor in Colombia - Placing splintering urbanism in the context of transport networks', *Journal of Transport Geography*, 51, pp. 180-192.

Santé, I., García, A.M., Miranda, D. and Crecente, R. (2010) 'Cellular automata models for the simulation of real-world urban processes: A review and analysis', *Landscape and Urban Planning*, 96(2), pp. 108-122.

Silva, E.A. and Clarke, K.C. (2002) 'Calibration of the SLEUTH urban growth model for Lisbon and Porto, Portugal', *Computers, Environment and Urban Systems*, 26(6), pp. 525-552.

Sridhar, K.S. (2007) 'Density gradients and their determinants: Evidence from India', *Regional Science and Urban Economics*, 37(3), pp. 314-344.

Stevens, D., Dragicevic, S. and Rothley, K. (2007) 'iCity: A GIS–CA modelling tool for urban planning and decision making', *Environmental Modelling & Software*, 22(6), pp. 761-773.

Straatman, B., White, R. and Engelen, G. (2004) 'Towards an automatic calibration procedure for constrained cellular automata', *Computers, Environment and Urban Systems*, 28(1), pp. 149-170.

UN-Habitat (2012) 'State of Latin American and Caribbean cities', , HS/064/12E.

USGS (2017) 'Panchromatic Image Sharpening of Landsat 7 ETM+', , 2017.

van Vliet, J., Hurkens, J., White, R. and van Delden, H. (2012) 'An Activity-Based Cellular Automaton Model to Simulate Land-Use Dynamics', *Environ Plann B Plann Des*, 39(2), pp. 198-212.

van Vliet, J., Naus, N., van Lammeren, R.J.A., Bregt, A.K., Hurkens, J. and van Delden, H. (2013) 'Measuring the neighbourhood effect to calibrate land use models', *Computers, Environment and Urban Systems*, 41, pp. 55-64.

Verburg, P.H., de Nijs, T.C.M., Ritsema van Eck, J., Visser, H. and de Jong, K. (2004) 'A method to analyse neighbourhood characteristics of land use patterns', *Computers, Environment and Urban Systems*, 28(6), pp. 667-690.

Verburg, P.H., Schot, P.P., Dijst, M.J. and Veldkamp, A. (2004) 'Land use change modelling: current practice and research priorities', *GeoJournal*, 61(4), pp. 309-324.

Wang, F., Hasbani, J., Wang, X. and Marceau, D.J. (2011) 'Identifying dominant factors for the calibration of a land-use cellular automata model using Rough Set Theory', *Computers, Environment and Urban Systems*, 35(2), pp. 116-125.

White, R. and Engelen, G. (1997) 'Cellular Automata as the Basis of Integrated Dynamic Regional Modelling', *Environ Plann B Plann Des*, 24(2), pp. 235-246.

Zheng, H.W., Shen, G.Q., Wang, H. and Hong, J. (2015) 'Simulating land use change in urban renewal areas: A case study in Hong Kong', *Habitat International*, 46, pp. 23-34.