Building a land use component in a CA based METRONAMICA Model for Jeddah City, Saudi Arabia

Amir Khan Mohmand

February, 2010

Building a land use component in a CA based METRONAMICA Model for Jeddah City, Saudi Arabia

By

Amir Khan Mohmand
February, 2010

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation, Specialisation: Urban Planning and Management

Thesis Assessment Board
Prof. Dr. Ir. M.F.A.M. van Maarseveen (Chair)
External Examiner: Ir. J. Van Vliet (RIKS)
Supervisors: Dr. J. (Johannes) Flacke
Ir. M.J.G. Brussel

INTERNATIONAL INSTITUTE FOR GEO-INFORMATION SCIENCE AND EARTH OBSERVATION
ENSCHEDE, THE NETHERLANDS

Disclaimer

Abstract

Sustainable urban development has had important impacts on national policies of urban planning in developing countries. To monitor and manage urban growth in a sustainable manner, computer-based models have been developed to describe urban growth patterns. The validity of any cellular automata model depends very much on a good set of transition rules. This study focussed on the available methods that derive transition rules. Two methods, one enrichment factor method and combination of both enrichment factor and weights derived by analytical hierarchy process were used to shape the transition rules. Enrichment factor is used to characterize the neighbourhood of a location in a land use map. This neighbourhood characteristic are calculated for every cell in a series of enrichment factor calculations for the different land use types. To derive AHP weights for the sub factors affecting suitability and accessibility, interviews were conducted from the decision makers and experts of the municipality of Jeddah. Pair wise comparison was used to get the average weights for soil, slope and elevation affecting the suitability and for major roads, primary roads and secondary roads affecting the accessibility. Both the methods were tested to build a land use component in a CA based METRONAMICA Model for Jeddah City. At the first instance, Enrichment factor values were applied in the model and calibrated for the time period between 1980 and 1993. Kappa statistics were used for comparison of simulated and actual land use maps of 1993 of Jeddah city. The overall kappa after calibration was 0.70. Individual kappa for residential was 0.65, commercial was 0.40 and for industrial was 0.45. The calibrated rules were applied for the next time step between 1993 and 2007 to validate the land use model. Individual kappa for residential was 0.72, commercial was 0.47 and for industrial was 0.55. The overall kappa after validation was 0.77. In the second method, model was run with accessibility weights and actual enrichment factor rules. The overall kappa after simulation was 0.72. Individual kappa for residential is 0.69, for commercial was 0.57 and for industrial was 0.48. When model is validated between 1993 and 2007, the overall kappa was found 0.81. Individual kappa for residential was 0.74, commercial was 0.61 and for industrial was 0.69. From the results it is concluded that a combination of both enrichment factor and weights derived for accessibility through AHP is a better method for modelling land use change in Jeddah city.

Acknowledgements

I really appreciate the opportunity provided by ITC to participate in this ITC capacity building programme. This precious study experience taught me not only a lot of scientific knowledge, but also development of an attitude about how to be a junior scientist. During my stay in Netherlands, I made many good friends and explored a lot of exotic culture.

I am very grateful to my first supervisor Dr. J (Johannes) Flacke. His great knowledge and serious scientific attitude helped so much on the whole way from preparing the proposal to the completion of the thesis. I would also like to send my sincere thanks to PhD student Muhammad Aljoufi for being a part in his research project. I would like to give my earnest gratitude to my second supervisor Mark Brussel and especially the initiator of this research Ir. Dr. Mark Zuidgeest for their support in this research. I am really thankful Muhammad Karimi PhD student, who gave me knowledge about CA modelling. I extend my gratitude Amjad PhD student and his son Ali for their support. One of the important aspects of this research, at least to me, was to visit the RIKS company which could not materialized may be because to time constraint. I am still thankful to RIKS for allowing me to use their model.

My sincere thanks also go to the students and teachers of UPM department for their assistance in this research work.

I am very much grateful to all my friends and country fellows. I can’t express my love and feelings in words to thank my dear parents, brothers, wife and beloved daughter Nawaal.

I dedicate this piece of work to my dear family

Table of contents

List of figures

List of tables

1. Introduction

1.1. Background and Justification

Land use change is a complex dynamic process. Various factors interact at different spatial scale (Lay 2000). Some of these contributory factors are population growth, economic activities, employment, migration and policies. Resources are continuously depleting like imbalance between land demand and supply, per-capita agricultural land decreases, greenhouse gases are increasing and economy is shrinking. Due to these enormously pressing resources and environmental problems in our society, interest in sustainable urban development has increased rapidly. Sustainable urban development has had important impacts on national policies of urban planning in developing countries which are undergoing rapid urbanization (Li and Yeh 2000). To monitor and manage urban growth in a sustainable manner, computer-based models are developed for describing urban growth patterns (Lee 1999; U.S. EPA. 2000).

One of them is cellular automata (CA) modelling technique. The study of Cellular Automata goes back to the late 1940’s when von Neumann and Ulam gave the idea of Cellular automata (Fredkin 1990). A cellular automata is a discrete dynamic system composed of a set of cells in a one or multidimensional lattice. The state of each cell in the regular spatial lattice depends on its previous state and the state of the cells in its neighbourhood. The cells change their state according to a set of rules called the transition rules. Actually the state of a cell at a given time depends only on its own state in the previous time period and the states of its nearby cells in the previous time. All CA models are synchronous processors that mean cells change their state in discrete time steps simultaneously. The overall behaviour of the system is determined by the evolution of the states of all cells which is the result of multiple interactions (Batty and Xie 1994; Couclelis 1997; Li and Yeh 2000).

After 1960’s and 1970’s CA gained gradually more attention in several scientific disciplines, particularly in Physics, Mathematics, Computer Science and Biology (Engelen, White et al. 1997). Currently a rapidly growing body of knowledge on the behaviour and capabilities of CA’s exists and many characteristics of real world are incorporated into it. For examples the applications of CA include among others: the modelling of urban and regional development (White and Engelen 1993; Batty and Xie 1994; White and Engelen 1994), spread of forest fires and spread of epidemics etc.

The application of CA in urban modelling can give insights into a wide variety of urban phenomena (White and Engelen 1993; Batty and Xie 1994; Wu 1998; Li and Yeh 2002). It has the strong structural capability of simulating urban patterns and validation of results to transform the policy decision into a better future (Deadman, Brown et al. 1993). Cellular automata understand the process by which cities and regions change. They are explicitly dynamic, representing changes that have been occurred over time. To experiment with the future, CA explores it through defining various scenarios and test policy and planning options before they are implemented.

Many urban simulation models has been based on Cellular automata and are well suited for modelling land use and land cover patterns at a local and regional scale. METRONAMICA is a CA based land use change modelling tool, which works at local and regional level. Regional models are used to model the levels of activity in different socio-economic sectors. Metronamica is composed of three types of project configurations. These are Metronamica single layer (SL), Metronamica multi layer (ML) and Metronamica land use transport (LUT).

1.2. Research Problem

In Saudi Arabia, the spatial expansion of cities has been triggered by the government’s public policies for urban development. In the last four decades, government initiatives of grant policy and interest free loans resulted in the enormous expansion of major cities such as Riyadh and Jeddah. Through these policies, residential plots were distributed free of cost to the general public to house the rapidly growing urban population in a short span of time. To tackle this problem, master plans were prepared for a number of rapidly growing cities. But in a short period of time they went futile. Lack of institutional support and absence of viable planning and decision making systems, further worsened the problem of urban growth in the major cities including Jeddah. This in turn, stimulated the rapid expansion of road network and other utilities with high land values (Al-Hathloul and Mughal 2004; Garba 2004).

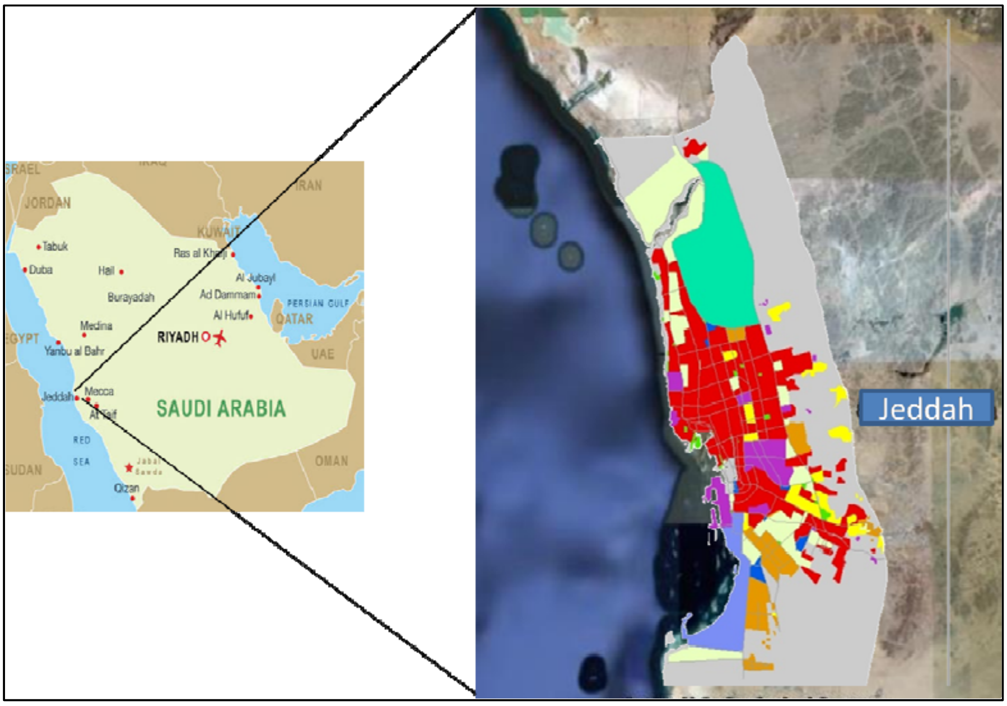


Figure 1: Jeddah city land use map overlay on Google Earth image

The area of this study is the city of Jeddah. The city of Jeddah is the most cosmopolitan of all Saudi Arabia's cities. It lies at latitude of 21°32'37" N and longitude of 39°10'23" E. It is situated on the Eastern coast of Red Sea. Over a period of twenty five hundred years, the city transformed from a small fishing village to a modern world class city. Most of the infrastructural development took place since 1948. From 1948-61, was seen as the period of unplanned urban growth along the coast in the North South direction and valleys development in the East. From 1961- 1970 detail studies were carried out and master plans were prepared. The third stage from 1971-85 was the most significant period in the development of Jeddah in particular. National development plans and action plans were introduced to direct all sectors of the economy in a coordinated manner.

Booming up of oil industry since 1970s, gave an impetus to unplanned city growth of Jeddah. The local government initiated new plans and restrictions to coup up with the situation. Government incentives to private sectors for infrastructural development, free housing and interest free loans complemented to urban deterioration even more (Al-Hathloul and Mughal 1991). The city became a magnetic pole, especially attractive for immigration of both skilled and unskilled workers. The population of the city grew rapidly from 381,000 in 1971 to about 600,000 in 1974 and to just over 1,000,000 by 1983 (Mohammed, Salagoor et al. 2002). According to the recent National Population Census of 2004, the population of the city has grown more than three millions. This explosion of population necessitated the need not only for provision of facilities and services but also for planned development (Mohammed, Salagoor et al. 2002). The policy for retail development also proved to be very unsuccessful and one of the main causes of urban deterioration of Jeddah city (Daghistani 1993). One of the current problems in Jeddah is the lack of coordination between the land use planning department and transport department. All these problems coupled and gave birth to a big management problem. Hence local authorities took up the cudgels of devising strategies for viable future planning and management and its effective implementation through integrated planning and decision support systems.

This research tries to understand the urban growth that has taken place in the past by developing a CA based land use change model using METRONAMICA. For the development of such a model, transition rules play a central role in determining the future state of a land use. The research problem is to evaluate methods for deriving transition rules for CA modelling and applying it to the case of Jeddah city. Such model could provide valuable information about the past and an understanding of future urban development to support strategies aimed at enabling the public sector to better deal with the management of urban growth.

1.3. Research Objective

The research objective is to evaluate different methods to derive transition rules for CA modelling and applying it to the case of Jeddah city.

1.4. Sub-objectives

1. Understanding spatial temporal urban growth that has taken place in Jeddah city.
2. To analyze different methods for deriving transition rules for a CA model.
3. To apply and evaluate different methods of transition rule to the case of Jeddah.
4. To calibrate and validate the CA model for city of Jeddah.

1.5. Research Questions

1. Understanding spatial temporal urban growth that has taken place in Jeddah city.
   - 1.1. How urban growth took place in Jeddah City over time?
   - 1.2. What are the driving factors for urban growth in Jeddah city?
2. To analyze different methods for deriving transition rules for a CA model.
   - 2.1. What is transition rule and how it is important in land use change?
   - 2.2. What approaches have been used in the existing CA models?
3. To apply and evaluate different methods of transition rule elicitation to the case of Jeddah.
   - 3.1. How transition rules can be derived using enrichment factor?
   - 3.2. What are the simulation results of the combined method of Analytical hierarchy process derived weights and enrichment factor?
4. To calibrate and validate the model for city of Jeddah
   - 4.1. Which method is more accurately simulating the urban growth in Jeddah?

2. Literature review

Cities are among the most complex structures created by people. Cities are highly complex social, economic, spatial synthesis, consisting of many components. These components will have different influence on urban development process, involves complex interrelationships among each other. In the past, people only take cities as one kind of natural or social phenomenon, which can be understood and predicted in traditional methods. While reality proved this shallow understanding about cities often shaded some real driving forces of urban development. This situation often resulted in some unsatisfactory or even wrong predictions of cities (Barredo, Kasanko et al. 2003).

In order to get a better understanding of complex urban system, people began to employ systematic analysis method, which were good at dealing with complex natural and social phenomena (Junfeng 2003). Urban system is good example of complex system, which consists of two complex processes, one is spontaneous development process, the other is self-organization process (Wu 2000).According to (Cheng 2003a), the urban development process is an entirely a complex process if we observe it from the spatial temporal characteristic of urban system (Batty, Xie et al. 1999). Barros and Sobreia (2002) regard city as a complex system, composed of many components that interact with each other for example land use, transportation, culture, population, policies economy and so on.

Urban development can be divided into two categories. The spontaneous growth is a homogeneous and scattered spatial pattern, which contains more random components, whereas self–organizational growth is the manifestation of spatial agglomeration pattern, which is associated with socio-economic activities (Cheng 2003a). Understanding the dynamic process of urban growth, i.e. from past to present to future involves various socio-economic and physical and ecological components at varied spatial and temporal scales, which result in a complex and dynamic system. Urban growth is one of the complex spatial changing phenomena in urban system. It is the transformation of the vacant land or natural environment to construction of other urban fabrics like residential, industrial and infrastructure development. It mostly occurs in the fringes of urban areas (Shenghe and Sylvia 2002).

2.1. Land use Models

In essence a land-use model is a mathematical, logical or mechanical representation of a relationship, process, system, or sequence of events related to urban land use. The ultimate use of the land use models is to perform structural analyses of the dynamics of urban land use patterns and their driving factors. Land use models create different scenarios to support spatial planning for various option of real world problems (Verburg, Dijst et al. 2001). In other words, through models we summarize the complex relationships of the real world of theory and experiments. Armstrong tried to show this relationship between theory and experiment through land-use models by including simulation visualizes in figure 3.

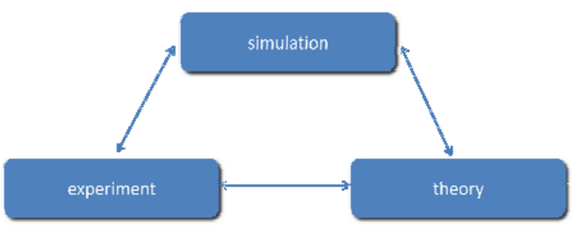


Figure 2: Relation between theory and experiment including simulation (Hagoort 2006)

Land-use models can be classified into three types, based on the purpose of modelling: descriptive, prescriptive and predictive (Webster 1993; Webster 1994).

1. Descriptive models foresee realistic near-future urban land-use transformation trends, as a function of the dynamics of the so-called driving factors of urban land-use change, under a set of conditions specified in one or more scenarios. In short, they follow pre-described actions or decisions to show their outcomes. The result of a simulation denotes the consequences of possible policies or autonomous developments to be applicable for specific areas to planners or decision makers.

2. Prescriptive models maximize theoretical urban land-use configurations to arrive at the best match under a set of goals and objectives. The simulated urban land-use configuration gives the 'best' set of inputs to obtain a particular urban land-use change trend. They outline which actions or decisions have to be taken to arrive at the pre-described desired results as best match as possible. Prescriptive models are used in situations where the objectives of a policy are known in advance. The results justify adjusting the goals identified by planners.

3. Predictive models reshape the descriptive model arrangements in such a way that particular variables at the end of an underlying sequence can be predicted from variables that played a role earlier in the sequence. The distinction between predictive and prescriptive models is dependent on whether it is actual behaviour or optimal or normative behaviour that is simulated in the modelling exercise (Batty, 1994)·

After presenting a broad view, a general operating structure of the land use models can be shown in figure (4) which contains three parts, the demand, supply and allocation.

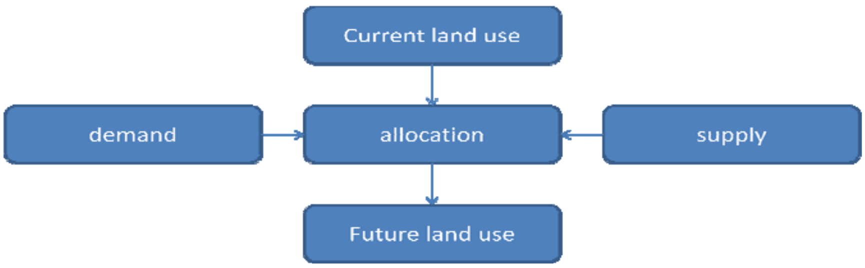


Figure 3: General land use model structure (Hagoort 2006)

The demand may contain claims of certain sectors or stakeholders for a certain area and particular land use. Usually general growth rates are used to derive demand for an area needed for a land-use type on a certain scale. The claims and demands are put exogenously into the models. By supply we mean to provide certain land-use types that fulfil the demand of a particular land use type for a specific purpose depending upon the suitability of a location. For example housing areas can be located in attractive areas based on the presence of slopes, special areas or soil type etc. Allocation determines the future state of an area based on the need to balance supply and demand. This also based on some facilitating and restrictive factors like slopes, special areas or soil type etc.
A variety of land-use change models are available through literature. They can be classified on the basis of modelling techniques, purposes, theories and types of land uses modelled and spatial and temporal levels of analysis (Hagoort 2006). For example, statistical and transition probability models, optimization models and linear programming, dynamic simulation models, agent based models and cellular automata (CA) which are commonly used to model land-use change (Hagoort and Geertman 2004). A broad survey of the various methods may also be found in (U.S. EPA. 2000; MODLUC 2002).

2.2. Categories of land-use change models

Based on the work of (Briassoulis 2000), Hagoort has given a detail overview of five main categories of models based on modelling tradition to which they belong.

Table 1: Categories of land-use change models by modelling tradition

- Statistical and econometric models;
    - Statistical models;
    - Econometric models,
- Spatial interaction models;
- Optimization models;
    - Linear programming models;
    - Dynamic programming models;
    - Goal programming, hierarchical programming, linear and quadratic assignment problem, nonlinear programming models;
    - Utility maximization models;
    - Multi-objective/multi-criteria decision-making models.
- Integrated models;
    - Econometric-type integrated models;
    - Gravity/spatial-interaction-type integrated models;
    - Simulation-based integrated models;
    - Input-output-based integrated models,
- Other modelling approaches.
    - Natural-sciences-oriented modelling approaches;
    - Markov modelling of land-use change
    - GIS-based modelling of land-use changes.

(Hagoort 2006)

However these traditional models have weaknesses despite the fact they are quite successful in the field of urban planning and management. Torrens 2001 has identified six major weaknesses. These are:

- A centralized CBD approach of older models does not exist nowadays.
- There is a lack of consideration for all the changes that have taken place between large time spans for which the model is setup.
- These models use information at a more aggregate level. This affects a model's simulation capabilities and hinders its ability to model reality more closely.
- These models are very complicated and require expert knowledge to use and draw conclusions about a land use change.

- The operational connection between the micro and macro scales are poorly integrated which reduces their flexibility at low scale urban processes.
- There is a lack of realism in the sense that both micro and macro scale processes (top-down as well as bottom-up) should be taken into account. In reality they cannot explain how the city behaves.

Land use changes can best be modelled by cellular automata (Torrens 2001). Urban planners and stakeholders are always in need to understand the past urban development and its main driving factors. The urban models help them to analyze and understand the city growth and the interaction of land uses and transport. Cellular Automata are dynamic which means representing change over time, and flexible, as well as intuitively useful and behaviourally realistic. It provide a mechanism for linking micro-approaches and macro-approaches (Hagoort 2006). It is also suitable for representing scattered activities on separate spatial units of urban infrastructure and for defining the spatial interactions between land uses at certain locations, conditions at these locations and the land-use types in the neighbourhood (Torrens 2001).

Cellular automata models are particularly useful in simulating urban spatial change (Yi 2009). It takes into account many factors like suitability, accessibility, zoning and future policy guideline for growth prediction over a period of time in past or future in discrete time steps. In the case of Jeddah, there was no such land use change model available. Such model, under the auspices of Metronamica, would help the planners and stakeholders understand the past urban development of the city and plan the future urban development in a sustainable manner. The scope of this research is the use of CA in Land Use Change modelling. Therefore we shall focus in literature review on Cellular Automata.

2.3. Driving forces of urban growth

In reality, urban growth is driven by many factors such as population growth, development of the urban economy, investment in infrastructure, increase in income, etc (Shenghe and Sylvia 2002). The common driving factors used in urban CA models are travel time or distance to road networks, railway or subway, the travel time or spatial distance to the town centre or sub-centres, the proximity to protected utilities, physical properties such as altitude, slope or soil and geology type, the proximity to socio-economic services, commercial or industrial districts, environmental suitability (Al-Ahmadi, See et al. 2009). For example the growth of a non urban cell into urban could depend on the accessibility to transport networks, topographical constraints and zoning regulations. For example, one driving factor is accessibility which is the capability to reach particular places by transport relative to one another like accessibility to employment centres and socio-economic services and accessibility to the town centre. It is a common factor to be included as input in current urban CA models (Wu 1998; White and Engelen 2000). Topography is one of the important driving factors of urban growth which include slope, elevation and aspect. These are the important inputs to the model. The slope of the land surface is important because it determines the use of land and cost of development (Al-Ahmadi, See et al. 2009). Soil is also an important determinant in the suitability calculations of land.

2.4. A simple Urban CA

There are many different techniques of land-use change modelling, together with their advantages and disadvantages. In recent years there has been one technique that occupied the land-use modelling community, the so-called Cellular Automata (CA). CA is an artificial process for locating urban activities of an area based on simple rules that shape the spatial patterns of cities (Batty 1997). Cellular Automata, which are self organization complex systems in themselves, are inherently spatial since they are defined on a two-dimensional grid and therefore very suitable for the simulation of land use dynamics.

CA was first devised by John von Neumann and Stanislaw Ulam in the 1940s as a framework to investigate the logical foundation of life. One can say that the “cellular” come from Ulam and the “automata” come from von Neumann (Rudy 1999).

The history of research and design cellular automata can be traced back to the beginning of digital computation. Alan Turing, an English mathematician, first stated that computers and their software can be used to produce rules that could reproduce themselves (Batty 1997). Stanislaw Ulam, a polish born American mathematician, used a simpler lattice network for studying growth of crystals in the 1940s. John Von Neumann, Ulam’s colleague at Los Alamos National Laboratory faced difficulty in building his self replicating robot. Ulam suggested that simple cellular automata could be found in sets of local rules that generated mathematical patterns in two-dimensional and three-dimensional space (Batty, Couclelis et al. 1997). Then Von Neumann found his four cell neighbourhood (the cell in question and its four surrounding neighbours) on a 2-dimensional cell space which is called Neumann neighbourhood. Around 1950s, his work was a beckon of light in the field of Cellular Automata (Batty 1997).

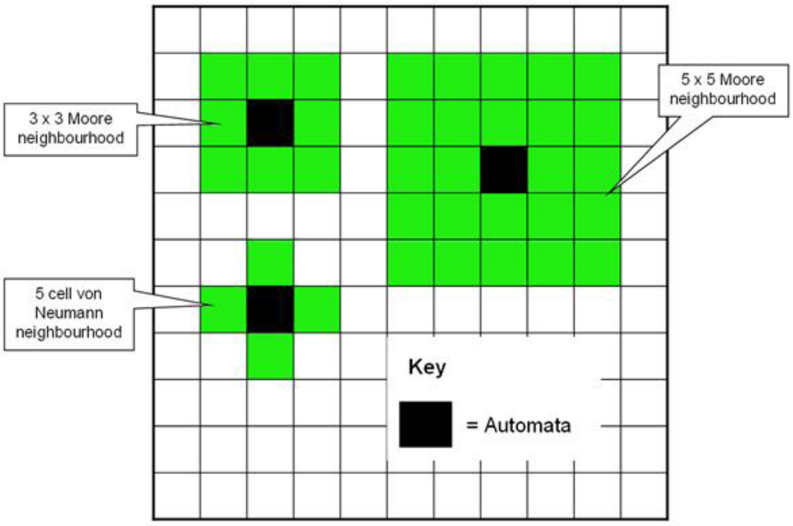


Figure 4: Neumann and Moore neighbourhood (Torrens 2006)

Based on their research and discoveries, the first CA application was formed by John Horton Conway as his famous “Game of Life”. It is a 2-dimensional lattice with two possible cell sates – dead or alive, based on an eight-cell neighbourhood. The eight-cell neighbourhood is called the Moore’s neighbourhood (Liu 2008).

After successive generations, Conway described his Game of Life, explained here under:

Survival of a cell

A live cell with two or three live neighbour cells will remain alive in the next generation.

Death of a cell

A live cell with two or more than three live cells dies either of isolation or of overcrowding.

Birth of a cell

A dead cell with exactly three live cells becomes alive in the next generation.

Conway’s Game of Life generated complex patterns using these simple rules and is of great interest to the researchers in the field of cellular automata. In 1983, Stephen Wolfram during his research investigated the simplest one dimensional cellular automata, which he termed as elementary cellular automata (Wolfram 1994; Wolfram 2002). Each cell in the elementary cellular automata has only two possible values or states of either 0 or 1, and the transition rules depend only on the nearest neighbourhood values (Liu 2008).

2.4.1. Components of a CA model

2.4.1.1. The cell

The cell is the basic spatial unit in a CA model. These cells are commonly arranged in a two-dimensional grid in CA which is used for modelling urban growth and land-use change. A three dimensional arrangement like honeycomb could be possible to represent the building heights of the urban structure but may not be feasible due to the complex nature of CA and yet to be explored in urban modelling (Liu 2008).

2.4.1.2. The state

The state of the cells shows its property whether it's a number like 0 or 1 or the land use type. Each cell can take only one state from a set of states at any one time. In urban-based cellular automata models, the states of cells may represent the types of land use or land cover, such as urban or rural, or any specific type of land use.

2.4.1.3. The neighbourhood

The neighbourhood can be defined as the cells present in the immediate vicinity of a cell with which the cell in question interacts. There are Two basic types of neighbourhoods in a two-dimensional space, the von Neumann Neighbourhood (four cells and the cell in question), and the Moore Neighbourhood (eight cells and the cell in question) as shown in figure 2.5. Other kinds of neighbourhoods, such as a circular neighbourhood lying at Euclidian distance from the cell in question, have also been used in urban modelling (White and Engelen 1994; White, Engelen et al. 1997).

2.4.1.4. The transition rule

The transition rule defines how the state of one cell chances in response to its current state and the states of its neighbours. The validity of any cellular automata model depends very much on a good set of transition rules, that is why it is the heart of CA (White 1998). A simple CA rule that define the cell transition could be written in IF-THEN statement as;

IF something happens in the neighbourhood of a cell
THEN something else will happen to the cell at the next time step

The time

Time specifies the temporal dimension in which a cellular automaton exists. According to the definition of cellular automata, the states of all cells are updated simultaneously at all iterations over time at every location (Torrens 2001). If we apply the above rule once, the temporal space is denoted by T, then the cell will change its state in the next time step T+1.

2.4.2. The Transition rules

CA is parallel synchronous machines composed of automata in which more than one automaton is active at any given time. The neighbourhood is the immediate adjacent region- a collection of cells, which provide input to the automata in question. In other words neighbourhood in an urban CA determines the activity in the region e.g., market catchment areas, commuting watersheds, etc. But the real driving force behind CA is transition rules. These are simply a set of conditional statements – the rules that specify the behaviour of cells as CA progresses over time. A rule can be devised to mirror how phenomena in real cities operate (O'Sullivan and Torrens 2001) .

"Rules express how they value the presence of other functions and land uses/land covers in their neighbourhood (on a relative scale)" (Engelen 2002).

"Rules establish how the different land-uses react on each other, if they attract or repulse each other in relation to the distance" (O'Sullivan and Torrens 2000).

"Rules specify the interaction between land-use types that are located in each other's neighbourhood" (Verburg, de Nijs et al. 2004).

"Rules define the influence that neighbouring land uses exert on each other in the form of push-pull forces" (Verburg, de Nijs et al. 2004).

Transition rules are decision rules or transition functions of a cellular automata model which may be deterministic or stochastic. It specifies the states of a cell before and after updating based on its neighbourhood conditions and guides the dynamic process of CA (Lay 2000; Silva and Clarke 2005).

The main problem of CA is the validity of the transition rules, which shows the behaviour of different land uses. In other words, a rule expresses how a particular land use values the presence of another land use in its neighbourhood on a relative scale and quantifies the relationship between individual cells of a land use as a function of distance.

2.4.3. The neighbourhood effect

In recent land use models the cell state represents the predominant land use present at this cell at a certain moment such as: residential, industrial, or forest etc. The neighbourhood of a cell describes which surrounding cells influence the transition of a cell state for a certain cell in the next time step. The cells further away have a smaller influence than cells closer to the centre cell neighbourhood effect represents the attraction (positive) and repulsion (negative) effects of the various land uses and land covers within the neighbourhood. In general, cells that are more distant in the neighbourhood will have a smaller effect. Thus each cell in a neighbourhood will receive a weight according to its state and its distance from the central cell (RIKS 2005).This effect is called distance decay effect as shown in figure 6.

$$^{t}R_{f,c} = \sum_{c' \in D(c)} w_{f,^{t}f(c')}\left(d\left(c,c'\right)\right)$$

$^{t}R_{f,c}$ The neighbourhood effect in cell *c* for land use *f* at time *t*.

$^{t}f\left(c\right)$ The land use occupied by cell *c* at time t

$d\left(a,b\right)$ The Euclidian distance between cell *a* and cell *b* see fig. 7.

$w_{f,f'}\left(d\right)$ The influence function, expressing the strength of the influence of a cell with land use *f'* on land use *f* for each distance *d* in the CA neighbourhood

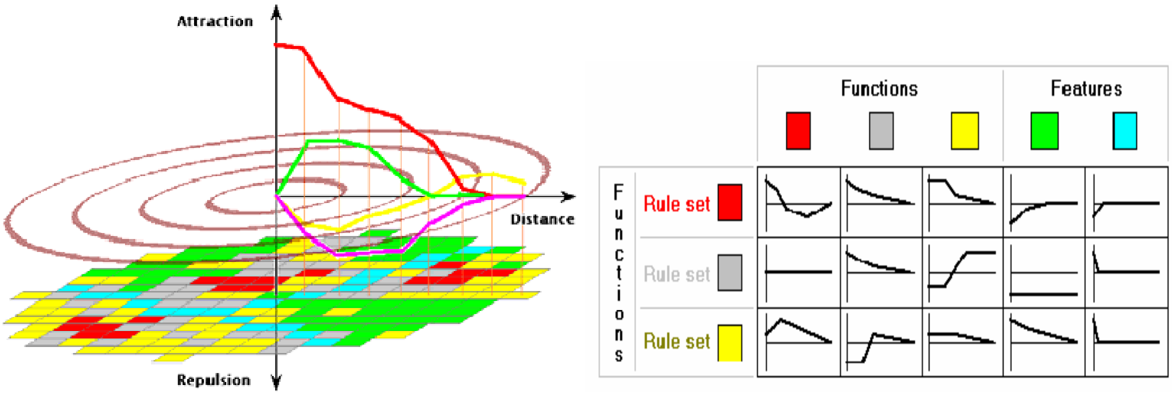


Figure 5: Neighbourhood effect and rules (RIKS 2005)

2.4.4. The cell size

Through literature, we found that different cell resolutions were applied which depended on the application. White and Engelen (1993) used a 500meter cell resolution in a constrained cellular automata model to simulate the urban land-use patterns in four U.S. cities, including Atlanta, Cincinnati, Houston and Milwaukee. Later on, for the same city of Cincinnati of Ohio, they changed the resolution to 250 meter to simulate urban land-use change (White, Engelen et al. 1997; White and Engelen 2000). Wu (1996) used a 28.5meter cell resolution for simulating urban dynamics in China's Guangzhou city. He also used 200 meter cell resolution for the same area (Wu 1996). However, Clark and Gaydos applied various cell resolutions in the in the Washington Baltimore region ranging from 210, 420, 840 to 1680 meter. They concluded that cell resolution only affected the road and slope factors in the model and no effect on other rules or factors (Clarke and Gaydos 1998). The idea behind coarse resolution is only to reduce computation time and no justification is evident from the work of these modellers (Liu 2008). The reason for choosing the cell scale was based on either the data availability or reducing the computation time of cellular automata.

2.4.5. The neighbourhood size

Like the cell size, the neighbourhood size is also the choice of the modellers. Different neighbourhood sizes have been applied to urban models. For instance, White and Engelen (1997, 1994) used a circular neighbourhood of 113 cells whose impact is based on distance decay rule on a central cell. Whereas Clarke and Gaydos (1998), Wu (I998a,b,c 1996) and Clarke, Hoppen, and Gaydos (1997) used the Moore Neighbourhood, which is small, as it only consists of nine cells including the cell in question. Mostly large neighbourhoods are used in urban cellular models (Batty and Xie 1994). One reason for this could be the distance decay effects of the surrounding cells on central cell (White and Engelen 1994; Wu 1996). In Metronamica, a large circular neighbourhood of 196 cells has been used (RIKS 2005). But the size of the neighbourhood in cellular automata-based urban models has not been fully justified yet (Liu 2008).

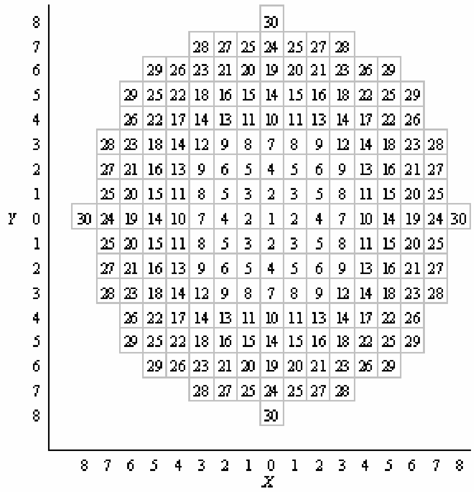


Figure 6: A circular neighbourhood around the cell out to a radius of eight cells where X and Y represent the horizontal and vertical distance between the cells (RIKS 2005)

2.5. CA models based on various transition rules

Many approaches have been used for deriving transition rules for cellular automata models (Liu 2008). The SimLand model developed by Wu.F (1998) for Guangzhou city in China uses AHP derived transition rules. The AHP method was used to include decision makers preferences and derive behaviour oriented transition rules (Wu 1998). Results were derived using pairwise comparison developed by Saaty method. The results were then used in the cellular automata model to configure the transition rules (Wu 1998). The artificial neural network (ANN) approach was used by Li and Yeh (2001, 2002) to generate and calibrate the model's parameters automatically. The ANN is actually based on biological neurons network. A three layer neural network with multiple output neurons was designed to conversion probabilities for multiple land uses. The model results were then used in transition functions of the cellular automata (Li and Yeh 2002). Sui and Zeng (2002) used multiple regression analysis to calculate weights of parameters to be used in the cellular automata model (Sui and Zeng 2001).

2.6. Available methods to derive Transition rules

Recently, various attempts have been undertaken to derive neighbourhood rules theoretically and empirically, though they have not proved to be all that much satisfactory (Sui and Zeng 2001). However these methods have been applied by many modellers and experts around the world. An overview of some of the methods is given here under:-

2.6.1. Enrichment factor

Enrichment factor is a method for exploring and quantifying the neighbourhood characteristics of land use and the interaction of land use types. The results are then used to help defining the transition rules for a cellular automata model. Neighbourhood effects alone are, however, not only the location factor suitable for describing the spatial pattern of land use. Other factors, such as accessibility, environmental suitability, spatial policies etc. do also influence the pattern of land use (Verburg, de Nijs et al. 2004). It is not possible to directly translate the neighbourhood characteristics into transition rules for cellular automata but it is used as a starting point for defining the transition rules in the CA model.

Advantages

This method is data driven, suitable for multi-scale analysis, easy to understand and explain the results.

Disadvantages

All neighbourhood effects identified with the enrichment factor do not have a causal explanation, because some of the observed neighbourhood effects are the indirect result of other interactions (Verburg, de Nijs et al. 2004).

2.6.2. Regression Analysis

Transition rules can be elicited using regression analysis for urban CA models. In this method, the modeller identifies the possible influence factors, which affect the transition of cells from one state to another depending upon its neighbourhood which include neighbourhood effect, suitability effect, and accessibility effect. Then the effects of different factors are measured. For example, the neighbourhood effect can be measured by the ratio of developed cells to all cells in the neighbourhood. The suitability of cells can be calculated by land suitability analysis and the accessibility of cells can be measured by the distance to different areas like business centre, employment centers, road network, etc using GIS software. Next different land use maps are overlaid and changing areas will be identified. Then random samples will be selected from these changing areas. To find the explanatory coefficients to the different influence factors, multiple regression analysis is used. These coefficients are then used as inputs in the transition potential rule to calculate the changing potentiality of different cell (Jiao and Boerboom 2006).

General statistical methods may have some limitations when spatial factors and model structures are too complicated. They are invalid when spatial factors correlate with each other. They also have difficulties in handling poor and noisy data (Li and Yeh 2002), but logistic regression method is specially suitable for CA models because it includes multiple land use types and a flexible neighbourhood (Verburg, de Nijs et al. 2004).

Advantages

This method is data driven, say, like suitability, zoning and accessibility, and even land use, as these can be done with only data. It is spatially explicit and suitable for multi-scale analysis, and gives much deeper understanding of the forces driving the growth and the formation of the urban spatial pattern (Hu and Lo 2007). The pattern of Jeddah city is like American cities and regression analysis has already been used there like Hu and Lo (2007) used it in the city of Atlanta (USA). Hence it can be tested in Jeddah.

Disadvantages

The drawback of regression analysis method is its lack of temporal dynamics. It is a static approach and not a dynamic, though it can give qualitative insights.

2.6.3. Artificial Neural Network Method

Li and Yeh (2001, 2002) have developed their models based on back propagation (BP) neural network which is good at capturing non-linear characteristics and strong in prediction. They used neighbourhood effects, accessibility effects and suitability effects. Then, some methods were used to measure the above three effects. In the third step, a neural network was formed. In the fourth step, a land use change map will be formed by overlaying historical data. It requires at least two land use change maps, one will be used to select some random samples to train the network, and the other map will be used to test the formed network. After that, a mature neural network will be formed that can act as the transition potential rule in urban CA models. Finally, the total land consumption in a given period will be calculated from the historical data and be used to control the whole iteration times of the formed network (Jiao and Boerboom 2006).

Advantages

Neural Network reduces the tedious work of defining parameter values, transition rules and model structures. Training data from the GIS can be easily used to obtain parameter values by calibrating the model. The model has the advantages of handling incomplete and erroneous input data (Li and Yeh 2002). This is a data-driven method, but good at generating urban configurations (Jiao and Boerboom 2006).

Disadvantages

The fundamental issue in using ANN approach is that it is basically a black box type of model in the sense that the modeller and the users do not know what is happening inside this black box, therefore it does not give explicit method to understand the urban change (Jiao and Boerboom 2006; Liu 2008). High interpretative ability is required for the procedure and not only that some data go in and out. It hinders the application in rules elicitation and hence lacks understanding of the real urban development process (Jiao and Boerboom 2006). Li and Yeh acknowledge that it works well for one land use conversion (urbanization) but that it is very hard for a combination of land use types as in case of Jeddah we have more than one land use conversions.

2.6.4. Visual Observation (Trial and Error) Method

Many researchers have used visual observation to calibrate their models and elicit the transition rules. Research Institute of Knowledge System (RIKS), in the Netherlands developed their CA based models using this method. RIKS models include two parts, the macro model and the micro model. Macro model is used to define cell state change based on the outside land demand while micro model is used to calculate the transition potential of each cells. Visual observation method is used to determine the interactions of the different land use types within the neighbourhood, which can be represented by distance curves. On X-axis the distance (8 cells) shows that how far the tested cell from the central cell and Y axis is shows the interaction (attraction or repulsion) between tested cell and central cell. After adjustment of these curves, the simulation results are compared with the real land use maps and try to readjust the parameter. Then a new simulation is run, which is then compared again with the real land use map. If the accuracy is increased then the parameter will be modified in the same direction as the first modification or vice versa. However this process is time consuming and shows uncertainty (Jiao and Boerboom 2006).

Advantages

This method is already used in Metronamica software and transition rules are obtained by calibration (trial and error) of parameter values. Modeller can explain the factors and their influence separately.

Disadvantages

Deriving transition rules by calibration (trial and error methods) is time consuming and full of uncertainty. It is based on expert knowledge to come up with good rules.

2.6.5. Analytical Hierarchy Process (AHP)

Analytic Hierarchy Process (AHP)is a powerful and comprehensive methodology designed to facilitate sound decision making by using both empirical data as well as subjective judgments of the decision-maker. The AHP is a mathematical theory first developed at the Wharton School of the University of Pennsylvania by one of Expert Choice's founders, Thomas L. Saaty.

Wu and Webster (1998), in their model tried to elicit behaviour-driven transition rules by using analytical hierarchy process (AHP) and multi-criteria evaluation (MCE). They involved the characteristics in people’s decision making process and elicited transition rules that can be quantified by pairwise comparison and weighting process such as cost distance to city centre, to major industrial districts or to new railway stations, etc. In this method, factors that affect the land development in a case study area are identified. Then a hierarchy criterion is developed to define the relationship between these factors. The importance of these factors is determined by pairwise comparison, a weighting process to establish importance of the different factors. These weights are then used in a cellular automata model to calculate the transition potential for shaping the transition rules.

Advantages

It is also a data-driven methods, but easier to reflect peoples’ decision making priorities. It can generate different urban development scenarios based on different decision making processes. The AHP assists with the decision making process by providing decision-makers with a structure to organize and evaluate the importance of various objectives and the preferences of alternative solutions to a decision.

Drawbacks

As this method focuses on elicitation of transition rule for CA models, questions, such as how to reach suitable decision makers and deal with their ideas about modelling of urban development arises.

3. Materials and methods

3.1. Data Used

The following data is used in this research.

Primary data

- Satellite images (Spot 20 meter for year 1993)
- Satellite images (Spot 2.5 meter for year 2007)
- Aerial photograph 1964, 1970 and 1983
- Master plans (1963, 1973, 1980, 2000, 2002)
- Aster DEM of Jeddah region

Secondary Data

- Population data
- Land use maps of 1970, 1980, 1993 & 2007
- Road map (major, primary and secondary)
- zoning maps
- Suitability map
- Region map
- Soil map
- Slop map
- Elevation map

3.2. Softwares

Arc GIS 9.2, METRONAMICA, Map comparison Kit (MCK), MATLAB and Expert choice are used for this study. Arc GIS 9.2 is used for data preparation and analysis. Metronamica modelling environment is used for urban growth simulations. Map comparison Kit (MCK) is used for comparing categorical maps. Community Viz is used for creating suitability maps. Zoning maps are prepared in Arc GIS. Microsoft office (Word and Excel) is used for reporting. Enrichment factor is calculated using MATLAB.

3.3. Data preparation

Aerial photograph of 1970 was used for on screen digitization and the land use map of 1970 was prepared. For the allocation of land uses we used master plan of 1973 and the local knowledge. The master plan shows both the existing and planned future development of the city. In master plan, some areas shown as developed, may not be developed on ground, in that case we interpreted the image visually. Aerial photograph of 1980 was used for on screen digitization and the land use map of 1980 was prepared. For the allocation of land uses we used master plan of 1980 and the local knowledge and the image itself. Spot Satellite image with 20 meter resolution was used for digitizing land use maps of 1993. Land uses were allotted using the master plan of 1986 and by visual interpretation of the image. Similarly we digitized the 2007 land use map using the spot satellite image of 5 meter

resolution. Land uses were allotted using the master plan of 2009 and by visual interpretation of the image.
Road maps of 1970, 1980, 1993 and 2007 were prepared by digitizing from the respective aerial photographs and satellite images.
The slope map was derived from Aster DEM using Ilwis. Elevation map was prepared from the same Aster DEM in Arc GIS. The soil map was obtained from the Jeddah municipality. A region map was created using the districts boundary which will be used as the simulation area in the Metronamica. The quality of the land use maps is the same as the master plans of Jeddah city for respective time periods. As explained above, after digitization from the respective aerial photographs and satellite images, land uses were allotted as per the land uses described in the master plans. There is no other reference data about the land uses is available during this research.

3.4. Research Method

A visual representation of the research steps can be found in figure (8) below. The framework is explained stepwise in this chapter. Some part is explained in the literature review. This research is composed of data preparation and the spatial temporal analysis part, to arrive at coherent data for choosing simulation period. The next part is the modelling, which includes defining transition rules for the CA based Metronamica model. The last part is the calibration of rules and validation of results for the next time step. Finally this research focuses on comparison of results of two relevant methods.

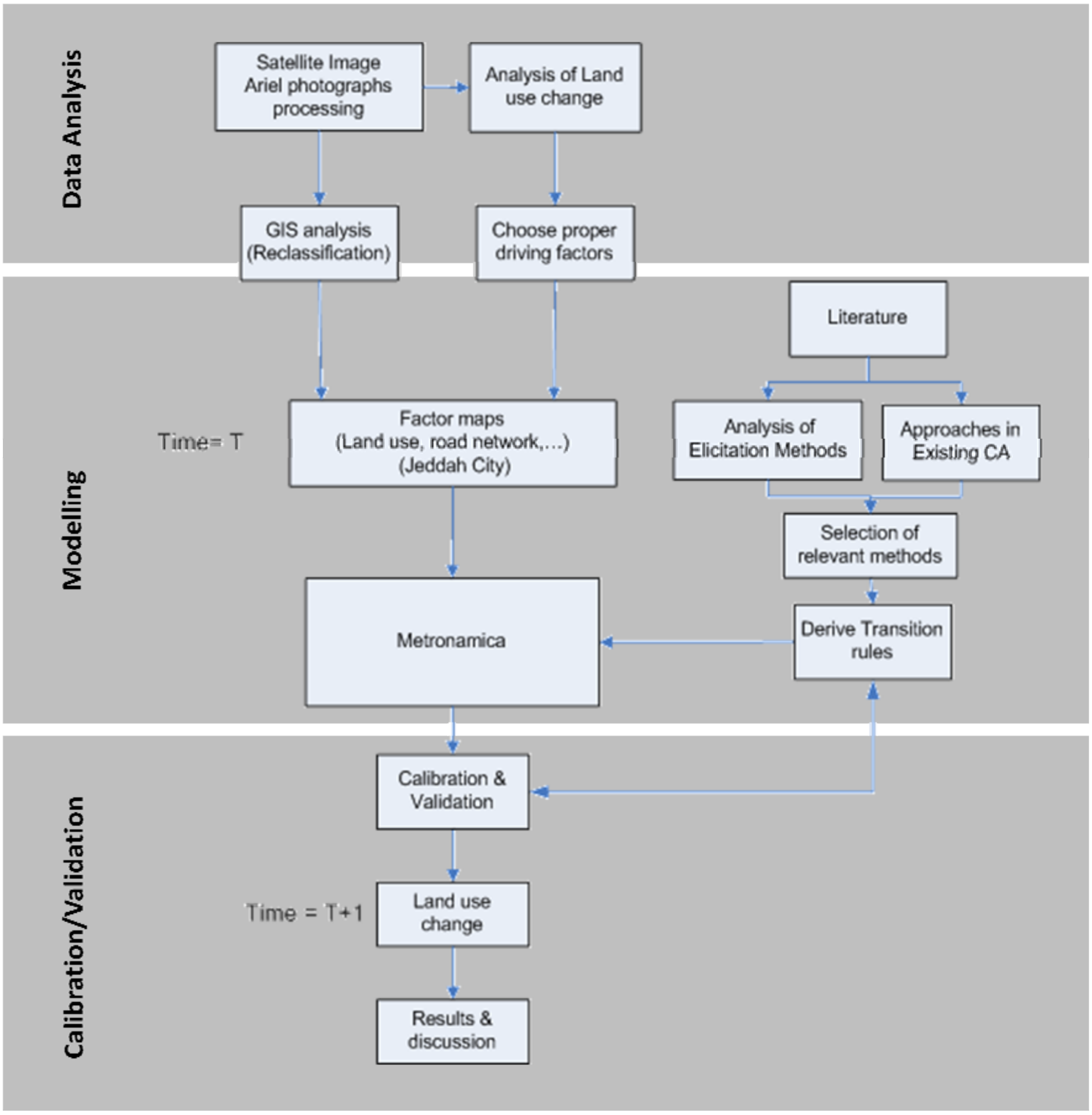


Figure 7: Conceptual Frame work

3.5. Transition rules elicitation methods used

As it is evident, transition rules are very central to cellular automata models. On the basis of neighbourhood interaction of different land uses types, certain rules need to be defined to accurately simulate the phenomenon of urban growth being the heart of CA model. An over view of advantages and disadvantages of different methods (already discussed under section 2.6) are given here under.

Table 1: Advantages and disadvantages of different methods

Method	Advantages	Disadvantages
Enrichment factor	data driven, suitable for multi-scale analysis, easy to understand and explain	May not always represent true behaviour. Depends upon data reliability
Regression Analysis	data driven, suitable for multi-scale analysis, easy to understand and explain	static approach not dynamic, May not always represent true behaviour. Depends upon data reliability
Artificial Neural Network	Can easily define parameter values,	Modeller and user do not know

Method	can handle incomplete data, generates good urban configurations	what is happening- a black box method. No room for human Interventions
Visual Observation	Modeller can explain the factors used, allows to experiment and study effects of factors	Time consuming and is Subjective, Full of uncertainty
Analytical Hierarchy Process	Includes human interventions and decision making process. Good for rules elicitation	Hard to reach all decision makers.

Based on the aforementioned methods, their advantages and disadvantages, two methods are thus proposed to be adopted, one that utilizes enrichment factor for calculating neighbourhood interactions of land uses and the other method is the combination of weights derived by analytical hierarchy process (AHP) and enrichment factor to shape the transition rules. The enrichment factor is chosen to be more suitable because it is data driven, suitable for multi-scale analysis, easy to understand and explain. Analytical Hierarchy Process includes human interventions and decision making process therefore very closely represents the urban development.

Both methods will be tested for the case of Jeddah city and results (land use maps) will be compared using Kappa statistics, as it is the commonly used method for map comparison (Van Vliet 2009).

3.5.1. Enrichment factor

Enrichment factor is used to characterize the neighbourhood of a location in a land use map which is defined as the occurrence of a land use type in the neighbourhood of a location relative to the occurrence of this land use type in the study area as a whole. The values in a neighbourhood rule are the cumulative effect of attractions and repulsions of land uses over distance. Enrichment factor has already been applied to arrive at more theoretically and empirically justifiable neighbourhood rules as well as spatial-temporal specific neighbourhood rules (Verburg, de Nijs et al. 2004).
Formula for calculating Enrichment factor is given as under.

$$F_{i,k,d} = \frac{n_{k,d,i} / n_{d,i}}{N_k / N}$$

F *ikd* is the enrichment factor of a neighbourhood *d* of location *i* with land use type *k.*
nk,d,i is the number of cells of land use type *k* in the neighbourhood *d* of cell *i*, *n d,i* is the total number of cells in the neighbourhood while *Nk* is the number of cells with land use type *k* in the whole raster and *N* is all cells in the whole raster.

This neighbourhood characteristic are calculated for every cell i in a series of enrichment factors for the different land use types (k). The calculation is repeated for different neighbourhoods located at different distances (d) from the grid cell to quantify the influence of distance on the relation between land use types. In this study we have used circular neighbourhood (White and Engelen 2000) at a distance d from the central cell as neighbourhoods. The shape of the neighbourhood can also influence the results obtained. The square shape causes differences in the distance between the neighbourhood and the central cell that has no theoretical validity (Verburg, de Nijs et al. 2004)

If *K* is the land use commercial and the distance from the central cell is d=3, then $F_{k,d}$ = 3.612046

This number (3.612046) is the average of enrichment factor of pixels (i). Here we only look at cells for which the land use type has been changed between 1980 and 1993. All the values given in appendices 3,4 and 5. Further discussion is given in section 5.2.

3.5.2. Analytical Hierarchy process method

In this method, a combination of three factors that are suitability, accessibility and neighbourhood is proposed. Analytical Hierarchy process will be used for pairwise comparison and weighting to quantify these factors. Expert interviews were arranged at the municipality level in Jeddah city to get the opinion of decision makers and experts about the importance of driving factors of urban growth.

In this method, a combination of three factors that are neighbourhood effect, suitability effect, and accessibility effect is used. The sub factors, also some time called as sub rule classes are slope, soil and elevation to measure suitability effect, distance to major roads, distance to secondary roads and distance to primary roads to measure accessibility. Pairwise comparison is used for averaging the weights of these factors and sub factors.

Neighbourhood: Neighbourhood effect will be calculated using development density (development density means the ratio of developed cells to the total cells in the neighbourhood. We can use this method to reflect neighbourhood effect.

Suitability: The term suitability is used to describe the degree to which an area is fit to support a particular land use function and the associated activity.

Based on literature review, the criteria for measuring the suitability effect are measured by factors like soil, slope, and elevation.

Soil: The soil shows the geology of an area such that whether it is fit for a particular development in future or not. For example hard rock is fit for multi story buildings as the chances of settlement or displacement is negligible. On the other hand, in clayey soil there are always chances of uneven settlement.

Elevation: It is also an important factor in determining the suitability of a particular land use for example residential areas at higher elevation may have inadequate facilities and also construction cost is higher. From the aesthetic sense, may be some people like to be at moderately high elevations.

Slope: The slope of the land surface is important because it determines the use of land and cost of development (Al-Ahmadi, See et al. 2009)

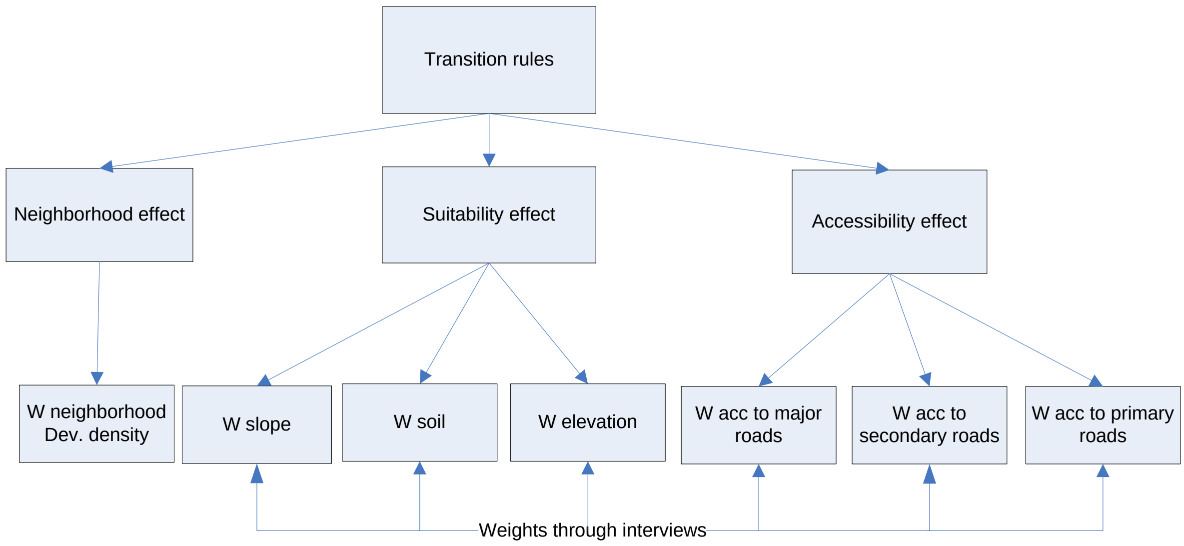


Figure 8: Criteria hierarchy structure of transition potential rule elicited by AHP method

Accessibility: Accessibility is the ability to get from place to place, the connectivity of a place with other places. Accessibility is the spatial constraint on urbanization. Better accessibility results in lower transportation (travel) costs. From a spatial perspective accessibility is the major factor in urban development. Accessibility effect can be measured by distance to major road, distance to secondary roads, and distance to primary roads etc.

Interviews

Interviews were conducted from the decision makers and urban planning experts about the importance of factors driving the urban growth in the city of Jeddah. Interviewees were requested to give weights of importance to these factors for the purpose of future planning using the 9 point Saaty scale 1980 given below, using the example of table 2.

9 point scale for pairwise comparison (Saaty, 1980)

Intensity of Scale	Meaning
1	Equal Importance
3	Weak
5	Moderate
7	Strong
9	Absolute
2,4,6,8	intermediate preferences

Table 2: Weights table

	suitability	Accessibility	Neighbour hood
suitability	1	2	5
Accessibility		1	3
Neighbour hood			1

For example in the table above, reading from the direction of column to row:

The weight of suitability to suitability is 1, which means equally important. Weight of accessibility to suitability is 2, which means accessibility is intermediately (between 1 and 3) more important than suitability. Weight of neighbourhood to suitability is 5, which means neighbourhood is moderately more important than suitability. The weight of accessibility to accessibility is 1, which means equally important. Weight of neighbourhood to accessibility is 3, which means neighbourhood is weakly more important than accessibility. The weight of neighbourhood to neighbourhood is 1, which means equally important.

Then the interviewers were asked to fill blank spaces (not in grey spaces) the relative weights of importance to the following sub factors of land use class, residential , Commercial and industrial, public places and green areas. Similarly they were also asked to fill in the blank spaces relative weights of importance to the suitability and accessibility factors. All the interview tables are given in appendices 6 and 7. The derived weights are then used in cellular automat model to calculate the transition potential of cells and then help in shaping the transition rules.

3.6. Calibration and validation

Calibration is carried out in order to improve the models goodness of fit through the adjustments of its parameters and to fit what is observed. For this procedure two sets of land use data are required, one for the start of the simulation period (time t) and one for the end of the simulation period (time t+1). We can do a calibration by simulating land use changes from time t to time t+1 and comparing the result with actual land use data (Van Vliet 2009).

Land use map of 1980 is selected as base year and the model will be run till 1993 because the city of Jeddah has experienced gradual urban growth during this period. The results will calibrated with the land use data of 1993. Then model will be run till 2007 and its results will be validated with the land use data of this year. The assessment of model can be performed by visual interpretation and Kappa statistics. Validation is the assessment of the models' quality or results to increase model's credibility, provide awareness of uncertainty, make it acceptable for use and confirm truth.

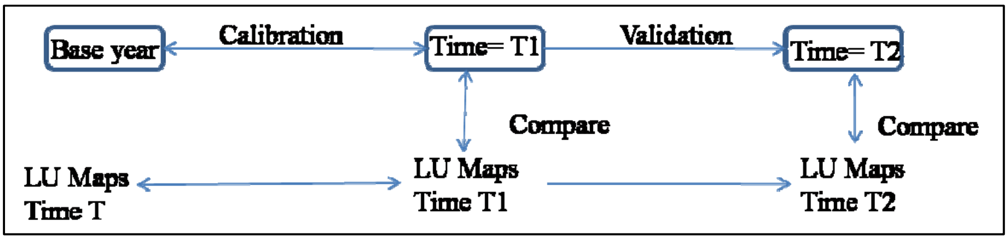


Figure 9: Calibration and Validation framework

3.7. Map comparison using Kappa statistics

In order to assess the explanatory power of land use maps, different ways of comparisons are possible. Visual interpretation is nevertheless a good way to assess simulation results, however it is highly subjective and can be performed by experts (Power, Simms et al. 2001; Hagen 2002).

Kappa statistics

The Kappa comparison method is based on a straightforward cell-by-cell map comparison (Van Vliet 2009), which considers for each pair of cells on the two maps whether they are equal or not. This results in a comparison map displaying the spatial distribution of agreement. Kappa statistics are very useful since they indicate the level of similarity of two maps in one number and uses pair wise comparisons. For calculating the Kappa values, the Map Comparison Kit will be used.

Kappa statistics are often applied to analyze observed and predicted results (RIKS 2005). It expresses the agreement between two categorical datasets. Its values range from -1 (for no agreement) to 1 (indicating a perfect agreement). The value 0 represents full agreement that can be expected by chance. Since land use maps are essentially categorical data sets, Kappa is well suited to compare a pair of land use maps. Therefore it is commonly used for the accuracy assessment of results of land use change models including Metronamica (Van Vliet 2009).The basis of Kappa calculation is the contingency table also called as confusion matrix. The table details how the distribution of categories in map A relates to that of map B. The cells shows a value which is the fraction of the cells in the map A, denoted by the category specified in the matrix row, and in map B by the category specified in the matrix column. For example, a value of 0.25 for t12 would indicate that 25 percent of the mapped area is of category 1 in map A and category 2 in map B (Hagen 2002).

This matrix tabulates for each pair of classes how often it occurs. A general form of contingency table is given here under. It is well established as a cornerstone of similarity assessment (Foody 2002). The last row and column give the column and row totals. Each row total represents the total fraction of cells of the related category in map A. Similarly each column total represents the total fraction of cells of the related category in map B. All fractions together makes up the whole map, hence the total sum is equals to 1 (Hagen 2002).

Table 3: Contingency table (Hagen-Zanker and Martens 2008)

Map A \ Map B	1	2	$\cdots$	c	Sum
1	t_{11}	t_{12}		t_{1c}	t_{1+}
2	t_{21}	t_{22}	$\cdots$	t_{2c}	t_{2+}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
c	t_{c1}	t_{c2}	$\cdots$	t_{cc}	t_{c-}
Sum	t_{+1}	t_{+2}	$\cdots$	t_{+c}	t_{++}

The following equations express how Kappa is calculated from the contingency table:

$$K = \frac{P(A) - P(E)}{1 - P(E)}$$

$$P(A) = \frac{1}{t_{++}} \sum_{i=1}^{c} t_{ii}$$

$$P(E) = \frac{1}{t_{++}} \sum_{i=1}^{c} \left(t_{i+} t_{+i} \right)$$

$$P(\max) = \sum_{i=1}^{c} \min\left(t_{i+} t_{+i} \right)$$

where K is Kappa and a value of 0 corresponds to the expected level of agreement, identical maps get value 1 and the lowest Kappa score is -1 (Hagen-Zanker and Martens 2008).

P(A) is the fraction of agreement and

P(E) is the expected fraction of agreement.

tij is the number of cells of class i in map A and class j in map B. ti+ is the number of cells of class i in map A. t+i is the number of cells of class j in map B. t++ is the total number of cells.
Kappa statistic covers up two types of similarities, similarity of quantity and similarity of location. Kappa quantity refers to the total number of cells taken in by each category found in the histogram and location refers to the spatial distribution of the different categories on the map. In order to recognize to which extent similarity of location and quantity are represented in the Kappa statistic, it is divide into two statistics: Kappa Histo (or KHisto) and Kappa Location (or KLoc).
KLocation is calculated according to Equation:

$$\text{Klocation} = \frac{P(A) - P(E)}{P(\max) - P(E)}$$

KHisto is defined by Equation

$$\text{Khisto} = \frac{P(\max) - P(E)}{1 - P(E)}$$

So Kappa is now defined as the product of two factors, the Klocation, which is a measure for the similarity of spatial allocation of categories of the two compared maps and the KHisto, which is a measure for the quantitative similarity of the two compared maps.

Kappa = KHisto * KLoc

Kappa, as well as KLoc and KHisto are calculated on the basis of the Contingency Table, which details the cross-distribution of categories on the two maps.
Interpretation of Kappa is explained here under.

	Poor	Slight	Fair	Moderate	Substantial	Almost perfect
Kappa	0.0	.20	.40	.60	.80	1.0

(Anthony J. Viera and Joanne M. Garrett 2005)
Kappa Agreement

< 0 Less than chance agreement	0.01–0.20 Slight agreement
0.21– 0.40 Fair agreement	0.41–0.60 Moderate agreement
0.61–0.80 Substantial agreement	0.81–0.99 Almost perfect agreement

4. Spatial Temporal growth of Jeddah city

Remote sensing techniques are extensively used in mapping urban land use/land cover, urban growth trends, and to elaborate the changes in land use/land cover (Pathan, Sastry et al. 1993). It can provide useful information about land use patterns and their changes and capable of detecting and measuring a variety of elements relating to the morphological traces of cities (Yeh and Li 2001).

Jeddah is a coastal city and an important business centre of the Kingdom of Saudi Arabia. More than 90% of Jeddah was developed after 1948 because of the expanding Saudi economy which has surplus housing and infrastructure at present (Al-Hathloul and Mughal 1991).

The City has grown from a small primitive fishing settlement of 100 ha to a modern beautiful metropolis covering an area of 30800 hectares in 1987. The shape of the city has been dictated by the topography and the attraction of the water front on the Red Sea coast. However, some policies and decisions have also been instrumental in giving the city its present shape for example the sitting up of the new airport to the North. The airport is a very strong determinant of a busy communication corridor parallel to the coastline. The location of the new Jeddah-Mecca expressway to the east of the city has acted as a constraint on the growth of the city in that direction.

The spatial temporal analysis is carried out using the aerial photographs of 1970, 1980 and spot image of 20 meter resolution. After digitization, the land uses were allocated using the master plans of the corresponding time period. It is worth mentioning that master plans are the documents for possible future development and may not exist on ground. In that case we took help from the image or photograph whether a particular built up area exists or not on the ground. The land uses are purely based on master plans.

4.1. Driving factors of urban growth in Jeddah

To know about the urban growth of Jeddah city over the past forty years, an analysis of data is important in the sense which years to be selected for modelling the growth. How the land use has changed? Which classes of land use have changed the most and where are these changes? Whether there is a homogeneous development in the entire period or if there are some breaks? Hence a period of homogeneous development should be taken for modelling. The choice of driving factors is based on the characteristics of urban growth in Jeddah city based on the population density, historical documents, master plans, zoning maps (policies), topographic constraints, scientific literature and data analysis.

4.1.1. Population growth

Population in Jeddah has increased very quickly and this development has stimulated the growth of construction in Jeddah. This growth, has led to an increase in the number of residential moves either between neighbourhoods within the city or within neighbourhoods. This trend of citizens changing houses has contributed to significant expansion in construction and rapid growth. This phenomenon is more evident in the north of Jeddah. The exceptionally high rate of migration, attractiveness for living due to proximity to the water and flexible housing supply from time taken to plan and build new house have made Jeddah municipal management very problematic. The huge population growth has influenced the social development in the city of Jeddah in particular and in the Kingdom of Saudi Arabia in general (Al-Otaibi 2006).

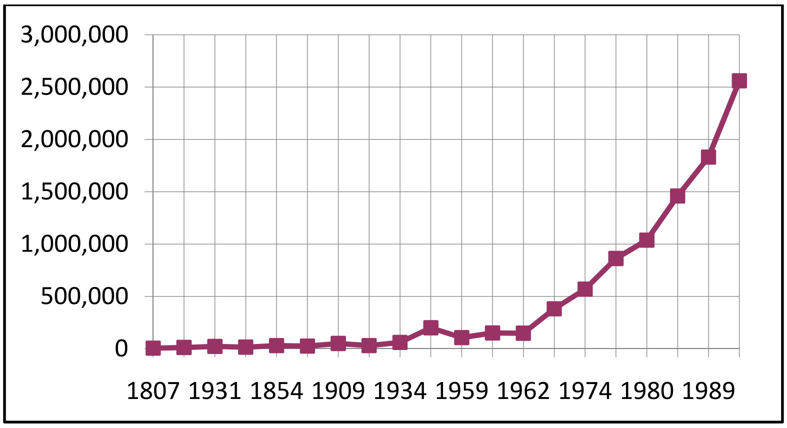


Figure 10: Population growth of Jeddah city through ages Source (Al-Otaibi 2006)

4.1.2. Slope

The mountains to the East of Jeddah have restricted its growth in the Eastern direction. Therefore the city grew in the North South direction.

4.1.3. Establishment of new airport

One of the main causes behind the growth of Jeddah city in the North direction is the location of new international airport. The roads leading towards the airport are the main hub of development.

4.1.4. Roads

The analysis of data shows that road infrastructure is one of the main driving factors of urban growth in Jeddah city. The analysis shows that Makkah road in the East and Madina road in the North support the main development of Jeddah city. The main feature of the urban development after 1970, are the Eastern and Northern development patterns, aided by the Madina-airport-Jeddah-Makkah expressway. The expressway provided an impetus for development. The industrial areas are mostly located in the south of city where as the new residential are constructed in the North which could be because of the new location of airport.

4.2. Land use change between 1970 and 1980

In 1970 the total built up area was 3899 hectare whereas 1980 it was increased to 29856 hectares which was near to the old airport, ring road and there was a huge flux of development to the North side and towards the Eastern corridor along Makkah Road.

Table 4: Land uses and areas for the years 1970, 1980, 1993 and 2007

Land use	Area(ha)2007	Area (ha)1993	Area (ha)1980	Area (ha)1970
Residential	21357	16030	8851	2339
Commercial	678	662	307	302
Industrial	4420	2655	1859	30
Informal settlement	3035	1615	1793	224
Public place	3902	2535	3128	678
Green area	523	299	396	1
Port	3663	3422	3364	192
Undeveloped	66996	74359	81213	119997

Vacant	10501	13684	14286	1436
Airport	10159	10159	10159	133

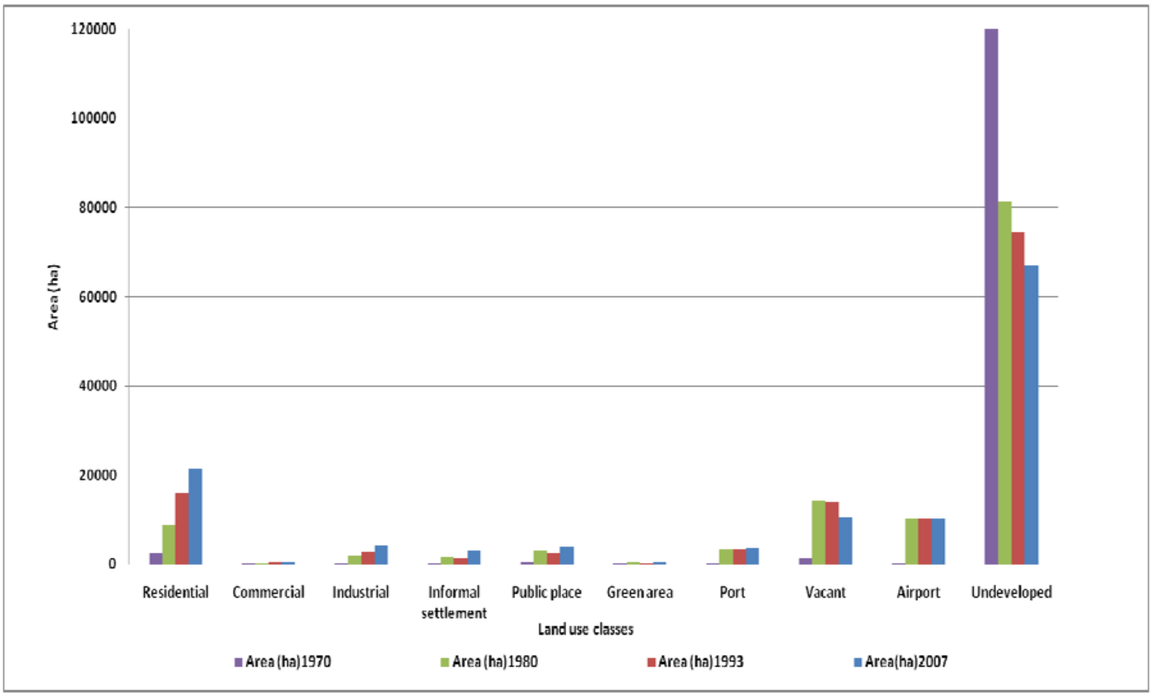


Figure 11: Graph showing the comparison of land uses during 1970, 1980, 1993 and 2007

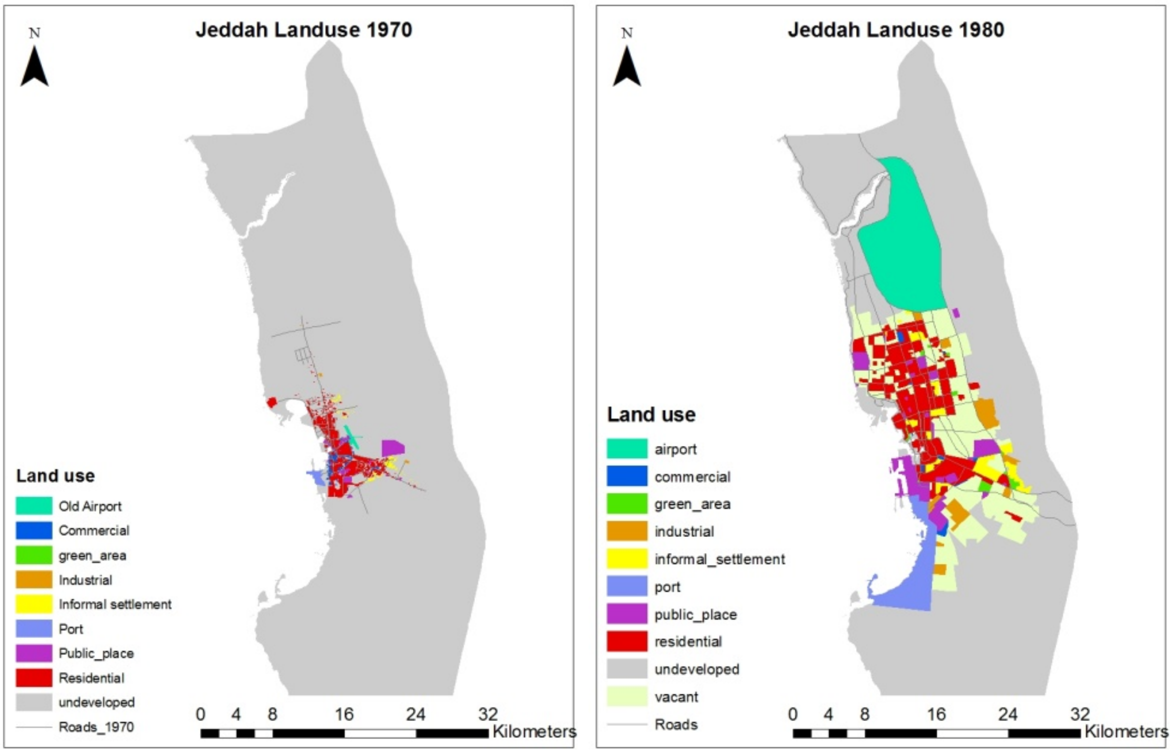


Figure 12: Maps showing Land use of Jeddah in 1970 and 1980

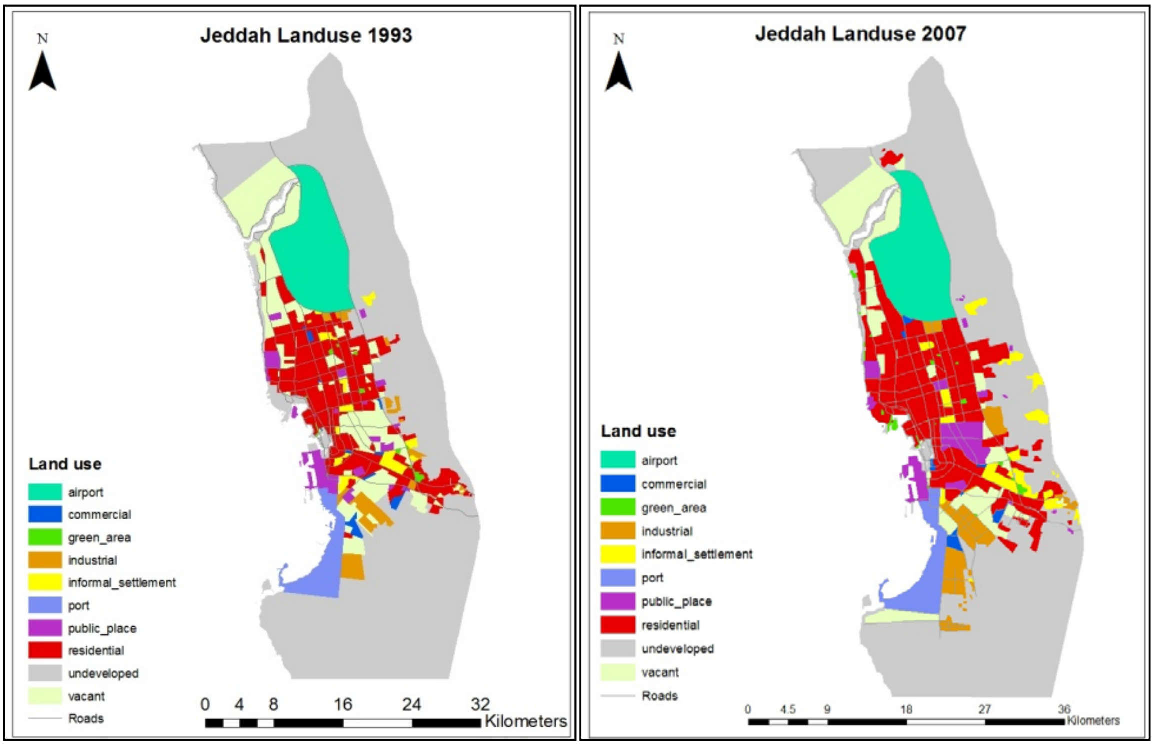
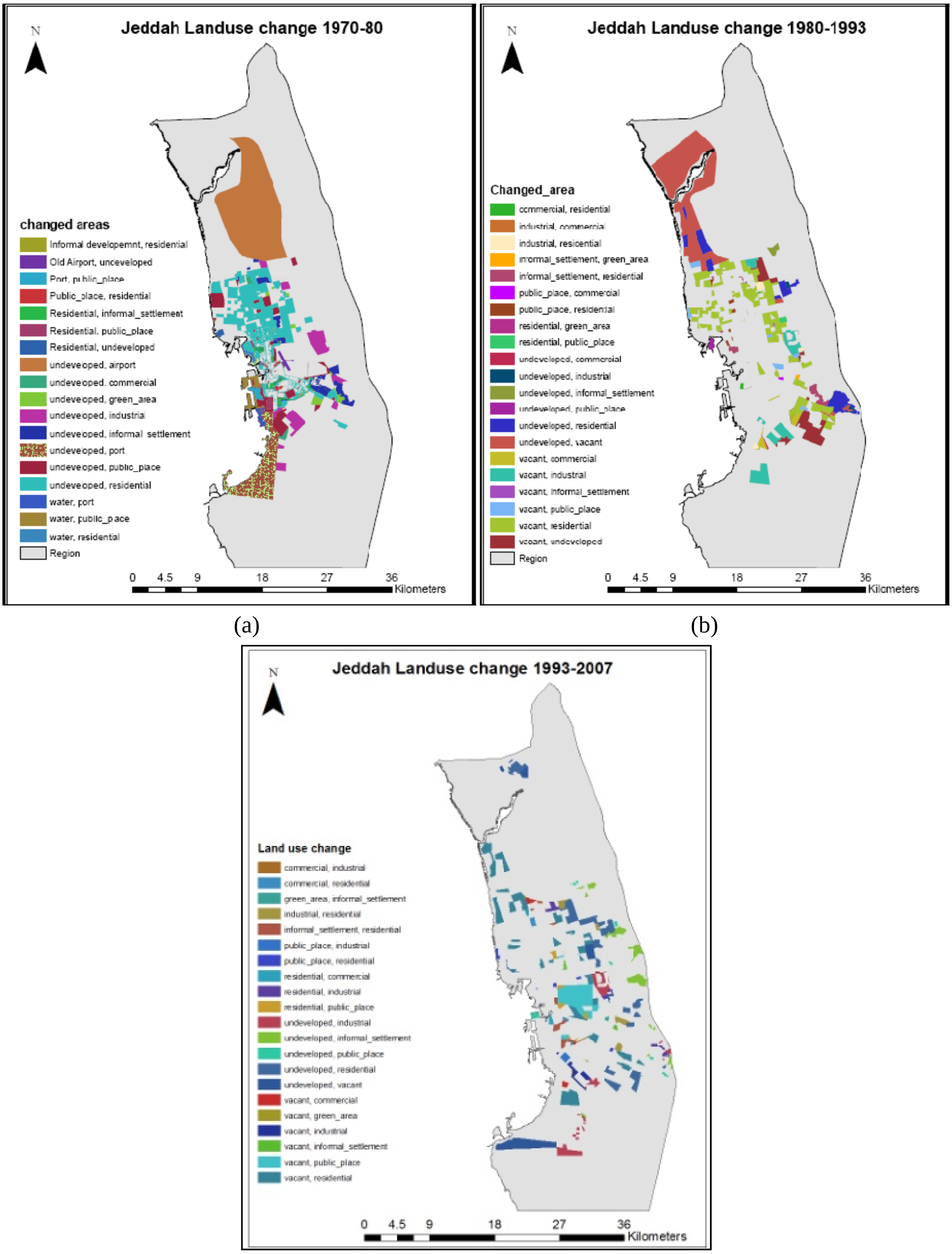


Fig: 13 Maps showing land use of Jeddah in 1993 and 2007



(a) (b)

(c)

Figure 14: Maps showing growth change of Jeddah (a) 1970-80, (b) 1980-93 and (c) 1993-2007

By looking at the analysis from the aerial photographs between 1970 and 1980, the above maps shows that most of the informal areas remain as informal while some 89 hectares changed to planned residential out of total 224 hectares. Commercial area remained the same. Location of new airport changed and therefore became vacant. Residential areas mostly remained the same but about 257 hectares changed to public place. Some water areas had been reclaimed and converted to port, public places and residential areas.

4.3. Land use change between 1980 and 1993

In the time period between 1980 and 1993 most of the vacant areas were developed into commercial, residential and public places. We also see some small informal areas pop up. Vacant areas near to existing industrial also converted to industrial which was about 1613 hectares. By vacant areas here we mean those areas either inside the city areas or those bounded by one side or very close to the existing built up figure 16 (b).

4.4. Land use change between 1993 and 2007

Looking at the change between 1993 and 2007 figure 16 (c), there is very little change in commercial areas, because it is not easy to recognize commercial areas. After digitization, the land uses were allocated using the land use map of 2009 as a reference, but we could not observe many changes in commercial areas which do not look realistic even after 14 years in such a fast developing city. Residential areas show substantial growth over this period from 16030 hectares to 21357 hectares. These areas developed along the Al Haramain expressway in vacant areas near to the existing residential. Residential also expanded along the Al Haramain expressway and Makkah road in the south eastern part. The informal growth is one of the prominent land uses that grew over during this time. These are visible in the Eastern part of Jeddah city that increased from 1615 hectares in 1993 to 3035 hectares in 2007. The growth of industrial areas is very prominent and sufficient in the southern part of Jeddah city close to the existing industrial areas showing an increase from 2655 hectares in 1993 to 4420 hectares in 2007. Public places show substantial increase from 2535 hectares in 1993 to 3902 hectares in 2007 which include health, educational facilities and other government offices and installations meant for the public. It means that with increase in residential, government also provided public amenities.

5. Results and discussion

5.1. Model set up

METRONAMICA is a CA based modelling environment used to develop land use change models. It is a generic and very flexible modelling environment which can be applied in a variety of spatial and temporal resolutions (RIKS 2005) as already tested in a number of projects.

In this study, METRONAMICA single layer (S.L) model is used, containing the land use model as a single model (RIKS 2009) which represents the whole Jeddah city as one single region. Nevertheless it has the constrained CA characteristic and uses a simple file specifying the number of cells required to be in each state at each time period (White et al., 2004).The cell space of the model consists of a rectangular grid of square cells with a resolution of 100 meter. The grid size contains 755 rows and 408 columns. Sitting up a new project in METRONAMICA application, requires many input maps and information. In the next sections, those requirements and information required are described in detail.

5.1.1. Land use maps

Land use maps of 1980 and 1993 already shown in figure 13 and 14 are used as base year and target year maps respectively. They are of the same dimensions and contain the following classes given below. The first category contains vacant areas present in the centre of the city or those bordering the existing built up areas. The second land use class is the undeveloped area which also belongs to vacant category. It comprises of non built up areas lying within the modeled area but outside the built up areas. Residential, industrial and commercial land use classes are the functions which are to be simulated. The feature category of the model includes the airport, green areas, informal settlement, port, public palace and outside simulation areas. These are not meant for simulation.

Land use classification used in Metronamica

Land use Id	Land use class	Category
0	Vacant	Vacant
1	Undeveloped	Vacant
2	Residential	Function
3	Commercial	Function
4	Industrial	Function
5	Airport	Feature
6	Green area	Feature
7	Informal settlement	Feature
8	Port	Feature
9	Public place	Feature
10	Outside simulation area	Feature

5.1.2. Suitability Maps

Suitability maps were prepared using Community-Viz planning support system. A tool called the Scenario 360 – an extension of Arc Map software. The suitability factors like the soil, slope and elevation of Jeddah city were used for suitability calculations. The criteria used for slope and elevation are that the lower the slope and elevation, the better the areas are for development. Most of soils are suitable. Also the reclaimed soils were taken as suitable. Only soil with very hard rocks and sand dunes were unsuitable. The suitability of soil is based on the data provided by the municipality of Jeddah. Same factors were used for all three types of suitability layers - the residential, commercial and industrial. During the preparation of suitability maps, the weights derived through AHP method for various factors for suitability were also used.

5.1.3. Zoning maps

Zoning maps were created from the land use maps using Arc map. Air port, port, public areas, and green areas are included as restricted areas. The green colour shows the areas allowed for development and the red colour areas are not allowed for development. For all three land use types- residential commercial and residential, zoning restriction were the same and same map is used in the model. As we did not have information regarding which areas are to be developed in phases during the simulation period, zoning maps has only two classes, allowed and not allowed.

5.1.4. Region Map

The whole process in Metronamica corresponds with a region map for the study area. For this purpose we used a shape file of district boundaries of Jeddah city and made a union with a polygon shape file in order to have the same extents for our maps going into Metronamica. The district boundary is rasterized with value 1 which means simulation allowed and the value 0 means outside simulation area. As we are working with Metronamica single layer there for, it required only one region map.

5.1.5. Accessibility Map

Metronamica single layer uses one road network map in a shape file format. To that end the road network map consists of three classes. The major roads which comprise of hi ways, one coming from the East side and another one passing through the city centre following the airport on western and Eastern side. There seems no obvious difference between the secondary and primary roads with the only difference that secondary roads run parallel to major roads in the North South direction while primary roads run approximately perpendicular to them.

5.1.6. Land use demand

For the initial set up of model, the land use demand used, is the number of residential. Commercial and industrial cells in the 1980 and 1993 land use maps. The demand is given here under:

Land use	1980 demand (cells)	1993 demand (cells)
Residential	8863	16051
Commercial	302	657
Industrial	1849	2637

5.2. Enrichment factor results

Enrichment factor is a method for exploring and quantifying the neighbourhood characteristics of land use and the interaction of land use types. The results are then used to help defining the transition rules for a cellular automata model. Neighbourhood effects alone are, however, not only the location factor suitable for describing the spatial pattern of land use. Other factors, such as accessibility, environmental suitability, spatial policies etc. do also influence the pattern of land use (Verburg, de Nijs et al. 2004). It is not possible to directly translate the neighbourhood characteristics into transition rules for cellular automata but it is used as a starting point for defining the transition rules in the CA model. All neighbourhood effects identified with the enrichment factor do not have a causal explanation, because some of the observed neighbourhood effects are the indirect result of other interactions (Verburg, de Nijs et al. 2004).

Results for all neighbourhoods are given for three major land use types in Figure (17). These graphs present the average enrichment factor ($F_{i;k;d}$) as a function of the distance. The enrichment factor is presented at a logarithmic scale to obtain an equal scale for land use types that occur more than average in the neighbourhood (enrichment factor >1) and land use types that occur less than average in the neighbourhood (enrichment factor <1). If the value for a land use type exceeds 1, this indicates an enrichment of the neighbourhood. If the value is between 0 and 1, a less than average occurrence of the land use type is present in the neighbourhood. For instance if the neighbourhood of a certain grid cell contains 15% industrial whereas the proportion of industrial in the country as a whole is 5% we characterize the neighbourhood by an enrichment factor of 3 for industrial. When the proportion of a land use type in the neighbourhood equals the national average, the neighbourhood is characterized by an enrichment factor 1 for that land use. The average neighbourhood characteristic for a particular land use type is calculated by taking the average of the enrichment factors for all grid cells belonging to a certain land-use type. They are shown as a function of distance when they are calculated for different neighbourhoods. Eight specific neighbourhoods are used, each of 100 x 100 meters resulting in a maximum distance of 800 meters, as shown below in figure (16).

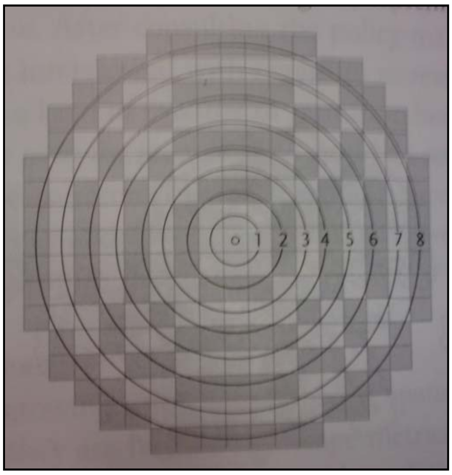


Figure 15 : Circular neighbourhood

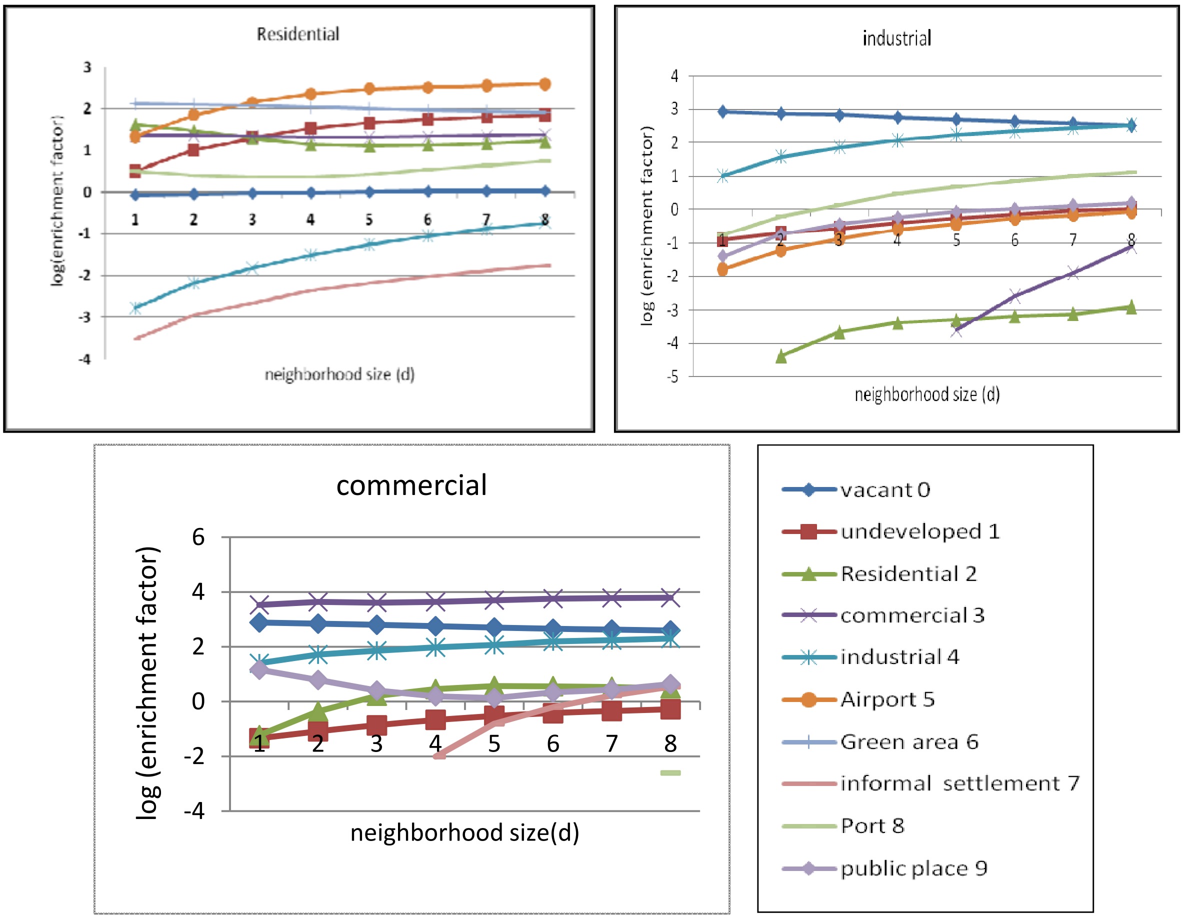


Figure 16: The logarithm of the enrichment factor

The logarithm of the enrichment factor (log (F *i;k;d*) as a function of the distance d of the neighbourhood from the central cell (see Fig.17). Each of the three graphs indicates the influence (neighbourhood characteristics) of other land uses on Residential, commercial and industrial land use.

5.3. Analysis of enrichment factor

5.3.1. Residential

Here we discuss the neighbourhood interactions obtained through the enrichment factor, between residential land use and all major land uses, foreseeing from residential point of view and explain their net interaction over distance.

Residential

The graph for residential areas shows that in its neighbourhood, the attraction for residential areas is positive and higher at a distance of 100 meters but starts decreasing at 200 meters till 500 meters distance. Again it starts attracting residential areas at far distances. But if we look at the influence line no significant variation exists, hence we can say that people in Jeddah wants to live close to each other for instance they have relatives and especially large families after division want to live close to their ancestors. Also they want to benefit from the existing infrastructure, public facilities, and commercial facilities could be the possible thresholds.

Commercial

The commercial areas in the neighbourhood of residential areas show a uniform positive attraction at all calculated distances. It means that people wants to live closer to commercial facilities to fulfil their daily needs. Hence their influence on residential areas is uniform throughout the distance.

Industrial

The industrial areas in the neighbourhood of residential areas show a uniformly decreasing repulsion at all calculated distances. Some of the negative effects of industrial areas are their unattractive location, noise, parking, congestion, smell, and all other dangers associated with them. However this influence is constantly decreasing but could not become an attraction even at 800 meters distance but looks like that it may become neutral after 1 km distance.

5.3.2. Commercial

Here we discuss the neighbourhood interactions obtained through the enrichment factor, between commercial land use and all major land uses, foreseeing from the commercial point of view and explain their net interaction over distance.

Residential

The graph for commercial areas shows that in its neighbourhood, residential area is negative showing repulsion starting from 100 meter and becomes neutral at 200 meter distance from the commercial area. It means that residential areas do not attract commercial areas between these distances. The repulsion within two hundred meters may be that people wants to live at farther distance from the busy markets because there is lack of parking facility and probably the congestion and noise near the markets and also people go homes early and then they don’t want any disturbance from the commercial markets. The residential areas then attract commercial from 200 meter onwards and increases till 500 meters distance, but declines we go farther. Residential areas are great assets for commercial which increases the customer base for commercial retail.

Commercial

In Jeddah commercial areas attract commercial within its neighbourhood at all calculated distances because these benefit from each other. Customers visit all the commercial retail and offices within walking distances. The private companies like to be close to other company even if they don’t know each other. Provision of basic amenities required for commercial areas provided by the municipality is also encouraged by government.

Industrial

There is an increasingly constant attraction between commercial and industrial areas. The industrial areas want closer to the commercial for their goods delivery and facilitate their employees to have easy access to commercial retail.

5.3.3. Industrial

Here we discuss the neighbourhood interactions obtained through the enrichment factor, between industrial land use and all other major land uses, foreseeing from the industrial point of view and explain their net interaction over distance.

Residential

Because of the noise and air pollution caused by industries, people do not like to live close to industrial locations. The graph for industrial areas shows that in its neighbourhood, residential area is negative showing repulsion throughout the calculated distances and do not like industrial close by this effect decreases with increase in distance. However neutral effect may be likely to occur beyond 1 km distance.

Commercial

Commercial areas show negative influence and there is repulsion between the two land uses. The calculated effect of commercial on industrial at a distance of 500 meter starts decreasing rapidly till 800 meters distance but seems to become neutral after 1 km distance. In spatial planning commercial areas are located near the industrial areas and vice versa. But in the case of Jeddah it seems that commercial areas are located far from commercial areas.

Industrial

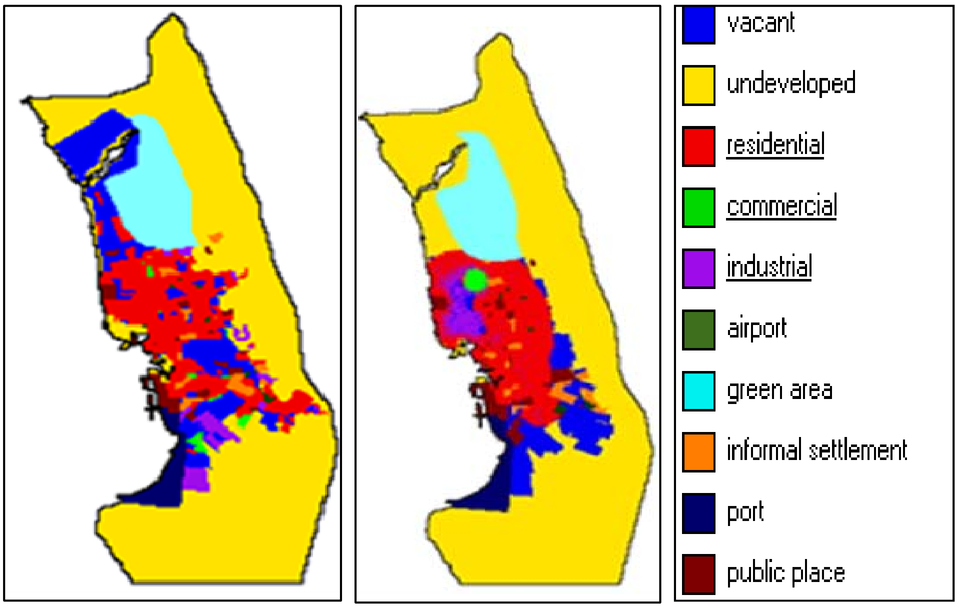
Industrial areas always attract other industries, the reason could be that road infrastructures and other amenities are provided by the local government or municipality in accordance with future expansion. Industrial plots in the existing industrial estate are cheaper and also enjoys some government tax waive off. Also to minimize nuisance and concentrate the traffic they are built close to each other. From the graph it is evident that industrial areas attract industrial having low attraction at near distance that increases with increase in distance.

However the enrichment factor outcome gives us a start to shape the transition rules but cannot be used directly as the calibrated rules in the model, because the neighbourhood effects found, are calculated over a period of 13 years. These effects are however non-linear over time and therefore not derived on yearly basis. So the calibration can be done through a qualitative approach to improve the simulation results and this can be based on the knowledge acquired through spatial temporal changes happened during the simulation period.

5.4. Simulation1 with actual enrichment factor

In Metronamica, the rules table is given where the calculated enrichment factor are put up to 8 cells. In simulation 1 the actual values of log enrichment factors were used as transition rules. The simulated and actual land use map of 1993 is given in figure (18) below. The comparison results show an overall agreement shown by Kappa of 0.64666. Kappa location of 0.69659, shows that 69 percent of cells are located at the right location. Kappa histo of 0.92833 identifies that 92 percent cells in total have been allocated during simulation table 5 and 6.

But the kappa statistics of simulated land uses show that in simulation 1 (before calibration), 51 % of residential cells have full agreement. Commercial areas show 12 % agreement because in southern part of the city there are commercial areas in the actual land use map but disappeared during simulation. Industrial areas show total disagreement. Because there are no industrial areas in the city centre but they are popped up during simulation. Another reason is that in the southern part of the city there are industrial areas but they vanished during simulation.



Actual land use 1993 simulated land use 1993

Figure 17: Results with actual enrichment factor

5.5. Calibration and assessment of the model

Simulation 2

To simulate known historical land use changes, Land use models need to be calibrated. Calibration is the process of adjusting parameter values to improve a model's goodness of fit. Another process is the validation which is the assessment of this goodness of fit with a dataset of another time period. Essentially, calibration and validation are iterative processes that works well when good calibration results are obtained (Van Vliet 2009). During the calibration process, transition rules were manually adjusted keeping in view the actual situation of land uses on ground in Jeddah city.

For the assessment of the method, an objective and reliable method is required to explain the model results after simulation. Land use maps among other aspects, are assessed at the pixel level. To this effect, contingency table is the base for the pixel by pixel similarity assessment which leads us to the calculation of Kappa coefficient of agreement (Van Vliet 2009). For the accuracy assessment of results of spatial models (Monserud and Leemans 1992) and remote sensing image classification (Foody 2002) , it is the most commonly used method.

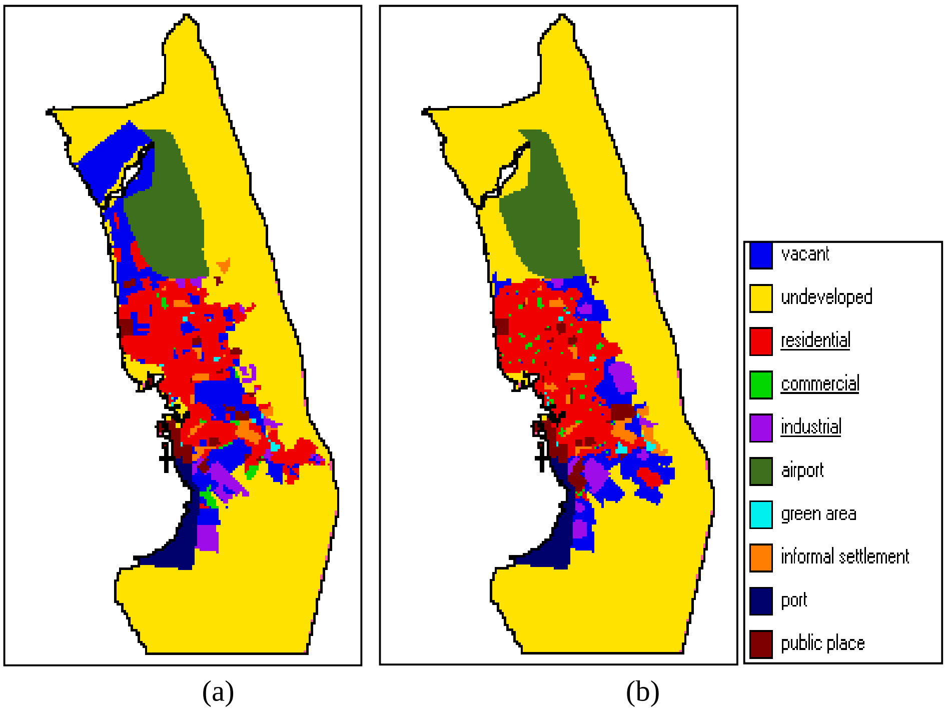


Figure 18: Actual land use map of 1993 (a) simulated map after calibration (d)

Table 5: similarity assessment results for sim 1 and sim 2, using kappa, Kloc & Khisto

Comparison	Kappa	KLocation	KHisto
result with actual enrichment sim 1	0.64666	0.69659	0.92833
result after calibration sim 2	0.70694	0.77281	0.91476

Table 6: Kappa statistics for major land use, simulation 1(before calibration) and 2(after calibration)

Comparison	Residential	Commercial	Industrial
Kappa before calibration Sim 1	0.5103	0.12783	-0.02157
Kappa after calibration sim 2	0.66222	0.29002	0.54713

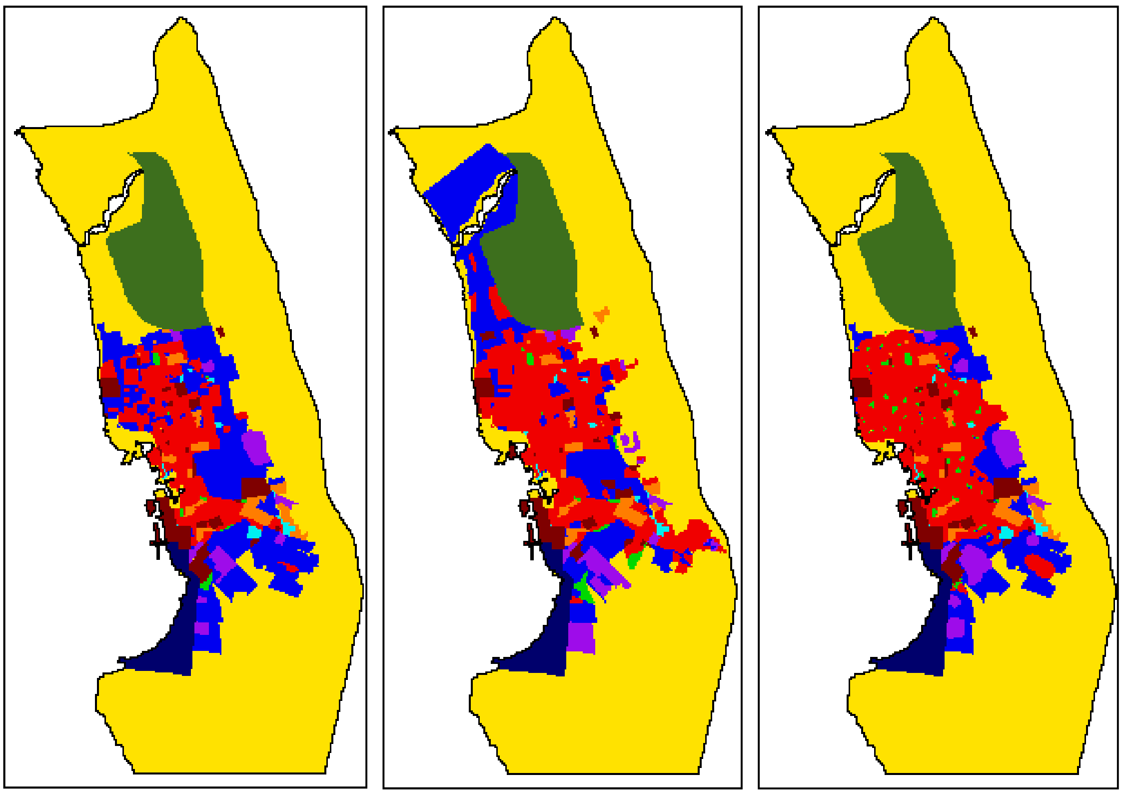
Table 5 above shows the comparison of two simulations using the actual values of enrichment factor as transition rules and then the calibrated rules. Looking at the statistics, both simulations show substantial agreement between the actual land use map of 1993 and the simulated land use map for the same year.

The kappa value for the land use class residential is higher than commercial and industrial found after the calibration. In this calibration, the expansion of residential areas is only based on the neighbourhood of the cells which results in a large and rounded cluster. But in reality, some large and smaller complex shapes are there in the actual land use map 1993 especially in the Eastern part along the Makkah road, but not very close to existing residential areas. However the large cluster after the calibration shows that a substantial number of the residential cells are exactly at the right location which results in high kappa values of 0.66. Much of the industrial and commercial areas are wrongly allocated that's why we have lower kappa values. Hence to increase the kappa, sufficient number of cells has to be exactly located at right place.

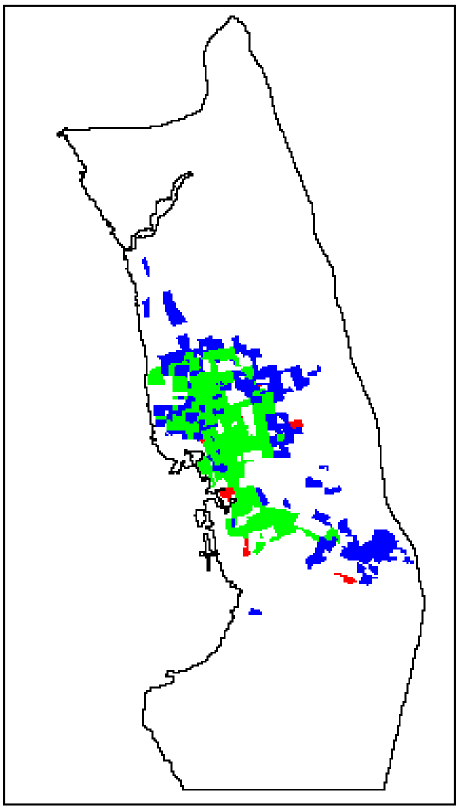
5.6. Visual interpretation after calibration

The predicted land use change can be compared with the observed actual change from 1980 to 1993. Through visual interpretation, a clear picture could be derived about the capability of the model that how it predicted the change. Figure 20 gives a clear idea about the actual land use changes from 1980 to 1993 for major land uses of our interest, the residential, commercial and industrial for the same period of simulation.

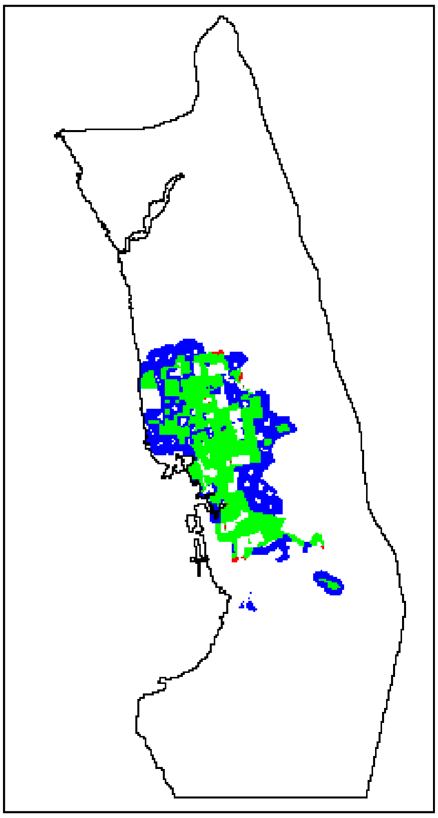
From the maps shown in the fig. 20, it is evident that the model performed well and created the same structures of change as the actual change for the residential land use class. This is also supported by the kappa for residential class which is 0.66 (see table: 6) showing substantial agreement. Furthermore, the predicted cluster sizes of residential areas look somehow similar to those of the actual change except for south eastern part along the Makkah road. That is the reason Kappa shows a low value. The commercial areas are match poorly as compared to the actual change. This is also indicated by the lower kappa values. The industrial areas show quite good similarity when compared to the actual change between 1980 and 1993. This is also indicated by a kappa value of 0.55 for industrial.



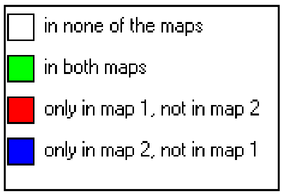
Map 1(1980) Map2 (1993 Map3 (1993

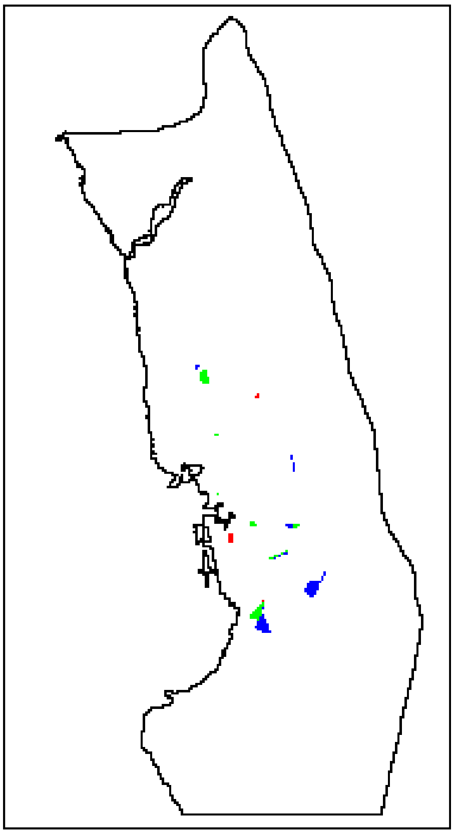


Real change Residential

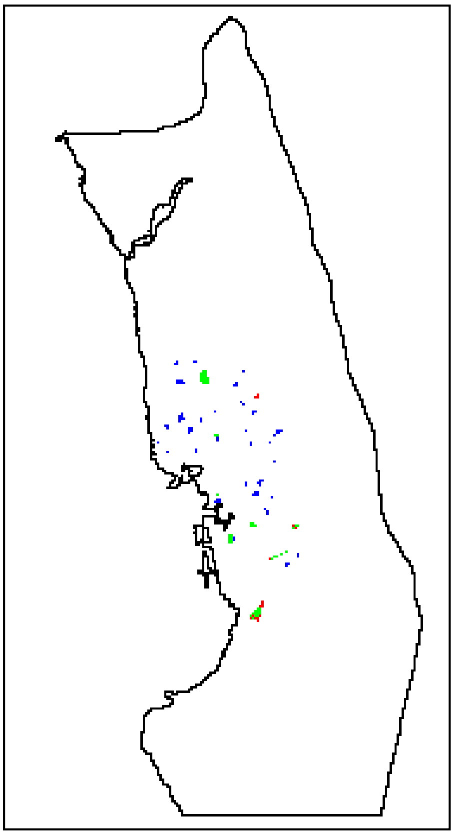


Predicted change (residential)

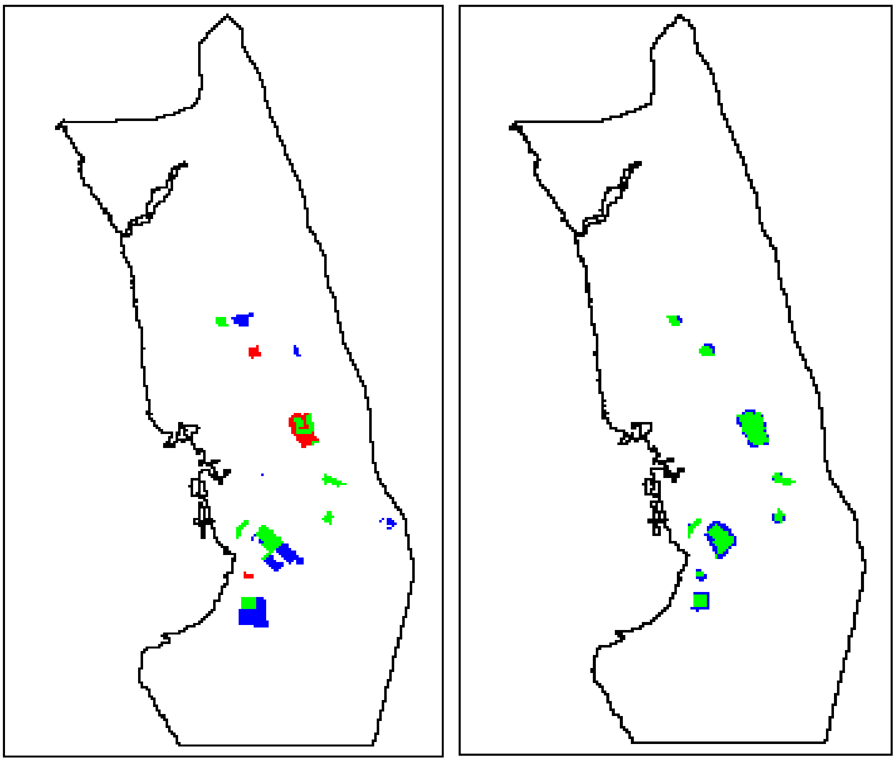




Actual change Commercial



Predicted change Commercial



Actual change industrial Predicted change industrial

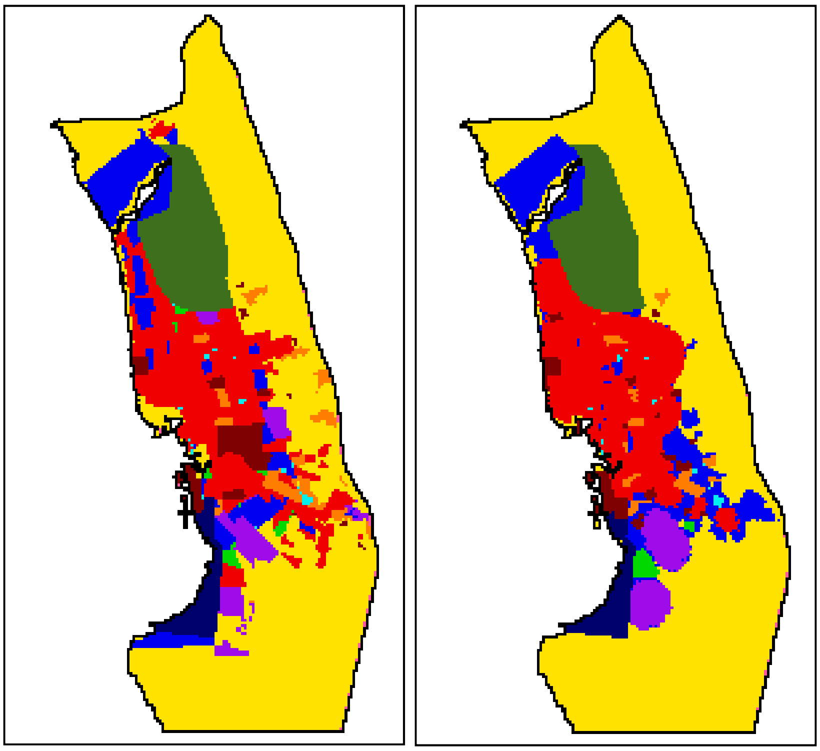
Figure 19: The actual and the predicted land use change for the period 1980 - 1993

5.7. Model validation

The land use maps of 1993 and 2007 have been used as the base year map and target map respectively for the validation period. The number of land use classes and their respective categories are the same as used during the calibration period, explained in section 5.1.1 above.
For setting up the model for the validation period, the land use demand used, is the number of residential. Commercial and industrial cells in the 1993 and 2007 land use maps as shown below.

Land use	1993 demand (cells)	2007 demand (cells)
Residential	1605	21341
Commercial	657	685
Industrial	2637	3874

The validation of rules show significant results as determined by comparing the results with that of the results of calibration period between 1980 and 1993 (see table 7 and 6). The residential and industrial areas are quite significantly simulated showing a higher kappa than that of the calibration period for two land uses. However, commercial areas show the same results because of very low change between the validation period. This is due to the reason that during map preparation of 2007, it was difficult to recognize commercial areas from the image of 2007. Also it was less shown in the reference map.



(a) (b)

Figure 20: Actual land use map 2007 (a); Simulated land use map 2007 (b)

Table 7: Kappa statistics for calibration

Calibration			
	Residential	Commercial	Industrial
Kappa	0.69329	0.46862	0.52742
KLoc	0.69329	0.46862	0.52742
KHisto	1	1	1

Table 8: Kappa statistics for validation

validation			
	Residential	Commercial	Industrial
Kappa	0.71932	0.46897	0.5474
KLoc	0.71942	0.46931	0.58793
KHisto	0.99986	0.99927	0.93107

The comparison of the results between the calibration and validation period shows a slight increase of performance over the validation period for the simulated land uses. This shows that the over calibration of model was not done between 1980 and 1993, others wise validation would have given lower results. However, most CA land-use models are not likely to perform very well in validation (Pontius, Huffaker et al. 2004).

5.8. Method combining AHP derived weights for accesibility and enrichment factor

Analysis of weights

The preferences are obtained from the decision makers and experts of Jeddah municipality in the form of weights. The average AHP weights for suitability, accessibility and neighbourhood for the major land use classes are obtained through pairwise comparison (see table 9). These weights show similarity of preferences for all mentioned land use types. However, for the accessibility classes, major roads have the highest importance than secondary and primary roads.

The simulation was run for the period between 1980 and 1993 with the combination of actual enrichment factor derived rules and the weights derived through analytical hierarchy process for the accessibility only (see table 9). The accessibility has three types of roads and their weights are given below. We used same weights for residential commercial and industrial land uses. It is worth mentioning that the transition potential formula which uses weights for suitability, accessibility and neighbourhood is not editable in this version of Metronamica, therefore we could not use the full AHP weights in transition potential calculation. However the derived weights for suitability were used in the preparation of suitability maps which is an important input in Metronamica model.

Simulations were run in combination with the enrichment factor derived transition rules before and after the calibration. The weights for three types of road classes obtained through AHP are given as under.

Major roads = 0.65

Primary roads = 0.085

Secondary roads = 0.265

Table 9: Average Analytical hierarchy process (AHP) Weights of different factors obtained through pairwise comparison

S= suitability, N=neighbourhood, A= accessibility

	Residential			industrial			commercial			public places			green areas		
Factors	**S**	**A**	**N**	**S**	**A**	**N**	**S**	**A**	**N**	**S**	**A**	**N**	**S**	**A**	**N**
AHP weights	0.722	0.202	0.076	0.721	0.186	0.093	0.758	0.176	0.066	0.748	0.189	0.064	0.705	0.214	0.081

	suitability			accessibility		
Factors	**slop**	**soil**	**elevation**	**Major road**	**Secondary road**	**Primary roads**
AHP weights	0.723	0.224	0.052	0.65	0.265	0.085

The results shows that when we simulate without any accessibility weights (zero weights), kappa for residential was 0.51, for commercial it was 0.13 and for industrial it was -0.02. Whereas using the above weights for accessibility, it gives the kappa for residential as 0.69, for commercial it is 0.57 and for industrial it comes to 0.48. This means that proper use of weights has a quite significant effect on the simulation results. Unlike the large clusters, it predicts quite sensible and smooth patterns.

Table 10: Simulation 1 Accessibility weights (ahp) combined with actual enrichment factor over all Kappa = 0.72

Land use	Residential	Commercial	Industrial
Kappa	0.6911	0.56698	0.48205
KLoc	0.6911	0.56698	0.48205
KHisto	1	1	1

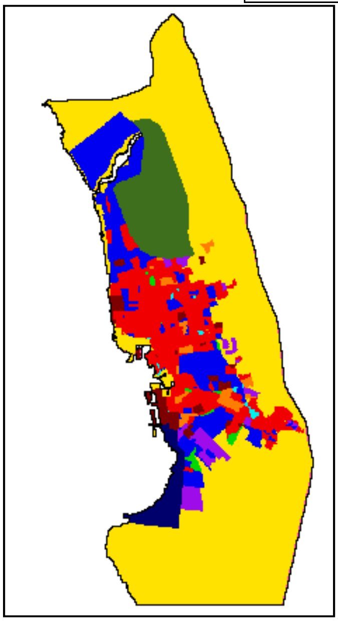
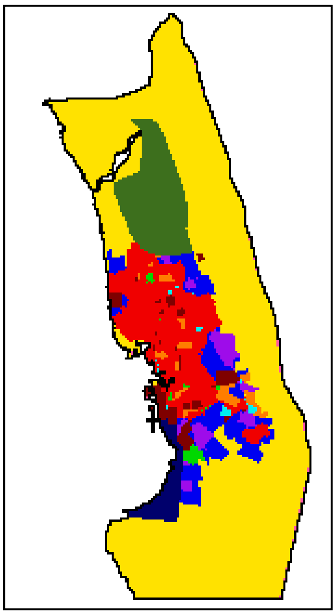

Actual Map1993

Simulated (1) Map 1993

Figure 21: Simulation 1 results

5.8.1. Validation results

The validation of the second method over the next time step between 1993 and 2007 is carried out. The similarity assessment was done using the kappa statistics. Results for individual land use classes are show in table 12 below. The agreement for the residential and industrial land use classes is quite high when compared with the actual land use of 2007. Agreement for commercial areas is a bit lower.
The overall kappa of the model is 0.81, which show the validity of the model using the second method.

Table 11: Kappa statistics for validation

	Residential	Commercial	Industrial
Kappa	0.74212	0.61247	0.69525
KLoc	0.74212	0.61247	0.69525
KHisto	1	1	1

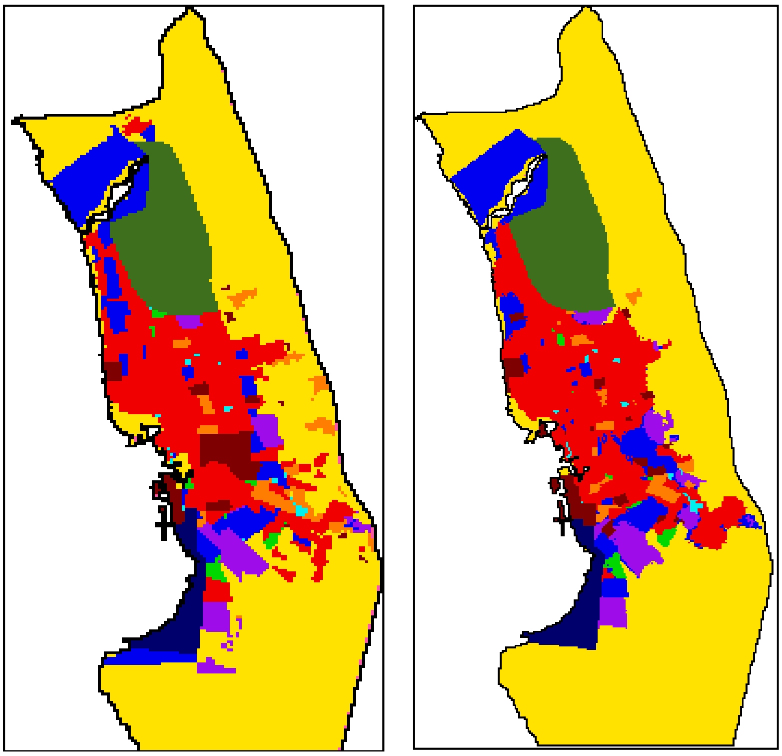


Figure 22: Simulation results for validation (Method 2)

5.9. Comparison of methods

By carefully analysing the results of both methods over the calibration and validation period, it is observed that a combination of both methods shows better results than only the enrichment factor derived rules (see table 13-16). However we used only the accessibility weights in the model, it would be more interesting if the model was flexible to include the AHP derived weights for suitability and neighbourhood as well.

Table 12: Method 1 calibration results 1980-1993

Calibration			
	Residential	Commercial	Industrial
Kappa	0.69329	0.46862	0.52742
KLoc	0.69329	0.46862	0.52742
KHisto	1	1	1

Table 13: Method 1 validation results 1993-2007

validation			
	Residential	Commercial	Industrial
Kappa	0.71932	0.46897	0.5474
KLoc	0.71942	0.46931	0.58793
KHisto	0.99986	0.99927	0.93107

Table 14: Method 2 calibration results 1980-1993 (Over all kappa .72)

Land use	Residential	Commercial	Industrial
Kappa	0.68195	0.365	0.2101
KLoc	0.68195	0.365	0.2101
KHisto	1	1	1

Table 15: Method 2 validation results 1993-2007 (Over all kappa 0.81)

	Residential	Commercial	Industrial
Kappa	0.74212	0.61247	0.69525
KLoc	0.74212	0.61247	0.69525
KHisto	1	1	1

However we used only the accessibility weights in the model, it would be more interesting if the model was flexible to include the AHP derived weights for suitability and neighbourhood as well.

6. Conclusion and recommendations

6.1. Conclusions

The main objective of this research was to derive transition rules by two methods. One method was to calculate the average enrichment factor for different land use types and quantify the neighbourhood characteristics of land use in Jeddah city. The results were then used to help defining the transition rules for the METRONAMICA model. It is not possible to directly translate the neighbourhood characteristics into transition rules for cellular automata but it is used as a starting point for defining the transition rules in the CA model (Verburg, Schot et al. 2004). The initial rules derived through the enrichment factor were used in the model. Simulation results as a whole were good but the results for the three function land uses were not satisfactory. However, after manual calibration of the rules the results improved showing an increase in overall Kappa statistics as well as in the individual function land uses. The second method was a combination of the AHP derived weights and the actual enrichment factor values. This method performed better than the enrichment factor alone. Even the validation results were quite close to the actual land use map.

The methods used for deriving transitional rules were successfully applied in this study however the methods which include AHP experienced the Software limitations and were found not fully applicable in Metronamica model version 4. This method needs its own model to be adjusted for incorporating the weights derived by AHP method for calculation of transition potential rules and further simulations.

6.2. Recommendations

The following recommendations are put forward after careful observation.

- It seems quite difficult to predict the land use classes like commercial and industrial to the possible extent. Hence high resolution data is required for modelling.
- It is very hard to efficiently model the land use at a local level because it needs detail and accurate land use data on a large scale.
- From the results it is concluded that a combination of both enrichment factor and weights derived for accessibility through AHP is a better method for modelling land use change in Jeddah city.
- The transition potential algorithm should be editable in the model.
- Knowledge about the area should got at hand before modelling
- Neighbourhood configurations of simulated maps depend not only on the transition rules, but also on research area, spatial and temporal resolution, and neighbourhood type.

References

Al-Ahmadi, K., L. See, et al. (2009). "*Calibration of a fuzzy cellular automata model of urban dynamics in Saudi Arabia*." Ecological Complexity **6**(2): 80-101.

Al-Hathloul, S. and M. A. Mughal (1991). "*Jeddah*." Cities **8**(4): 267-273.

Al-Hathloul, S. and M. A. Mughal (2004). "*Urban growth management-the Saudi experience*." Habitat International **28**(4): 609-623.

Al-Otaibi, A. M. (2006). *HOUSING SUPPLY AND DEMAND IN NORTHERN JEDDAH: PREFERENCES AND ASPIRATIONS*. School Of Architecture, Planning and landscape, University of Newcastle upon Tyne, Newcastle United Kingdom. **PhD**.

Anthony J. Viera, M. and P. Joanne M. Garrett (2005). "*Understanding Interobserver Agreement: The Kappa Statistic*
" Family Medicine **37(5)**: 360-3).

Barredo, J. I., M. Kasanko, et al. (2003). "*Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata*." Landscape and Urban Planning **64**(3): 145-160.

Batty, M. (1997). "*Cellular Automata and Urban Form: A Primer*." Journal of the American Planning Association **63**(2): 266-274.

Batty, M., H. Couclelis, et al. (1997). "*Special issue-urban systems as cellular automata*." Environment and Planning B: Planning and Design **24** 159 - 305

Batty, M. and Y. Xie (1994). "*From cells to cities*." Environment and Planning B: Planning and Design **21**: 531-548.

Batty, M., Y. Xie, et al. (1999). "*Modeling urban dynamics through GIS-based cellular automata*." Computers, Environment and Urban Systems **23**(3): 205-233.

Briassoulis, H. (2000). *Analysis of land use change: thcoretical and modelIing approaches*, . Department of Geography, University of the Aegean, Lesvos, Greece and Regional Research Institute, West Virginia University, United States.

Cheng, J. (2003a). *Modelling Spatial & Temporal Urban Growth*. Faculty of Geographical Sciences. , Utrecht, The Netherlands, Utrecht University; ITC. **Doctoral:** 214.

Clarke, K. C. and L. J. Gaydos (1998). "*Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore*." International Journal of Geographical Information Science **12**(7): 699-714.

Couclelis, H. (1997). "*From cellular automata to urban models: new principles for model development and implementation*." Environment and Planning B: Planning and Design **24**(2): 165-174.

Daghistani, A.-M. I. (1993) "*A CASE STUDY IN PLANNING IMPLEMENTATION JEDDAH, SAUDI ARABIA*." **Working paper No 5**.

Deadman, P., R. Brown, et al. (1993). "*Modelling Rural Residential Settlement Patterns with Cellular Automata*." Journal of Environmental Management **37**(2): 147-160.

Engelen, G. (2002). *Presentation at MODLUC, Advanced Study Course on Modelling Land-Use Change, European Commission, Environment and Climate Programme, University of Louvain-la-Neuve Belgium*. , RIKS, Maastricht. .

Engelen, G., R. White, et al. (1997). Integrating Constrained Cellular Automata Models, GIS and Decision Support Tools for Urban Planning and Policy Making. *Decision Support Systems in Urban Planning*. London, E&FN Spon**:** Pages 125-155.

Foody, G. M. (2002). "*Status of land cover classification accuracy assessment*." Remote Sensing of Environment **80**: 185-201.

Fredkin, E. (1990). "*Digital mechanics: an informational process based on reversible universal cellular automata*." Phys. D **45**(1-3): 254-270.

Garba, S. B. (2004). "*Managing urban growth and development in the Riyadh metropolitan area, Saudi Arabia*." Habitat International 28 (2004) 593–608.

Hagen-Zanker, A. and P. Martens (2008). Map Comparison Methods for Comprehensive Assessment of Geosimulation Models**:** 194-209.

Hagen, A. (2002). *Multi-method assessment of map similarity*. 5th AGILE Conference on Geographic Information Science Palma (Mallorca, Spain, Research Institute for Knowledge Systems Maastricht, The Netherlands.

Hagen, A. (2002). *Multi-method assessment of map similarity*. 5th AGILE Conference on Geographic Information Science, . Palma (Mallorca, Spain) RIKS bv.

Hagoort, M. and S. Geertman (2004). *Transition Rules for Modelling Land-Use Change* 7th AGILE Conference on Geographic Information Science, Parallel Session 2.2- "Urban Modelling I". Heraklion, Greece.

Hagoort, M. J. (2006). *The Neighbourhood Rules: Land-use Interactions, Urban Dynamics and Cellular Automata Modelling.* KNAG/Utrecht University, Utrecht, The Netherlands

Hu, Z. and C. P. Lo (2007). "*Modeling urban growth in Atlanta using logistic regression*." Computers, Environment and Urban Systems **31**(6): 667-688.

Jiao, J. and L. Boerboom (2006). Transition Rule Elicitation Methods for Urban Cellular Automata Models. *Innovations in Design & Decision Support Systems in Architecture and Urban Planning*: 53-68.

Junfeng, J. (2003). *Transition Rule Elicitation for Urban Cellular Automata models (Case study: Wuhan, China)*. ITC, Enschede.

Lay, J.-G. (2000). "*A Land Use Change Study using Cellular Automata*." GIS Development.

Lee, J. (1999). *DEVELOPMENT OF A COMMUNITY-ACCESSIBLE URBAN SPRAWL IMPACT ASSESSMENT SYSTEM IN NORTHEAST OHIO 15-COUNTY REGION FOR THE EMPACT PROJECT*, US EPA.

Li, X. and A. G.-O. Yeh (2000). "*Modelling sustainable urban development by the integration of constrained cellular automata and GIS*." International Journal of Geographical Information Science **14**(2): 131-152.

Li, X. and A. G.-O. Yeh (2002). "*Neural-network-based cellular automata for simulating multiple land use changes using GIS*." International Journal of Geographical Information Science **16**(4): 323-343.

Liu, Y. (2008). *Modelling Urban Development with Geographical Information Systems and Cellular Automata*, CRC Press.

MODLUC (2002). *Advanced Study Course on Modelling Land-Use Change*, . European Commission, Environment and Climate Programme. Belgium, University of Louvain-la-Neuve.

Mohammed, S. A., J. Y. Salagoor, et al. (2002). "*JEDDAH URBAN GROWTH AND DEVELOPMENT PROCESS : THE UNDERLYING FACTORS*." Scientific Journal of King Faisal University (Basic and Applied Sciences) **3**(1).

Monserud and Leemans (1992). "*Comparing global vegetation maps with the Kappa statistic*." Ecological modeling **62**: 275 - 293.

O'Sullivan, D. and P. M. Torrens (2000). *Cellular models of urban systems. Working paper*, Centre for Advanced Spatial Analysis (UCL), London, UK. CASA Working Papers. **(22)**.

O'Sullivan, D. and P. M. Torrens (2001). *Cellular Models of Urban Systems*. London, Springer-Verlag.

Pathan, S. K., S. V. C. Sastry, et al. (1993). "*Urban growth trend analysis using GIS techniques—a case study of the Bombay metropolitan region*." International Journal of Remote Sensing **14**(17): 3169-3179.

Pontius, R. G., D. Huffaker, et al. (2004). "*Useful techniques of validation for spatially explicit land-change models*." Ecological Modelling **179**(4): 445-461.

Power, C., A. Simms, et al. (2001). "*Hierarchical fuzzy pattern matching for regional comparison of land use maps*." International Journal of Geographical Information Science **15**(1): 77 - 100.

RIKS, b. (2005). "*Model description*." Research Institute for Knowledge Systems, Maastricht, The Netherlands.

RIKS, b. (2009). "*Metronamica user manuel*." Research Institute for Knowledge Systems BV Maastricht, The Netherlands.

Rudy, R. (1999). *Seek! Selected Nonfiction by Rudy Rucker*. New York, Four Walls Eight Windows.

Shenghe, L. and P. Sylvia (2002). *Spatial Patterns and Dynamic Mechanisms of Urban Land Use Growth in China: Case studies in Beijing and Shanghai*. IR-02-005, International Institute for Applied Systems Analysis Laxenburg, Austria.

Silva, E. A. and K. C. Clarke (2005). "*Complexity, emergence and cellular urban models: lessons learned from applying SLEUTH to two Portuguese metropolitan areas*." European Planning Studies **13**: 93-115.

Sui, D. Z. and H. Zeng (2001). "*Modeling the dynamics of landscape structure in Asia's emerging desakota regions: a case study in Shenzhen*." Landscape and Urban Planning **53**(1-4): 37-52.

Torrens, P. M. (2001). *Can geocomputation save urban simulation? Throw some agents into the mixture, simmer and wait... .* Centre for Advanced Spatial Analysis University College London (UCL). Working pape. **32**.

U.S. EPA. (2000). *Projecting land use change, EPA/600/R-00/098 U.S., Environmental Protection Agency, Office of Research and Development*. Cincinnati, OH.

Van Vliet , J. (2009). *Assessing the Accuracy of Changes in Spatial Explicit Land Use Change Models*. 12th AGILE International Conference on Geographic Information Science. Leibniz Universität Hannover, Germany, Research Institute for Knowledge Systems Maastricht, The Netherlands.

Verburg, P. H., T. C. M. de Nijs, et al. (2004). "*A method to analyse neighbourhood characteristics of land use patterns*." Computers, Environment and Urban Systems **28**(6): 667-690.

Verburg, P. H., M. J. Dijst, et al. (2001). "*Modelling approaches for the analysis of the dynamics of land uses: perspectives for a new generation of integrated models* " Network Working Paper Series, Uthrecht University.

Verburg, P. H., P. P. Schot, et al. (2004). "*Land use change modelling: current practice and research priorities*." GeoJournal **61**(4): 309-324.

Webster, C. J. (1993). "*GIS and the scientific inputs to urban planning. Part 1: description*." Environment and Planning B: Planning and Design **20**(6): 709-728.

Webster, C. J. (1994). "*GIS and the scientific inputs to planning. Part 2: prediction and prescription*." Environment and Planning B: Planning and Design **21**(2): 145-157.

White, R. (1998). "*Cities and cellular automata*." Discrete Dynamics in Nature and Society **2**(2): 111-125.

White, R. and G. Engelen (1993). "*Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land use patterns*." Environment and Planning A(25): 1175-1199.

White, R. and G. Engelen (1994). "*Cellular Dynamics and GIS: Modelling Spatial Complexity*." Geographical Systems **1**: 237-253.

White, R. and G. Engelen (2000). "*High-resolution integrated modelling of the spatial dynamics of urban and regional systems*." Computers, Environment and Urban Systems **24**(5): 383-400.

White, R., G. Engelen, et al. (1997). "*The use of constrained CA for high-resolution modelling of urban land-use dynamics* " Evironment and Planning B **24**: 323-343.

Wolfram, S. (1994). *Cellular automata and Complexity:Collected Papers.* Reading, MA: Addison-Wesley.

Wolfram, S. (2002). *A New Kind of Science*. Champaign, Illinois, Wolfram Media.

Wu, F. (1996). "*A linguistic cellular automata simulation approach for sustainable land development in a fast growing region*." Computers, Environment and Urban Systems **20**: 367-387.

Wu, F. (1998). "*An experiment on the generic polycentricity of urban growth in a cellular automatic city*." Environment and Planning B: Planning and Design **25**: 103-126.

Wu, F. (1998). "*SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules*." International Journal of Geographical Information Science **12**(1): 63 - 82.

Wu, F. (1998). "*Simulating urban encroachment on rural land with fuzzy-logic-controlled cellular automata in a geographical information system*." Journal of Environmental Management **53**(4): 293-308.

Wu, F. (2000). *A Parameterised Urban Cellular Model Combining Spontaneous and Self-Organising* P. A. a. D. Martin. London, Taylor & Francis: London.

Yeh, A. G.-O. and X. Li (2001). "*Measurement and monitoring of urban sprawl in a rapidly growing region using entropy*." Photogrammetric engineering and remote sensing **67**(1): 83-90.

Yi, Q. I. (2009). *A Practice of Cumulative CA Model for Urban Spatial Development Simulations Using Vector Datasets*.

Appendices

Appendix 1: AHP weights table based on Saaty scale

	Residential			industrial			commercial			public places			green areas		
No of interview	A-S	N - S	N - A	A-S	N - S	N - A	A-S	N - S	N - A	A-S	N - S	N - A	A-S	N - S	N - A
1	7	6	4	9	3	3	9	6	3	9	7	3	9	3	3
2	7	3	3	9	3	3	7	6	4	5	7	7	5	9	7
3	3	7	7	7	3	3	6	5	5	7	7	5	5	6	1
4	5	3	3	9	5	3	7	7	5	7	9	6	6	7	6
5	3	3	7	7	3	3	7	6	3	7	5	2	7	5	3
6	3	7	5	5	3	3	5	4	4	5	6	4	3	7	5
7	7	3	3	3	3	3	6	7	5	6	7	5	3	3	5
8	9	7	4	7	9	3	9	3	5	9	7	5	5	9	4
9	9	9	3	9	9	3	9	9	3	7	9	9	9	4	7
10	5	3	3	5	7	3	5	9	7	8	9	6	5	3	3
11	9	4	3	9	4	5	7	9	6	8	5	6	5	3	4
12	3	5	7	1	3	3	7	6	6	7	5	4	3	5	5
13	7	9	5	3	9	7	9	7	3	9	4	5	2	9	7
14	7	7	3	7	3	3	9	7	5	9	7	5	7	7	1
15	3	7	5	7	3	5	9	7	5	3	7	7	7	5	7
16	5	9	5	7	5	5	7	7	5	5	7	7	5	5	5
17	3	7	6	9	3	3	9	6	5	3	7	6	3	3	3
18	7	6	3	7	3	3	7	7	6	9	5	3	7	5	3
19	9	5	3	9	3	3	9	6	3	9	7	3	7	3	3
AHP wt	0.722	0.202	0.076	0.721	0.186	0.093	0.758	0.176	0.066	0.748	0.189	0.064	0.705	0.214	0.081

	suitability			accessibility		
No. of interviews	soil - slop	elevation - slope	elevation - soil	SR - MR	PR - MR	PR - SR
1	3	9	9	3	3	3
2	3	9	9	3	7	9
3	7	9	5	1	4	4
4	4	9	8	4	5	7
5	6	9	9	3	7	6
6	3	5	5	5	3	5
7	7	9	8	7	7	6
8	5	9	3	2	2	3
9	9	9	9	5	3	3
10	5	9	9	5	3	4
11	8	9	9	5	3	4
12	5	9	7	4	6	5
13	4	9	1	3	5	4
14	7	9	7	3	9	7

15	5	7	3	5	9	7
16	5	5	5	5	3	7
17	6	9	9	6	4	4
18	3	9	9	4	3	3
19	5	9	9	3	3	5
AHP wt	0.723	0.224	0.052	0.65	0.265	0.085

Appendix 2: Contingency table simulation 2 (calibrated rules Kappa 0.69)

Map 1 \ Map 2	**vacant**	**undeveloped**	**residential**	**commercial**	**industrial**	**airport**	**green area**	**informal settlement**	**port**	**public place**	**outside simulation area**	**Sum Map 1**
Vacant	2100	1823	2362	230	1211	0	1	78	0	178	0	7983
Undeveloped	6414	70502	1634	5	108	0	0	138	0	81	2	78884
Residential_	3103	1250	11221	86	106	0	7	28	0	250	0	16051
Commercial_	67	38	281	271	0	0	0	0	0	0	0	657
Industrial_	967	398	15	45	1212	0	0	0	0	0	0	2637
Airport	0	0	0	0	0	10161	0	0	0	0	0	10161
Green area	134	0	0	0	0	0	260	0	0	0	0	394
Informal settlement	0	0	392	0	0	0	29	1366	0	0	0	1787
Port	0	0	0	0	0	0	0	0	3245	0	0	3245
Public place	918	10	146	20	0	0	0	0	0	2003	0	3097
Outside simulation area	0	0	0	0	0	0	0	0	0	0	0	0
Sum Map 2	13703	74021	16051	657	2637	10161	297	1610	3245	2512	2	124896

Appendix 3: Log Enrichment factors for residential

Residential2	vacant 0	2.569301	2.502091	2.443659	2.377767	2.322758	2.276422	2.236695	2.201831
Residential2	undeveloped 1	-0.066	-0.04337	-0.02616	-0.00691	0.007433	0.021101	0.028623	0.03563
Residential2	Residential 2	0.499094	1.019478	1.300312	1.523581	1.657955	1.745194	1.805639	1.844764
Residential2	commercial 3	1.621958	1.462626	1.299343	1.144619	1.120423	1.140595	1.180642	1.224764
Residential2	industrial 4	1.354929	1.345673	1.331989	1.318666	1.317694	1.323102	1.344814	1.37747
Residential2	Airport 5	-2.77272	-2.17462	-1.82184	-1.51143	-1.24891	-1.05511	-0.87889	-0.73254
Residential2	Green area 6	1.335899	1.852808	2.151714	2.352725	2.481648	2.508539	2.552872	2.590436
Residential2	informal settlement 7	2.122466	2.095092	2.071674	2.03632	2.000028	1.972007	1.941233	1.915681
Residential2	Port 8	-3.517	-2.95489	-2.64373	-2.35468	-2.18082	-2.01176	-1.89086	-1.764
Residential2	public place 9	0.499665	0.40244	0.357476	0.367969	0.430187	0.538086	0.644018	0.743109
Residential2	outside simulation 10			-10.0872	-7.44809	-6.49836	-5.97737	-5.49947	-5.12353

Appendix 4: Log enrichment factors for Commercial

commercial 3	Vacant 0	2.886705	2.84452	2.80136	2.749906	2.702274	2.659623	2.627832	2.595565
commercial 3	Undeveloped 1	-1.33819	-1.08503	-0.86988	-0.66677	-0.52117	-0.4257	-0.34443	-0.28246
commercial 3	Residential 2	-1.22922	-0.35581	0.21114	0.453398	0.5643	0.556346	0.53246	0.483352
commercial 3	commercial 3	3.526984	3.636076	3.612046	3.639561	3.699466	3.754975	3.781409	3.788629
commercial 3	industrial 4	1.398891	1.718121	1.858423	1.978453	2.066062	2.196629	2.239573	2.299677
commercial 3	Airport 5								
commercial 3	Green area 6								
commercial 3	informal settlement 7				-2.00742	-0.80918	-0.20242	0.220441	0.542024
commercial 3	Port 8								-2.60585
commercial 3	public place 9	1.164667	0.78764	0.40184	0.190228	0.127326	0.335789	0.4323	0.6366
commercial 3	outside simulation 10								

Appendix 5: Log enrichment factors for Industrial

industrial 4	vacant 0	2.924097	2.869223	2.818427	2.754842	2.693132	2.63537	2.575184	2.515487
industrial 4	undeveloped 1	-0.90661	-0.71049	-0.55538	-0.38899	-0.24896	-0.14133	-0.04435	0.035416
industrial 4	Residential 2		-4.35044	-3.65729	-3.38536	-3.30062	-3.18729	-3.11206	-2.90352
industrial 4	commercial 3					-3.61028	-2.58066	-1.88752	-1.10476
industrial 4	industrial 4	1.017894	1.577776	1.857581	2.080483	2.239743	2.343144	2.450048	2.534309
industrial 4	Airport 5	-1.77906	-1.21627	-0.88297	-0.59147	-0.4279	-0.2676	-0.16364	-0.05977
industrial 4	Green area 6								
industrial 4	informal settlement 7								
industrial 4	Port 8	-0.76352	-0.19051	0.148992	0.478405	0.680384	0.869784	0.998236	1.127926
industrial 4	public place 9	-1.39556	-0.74015	-0.43578	-0.24288	-0.07486	0.0109	0.11916	0.188876
industrial 4	outside simulation 10								-9.8636

Appendix 6: Sub rules classes for residential, commercial, industrial, public places, green areas used for interviews (total 5 tables)

	suitability	Accessibility	Neighbourhood
suitability	1		
Accessibility		1	
Neighbour hood			1

Appendix 7: Factors for suitability and accessibility used for interviews

Suitability

	Slope	soil	Elevations
slope	1		
soil		1	
Land value			1

Accessibility

	Distance to major roads	Distance to secondary roads	Distance to primary
Distance to major roads	1		
Distance to secondary roads		1	
Distance to urban centre			1

Appendix 8: Method 2 validation kappa statistics

Kappa	0.81224
KLocation	0.84877
KHisto	0.95697
Fraction correct	0.87952