2.1 Introduction

In this chapter, it is intended to bring a summary about theoretical and fundamental fraction of agent-based modelling and how to design it according to the standard definitions. After this overview, the relationship between land change matter and the change drivers will be identified in terms of environmental and socio-economically investigation. Therefore, the appropriate and the most useful tools to implement the aim of this research will be depicted. It begins with the definition of the terms “land use” and “land cover” to outline their differences (Lambin et al. 2007). Land use/cover changes have various causes and consequences (i.e. loss of biodiversity, climate change, pollution, etc.) in the life cycle, which will be addressed briefly.

2.2 Land Use/Cover Change

The terms Land use and Land cover are not technically synonymous; hence, we draw attention to their unique characteristics to differentiate between them. The terms land use and land cover will be clarified in this chapter. There are different definitions of land cover and land use among the relevant scientists. Therefore, a brief explanation about these two terms is provided in this section from the Encyclopaedia of Earth. In general, the term land use and land cover change (LULCC) identifies all kinds of human modification of the Earth’s surface. Land cover refers to the physical and biological cover over the surface of land, including water, vegetation, bare soil, and/or artificial structures (Ellis and Pontius 2006).

Land use has a complicated expression with different views compared with the term land cover. In fact, social scientists and land managers characterise this term
more general to involve the social and economic purposes. Natural science researchers classify the term land use in different aspects of human activities upon lands such as farming, forestry and man-made constructions.

Turner et al. (1995) believe *Land use involves both the manner in which the biophysical attributes of the land are manipulated and the intent underlying that manipulation—the purpose for which the land is used.* Lambin et al. (2007) differentiate between *land cover* (i.e. whatever can be observed such as grass, building) and *land use* (i.e. the actual use of land types such as grassland for livestock grazing, residential area). In fact, the term *land use/cover* will be used chiefly in this thesis, referring to the land cover and the actual land use.

2.3 Land Use/Cover Change Causes and Consequences

LUCC can occur through the direct and indirect consequences of human activities to secure essential resources. This may first have occurred by means of burning of areas to develop the availability of wild game and it accelerated with the birth of agriculture, resulting in extensive clearing such as deforestation and earth’s terrestrial surface management that takes place today (Ellis and Pontius 2006). Land-use/cover change is known as a complex process which is caused by the mutual interactions between environmental and social factors at different spatial and temporal scales (Valbuena et al. 2008; Rindfuss et al. 2004).

More recently, industrial activities and developments, the so-called industrialisation, has encouraged the concentration of population within urban areas. This is called urbanization, which includes depopulation of rural regions along with intensive farming in the most productive lands and the abandonment of marginal lands (Ellis and Pontius 2006). Land use changes are increasingly known as the consequence of actors and factors’ interactions (Bakker and van Doorn 2009). These conversions and their consequences are obvious around the world and it has been becoming a disaster around the metropolitan areas in developing countries.

2.3.1 Loss of Biodiversity

Biodiversity has been diminishing considerably by land change. While lands change from a primary forested land to a farming type, the loss of forest species within deforested areas is immediate and huge (Ellis and Pontius 2006). According to Ellis and Pontius (2006):

> The habitat suitability of forests and other ecosystems surrounding those under intensive use are also impacted by the fragmenting of existing habitat into smaller pieces, which exposes forest edges to external influences and decreases core habitat area.
2.3.2 Climate Change

Land use and cover change matters play a significant role in climate change at different scales such as regional, local and global scales. At global scale, LUCC is accountable for releasing greenhouse gases to the atmosphere, thus leading to global warming. LUCC is able to increase the carbon dioxide balance to the atmosphere by disturbance of terrestrial soils and vegetation. Furthermore, LUCC undoubtedly plays an essential role in greenhouse gas emissions.

2.3.3 Pollution

Tree harvesting, land clearing and other forms of biomass damage to the environment arising from land change are able to increase the pollution percentage of the environment. Vegetation removal makes soils vulnerable to a massive increase in windy and water soil erosion forms, particularly on steep topography. When accompanied by fire, also pollutants to the atmosphere are released. Soil fertility degradation within time is not the only negative impact; it does not only cause damage to the land suitability for future farming, but also releases a huge amount of phosphorus, nitrogen, and sediments to aquatic ecosystems, causing multiple harmful impacts. All of these issues drive water, soil and air pollution at large scale. Besides, other agricultural activities such as using herbicides and pesticides also release toxics to the surface waters, which sometimes remain in the top soil.

2.3.4 Other Impacts

Other environmental impacts of LUCC include the destruction of stratospheric ozone by oxide release from agricultural land and altered regional and local hydrology. Moreover, the most urgent concern for a great part of the human population and most governments is the long-term supply and production of food and other fundamentals required in the future Pontius and Chen (2006).

2.4 Driving Forces of the Land Use/Cover Changes

Assessing the driving forces behind LUCC is essential if previous patterns can explain and be utilised in forecasting future patterns. Land use and cover change can be caused by multiple driving forces that control some environmental, social and economic variables. These driving forces can contain any factor which influences human activities, including local culture, economic and financial
matters, environmental circumstances (i.e. greenness, land quality, terrain situation, water availability and accessibility to recreation), current land policy and development plans, and also interactions between these factors. Therefore, these drivers have to be found to pursue these controlling variables. The driving forces will be utilised in order to manage land change.

Investigation of interrelations between the drivers of land change needs a strong knowledge about methods and effective variables, as well as land policy (Ellis and Pontius 2006). LUCC is frequently addressed through various selected biophysical and socioeconomic variables. In order to facilitate simulation, driving factors are mostly considered exogenous to the land use system (Verburg et al. 2004). Associations between driving forces and LUCC could be addressed qualitatively and quantitatively by means of appropriate approaches.

2.5 Land Use/Cover Change Simulation

Spatially-explicit models, which consider social and environmental causes and consequences, can be the most appropriate form of existing models to simulate land changes. These approaches are capable of checking relationships between environmental and social variables. Integration of existing geographical data and advanced GIS functionality, as well as the ABM functionalities allow this research to achieve the proposed objectives. Considering this, LUCC can be affected remarkably by political and economic decisions. However, the traditional models are not capable of considering all these variables (Ellis and Pontius 2006). These geospatial models can result in precise outcomes that help land managers and policymakers towards a better landscape administration and sustainable land management.

It does not seem simple to compare the performance of the numerous models of LUCC modelling, because they are created from different fundamental bases. For instance, the GEOMOD model simulates change between two land categories, whilst others, such as the Markov chain model and the cellular automata-Markov model simulate change among several categories. Nonetheless, by developing multiagent-based systems (MABS) lately, research is improving these methods to achieve better outcomes. Also, some models use raster data, while others are in vector format. Even in the case of all researchers using the same model, comparison among model performance would still be complicated because researchers usually focus on one study area and do not make a global use approach (Pontius and Chen 2006).

Pontius and Chen (2006) believe that,

it is complicated to separate the quality of the model from the complexity of the landscape and the data.

As an example, if a model does not perform strongly, it does not necessarily imply that the conceptual foundation of that model is weak, but it could mean that
the event of land change in that particular study area is complex or the data is inaccurate. However, if a model performs properly, it is difficult to recognize whether theoretical basis of that model is strong, or that land change case in the study area is particularly uncomplicated, or the used data is extremely uncertain.

Perhaps most importantly, there is not yet a global agreement about methods to determine the performance of LUCC models; therefore, two users who performed the same model on the same landscape and data situation might evaluate one simulation execution differently depending on the criteria used for evaluation (Pontius Jr. and Chen 2006). Land-use change modellers might conclude that the intellectual basis of the validation of the models has some weaknesses (Kok and Veldkamp 2001; Pontius Jr. et al. 2001; Pontius and Schneider 2001; Pontius et al. 2004).

2.6 Land Use Change Trend

Change in economy and spatial distribution of population can occur through conversion from one land use to another, for instance, converting farming lands into residential, industrial, commercial or recreational use. The land owners play a key role in whatever will take place at the land and, therefore, their decisions identify the direction and quantity of changes (Ettema et al. 2007).

Therefore, different types of land owners (e.g. farmers, developers, private individuals, government) decide in a different way according to their type and their parameters. The owners have to supply the financial investment of land change, thus, their awareness of the economic situation can control the speed of the changes. At each time step, the landowner can decide the following decisions:

- Leave the land at current circumstances;
- Develop the land by changing the land usage and exploit it;
- Develop the land by changing the land usage and sell it;
- Sell the land to another owner.

However, the options vary for some owners. For instance, a farmer is not able to develop his land into a residential area, if he does not have the required investment power and skills. Moreover, all actions may not be allowed given planning regulations. Ettema et al. (2007) differentiate between three different types of owners with preferences: farmers (preferences: exploit, sell or buy), government (preferences: maintain, sell to farmer, sell to developer or develop and maintain) and developers (preferences: develop and sell, redevelop and exploit, sell).

Eventually, the decision, which will be most likely made, totally depends on the expected value of each option to the owner. In case of commercial owners, utility will match with profitability: the action will be taken that delivers the highest profit. In case of governmental part, also social benefits might play a significant role, whereas in the farmers’ case, personal and emotional reasons may influence their decision. The market price is a valuable index in deciding whether or not to sell a land with or without developing it (Ettema et al. 2007; Koomen et al. 2007).
2.7 Predicting Future Land Use Patterns

As an essential part of their profession, land use planners envision and forecast alternate future land use and activity patterns in order to change the status quo (Brail and Klosterman 2001). Assessing, forecasting, and evaluating future land change is a complex set of tasks and, hence, it has to be performed after a deep scientific knowledge of the extent individuals, characters, as well as consequences of land transformation have been gathered (Meyer and Turner 1994). A typical land use planning process requires the landscape planners to realise, classify, and investigate the current circumstances in order to project future probable development patterns, and propose plans based on available information (Brail and Klosterman 2001). According to Brail and Klosterman (2001), planners usually approach this task in two ways, a predominant or traditional approach and an analytical approach. The traditional approach foresees a future land use outcome and then prioritises present-day policies required to achieve that outcome. The analytical approach simulates alternate current strategies and compares their consequences.

A recent pervasive approach to consider and simulate human decisions in LUCC is the use of multi-agent systems (MAS) (Parker et al. 2003; Matthews 2006; Robinson et al. 2007; Valbuena et al. 2008). MAS are defined as modelling tools that allow entities to make decisions according to the predefined agents, and the environment also has a spatial explicit pattern. In fact, agents in the system might represent groups of people or individuals, etc. (Valbuena et al. 2008; Sawyer 2003; Bonabeau 2002; Crawford et al. 2005). Agents can be designed with different characteristics which will be explained later in this chapter.

2.8 Simulating Sprawl

Urban sprawl is fairly a contemporary theme in urban studies. Torrens (2006a) noted that;

Suburban sprawl is among the most important urban policy matters facing contemporary cities.

Spatial simulation is able to support sprawl associated research by means of what-if experimentation environments. Sprawling cities are being considered as complex systems and this justifies use of geosimulation to accommodate the space–time dynamics of numerous interacting entities. Automata are compatible tools to represent such systems, but they can be improved to capture uniquely geographical traits of phenomena such as sprawl. Therefore, the development of a model for the geographic dynamics simulation of urban sprawl is explored (Torrens 2006a).
2.9 Approaches to the LUCC Modelling

There are plenty of models concerning land use/cover change modelling. Despite their differences they basically rely on a limited number of methods and assumptions. Those models include economic models (Irwin and Geoghegan 2001), spatial interaction, cellular automata (Yang et al. 2008), statistical models (Veldkamp and Lambin 2001), optimisation techniques (e.g. Ducheyne 2003), rule based models, multi-agent models (e.g. Torrens 2006b), and microsimulation (e.g. Timmermans 2003).

This subsection aims to bring an overview of traditional and current LUCC modelling techniques and eventually, will suggest multi-agent-based systems as a complementary tool. Briefly, the strengths and weaknesses of some models will be discussed here. This appraisal is not in-depth and only presents the best methods which can be complemented by MAS models:

- Equation-Based Models,
- System Models,
- Statistical Techniques,
- Expert Models,
- Evolutionary Models,
- Economic Principles,
- Spatial Interaction,
- Evolutionary Algorithms,
- Genetic Algorithms,
- Optimisation Techniques,
- Cellular Models,
- Hybrid Models,
- Multi-Agent Models,
- Microsimulation.

2.10 Agent-Based Modelling and Geosimulation Terminology

Macal and North (2006) believe that “There is no universal agreement on the precise definition of the term ‘agent’, although definitions tend to agree on more points than they disagree”. It seems very complicated to extract agent characteristics from the literature in a consistent and constant perspective, because they are utilised in different ways (Bonabeau 2002).

Agent-based modelling (ABM) is able to simulate the individual activities by measuring their behaviour and results over time for developing models of cities (Crooks 2006). Crooks (2006) explains cities as follows:

Cities are complex systems, with many dynamically changing parameters and large numbers of discrete actors. The heterogeneous nature of cities, make it difficult to
generalize localized problems from that of city-wide problems. To understand cities’ problems such as sprawl, congestion and segregation, we need to adapt a bottom-up approach to urban systems, to research the reasoning on which individual decisions are made. As cities are highly dynamic, both in space and time and secondly, as cities operate on a cross scale basis, propagating through urban systems from interactions between individuals in space to regional scale geographies. For example, it is easier to conceptualize, and model how individual vehicles move around on a road network, where each car follows a simple set of rules. For instance if there’s a car close ahead, it slows down, if there’s no car ahead, it speeds up and how this can lead to traffic jams without any obvious incident.

Human agents are becoming increasingly significant in land use simulation, despite the fact that traditional environmental and economic models presume one main agent aiming at optimisation in financial conditions (Bakker and van Doorn 2009; Irwin and Bockstael 2002). A variety of MAS models has been developed for land use dynamics modelling that will be mentioned in this chapter (and so far, these models have mostly been performed rule-based (Ligtenberg et al. 2004; Bousquet and Le Page 2004; Berger 2001; Bakker and van Doorn 2009). Certainly, it is vital to represent the agents’ intentions and behaviours with respect to decision making, realistically.

2.10.1 Agents and Agent-Based Models

An agent can be defined according to Russell and Norvig (2009) as follows:

The concept of an agent is meant to be a tool for system analyzing, not an absolute classification where entities can be defined as agents or non-agents.

For instance, a number of experts take into consideration any sort of independent components (e.g. software, individual, etc.) an agent, while some others believe that a component’s behaviour needs to be adaptive in order to be considered an agent, where the term agent is reserved for components that can learn through their environments and change their behaviours accordingly (Macal and North 2005). Nevertheless, several common features exist for most agents (Wooldridge and Jennings 1995; Castle and Crooks 2006)—extended and explained further by Franklin and Graesser (1996), Epstein (2007), and Macal and North (2005).

Therefore, the following characteristics can be defined according to the definitions by Benenson and Torrens (2006).

- **Autonomy**: Agents are independent and autonomous units that are capable of information processes and exchanging them with other agents to independently make decisions. They are also capable of being interactive with other agents and this does not necessarily influence their autonomy (Castle and Crooks 2006; Smith et al. 2007; Benenson and Torrens 2004).
• **Heterogeneity:** Agents can exist and act as groups, but they are constructed through a bottom-up way and combinations of similar autonomous individuals.

• **Mobility:** The mobility of agents is particularly a practical characteristic for spatial simulations. Agents can move around the space within a model.

• **Adaptation and Learning:** Agents are flexible to be adaptive to produce Complex Adaptive Systems (Holland 1996). Agents can be designed to change their locations depending on their current state, following their designed memory (Smith et al. 2007).

• **Activity:** Agents have to be active since they perform independent impacts in a Geosimulation. The following active features can be identified:
  
  – **Pro-active (i.e. goal-directed):** Agents are often considered goal-directed elements, following goals to be accomplished with respect to their behaviours. For instance, agents in a geographic environment can be designed to discover a set of spatial manipulations to achieve an aim within a certain limitation (e.g. time), while evacuating a building during an urgent situation.
  
  – **Reactive (i.e. perceptive):** Agents can be developed to have a consciousness of their surroundings to draw a ‘mental map’ by means of prior knowledge; thus, providing them with an awareness of other entities, obstacles, or required destinations within their environment.
  
  – **Bounded Rationality:** In social sciences, a dominant type of modelling based on rational-choice paradigm has to exist. Rational-choice models commonly assume that agents are perfectly rational optimisers with easy access to gathered information, foresight, and infinite analytical capability. These agents are therefore able to solve deductively complex mathematical optimization matters.
  
  – **Interactive (i.e. communicative):** Agents communicate to each other, extensively. For instance, agents can enquire other agents and the environment within a neighbourhood, searching particular attributes, with the ability to disregard an input which does not match a desirable threshold.

Agent-based models consist of several interactive agents placed within a simulation environment. Relationships between the existing agents are formulated, linking agents to other agents within a system. Relationships can be specified in a number of ways, from simply reactive (i.e. agents only accomplish events when activated to do so by external stimulus e.g. behaviour of another agent), to goal-directed (i.e. seeking a particular purpose). In some cases, the action of predefined agents can be programmed to occur synchronously (i.e. each particular agent executes events at each discrete time point), or asynchronously (i.e. agent reactions are planned by the actions of other agents and/or with reference to a predefined time) (Showalter and Lu 2009).

According to Castle and Crooks (2006),

Environments define the space in which agents operate, serving to support their interaction with the environment and other agents. Agents within an environment may be spatially explicit, meaning agents have a location in geometrical space, although the agent itself may be static. For example, within a building evacuation model agents would be required
to have a specific location for them to assess their exit strategy. Conversely, agents within an environment may be spatially implicit; meaning their location within the environment is irrelevant. For instance, a model of a computer network does necessarily require each computer to know the physical location of other computers within the network.

In simulation environments, agent-based models can be used as experimental media for performing and monitoring agent-based simulations. They can be pictured as a miniature laboratory, where the characteristics and behaviours of agents can be transformed and the consequences observed over multiple simulation runs. As a matter of fact, the ability of individual actions simulation upon various agents and measure the resulting system behaviour and consequences over time means agent-based models can be employed profitably in order to investigate processes that operate at various scales. In fact, the roots of ABMs lie within the individuals’ behaviours simulation and human decision-making (Bonabeau 2002).

Furthermore, Bazghandi and Pouyan (2008) state that ABM is not the same as object-oriented simulation, although the object-oriented paradigm provides a suitable medium for the development of agent-based models. Consequently, ABM systems are invariably object-oriented.

Considering that agent-based models express the behaviours and interactions of a system’s constituent parts from bottom to top, they are the canonical approach for modelling emergent phenomena (Bonabeau 2002). Bonabeau (2002) has categorised a number of conditions that ABMs are practical for capturing emergent behaviour.

### 2.11 Characteristics of the Geosimulation Model

Geosimulation differs from cellular automata in one particular respect: individual automata are basically free to move around, i.e. they are not fixed agents and their movements do not have to take place cell by cell. This feature has obvious consequences for the representation of spatial systems (Longley and Batty 2003); therefore, this topic will be explained in detail within this chapter. Figure 2.1 represents a schematic view of characteristics of a multi-agent system.

#### 2.11.1 Management of Spatial Entities

A basic aspect of geosimulation regards the characterisation of spatial entities that form the building blocks of a simulation model. In fact, urban simulation models have identified units of urban systems (e.g. real estate, land, individuals, etc.) by aggregation of geographic zones, tracts, and socioeconomic groups. These collect units that are spatially modifiable (Openshaw 1983).
2.11 Characteristics of the Geosimulation Model

2.11.2 Management of Spatial Relationships

The second aspect of geosimulation relates to the portrayal of spatial relationships in models. For instance, we can consider this in the framework of geospatial interactions; their representation in traditional urban simulation has been limited to flows between aggregate units. Geosimulation models consider interactions as an outcome of the behaviour of elementary geographic objects. In this way, geosimulation models have the potential to represent spatial interaction of a much wider spectrum of forms, including traditional and far-distance migration (Crooks 2006).

2.11.3 Management of Time

The third distinctive characteristic of them relates to the action of time in models. Urban systems convert over time, and diverse phenomena happen at different time scales. Benenson and Torrens (2004) believe that

Geosimulation models treat time through intuitively justified units such as housing search cycles. Objects’ temporal behaviour can be considered as either synchronous, when all objects change simultaneously, or asynchronous, when they change in turn, with each observing the urban reality as left by the previous one.

2.11.4 Direct Modelling

Disappointment with the appearance of urban simulation as a new field of study in the 1970s was an expectation of what urban simulation models need to accomplish in reality (Crooks 2007a, b; Batty 2005). One of the goals of the geosimulation approach is to move towards the creation of “tools to think with” (Benenson and Torrens 2004). Benenson and Torrens (2004) note that

![A schematic view of a multi-agent system](image)
Realistic descriptions of objects’ behaviour in ways that were not previously obtainable, either technologically or intellectually, makes these worthwhile and, further, allows for direct relation between conceptual and real-world modelling. The idea underlying geosimulation is that the same model can be applied to abstract real-world phenomena; if modelled phenomena are an abstraction of real-world phenomena, why should modelled objects differ from their counterparts in the real world? The Geosimulation approach is supported by several key developments in the geographical sciences and other fields, particularly mathematics, computer science, and general system studies. The cornerstone of the geosimulation approach, however, is the automaton, which has been widely used in simulation and features prominently in geosimulation toolkits.

2.12 The Basic of Geosimulation Framework: Automata

The description of objects’ behaviour in the geosimulation framework is based on the idea of automata. Simply stated, an automaton is a processing mechanism with characteristics that change over time based on its internal characteristics, rules, and external input. Automata are used to process information input into them from their environs with the characteristics altering according to rules that govern their reaction to those inputs. Levy (1992) explains automata as below.

An automaton is a machine that processes information, proceeding logically, inexorably performing its next action after applying data received from outside itself in light of instructions programmed within itself.

Automata are a practical concept of “behaving objects” for many causes, but chiefly because they provide an efficient formal mechanism for representing their fundamental properties: behaviours, attributes, relationships, environments, and time.

Formally, a finite automaton $A$ can be represented by means of a finite set of states $S = \{S_1, S_2, S_3, \ldots, S_N\}$ and a set of transition rules $T$.

$$A \sim (S, T)$$  \hspace{1cm} (2.1)

Transition rules define an automaton’s state, $S_{t+1}$, at time step $t + 1$ depending on its state, $S_t(S_t, S_{t+1} \in \{S\})$, and input, $I_t$, at time step $t$:

$$T : (S_t, I_t) \rightarrow S_{t+1}$$  \hspace{1cm} (2.2)

2.13 Cellular Automata versus Multi-Agent Systems

Geosimulation requires a geospatial structure for modelling urban systems, one as formulated on the basis of objects located in space. Ideally, such an approach allow for simulated geospatial entities to be considered as automata; moreover,
Cellular Automata and multi-agent systems concepts could ideally be combined by considering collections of interacting geographic automata. In this section of the chapter, it is intended to introduce such a framework, which considers geographic objects, interacting to form geographic automata systems and urban system as a whole are considered as the products of collective dynamics among multiple inanimate and animate geographic automata (Benenson and Torrens 2004). Figure 2.2 represents relation between cell-based GIS, CA modelling and MAS.

2.14 Geographic Automata System

The geographic automata system (GAS) framework joins CA and MAS directly reflecting a geographic and object-based (more particularly, automata-based) view of urban systems. This idea was introduced for the first time by (Benenson and Torrens 2004) as a new paradigm in natural studies for better and more accurate results.

2.14.1 Definitions of Geographic Automata Systems

There is a distinct class of automata, geographic automata systems (GAS), consisting of geographic automata of various types. In general, the states and transition rules characterise automata (Benenson and Torrens 2004).

Basically, the $G$ value in GAS can be defined as consisting of seven following components:

$$G \sim (K; S, T_S; L; M_L; N, R_N)$$

(2.3)

Here, $K$ represents a set of types of automata represented in the GAS and three pairs of symbols denote the other components, each one representing a specific
spatial or non-spatial characteristic and the rules identify its dynamics. The first pair denotes a set of states $S$, linked with the GAS. $G$ consists of a set of states $S_k$ of automata for each type of $k \in K$. A set of state transition rules $T_s$, determine how automata positions are supposed to change within time. The second pair represents location information. $L$ denotes the geo-referencing conventions that dictate the location of automata in the system and $M_L$ denotes the movement rules for automata, governing changes in their location (Benenson and Torrens 2003). Hence, changes in location and transitions of states for geographic automata depend on the automata and also, on input $(I)$, specified by the states of neighbours. The third pair specifies this condition. $N$ denotes the neighbours of the automata and $R_N$ represents the neighbourhood rules that manage how automata relate to the other automata in their vicinity.

### 2.14.2 Geographic Automata Types

GAS consists of different types of automata. Two main types of automata can be distinguished; non-fixed and fixed geographic automata. Fixed geographic automata stand for objects that do not move over time and thus have close analogies with CA cells. For instance, in the context of urban systems, a variety of urban items may be indicated as fixed geographic automata: building footprints, road networks conjunctions, parks, etc. Fixed geographic automata may be addressed by any of the transition rules outlined already, except rules of movement, $M_L$. Non-fixed geographic automata identify entities which move around over time. The full array of rules for GAS can perform with non-fixed geographic automata, including movement rules (Benenson and Torrens 2004).

### 2.14.3 Geographic Automata States and State Transition Rules

A number of state variables $S$ can be assigned to the individual geographic automata, that comprise a GAS, and these states explain the characteristics of the automata. Any variable can be employed to derive state values, including variables of geographic significance. Pointing to the non-fixed automata, location variables of relevance to the transition rules of the model might be initiated.

In fact, state transition rules are based on geographic automata of all forms of $K$. It seems necessarily vital mentioning that, in the framework of the GAS, CA is artificially closed, simply because cell state transition rules are driven only by cells (Benenson and Torrens 2004). In contrast, the states of urban infrastructure objects represented by means of geographic automata totally depend on the surrounding objects of that infrastructure, but are also driven by mobile geospatial automata (i.e. agents) that are responsible for controlling object states such as land value or land-use (O’Sullivan et al. 2003). This is a crucial concept for simulating human-
driven urban systems that show how individuals interact and are affected by the environments.

2.14.4 Geographic Automata Spatial Migration Rules

Geo-referencing conventions (L) administrate how geographic automata should be registered in space. For fixed geographic automata, geo-referencing is a straightforward process in most instances; these automata can be geo-referenced by recording their position coordinates. However, for non-fixed geographic automata, geo-referencing has to be dynamic and automata may move. Also, their location in relation to other automata, represented in simulated goals, destinations, opportunities, etc., may be dynamic in space and time (see Fig. 2.3). It is also essential noting that there are examples in which Georeferencing is dynamic for fixed geographic automata also, for example, when land parcel objects are sub-divided during simulation (Benenson and Torrens 2004).

2.14.5 Geographic Automata Neighbours and Neighbourhood Rules

Another element of GAS that requires explicit explanation is the set of neighbours of automata, N, and the rule set for determining the change in neighbourhood relationships between automata, R_N. Different type of neighbours is necessary for the application of transition rules state transition (T_s) and movement (M_L), which totally depend on characteristics of geographic automata and their neighbours.
In opposition to the static and symmetrical neighbourhoods utilised in usual CA models, geographical relationships between geospatial automata can change in space and time, thus, RN rules need to be formulated to account for geographic automata positions’ neighbours at each time point. Neighbourhood rules for fixed geographic objects can be defined easily comparatively, because the objects are static in space.

2.14.6 Types of Simulation Systems for Agent-Based Modelling

Generally, two types of simulation systems can be performed to develop agent-based models: either toolkits or software. Based on this, toolkits are simulation or modelling systems that provide a conceptual framework for designing ABMs which provide required libraries of software functionality that consist of pre-defined modules, routines and functions distinctively designed for ABM. The object-oriented prototype allows importing extra functionalities through other libraries, which are not supplied by the simulation toolkit, developing the capabilities of these toolkits (Crooks et al. 2008). The most interesting part of this approach is the capability of integration of GIS functionality from ArcGIS software libraries with an ABM context.

The development of agent-based models can be significantly facilitated by the utilisation of simulation and modelling toolkits. In fact, they are able to provide reliable templates for the design, accomplishment and visualisation of agent-based models, allowing modellers to concentrate on the content of research, rather than coding fundamentals required to run a simulation (Tobias and Hofmann 2004). In particular, the use of toolkits can decrease the burden of modellers challenged with programming matters of a simulation (e.g. GUI design, data import and export, visualisation and model representation). It is also crucial to improve the model’s trustworthiness and efficiency (Smith et al. 2007).

Unsurprisingly, there are limitations of using simulation/modelling systems to develop agent-based models; for instance, a considerable amount of effort has to be spent to realise how to design and implement a model (Crooks 2007a, b). Benenson et al. (2005) and Crooks (2007a, b) note that toolkit users are accompanied by the fear of discovering that a particular function cannot be used, will conflict, or is incompatible with another part of the model late in the development process.

2.15 Current Simulation Systems

Various environments are available in order to develop agent-based models. This section aims to review an overview of these systems:
1. Open source such as Swarm, MASON and Repast,
2. Shareware/freeware such as StarLogo, NetLogo and OBEUS,
3. Proprietary systems such as AgentSheets and AnyLogic (Bandini et al. 2009).

The mentioned systems need to fulfill the majority of the following criteria:

- retained and still being improved,
- broad range of users and also supported by strong user communities,
- accompanied by various demonstration models and in some instances the model’s programming source code made available,
- Capable of developing spatially explicit models and integration with GIS functionality.

Further information about each system, as well as identifying examples of spatial models that have been developed will be provided in this section. In this part of chapter, a brief introduction of all affordable toolkits will be presented in order to acquire a preliminary knowledge over mentioned prototypes. Certainly, the earliest and most well-known toolkit was SWARM, although many other toolkits more recently have been released. There are a variety of toolkits available for ABM at this time. However, variation between toolkits needs to be considered. For instance, their purpose, level of development, and modelling capabilities can vary. A review of the most user-friendly toolkits will be presented throughout this chapter.

2.15.1 ASCAPE

ASCAPE (Agent-Landscape) is one of the earliest toolkits associated with ABMs which has been developed by the Centre on Social and Economic Dynamics (CSED), Brookings Institution. ASCAPE is a research toolkit to support agent-based modelling and simulation. In fact, high-level frameworks support complex model designs, while end-user tools prepare it for non-programmers to investigate various aspects of model dynamics. This toolkit is written completely in Java, and runs on Java-enabled platforms. Models developed by this means can be easily published to the web for use with common web browsers (Batty and Jiang 1999; Epstein and Axtell 1996).

2.15.2 StarLogo

According to the StarLogo official website (2008);

StarLogo is a programmable modelling environment for exploring the workings of decentralized systems—systems that are organized without an organizer, coordinated without a coordinator. With StarLogo, you can model (and gain insights into) many real-life phenomena, such as bird flocks, traffic jams, ant colonies, and market economies.
StarLogo is a particular version of the Logo programming language. Also, it is practical to create drawings and animations by giving commands to graphics. It expands this idea by allowing users to control many graphic turtles in parallel. In addition, StarLogo makes the turtles’ world computationally active; therefore, it is possible to create the turtles’ environment by code. Turtles and patches can interact with one another. StarLogo is predominantly well-suited for artificial life investigations. In decentralised systems, orderly patterns can take place without centralised control. StarLogo has been developed to facilitate students, as well as researchers to extend new ways of understanding decentralised systems (Camazine et al. 2003).

### 2.15.3 NetLogo

NetLogo is a multi-agent programmable platform developed by the Centre for Connected Learning and Computer-Based modelling, Northwestern University, USA (Tisue and Wilensky 2004). NetLogo allows the users to access a large library of sample models and code examples that help users to start authoring models. NetLogo is being used by research labs and university lessons in social and natural sciences.

### 2.15.4 OBEUS

Object-Based Environment for Urban Simulation (OBEUS) is a software environment based on a GAS conceptual core. In fact, OBEUS has been established according to the basic components of GAS with respect to automata types. These are accomplished by means of three following root classes:

- Population that contains information regarding the population of objects of a given type k as a whole;
- Geo-Automata, acting as a container for geographic automata of a given type k;
- Geo-Relationship that facilitates specification of spatial relationships between geographic automata of the same or different types (Benenson and Torrens 2004).

### 2.15.5 AgentSheets

AgentSheets is another toolkit for construction of interactive graphical systems. It is a simulation system that allows modellers with partial coding skill to develop an agent-based model, because models can be developed through a GUI
(Repenning et al. 2000). Several demonstration models exist on the system website; for instance, Sustainopolis. The system lacks, however, functionality to dynamically chart simulation output, and agents are limited to movement within a 2-dimension lattice environment (Crooks 2007a, b).

### 2.15.6 AnyLogic

AnyLogic allows modifying a simulation model using several methods; system dynamics, agent based and discrete event (process-centric) modelling. Furthermore, it is also possible to combine different methods in one model. AnyLogic modelling language is an extension of UML-RT, a set of the best engineering practices have been verified successfully in the modelling of complex systems (Anylogic 2006).

### 2.15.7 SWARM

Swarm is one of the oldest agent-based modelling toolkits. Swarm has been originally written in Objective-C language, and then exported to Java (Getchell 2008). Nevertheless, the documentation and research papers on Swarm established many of the foundational concepts and ideas in ABM, and reading over these materials serves as an excellent introduction to the large and growing field of agent-based modelling.

### 2.15.8 MASON

MASON or Multi-Agent Simulator of Neighbourhoods/Networks is another simulation library in Java, designed to serve as the base class structure for custom Java simulations. It also includes a model library and suite of 2D and 3D visualisation tools, and is developed with an emphasis on speed and portability.

### 2.15.9 NetLogo

NetLogo is another ABM toolkit, which is not open source, and also designed for educational use, being based on a simple Logo-type language. It was initially developed in 1999 by Uri Wilensky, and it has been under continuous development thereafter, and has a large and broad user community.
2.15.10 Repast

Repast has been made based on Swarm, but executed in Java. Repast has several versions available; the current standard Repast version is 3. RepastPy is a simplified version of Repast, and introduced a friendly graphical user interface. Also it benefits from a subset of Python as its scripting language. RepastPy is faster and easier to employ than Repast, and is generally recommended as being a good version for creating prototype models. In fact, the Python scripts generate Java objects (Getchell 2008).

Repast.NET is another version of Repast based on the .NET runtime. The .NET runtime is flexible and powerful due to having a large number of functional libraries for handling nearly anything. It also has a stylish successor language to Java, C#, as well as the ability to run any language that can be linked to the .NET platform such as Python, Visual Basic, Ruby, etc. The software project source codes are not compatible with later versions of Visual Studio. Repast Simphony is the latest version of Repast, which is combined with the powerful Eclipse integrated development environment, and also automated connectors to additional tools such as R, VisAD, MATLAB, Weka, and iReport (Getchell 2008).

2.15.11 Agent Analyst Extension

In this section, it is intended to present an overview of the Agent Analyst extension, which is an agent-based modelling and simulation extension for the ESRI-ArcGIS. Agent Analyst integrates the functionalities of the Repast simulation environment with the strengths and flexibilities of ArcObjects and ArcGIS in spatial analysis. Agent Analyst has the capability to integrate ABMS with GIS. GIS modellers are able to simulate environment behaviours and processes as change and movement over time by means of this extension (e.g. simulate land use and cover changes, predator–prey communications). This can help ABMS modellers to integrate detailed real-world biophysical data to execute complex spatial processes, as well as study how behaviour is constrained by space and geography. In addition, ABMS models can include update GIS data feeds for circumstances (e.g. fire-fighting, disaster management). This extension allows modellers to create, customise, and perform Repast models through the ArcGIS 9.2 geoprocessing framework, including access through the ArcToolbox, Model Builder, and ArcMap. Additionally, the Agent Analyst GUI allows users to create agents, schedule simulations, establish mappings by ArcGIS layers, and specify the behaviour and interactions of each agent (Bertelle et al. 2009).
2.16 Selection of ABM Implementation Toolkit

As it has been stated before, numerous toolkits have been developed in order to achieve ABM interests. Several software packages are also available for developing agent-based models that can facilitate the implementation process. For instance, simulation software often negates the necessity of developing agent-based models through low-level programming languages (e.g. Java, C++, Visual Basic, etc.). However, this brings some restrictions for developers to design their customised framework. As an example, some ABM software may support particular environments (e.g. either raster or vector), or agent transition rules might be limited in forms such as Von Neumann or Moore type. Furthermore, modellers will be constrained to the use of functions provided by the software, particularly while the toolkit has been written in its own programming engine (e.g. NetLogo) (Castle and Crooks 2006).

Each model has to be tested in their particular environments and, since the aim of this dissertation is to run a geosimulation prototype within ArcGIS software, it was decided to employ GIS functions for this simulation. In other words, it is intended to code an ABM environment within GIS software. Table 2.1 is a sample of open source simulation toolkits comparing the specified criteria and their advantages or disadvantages.

Agents can be designed and parameterised in a variety of ways, depending on the goals of the MAS (Valbuena et al. 2008; Robinson et al. 2007; Janssen and Ostrom 2006), for instance, agents can represent governmental regulation, land developers actions and residents behaviours at different environmental and financial circumstances (Valbuena et al. 2008; Ligtenberg et al. 2004; Monticino et al. 2007). As a matter of fact, the decision-making process has to be particularly parameterised by decision rules. These rules can be defined according to the expert knowledge or/and other researches’ outcomes. Agent parameterisation with others’ findings is the more frequent approach in MAS (Berger and Schreinemachers 2006; Valbuena et al. 2008). Alternatively, the parameterisation of agents with expert knowledge and empirical data facilitates understanding the real LUCC process. Most researches benefit from empirical data to specify and parameterise agents relying on huge data gathering (e.g. Valbuena et al. 2008; Jepsen et al. 2006; Castella et al. 2005; Huigen 2004; Bousquet et al. 2001).

2.17 Designing a Multi Agent System

In order to design a multi agent system, we will pursue the following stages, which are mentioned here:

Step 1: Collection and analysis of required input data for multiagents geosimulation,
<table>
<thead>
<tr>
<th></th>
<th>Swarm</th>
<th>MASON</th>
<th>Repast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developers</td>
<td>Santa Fe institute/Swarm development group, USA</td>
<td>Centre for social complexity, George Mason university, USA</td>
<td>University of Chicago, department of social science research computing, USA</td>
</tr>
<tr>
<td>Date of inception</td>
<td>1996</td>
<td>2003</td>
<td>2000</td>
</tr>
<tr>
<td>Implementation</td>
<td>Objective-C/Java</td>
<td>Java</td>
<td>Java/Python/Microsoft .Net</td>
</tr>
<tr>
<td>language</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating system</td>
<td>Windows, UNIX, Linux, Mac OSX</td>
<td>Windows, UNIX, Linux, Mac OSX</td>
<td>Windows, UNIX, Linux, Mac OSX</td>
</tr>
<tr>
<td>Required programming experience</td>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
</tr>
<tr>
<td>Integrated GIS</td>
<td>Yes</td>
<td>None</td>
<td>Yes (e.g. OpenMap, Java Topology Suite, and GeoTools)</td>
</tr>
<tr>
<td>functionality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated charting/</td>
<td>Yes (e.g. R and S-plus statistical packages)</td>
<td>None</td>
<td>Yes (e.g. Colt statistical package, and basic Repast functionality for simple network statistics)</td>
</tr>
<tr>
<td>graphing/statistics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability of</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>demonstration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source code of</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>demonstration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tutorials/how-to</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>documentation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Step 2: Evaluation of existing geosimulation frameworks,
Step 3: Selection of an appropriate platform for geosimulation performance,
Step 4: Consideration of the land change drivers,
Step 5: Classification of agents,
Step 6: Specification of related factors and variables to each particular agent,
Step 7: Agents combination and reflection of agents’ interactions,
Step 8: Creation and performance of multiagents geosimulation,
Step 9: Collection and analysis of the multiagents geosimulation results arising from the prototype,
Step 10: Verification and validation of multiagents geosimulation results,
Step 11: Report of the multiagents geosimulation’ results,
Step 12: Visualisation of the multiagents geosimulation outcomes.

Figure 2.4 demonstrates a schematic workflow of geosimulation performance.

### 2.18 Fuzzy Decision Theory in Geographical Entities

Many phenomena exist with a degree of vagueness or uncertainty. Some terrestrial objects cannot be appropriately expressed with crisp sets. Kainz (2008) states that

In human thinking and language we often use uncertain or vague concepts. Our thinking and language is not binary, i.e. black and white, zero or one, yes or no. In real life, we add much more variation to our judgments and classifications. These vague or uncertain concepts are said to be fuzzy. We encounter fuzziness almost everywhere in our everyday lives.

Fuzzy set theory was developed by Zadeh (1965) and extended by many authors, notably by Dubois and Prade (1979), to model uncertainties to allow for a more general theory of uncertainty than probability theory models. There is often confusion in the semantics of uncertainty pertaining to probability, interval, fuzzy, and possibility. GIS applications have seldom, if ever, used possibilistic geographical analysis. There are many reasons for this. Perhaps the most significant reason is that fuzzy set theory, as distinguished from possibility theory, is not always clear. Second, since geographical entities are often fuzzy (boundaries are gradual or transitional in nature between geographical entities) the use of possible entities is frequently omitted. Third, since Zadeh (1965) developed possibility theory via fuzzy set theory, most authors do not make a distinction and consider possibility distributions the same as fuzzy membership functions (Lodwick 2007).

Fuzzy set theory incorporates some concepts that can be used to overcome some of the problems described above, modelling some types of uncertainty associated to geographical information, as well as its heterogeneity. They enable the development of an alternative data model that integrates characteristics of the object and the field data models, where the geographical information is represented with fuzzy geographical entities, which are geographical entities (GE) represented
Fig. 2.4 Schematic view of multiagents geosimulation implementation
2.18 Fuzzy Decision Theory in Geographical Entities

A geographic entity is characterised by an attribute and a geographical location. These two components of the geographic entities are intimately related, and therefore, the errors and uncertainty associated to each of them are also related. A fuzzy geographic entity (FGE), $E_A$, characterized by attribute “A”, is a geographic entity whose position in the geographical space is defined by the fuzzy set:

$$E_A = \{(x, y): (x, y) \text{ belongs to the GE characterised by attribute “A”}\}$$

With a membership function, every location in the space of interest is defined between 0 and 1. The membership value 1 stands for full membership, and the membership value 0 represents no membership and values in between correspond to degree of membership to $E_A$, decreasing from 1 to 0. To model the uncertainty or errors regarding the positioning of an attribute, two cases may be considered. One is when the attribute is defined using only a concept such as buildings, forest, or rivers, and the other case is when the attribute definition is based on measurable quantities. In the first case, the difficulty to position the attributes on the ground depends on the details given on its definition and upon the heterogeneity of those attributes in the region under study. The details used in the attribute definition should be such that it is clear what the attribute represents, but note that if too many details are given, the identification of all those details in the ground may complicate the operator’s work, since it may be difficult to identify, for example, in an aerial photograph, if a hut is made of wood or brick. So, the attribute definition should be adapted to the methods and sources of information available to identify the entities.

Whenever the operator has some difficulties in the classification, he may always assign a degree of uncertainty to the entity. For example, if there is some uncertainty whether a certain entity should be considered a building or not, a degree of uncertainty may be assigned to it. These degrees of uncertainty are subjective and only indicative in the sense that some difficulties in the classification were found. They are assigned to the entity as a whole, since it is an indivisible object. Note that, in this case, the grade of membership represents uncertainty in the attribute that should be assigned to the entire region. The outcome of this process is then a GE with a constant grade of membership to an attribute. These grades of membership translate degrees of membership to the attribute defining the GE and not uncertainty on the geographical space. They result from lack of data to assign the correct attribute or lack of attribute definition. In other situations, the uncertainty is not in the identification of the attribute corresponding to a certain GE, but in the
identification of its exact location. In this case, a fuzzy set may be used to express the entities’ location in the geographical space.

This corresponds to the identification of the core of the fuzzy set and its support, and therefore all the region between the support and the core belongs to the uncertainty region. If it is possible to differentiate further inside the uncertainty region, then \( \alpha \)-levels may be identified, if not, only an uncertainty region may be used.

This process generates fuzzy geographic entities, which correspond to the “egg-yolk” approach (Cohn and Gotts 1996). This approach considers that the geographic entities are formed by three regions: the interior, the frontier, and the exterior, where the frontier is represented not by lines, but by a region with any dimension or shape, and that may be considered homogeneous or heterogeneous (see Fig. 2.5). The “egg-yolk” representation is a simplified representation of fuzzy geographic entities and is convenient when the geographic entities are to be represented using the vector data structure and to establish neighbourhood relations between geographic entities with uncertain or fuzzy geographical position (Lodwick 2007).

2.18.2 Processing Fuzzy Geographical Entities

The use of fuzzy geographical entities in GIS environments requires operators capable of processing this type of entity. The instant approach to process FGEs is to convert them into crisp entities and use the usual operators to perform the necessary operations. Since the \( \alpha \)-levels of fuzzy sets are crisp sets, the easiest way to convert fuzzy GEs to crisp GEs is to substitute the entity by one of its \( \alpha \)-levels. By means of this approach, a variety of versions of the FGE may be achieved according to the needs of each application, selecting different \( \alpha \)-levels to represent it. In this research, we will use fuzzy sets to represent prediction process of land use classes over time (Lodwick 2007).
2.19 The Analytical Hierarchy Process Weighting

The analytic hierarchy process (AHP) is a very pervasive and commonly used application for decision-making matters (De Feo and De Gisi 2010). Indeed, the AHP was developed by Thomas L. Saaty in the 1970s. The AHP affords a comprehensive rational structure in order to solve the decision-making problems, as well as characterising and quantifying its elements and conduction of the related elements towards overall goals, plus evaluating alternative solutions (De Feo and De Gisi 2010; Forman and Selly 2001; Saaty 1977).

In fact, the AHP has special advantages in multi-variable evaluation. It has been utilised in various research fields, such as natural science, economy and society (Ramanathan and Ganesh 1995). AHP is also becoming a common tool of eco-environment quality evaluation, for ecological environment is a large and multi-layered system (Hill et al. 2005; Klungboonkrong and Taylor 1998; Li et al. 2007; Yedla and Shrestha 2003). GIS-based AHP is popular because of its strong capacity to integrate various types of heterogeneous variables and its simplicity to obtain the weights of appropriate variables. Therefore, this reasoning promotes its strength in criteria weighting (Hossain and Das 2010; Tiwari et al. 1999).

The AHP method splits a complex multi-criteria decision matter into a hierarchy and performs on the basis of a pair-wise comparison of the importance of different criteria and sub-criteria (De Feo and De Gisi 2010; Forman and Selly 2001; Saaty 1977). According to Saaty (1977) the AHP process is based on three main steps. The first step is to establish a hierarchical structure, where the first hierarchy of a structure is the goal. The final hierarchy deals with identifying alternatives, while the middle hierarchy levels appraise certain factors or conditions.

The second step computes the element weights of various hierarchies by means of three sub-steps. The first sub-step creates the pair-wise comparison matrix, then a pair-wise comparison is conducted for each element based on an element of the upper hierarchy that is an evaluation standard. The second sub-step computes the eigenvalue and eigenvector of the pair-wise comparison matrix. The third sub-steps execute the consistency test. The difference between the dominant eigenvalue of the pair-wise comparison matrix ($\lambda_{\text{max}}$) and the matrix dimension ($k$) is used in defining the inconsistency index, $II$ (Hsu et al. 2008; Karamouz et al. 2007; Saaty 1999):

$$II = [(\lambda_{\text{max}} - k)/(k - 1)]$$  \hspace{1cm} (2.4)

Moreover, the inconsistency ratio (IR) is defined as (Hsu et al. 2008; Karamouz et al. 2007; Saaty 1999):

$$IR = II / \text{CRI}$$  \hspace{1cm} (2.5)

The CRI presents the inconsistency index of the random matrix retrieved by calculating $II$ for a randomly filled matrix. If $IR < 10\%$, then the consistency criterion is acceptable. Otherwise, the decision-maker has to refine the pair-wise
comparisons. This procedure goes on until all the pair-wise comparisons satisfy the consistency criterion. The eigenvector of the pair-wise comparison matrix is used to estimate the relative weight of different choices. Finally, the third step of the AHP technique calculates the entire hierarchical weight. In reality, AHP generates an overall ranking of the solutions using the comparison matrix among the alternatives and information on the ranking of the criteria. Thus, the option with the highest eigenvector value is approved to be the first choice (Hsu et al. 2008; Karamouz et al. 2007; Saaty 1999).

In fact, the strength of AHP is that it allows for the creation of inconsistent relationships, besides affording the CR index as an indicator of the degree of consistency or inconsistency (Forman and Selly 2001). Thus, the AHP execution in this thesis will incorporate an opportunity to let the user define a satisfactory CR threshold value. In this thesis, it is intended to improve the performance of agent-based modelling by means of a combination of GIS, AHP and ABM; therefore, all socio-economic and environmental variables are combined according to their weights.

2.20 Moran’s Autocorrelation Coefficient Analysis

Moran’s autocorrelation coefficient denoted as Moran’s-\(I\) is an extension of Pearson product-moment correlation coefficient to a univariate series. Recall that Pearson’s correlation symbolised as \(\rho\) between two variables \(x\) and \(y\) both of extent \(n\) is:

\[
\rho = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}}
\]

Where \(\bar{x}\) and \(\bar{y}\) are the sample means of both variables. \(\rho\) measures whether, on average, \(x_i\) and \(y_i\) are associated. For a single variable, for instance \(x\), \(I\) will measure whether \(x_i\) and \(x_j\) with \(i \neq j\), are associated. Note that with \(\rho\) \(x_i\) and \(x_j\) are not associated since the pairs \((x_i, y_i)\) are assumed to be independent of each other.

Moran’s \(I\) formula is as following:

\[
I = \frac{n}{s_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}
\]

Where \(\omega_{ij}\) is the weight between observation \(i\) and \(j\), and \(s_0\) is the sum of all \(\omega_{ij}\):

\[
s_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij}
\]

The expected value of \(I\) under the null hypothesis of no autocorrelation is not equivalent to zero but given by \(I_0 = -1/(n - 1)\). The expected variance of \(I_0\) is
also known, and so we can make a test of the null hypothesis. If the observed value of $I$ denoted $\hat{I}$ is significantly greater than $I_0$, then values of $x$ are positively auto correlated, while if $\hat{I} < I_0$, this points out negative autocorrelation. This allows designing one or two-tailed tests in the standard way (Ellingson and Andersen 2002).

### 2.21 Accuracy Assessment and Uncertainty in Maps Comparison

Land use/cover change simulation models basically examine land change maps at two separated periods ($t_0$ and $t_1$) and then, by evaluation of the occurred changes within these two periods and change factors, an appropriate simulation model is performed. This performance will predict land change maps for future periods (i.e. $t_2$). This predicted map at point $t_2$ can be typically compared to a reference map (i.e. the map of reality) to estimate the model performance. Therefore, in case the result of this comparison demonstrates a high degree of similarity, then it can be proved that the model was successful to simulate the changes. Even though, this result cannot necessarily indicate that the model provided supplementary findings beyond what the scientist would have predicted without the model. As we examined in this research, and as some scientists believe for most of the LUCC models, the agreement between the reference map of $t_1$ and $t_2$ always appears to be better than the agreement between the predicted map of $t_2$ and its reference map, at the resolution at which the model was run. Hence, this causes alarm in the LUCC modelling amongst scientists.

#### 2.21.1 Calibration and Validation

In this section, it is intended to clarify the distinction between two terms of ‘Calibration’ and ‘Validation’. It is better to distinguish between these two terms, whereas, they will be used after the model execution. Nevertheless, in most of the read papers, it was complicated to note their distinctions.

Rykiel (1996) noted that

Calibration is the estimation and adjustment of the model parameters and constraints to improve the agreement between model output and a data set,

whereas validation is

a demonstration that a model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model.

Besides, in some cases, the terms of “modelled” and “simulated” were used in error. For instance, the word “modelled” means that the model was fitted and the
term “simulated” indicated that something was predicted. Furthermore, in some other cases, there is a lack of methodology reflection (Wu and Webster 1998).

These issues lead scientific communities to a misunderstanding of the model’s certainty. Two separated types of data should be utilised for the calibration and evaluation processes. There are some ways to break up the calibration data from the validation data. Separation through time is one of the usual ways. Hence, if the model aims to predict the change on the landscape after time $t_1$, then any information at $t_1$ or before $t_1$ is justifiable to use for calibration. For instance, a usual calibration method is to carry out statistical regression on the change quantity between point $t_0$ and $t_1$. The results of the regression are fitted estimates. The fitted parameters and the regression relationship might be used to extrapolate the change between point $t_1$ and $t_2$; thus, any information subsequent to time $t_1$ cannot be used in the calibration process. The validation process compares the predicted map of time $t_2$ to the reference map of time $t_2$. Separation through space is another general method to separate calibration information from validation information. In this scheme, the model uses data from one study site to fit the parameters, and then the fitted model is applied to a different site to predict change. Distinction between the calibration process and the validation process can help to guarantee that the model is not over-fitted (Pontius et al. 2004).

Before implementing the simulation process, it has to compare the resultant maps arising from modelling execution to ensure the validity of the model. In fact, there is no universal concord to evaluate the goodness-of-fit of validation (Rykiel 1996). Each particular model comprises a specific purpose, and hence, the criterion should be related to the purpose. Besides, scale is essential to consider throughout any comparison of maps, since results might be sensitive to scale and, also, certain patterns may be evident at only certain scales (Quattrochi and Goodchild 1997; Kok et al. 2001; Pontius et al. 2004).

### 2.21.2 Techniques of Validation for Land Change Models

According to the definition of the United States Geological Survey (USGS), regarding the accuracy of geospatial data;

The closeness of results of observations, computations, or estimates to the true values or the values accepted as being true.

Nonetheless, the term “truth” has a certain definition. Accuracy assessment is one of the most imperative and significant steps in remote sensing and map analysis. GIS and remote sensing outcomes are being used as basic inputs for local, national, and global decisions; therefore, precise and accurate outputs lead researches towards correct routes. New users have to be taught about the reliability of the maps which are produced from GIS and remote sensing tasks (Banko 1998).
2.21.2.1 Visual Interpretation

A visual interpretation can give the scientists a better general assessment of model performance. Pure visual comparison is vulnerable to the personal opinion of the user; therefore, one scientist might believe the results perfect, and another might identify them poor.

2.21.2.2 Kappa Coefficient

Statistical techniques of results comparison are helpful to discover patterns that the individual mind ignores and also to facilitate communication between scientists. The Kappa coefficient was set up to the remote sensing societies in the early 1980s by Congalton et al. (1983), Congalton and Mead (1983), Cohen (1960) and became a pervasive measurement index for classification accuracy (Huang et al. 2002). It was recommended as a standard by Rosenfield and Fitzpatrick-Lins (1986). The Kappa coefficient measures the overall agreement of a matrix. The ratio of the summation of diagonal values to the total number of pixel counts in the matrix, the Kappa coefficient considers non-diagonal elements as well (Rosenfield and Fitzpatrick-Lins 1986). In fact, the Kappa coefficient computes the fraction of agreement after elimination of the chance agreement from considerations. A Kappa of 0 arises while the agreement between actual and reference maps equals chance agreement, and Kappa increases up to 1 (Banko 1998; Lakide 2009).

2.21.2.3 Relative Operating Characteristic

Pontius et al. (2004) have suggested a comparison technique that considers the agreement between two pairs of maps. The first comparison performs between the reference map of point t_1 and the reference map of time t_2. The second comparison performs between the predicted map of point t_2 and the reference map of point t_2. Eventually, the procedure evaluates the first comparison in comparison with the second comparison.

Swets (1986, 1988) depicts the logic of the ROC, although other researchers in this field explain how to compute the ROC in the digital maps comparison (Pontius et al. 2004; Pontius and Batchu 2003). The relative operating characteristic is a statistical measurement to compare a Boolean variable versus a categorical variable. The ROC is able to compute the accuracy of the prediction at several diverse threshold levels. For each threshold domain, each cell of probability surface map is reclassified as either over or under the threshold (Pontius and Batchu 2003).

Pontius and Batchu (2003) believe that the ROC is an outstanding method for analysing propensity surface values. Moreover, the ROC compares two maps specification in two separate ways: in terms of location and also quantity. ROC achieves this by the goodness-of-fit calculation of the validation at various thresholds, thereafter, aggregating the information at all thresholds into one measure of agreement. Accordingly, this method is purported to have distinct values concerning
the goodness-of-fit of location versus the goodness-of-fit of quantity; therefore, modellers will be able to improve the predictive model capability.

2.22 Summary

This chapter started with a brief introduction about the chapter and then introduced land use and land cover terms (Sect. 2.2) before making the distinction between these two terms, thereby highlighting the LUCC causes and consequences (Sect. 2.3). LUCC driving forces (Sect. 2.4) and LUCC simulation (Sect. 2.5) were introduced thereafter. The typical methodologies for trend evaluation of land use change was also discussed (Sect. 2.6). Section 2.7 dealt with how to predict the upcoming land use patterns and Sect. 2.8 with how to simulate urban sprawl. A comprehensive discussion over the popular and existing approaches in LUCC studies was presented in Sect. 2.8. A comprehensive explanation of geosimulation methodology, its terminology and characteristics was also depicted (Sects. 2.10–2.12). Later, Sects. 2.13, 2.14 argued about the differentiation between cellular automata and the GAS model. The existing simulation environments were taken into consideration to pick the optimum system for this research (Sect. 2.15). The fundamentals of a fuzzy decision system, the AHP weighting system, and the Moran autocorrelation coefficient were explained (Sect. 2.16–2.20). The usual map comparison methods were explained in order to evaluate the certainty of output maps.

References

References

Ellis E, Pontius Jr RG (2006) land-use and land-cover change—encyclopedia of earth. Available at: http://www.eoearth.org/article/land-use_and_land-cover_change


Forman EH, Selly MA (2001) Decision by objectives: how to convince others that you are right. World Scientific, Singapore


Kainz W (2008) Fuzzy logic and GIS. University of Vienna, Available at: http://homepage.univie.ac.at/wolfgang.kainz/Lehrveranstaltungen/ESRI_Fuzzy_Lo
Lambin EF, Geist HJ, Ellis E (2007) Causes of land-use and land-cover change. In encyclopedia of earth
Pontius RG Jr, Batchu K (2003) Using the relative operating characteristic to quantify certainty in prediction of location of land cover change in India. Trans GIS 7(4):467–484
Pontius RG Jr, Chen H (2006) GEOMOD modeling, idrisi andes help contents. Clark University, Massachusetts
Quattrochi DA, Goodchild MF (1997) Scale in remote sensing and GIS, 1st edn. CRC Press, Boca Raton
References
Timmermans H (2003) The saga of integrated land use-transport modeling: how many more dreams before we wake up. In Keynote paper auf der 10th international conference on travel behavior research, Luzern
International conference on complex systems (ICCS 2004), Boston, MA, pp 16–21
Dynamic land use/cover change modelling
Geosimulation and multiagent-based modelling
Jamal, J.A.
2012, XVI, 140 p., Hardcover
ISBN: 978-3-642-23704-1