Chapter 2*

Understanding the Urban Growth System

Abstract

The rapid urbanisation and urban sprawl in particular in the developing world require a scientific understanding of complex urban growth patterns and processes. This knowledge is highly crucial to sustainable land management and urban development planning. Progress in modern remote sensing and GIS techniques has opened up great opportunities, and significant success has already been achieved in monitoring and managing fast urban growth. However, these techniques are still poor when it comes to supporting decision-making on sustainable development, as reasonable theories and methods have not been sufficiently and systematically developed to understand the complexity inherent in urban growth. Understanding the urban growth system is a prerequisite for modelling and forecasting future trends of urban land use/cover change and its ecological impacts. As urban growth involves various actors with different patterns of behaviour, we argue that scientific understanding must be based on elaborated complexity theory and a multidisciplinary framework. The theoretical analysis can provide a guideline for selecting modelling methods currently available in complexity modelling and in remote sensing and GIS environments. This chapter first proposes a conceptual model for defining urban growth and its complexity, in which spatial, temporal and decision-making complexity are distinguished as separate domains. Second, this chapter links the conceptual model with the major current methods of modern urban modelling, such as cellular automata, fractals, neural networks, multi-agent, spatial statistics etc. This confrontation enables the possibilities of various modelling methods to understand urban growth complexity to be indicated. Third, this chapter evaluates the operational implementation of representative methods based on criteria such as interpretability, data need and GIS embeddedness.

Key words: understanding, urban growth, complexity, modelling, methods

* Based on Cheng et al. (2003a) and Cheng et al. (2003b).
2.1 Introduction

Geography is not about collecting facts, expressed as a proposition in logic, but about understanding the causes – the processes in space and time which created these facts (Frank, 2000). The process is typically represented by the complex interactions between humanity and nature. Traditional differential equations were the only well-known formalism to describe processes that affect change in time and space (Frank, 2000). However, this method is only suitable for physical geography and not appropriate for topics from human geography. The latter has properties that distinguish it completely from the former. The importance of human geography is strongly linked with the risks of human decision-making at varied spatial and temporal scales. Over the last 20 years, complexity issues have deeply affected modelling approaches in geography (Occel, 2002).

In the field of urban planning, one of the important subjects of concern is to predict the trend of land use transition (Osaragi and Kurisaki, 2000). However, prediction without a scientific understanding of the system under study implies a certain degree of uncertainty due to the numerous unknown factors involved. This may result in risky decision-making in urban development planning and management. Wrong decision-making may cause severe economic and environmental losses, or even lead to large disasters. As a consequence, scientific decision-making has been the pursuit of urban development planning and management that is highly dependent on the reasonable understanding of the objects involved. Understanding needs modelling to analyse the complex relationships involved in the decision-making; it also needs an understanding of the properties of the problems being studied.

To date, quite a number of models have been developed and applied in wide scientific areas. But most of them have been criticised. This may indicate that most objects being modelled are not completely understood conceptually. Rakodi (2001) argues that one of the proposals for improving the quality of planning is an attempt to improve the understanding and analysis of the interrelated components of the urban development process in order to arrive at more appropriate priorities and sets of policies.

Looking through the history of modelling, it is quite clear that its progress is dependent on the advances in other areas such as system sciences (including mathematics, physics and chemistry), computer science and techniques, and various application domains. Progress in system sciences and computer science has brought about a new revolution in quantitative geography. The "quantitative revolution" in economics, geography and the social sciences reached the planning profession in 1960s (Wegener, 2001). The emergence of "the old three system theories" (general system theory, information theory and cybernetics) and computer techniques in the 1940s spurred the first modelling revolution, which is based on structural linear equations but is not spatially explicit. Famous paradigms include the Lowry urban development model (Lowry, 1964), the spatial interaction model (Wilson, 1970) and the input-output model (Leontief, 1970). It is persuasive that the big forward movement in remote sensing (RS), geographical information science (GIS) and system theories, especially the developing complexity and non-linear theories (the most promising science in the 21st century), is undoubtedly stimulating a new development wave of modelling. The
reasons are threefold. First, complexity theory brings hopes for re-understanding the systems or phenomena under study. A recent resurgence of interest in complexity issues is evident as new theories and methods have mushroomed in the last few decades (Wu and David, 2002). Second, new mathematical methods create new means to represent and quantify the complexity. Third, remote sensing and GIS guarantee the availability of data on various spatial and temporal scales.

We argue that scientific understanding must be based on complexity theory and a multidisciplinary framework. In the field of urban analysis and modelling, perhaps the most promising approach has been the application of systems theory and ecological theory to the analysis of urban evolution and the flows of materials through the urban environment (Kropp, 1998). However, approaches that capture the complexity of large urban systems, and efforts to integrate the various themes are rare (Kropp, 1998). The application of complexity theory in urban analysis (qualitative or quantitative) has been increasing recently – for example deterministic chaos, stochastic dynamics, artificial life, ecological and natural evolutionary dynamics, evolutionary and genetic programming, cellular automata, percolation theory, cellular games, agent-based modelling, and neural networks. However, the complexity of urban growth and its impacts on urban development planning and sustainable growth management have not been systematically researched.

Here, within the framework of complexity theory and in the environments of remote sensing and GIS, we attempt to answer these questions: What is the urban growth system? And why and how should the complexity of this complex system be understood? With this purpose in mind, this chapter first proposes a conceptual model to define the urban growth system and then another conceptual model to project the complexity of the urban growth system onto spatial, temporal and decision-making process dimensions. Second, this chapter links the conceptual model with the major current methods of modern urban modelling such as cellular automata, fractals, neural networks, spatial statistics, multi-agent etc. This confrontation makes it possible to indicate the possibilities of the various modelling methods to understand urban growth complexity. Third, this chapter evaluates the operational implementation of representative methods based on criteria such as interpretability, data need and GIS embeddedness. Finally, it ends with some discussion and conclusions.

### 2.2 Complexity of Urban Growth

Modelling urban growth aims to support urban development planning and sustainable growth management. Scientific planning and management must be based on the proper understanding of the dynamic process of urban growth, i.e. from past to present to future. Such understanding enables planners to experimentally simulate "what-if" decision-making based on various scenarios. However, the dynamic process involves various socio-economic and physical and ecological components at varied spatial and temporal scales, which result in such a complex and dynamic system. Consequently, it requires a systematic perspective to understand this complexity.
Chapter 2

2.2.1 Complexity

That the urban system is highly complex has become a well-recognised fact. Systems thinking has been widely accepted by urban planners and other decision-makers engaged in urban management and construction. While the concepts of "complexity" themselves are not new, the application of these concepts to socio-economic processes is a relatively new phenomenon. Advocates of complexity theory see it as a means of simplifying seemingly complex systems. Complexity often results from the non-linear interactions among complex system components, which frequently lead to emergent properties, unexpected dynamics and the characteristics of self-organisation becoming the basic properties of complex systems.

Non-linear relationships and feedback among all components at the same and different scales often lead to instability and unpredictability in large complex systems.

Emergence (as a phenomenon that high-level behaviours emerge naturally out of low-level interactions) implies that the behaviour of the small part is different in isolation than when it is part of the larger system. This description is often summarised as "a whole that is greater than the sum of its parts or in simple terms, much coming from little". Thus the collective behaviour of a complex system is dependent on the behaviour of all of its parts.

For example, Portugali and Benenson (1997), who have intensively studied the theoretical aspects of socio-cultural emergence during recent years, show the emergence of different forms of cultural and economic segregation as a consequence of the interactions between individuals and the city environment at the local and global levels.

Self-organisation (the spontaneous emergence of macroscopic non-equilibrium organised structure due to the collective interactions among a large assemblage of simple microscopic objects as they react and adapt to their environment) implies that the system organises itself from within and structures are not imposed from the outside – in other words, owing to purely internal dynamics instead of any external force. It requires an interaction with its environment and non-linear relations between its elements.

In a self-organising system (SOS), the local actions and interactions of individuals are the source of the higher-level organisation of the system into patterned ordered structures with recognisable dynamics. Since the origins of order in SOS are the subtle differences among components and the interactions among them, system dynamics cannot be understood by decomposing the system into its constituent parts. Self-organisation theory suggests that insignificant local interaction behaviour can lead eventually to a qualitatively different global structure (Wu, 1998a; Batty, 1995), which constitutes the basis of cellular automata and multi-agents theory. SOS theory has been applied for explaining many urban phenomena, such as spatial economies (Krugman, 1996) and urban evolution (Allen, 1997b; Haken and Portugali, 1995; Portugali, 1999; Schweitzer and Steinbrink, 1998). Order in the spatial structures or urban systems emerges from the structured responses of multitudes of individuals to outside forces and constraints (Benguigui et al., 2001b).
A branch of self-organisation theory, synergetics, attempts to illuminate explicit relationships between the behaviour of individuals (micro level) and evolving patterns (macro level). The approach is based on concepts such as order parameters, which typically represent macroscopic patterns, and the so-called slaving principle showing the relation to microscopic structures (Daffertshofer et al., 2001). The initial goal of synergetics was to understand how the emergence of a macroscopic system, showing a high degree of order, may be explained by the microscopic behaviours (Tannier and Frankhauser, 2001). The principles have been gradually popularised and applied to socio-economic and ecological systems. Haken and Portugali (1995) applied a synergetic approach to explain the self-organisation of urban settlement, based on a framework of pattern recognition within which the interplay between the material pattern of cities and the cognitive pattern of cities were conceptualised and subsequently analysed.

Complexity frequently takes the form of hierarchy, whereby a complex system consists of interrelated subsystems that are in turn composed of their own subsystems, and so on, until the level of elementary component is reached (Kronert et al., 2001). Hierarchy theory applies hierarchy to organise concepts and interpret various complexities. The theory examines closely the issues of scale, levels of organisation, levels of observation, and levels of explanation in a complex system characterised by hierarchical structures and interactions across levels. Hierarchy theory suggests that when a phenomenon is studied at a particular hierarchical level (the focal level, often denoted as Level 0), the mechanistic understanding comes from the next lower level (Level -1), whereas the significance of that phenomenon can only be revealed at the next higher level (Level +1) (Kronert et al., 2001).

The key to understanding hierarchical structure is scale. Scale is the central concept for describing and explaining the complex hierarchical organisation of the geographical world (Marceau, 1999). In a hierarchical system, higher levels (or smaller scale) set constraints or boundary conditions for lower levels. The latter operate much too rapidly to be of interest and can be ignored. In spatial analysis, the scope of scale can be threefold: spatial, temporal and decision-making (see chapter 5).

### 2.2.2 Complex system of urban growth

When we consider urban growth as a system, in particular a complex system, we need to uncover the universal and unique characteristics that it shares with and distinguishes it from other complex systems. This exploration is conducted by answering four relevant research questions. The first question is: *Where is urban growth occurring from a system perspective?*

As far as the type of urban development is concerned, it consists of physical expansion and functional changes. The former refers to the change in space (transition from non-built-up to urban), such as increasing the physical size of a built-up area, the latter to the change in major activities (land uses), such as residential or commercial function. Although the focus of this research is on the physical expansion, the functional aspects have to be taken into account in interpreting the causal effects of the former as both interact spatially and temporally. For example, the activities at a location may influence the change in space at
Another location; the activities in a period may impact on the change in space at another later period. As a result, space and activity should be the basic elements of any systems defined for understanding urban growth.

In figure 2.1, it is supposed that urban growth occurred in a specific period from time \( t_1 \) to \( t_2 \); apparently the evolution of urban growth is closely related to three systems – \( P \), \( U \) and \( N \). \( U \) itself is a highly complex social and economic system, as the concentration of considerable urban activities present at time \( t_1 \) shows. It offers current activities rather than space for urban growth to come. \( N \) is a typical physical and ecological system, including various ecological units (water body, forest etc.) and agricultural land. It primarily provides possible opportunities and potential for urban growth in space, instead of activities until time \( t_2 \). \( P \) is a spatial and conceptual system that results from a spatial planning scheme. It prepares organised space and activities for urban growth in the future. Urban growth is a temporally relative term. New development units will be administratively transformed from rural management into an urban built-up area after a certain term has elapsed since birth. For example, in figure 2.1 urban growth, being the transformed area from \( t_1 \) to \( t_2 \), will become a part of system \( U \) after \( t_2 \) from system \( N \) at time \( t_1 \). As the main topic of this research, new urban growth is treated here as an independent system within the specific period under modelling. Under such an assumption, urban growth \( G \) can be defined as a system resulting from the complex dynamic interactions (only from \( t_1 \) to \( t_2 \)) between the three systems \( (P, U \) and \( N) \). The thin arrows in figure 2.1 refer to the interaction between the three systems, and the thick arrows to the contributions to urban growth made by the three systems. System \( P \) contributes planning control and requirements to \( G \); system \( N \)
contributes developable land, and system $U$ contributes activities and stimulant factors to the growth of $G$.

A key to understanding urban growth is to understand the complex dynamic interactions. In terms of physics, system $U$ exerts "pull" forces on system $G$, which is attracted by a certain scale of urban social and economic activities. Conversely, system $N$ exerts "push" forces on $G$, which is excluded by the limitation and requirement of ecological protection or sustainable agriculture. Hence, $G$ results from the interaction between "push" and "pull" forces. We can say the interaction is open, non-linear, dynamic and emergent. Urban growth is a self-organised system.

The major decision-making in urban growth is related to plans, policies and projects. Projects are special land use or development proposals initiated usually by various levels of actors. Projects evolve in the context of various levels of policy and plans. Urban growth creates a new dynamic system, which comprises a quantity of projects constructed that are increasing with time from $t_1$ to $t_2$. It is an open system. In the course of urban development, it incessantly exchanges matter, energy and information with external physical and ecological systems (water, land), other regions and cities. It imports a variety of regulations/decision-making styles, investment from higher organisations, external investors, inhabitants and managers. Its non-linearity is indicated in the following aspects. In the spatial dimension, new development density (population density or land conversion) decreases non-lineally with the distance from the city centre and sub-centres. This is mostly represented by a negative exponential function (Clark, 1951) or an inverse power function (Batty and Kim, 1992). In the temporal dimension, new growth does not follow a linear trend but, in most cases, a logistic trend (Herbert and Thomas, 1997). The interactions among a huge number of factors have proved to have the unknown non-linear relationship, such as the famous interaction between transport and land use (Wilson, 1998).

The structure and function of each local project depend not only on its neighbouring projects but also its built-up environment, i.e. these new projects interact not only with each other but with developed areas, as well as spatially and temporally. These non-linear interactions result in globally ordered land use patterns. The order is typically indicated by a large-scale spatial agglomeration or by clustered patterns. From this, we can infer that urban growth is a typical self-organised system where the three systems are treated as a whole.

As a focus, this research only discusses the impacts of other systems on urban growth, as indicated by the one-way arrows (figure 2.1). In reality, the impacts of system $G$ on $N$ have been the major concern of landscape ecology, the interactions inside system $U$ being the major concern of urban land use change. Therefore, urban growth involves landscape ecology (pattern and process), urban planning (decision-making) and urban geography (activities and behaviours). We need an interdisciplinary instrument to understand these complex relationships. Complexity theory is undoubtedly an ideal tool to construct conceptual frameworks systematically.
Second, we need to answer the questions: *What should be understood in supporting urban development planning and management? And how can urban growth be represented for modelling purpose?* Traditional approaches to urban science as exemplified in the work of Christaller and others are based on the assumption that cities grow homogeneously in a manner that suggests that their morphology can be described using conventional Euclidean geometry. However, recent studies have shown that the complex spatial phenomena associated with actual urban systems are better described as a dynamic process consistent with growth in disordered patterns. The process of urban growth does not exist independently but rather coexists with pattern and behaviour. They interact mutually and comprise three interrelated conceptual subsystems that are crucial to the decision-making for urban planning and management. The work of Sui (1998) shows a need to understand urban form, process and policies in this new information society. When moving to urban growth, an emphasis should be given to pattern, process and behaviour.

As illustrated in figure 2.2, understanding urban growth can be summarised as five interweaving levels: policy, actor, behaviour, process and pattern. Policy is the level proven to be the most influential factor or driving force of urban growth on the macro scale. Pattern is the lowest level, which is a directly observable outcome. Process indicates the dynamics of urban growth, behaviour indicates the actions of the actors involved, and actors indicate the agents of behaviour. From policy to pattern, the qualitative degree is decreasing and the quantitative degree is increasing. As a result, modelling has to follow a ladder (figure 2.2), from pattern gradually to policy level. This ladder works in the opposite direction to the real urban growth hierarchy. On the one hand, in the terms of hierarchy theory (see previous section), understanding a single level must consider its lower and upper levels as they are comparatively closely linked.

![Figure 2.2 A ladder for modelling](image-url)
Consequently, to understand a process, one must take its pattern and behaviour into account. On the other hand, as actor and policy are interrelated, they principally impact on the decision-making units and processes, and are linked with behaviours. Thus, process, pattern and behaviour are becoming the key levels for modelling urban growth. A pattern is the temporal snapshot of a process, and behaviour is the decision-making source of a process.

(1) Pattern

What is the definition of pattern? How to classify and distinguish the patterns of urban growth? The Oxford English Reference Dictionary defines pattern as "a regular or logic form, order, or arrangement of parts such as behaviour pattern". Two key components are stressed in the definition: elements and the logical ordering among the elements. As such, spatial pattern focuses on the spatially ordering and temporal pattern on the dynamic ordering, i.e. logically ordering described from the perspectives of space and time respectively. However, the concept "pattern" varies with discipline in academic circles. In spatial sciences, pattern refers to a "regular arrangement of objects", which may be explained in terms of structures, processes and systems. It refers to the manner in which a phenomenon is arranged in time and systems. In landscape ecology, patterns refer to the spatial configuration of discrete landscape elements, which can be of different geometrical nature.

To summarise, pattern is a relative term, which is dependent on a specific system under study. Pattern is based on defined elements of the system. In this sense, urban growth patterns can be viewed from two standpoints: one is on the urban growth system itself, the other is as part of a larger system \( (G, U, P, N) \). The former only comprises new development units. The latter includes not only urban growth but also the three other systems \( U, P, N \). A development unit can be defined as any spatial entity that will be subject to change, albeit physical or functional, during the period to be modelled. The physical change means the appearance or disappearance of a new unit; the functional change indicates the new usage of a unit, such as change from industrial to commercial. The pattern in system \( G \) is called univariate as it focuses on the logic arrangement among the new development units. Landscape metric, point-pattern and spatial auto-correlation belong to this category, which contributes to the quantitative description of the spatial distribution of urban growth. In a larger system, elements include relevant spatial entities coming from three other systems, which stimulate or constrain the occurrence of new development units. They can be river, water body, railway line, slope, shopping centre, road network etc. Actually, this category aims to model the spatial relationship between \( G \) and \( P, N, U \), instead of \( G \)'s spatial distribution. It is called a multivariate pattern, which contributes to the quantitative description of interaction pattern between multiple systems.

Various types of "pattern" studies have been carried out in urban modelling, such as residential or settlement pattern (I-Shian, 1998), land use development pattern (Kiril, 1998; Yeh and Li, 1998), population and employment pattern (Ingram, 1998), development pattern of informal settlement (Mahmud and Duyar-Kienast, 2001), land development pattern (Wu and Yeh, 1997), and transport/land use interaction pattern (Susantono, 1998).
Wu and Yeh (1997) focus on the multivariate functional pattern (land use) between system $G$ and $P, N, U$ in a case study of Guangzhou city in China. They model not only spatial patterns but also temporal patterns in two periods. These models are very helpful in comparing and explaining the land development patterns under two distinct economic systems.

Pijanowskia et al. (2002) study the multivariate physical pattern (land cover) between system $G$ and $P, N, U$ in a case study of Michigan’s Grand Traverse Bay. They explore how factors such as roads, highways, residential streets, rivers, the Great Lakes’ coastlines, recreational facilities, inland lakes, agricultural density and the quality of views can influence urbanisation patterns in this coastal watershed. Artificial neural networks (ANNs) are used to learn the patterns of development in the region and test the predictive capacity of the model, while GIS is used to develop the spatial predictor drivers and perform spatial analysis on the results.

(2) Process

Space and time are well-known notions but in order to explain them they must be connected to other fundamental concepts such as change or process. Since relative space is inseparably fused with relative time, nothing in the physical world is purely spatial or temporal; everything is process. Change must be seen as a composite of processes that occur on a wide band of time scales in space. Therefore, the link between space and time is through the process itself, where specific processes determine specific temporal and spatial conceptualisation (Dragicevic et al., 2001). The process discussed here does not include the social and economic processes, which are the driving forces of physical and functional urban growth.

Process generally refers to the sequence of changes in space and time; the former is called the spatial process, the latter the temporal process. It should be noted that strictly speaking spatial and temporal processes cannot be clearly separated as any geographical phenomena are bound to have spatial and temporal dimensions or named a spatio-temporal process. Understanding change through both time and space should, theoretically, lead to an improved understanding of change and of the processes driving change (Gregory, 2002).

However, spatial processes are much more than any sequence of changes. Spatial process implies a logical sequence of changes being carried on in some definite manner, which lead to a recognisable result (Getis and Boots, 1978). Summing up, the key components of process are change and logical sequence. The former is defined by a series of patterns. The latter implies the understanding of process. In contrast with pattern, process contains a component of dynamics.

Pang and Shi (2002) propose a spatial system theory in which they define spatial process as a system containing two components: based on structure and movement (including add, delete, move, merge and subdivide operations). They actually correspond to pattern and sequence (a set of operations) of change. This is a generalised process for spatial modelling.
Understanding the urban growth system

in GIS. However, it is not suitable for urban growth as urban growth only includes rural-urban land cover conversion and not decline in land use in the inner city.

In landscape ecology, fragmentation is a common process related to landscape change, affecting both its structure and function. It causes the division of landscape elements into smaller pieces. In this domain, landscape pattern comprises various patches, which represent the diverse structure and function of landscape elements. Fragmentation of patches or patch dynamics (Wu and David, 2002) can be utilised to explain the ecological process of landscape pattern change. It is a spatial process in system $N$.

Landis and Zhang (2000) define spatial processes as those by which activities at one location affect or are affected by activities at another location. They identify four types of spatial processes that arise in urban activities: spatial diffusion and dispersal, exchange and transfer, interaction, and segmentation or percolation. The outcome of the spatial process refers to urban land use. Process is to understand the causal relationships of urban land use change; they are spatial processes in system $U$. Arbia (2001) classifies the spatial processes of individual firms into a birth process (new firms) and a growth process (existing firms) and proposes a model of economic activities on a continuous space. This classification aims to analyse the economic behaviour of individual firms. It is not an explicit spatial process.

Benguigui et al. (2001a) described city growth as a leapfrogging process, based on population growth in a case study of the Tel-Aviv Metropolis. With reference to three case studies (Beijing, Shanghai and Guangzhou), Gaubatz (1999) generalised the urban development process of Chinese cities after land reform was initiated in 1987 into three aspects: production of urban plans, urban renewal, and privatisation of the housing and real estate market. These two definitions are only given as a specific requirement of the analysis, not in any systematic way.

According to hierarchical theory, processes – in particular spatial processes – may be divided into two levels: global and local. The former takes the whole study area into account, the latter only a neighbourhood. For example, Mendonca-Santos and Claramunt (2001) defined explicitly spatial processes on two scales (landscape and class or local) in order to explain the change in landscape patterns. At the landscape level (global), spatial processes are identified as fragmentation, perforation, diversification and simplification. And at the class level (local), spatial processes are characterised by expansion, contraction, stability, invasion, domination and succession. They argued that different levels (landscape and class) have a specific time scale: evolution process on a local scale is likely to happen in a faster mode than the ones identified at the landscape level.

To sum up, the classification of process is very complex, and dependent on the specific requirements of the analysis. The same patterns can be explained from the standpoint of different processes. In urban growth, when we focus on urban growth system $G$, we may classify it as a spontaneous or self-organised process. The former is indicated by sporadic patterns and the latter is reflected by clustered patterns. This classification can be better linked to social and economic processes and also the decision-making processes of urban development planning. When we focus on the interaction between the three systems, we
can define the process as leapfrog, space fill-in, dispersed, scattered, road-influenced, spread etc. This classification in particular considers the interaction of urban growth with developed urban areas from a global perspective.

(3) Behaviour

Urban growth results from direct or indirect decisions to alter the current uses of land at various levels. The analysis of urban growth necessarily asks who decides to change the transition, and where, when and why. The factors that are taken into account relate to the particular decision-making units and processes.

Behaviour refers to the decision-making of actors. Spatial behaviour focuses on spatial decision-making, temporal behaviour on temporal decision-making. The key components of behaviour are decision making and the actor. In spatial science, examples include way-finding, travel mode, site selection, and land use allocation.

The actors from the three systems $U$, $P$, $N$ – individuals, households, businesses, developers, farmers, landowners, planners and governments – make decisions about their social and economic activities, and their spatial location and temporal scheduling, leading to changes in land cover and land use. These decisions affect, directly and indirectly, the physical and functional system $G$ through the conversion of land, the use of resources, and the generation of interaction.

For example, the projects for commercial use make choices about scale, location, cost and transport. Households make choices about employment, location, housing type, travel mode, and other lifestyle factors leading to varied spatial behaviours. Developers make decisions about investing in development and redevelopment. Governments make decisions about investing in infrastructures and services and adopting policies and regulations. Decisions take place at the individual and community levels through the economic and social institutions. The actors interact in three sub-markets: the job market, the land market and the housing market. These actors also interact in non-market institutions, including governmental and other non-profit and non-governmental organisations.

A variety of decision making and diverse actors create disparate spatial and temporal behaviours in the urban growth process. Urban growth is highly impacted or controlled by the major actors of urban construction, planning and management. Urban spatial structure can be described as a cumulative and aggregate order that results from numerous locally made decisions involving a large number of intelligent and adaptive agents. The behaviour of these agents is subject to their rules of action based upon new information. The local behaviour of multiple decision-makers can eventually lead to qualitatively different global patterns. Due to the number of actors involved in urban growth, the spatial behaviour of urban growth falls into various levels: individual behaviour, planners’ behaviour, developers’ behaviour etc. Spatial behaviour regarding urban growth includes site selection behaviour and spatial spread behaviour (such as scale, density, intensity). Temporal behaviour contains the speed of growth.
Decisions are made under constraints (space and time), and they reflect the attitudes, values and beliefs of people and of society. Therefore, behaviours are individually subjective and stochastic but follow global regulation in statistical terms. Meanwhile the basis for making decisions may change over time.

For instance, previous studies regarding the urban growth of Chinese cities (Gaubatz, 1999; Fung, 1981; Wu, 1998b) have shown that because of the determinant role of the state budget the state and work units were the main urban developers in the period before 1987. Urban planning principally contributed to site selection for industrial projects. Since the land reform initiated in 1987, however, with the retreat of state work units from urban construction, the comprehensive management of local governments and the new land-leasing system, the right of controlling urban space has been transferred from work units to local governments and then to external developers (Wu, 2000a). More actors are involved in the decision-making, with more vague functions (Han, 2000; Jiang et al., 1998; Zhang, 2000b). As a result, to understand the urban development process, the roles of various actors and their behaviour should be taken into account. Wu (1998b) argues that, in order to explain the complicated spatial structure and process of Chinese cities, one must understand the two points: capital and its movement, social actors/agents and their functions/roles.

2.2.3 Projection of complexity in urban growth

Much of our understanding of explicit dynamic processes will coincide with our ability to understand complex systems in general (Box, 2000). A third question is: What is the complexity of urban growth? or How should we look at its complexity?

(1) Sources and measurement of complexity

Contemporary urban growth is characterised by dispersal and decentralised patterns, especially in the USA (Gordon and Richardson, 1997). Restructuring has involved the decentralisation of jobs, services and residences from traditional urban centres to suburban settings and "edge cities" within expanded metropolitan areas (Garreau, 1991). The new urban regions are multi-centred, with more than one core (Fishman, 1990). This trend is the result of a variety of heterogeneity.

Kolasa and Pickett (1992) gave a more conceptual definition of heterogeneity: "a system is heterogeneous in time and/or space if a specific temporal interval and/or different location is characterised by different values". Homogeneity and heterogeneity can be defined as the "border" between individual levels of hierarchy.

Systems of any interest are composed of heterogeneous agents and objects, indeed their very richness comes from such heterogeneity (Batty and Torrens, 2001). Socio-economic events have an explicit heterogeneous spatiality and temporality. Their structure and function are defined in and by space, as well as in and by time, such that no two locations are alike. Maintaining heterogeneity may be critical for the movement of energy, matter and information within different social contexts. Or rather, heterogeneity is the source of complexity of any system.
In the field of landscape ecology, increasing attention is given to the importance of spatial heterogeneity in understanding the relationship between pattern and process (Turner, 1989). Hierarchical patch dynamic models are being developed to incorporate the effects of spatial heterogeneity on ecosystem dynamics. An increasing hierarchical order is often accompanied by an increase in heterogeneity (Kronert et al., 2001).

Urban growth consists of the various scales of new projects. Large-scale projects are characterised by heavy investment, long-term construction and the number of actors involved; examples include airports, industrial parks and universities. In contrast, small-scale projects are characterised by rapid construction, light investment and few actors; examples can be a private house and a small shop.

Urban growth results in various land uses with different levels of social, economic and environmental values. This is a higher dimension of heterogeneity, indicated in the attributes of spatial objects. For instance, a university accommodating many people has a high social value but a low economic value. Conversely, a sewage treatment plant accommodating few workers has a low social and economic value but a high environmental value. As a result, each unit of new development is assigned different values. They are the spatial entities carrying heterogeneous social, economic and environmental activities.

Consequently, urban growth comprises a large number of varied scale projects. The functional differences between them, and also between the new units and the other three systems, create a massive flow of matter, people, energy and information. They are the sources of the complexity inherent in urban growth. Our observation or assumption is that the spatial, temporal and decision-making heterogeneity of urban growth results from socio-economic-ecological heterogeneity. Such heterogeneity may originate from self-organised socio-economic processes. For example, the self-organised process to some extent can be explained by scale economy, multiple nuclei etc. The integration or interaction between these categories of heterogeneity creates complex patterns, behaviours and processes of urban growth.

As a first step towards decision-making support, quantitative measurement plays a crucial role, affecting the accuracy of modelling and further the risks of decision-making. To effectively measure the complexity of a system remains an unsolved issue even in complexity theory. In urban growth, such complexity can be threefold (or projected onto): spatial measurement, temporal measurement and decision-making measurement, which correspond to the three categories of heterogeneity. In the published literature, although numerous indicators are designed for the quantification required by any specific analysis such as proximity, accessibility, and density based on remote sensing and GIS techniques, they are still not rich enough to understand all aspects of multiple complexity. A major reason is that conceptual understanding of any specific complex system is still limited at present.
(2) Spatial complexity

Classic location theory utilises micro-economic concepts such as perfect competition and perfect rationality; but the over-simplified economic and spatial landscape it assumes is not sufficient to explain existing spatial processes where location choices depend on relationships rather than on an individual actor's choices (Besussi et al., 1998).

A frequently cited shortcoming of GIS and most spatial analysis tools is their difficulty in dealing with dynamic processes over landscapes (Box, 2000). This is not because of a lack of people thinking about dynamic processes in space, nor is it from a lack of talent or technology. It has more to do with the fact that space is inherently complex, and dynamic processes often become complex when they are regarded in a spatial context. As a result, the first step to spatial modelling is to recognise the spatial complexity in the study. Spatial complexity may include spatial interdependence, multi-scale issues and structural or functional complexity.

Spatial dependence is defined as a functional relationship between what happens at one point in space and what happens at a neighbouring point. In urban growth, spatial dependence is indicated by the impacts of neighbouring sites on land conversion of any site – which is the result of a causal relationship among neighbouring entities, e.g. interaction. The impacts can be twofold: positive (stimulation) or negative (constraint) from three systems \((U, P, N)\). Examples of positive impacts may include transport infrastructure or developed urban area; in particular low density fringe growth is highly dependent on transport infrastructure. Examples of negative impacts may be steep terrain and non-developable land such as deep lakes. The complexity lies in the following facts:

- The impacts are determined by an unknown number of factors and their spatial relationships are non-linear;
- The intensity of spatial dependence or neighbourhood size is spatially and locally varied;
- Land conversion includes probability (occurred or not), density (scale), intensity (floor number), function (land use) and structure (shape or morphology); each may have its distinct spatial dependence.

Urban growth involves a number of hierarchical structures. In the spatial dimension, \(U\) includes different levels of shopping centres and road networks; system \(N\) includes different levels of ecological units; system \(P\) contains different levels of urban planning (general plan, district plan and zoning plan). As a result, urban growth \(G\) may be related to more complex spatial hierarchies as interacting with three systems. From the perspective of land development, urban growth can be divided into different scales of projects, such as large-scale new development zones, key-point industrial zones or parks, middle-scale new residential areas, and small-scale shops. Spatial complexity resulting from the multi-scale issue lies in the following facts:
Urban growth pattern, process and behaviour and their relationships are spatially varied with different scales;

- The relationships between scale and various levels of urban development planning and land management are still fathomless;
- The spatial framework supporting multi-scale modelling is impacted by numerous institutional factors, especially in developing countries.

Patterns and processes have components that are reciprocally related, and both patterns and processes, as well as their relationships, change with scale. Different patterns and processes usually differ in the characteristic scales at which they operate. Scale issues are inherent in studies examining the physical and human forces driving land use and land cover changes (Currit, 2000). An understanding of how processes operate at various spatial scales and how they can be linked across scales becomes a primary goal when investigating these complex phenomena (Marceau, 1999).

In spatial science, structure is the physical arrangement of ecological, physical and social components, and function refers to the way the components interact (Zipperer et al., 2000). Urban growth involves both; structure is more linked with pattern and function rather than with process. The representation or semantics understanding of a spatial system is diverse. The spatial representation of structure and function may influence the spatial understanding of urban growth pattern and process. Its complexity lies in the following:

- The self-organised process of urban growth has complex spatial representation and understanding;
- The interaction between pattern and process is dynamic and non-linear.

### Temporal complexity

In the time dimension, the physical size of a city is increasing continuously, with a functional decline in some parts, such as the inner city. However, urban growth means only increasing the number of new units transformed from non-urban resources. In different countries, regions and cities, the speed, rate and trend of urban growth are very distinct. In developed countries (e.g. the USA, the UK), urban growth may be much more gentle than that in rapidly developing countries such as China and India. Urban growth is largely controlled or impacted by its economic development scale and environmental protection strategy. Or rather it is controlled by the systematic co-ordination between the three systems. For example, when system \( N \) is not influential and strong, more arable land might be encroached upon. Economic development is not predictive, in particular in the long term, due to numerous uncertain factors. The non-linear interactions between the three systems lead to a non-linear curve of urban growth. This results in patterns, processes and behaviours of urban growth that are temporally varied, i.e. temporal scale is a highly influential factor for understanding its dynamic process. In the longer term, urban growth might be considered uncertain and unpredictable or even chaotic. Urban systems are rather complicated and their exact evolution is unpredictable (Yeh and Li, 2001a). This means its development process is sensitive to unknown initial conditions such as war, natural disaster, and new policies of the central government. These conditions can not often be predicted,
Understanding the urban growth system

particularly in quantitative terms. If the system of interest is chaotic, the prediction of the values of all system variables is possible only within a usually short time horizon.

Generally, urban growth is in a state of disequilibrium, especially in most rapidly developing cities. In such cases, uncertainty becomes predominantly important, because in these systems spontaneous growth or any surprising changes that depart from observed past trends, indeed any form of novelty or innovation, open up the path of system evolution. This can be illustrated by the so-called "evolutionary drive", showing how the error-making of a particular type of individual in an initially pure population eventually diversifies the characteristics and behaviours of the population. This indicates that the urban development process contains stochastic components to a certain degree.

The temporal scales of various decision-making are also different. Large-scale projects such as shopping centres or industrial parks frequently take a few years, much longer than small-scale constructions such as a shop. It is likely that various levels of actors have different temporal scales of decision-making behaviour. Local government needs to have a series of procedures, such as public participation or interviews with local people, to support their democratic decision-making. Individuals or households are able to make much quicker decisions because their decision-making process is simple and the criteria for their decision objectives are also fewer.

From the perspective of urban planning and management, understanding the dynamic process of urban growth includes the temporal comparison of various periods. Such comparisons enable planners to modify or update their planning schemes in order to adapt to the changing environment. However, these comparisons are subjective and depend on numerous fuzzy criteria.

As a complex system, urban growth involves a certain degree of unpredictability, phases of rapid and surprising change, and the emergence of system-wide properties. Temporal complexity is specifically indicated in the following ways:

- Patterns, processes and behaviours of urban growth are temporally varied with scale;
- The dynamic process of urban growth is non-linear, stochastic or even chaotic in the longer term;
- Temporal comparison of urban growth is subjective and fuzzy.

(4) Decision-making complexity

Quantitative geographers increasingly recognise that spatial patterns resulting from human decisions need to account for aspects of human decision-making processes (Fotheringham et al., 2000). In particular, the urban spatial structure is viewed as a result of interlocked multiple decision-making processes (Allen and Sanglier, 1981a). Decision-making complexity is indicated in the unit and process of decision-making, and the actors or decision-makers.
The decision-making unit and process of large-scale projects are relatively more complicated than those of small-scale ones. They involve more actors or decision-makers. For example, in China, decision-making in an industrial park project may include investment sources, site location, development scale, time scheduling. Actors may include central government, local government, foreign investors, local developers and work units. However, a small shop only needs the decision-making of one private developer. Large-scale projects are limited in quantity and their decision-making is more certain and well planned if compared with others. The latter are large in quantity and their decision-making is more uncertain, dynamic and less organised. However, the collective behaviours of small-scale projects can be controlled or guided by various management and urban development policies. From the perspective of self-organising theory, all of these small-scale and large-scale projects are spatially and temporally self-organised into an ordering system. The decision-making behaviours of different functions of projects are also disparate, e.g. commercial and residential. Their differences are indicated in the various actors and the criteria for respective decision-making. Consequently, decision-making in urban growth is a completely multi-agent, dynamic and stochastic system.

As discussed above, urban growth involves various levels and scales of decision-making, from individual land rent to a government's master plan. Each actor has a distinguishing domain of decision-making and profit pursuit, which are frequently in conflict. The interactions between these actors are spatially and temporally varied. This is a typical multi-agent system spanning broad spatial and temporal scales.

Understanding the dynamic process of urban growth must be based on the linkage with the decision-making process as the final users of modelling are the various levels of decision-makers. However, the interaction between these actors is in essence non-linear, dynamic, and self-organised. The ability to realistically represent the behaviour of the key actors depends on the level of aggregation at which actors and their behaviours will be represented in the model. Real decision-makers are a diffuse and often diversified group of people who will make a series of relevant decisions and trade-offs over a period of time. Their decisions will depend on a broad range of characteristics, such as site characteristics, locational conditions and legal constraints. Furthermore, in the real world the costs and benefits of alternative decisions are both distributed and valued differently among these decision-makers. In addition it is important to note that these actors also learn through time. Hence, the interaction between the spatial, temporal and decision-making processes is much more complicated.

Summing up, decision-making complexity is specifically indicated as follows:

- Decision-making for urban growth is a multi-agent dynamic and stochastic system;
- Its spatial and temporal projection is a self-organised process;
- Decision-making behaviours are subjective and fuzzy.
(5) **An example in transport and land use interaction**

The pattern of urban development principally results from the accumulative effects of transport/land use interactions at different spatial and temporal scales. The term *interaction* implies a *feedback* mechanism between transport and land use systems. The land use system supplies the transport system with estimates of the location and volume of travel generators. The transport system affects the land use system through the notion of accessibility, often in a temporally lagged manner. As an integral part of such accessibility, changes in travel costs become part of the mechanism used to relocate labour, residence and other urban economic activities. Many empirical studies have shown that the interactions are complex, bi-directional, and difficult to sort out due to spatial and temporal scaling factors. In the temporal scale, the interactions can be distinguished and summarised as follows (Hanson, 1995):

- Short-term effects of land use on transport;
- Medium-term effects of transport on employment location;
- Long-term effects of transport on housing location.

The implication is that transport system changes, notably major infrastructure investment in new highways or rail transit lines, will need time to affect urban land use patterns. Once introduced, such land use patterns may also, but within shorter time frames, induce further changes in urban travel demand.

At the spatial scale, on the one hand, the link between transportation and land use may be stronger only when transport costs are significant, or when transport or development decisions significantly affect accessibility. These conditions are generally met in two very different circumstances: *heavily congested downtown areas* and *rapidly growing suburban areas*. On the other hand, the impact of highway investments today, with a mature highway system, may not be the same as in earlier periods. They have a decreasing impact.

### 2.3 Complexity Modelling

This section is going to answer the fourth question: *How can the complexity of urban growth be modelled (understood) and what are the strengths and weaknesses of each method from the perspective of complexity described above?*

In one philosophical tradition, understanding means the construction of models (Newell, 1997). There are a number of ways of classifying models of urban growth. For example, in terms of system completeness, models can be system-level or specific-level. The former takes all components of urban systems into account; the latter focuses only on a specific phenomenon or problem by using a limited number of components in the system under study, such as residential dynamics. In terms of dimension, they can be divided into spatial models, temporal models and spatio-temporal models. Different dimensions distinguish focus or emphasis and requirements of data. In terms of analysis objectives, they can be
pattern models, process models and behaviour models. With the general purpose of understanding the complexity of urban growth, we hereby attempt to classify them as cellular automata modelling, multi-agent modelling, neural network modelling, fractal modelling etc., according to the methods available for modelling complexity and non-linearity.

2.3.1 CA-based modelling

Cellular automata (CA) are dynamic discrete space and time systems. A classic cellular automaton system consists of a regular grid of cells, each of which can be in one of a finite number of \( k \) possible states, updated synchronously in discrete time steps according to a local identical interaction rule.

The idea of CA is closely associated with that of microscopic simulation in which the behaviour at a local scale gives rise to an emerging global organisation (Webster and Wu, 2001). Global structure in a CA system is often seen to emerge out of purely local interactions between cells. This is attractive because it matches our intuitive sense that much human spatial activity is not centrally planned or organised, but arises from the responses of various actors, residents, developers, planners, politicians and local circumstances (O'Sullivan, 2001). It also holds out some promise of deeper insight into the enduring mystery of the relationship between processes at the micro level and the macro level of geographical and economic activity.

As an effective bottom-up simulation tool, CA first offer a new way of thinking for dynamic process modelling, and second provide a laboratory for testing the decision-making processes in complex spatial systems. By mimicking the manner in which macro-scale urban structures may emerge from the myriad interactions of simple elements, CA offer a framework for the exploration of complex adaptive systems (Torrens and O'Sullivan, 2001). CA represent a modelling approach quite different from top-down and macroscopic approaches (Webster and Wu, 2001).

CA have many advantages for modelling urban phenomena, including their decentralised approach, the link they provide to the complexity theory, the connection of form with function and pattern with process, the relative ease with which model results can be visualised, their flexibility, their dynamic approach, and also their affinities with geographical information systems and remotely sensed data (Torrens and O'Sullivan, 2001). Perhaps the most significant of their qualities, however, is their relative simplicity.

The many applications of CA can be classified into three types: complexity and GIS theory, theoretically artificial urban studies, and empirical case studies. Research has shown the great potential of CA for discovering the complexity (in particular spatial complexity) of urban system or its subsystems.

The first type links CA with complexity and GIS theory, e.g. CA theory (Batty and Xie, 1994; Childress et al., 1996; Couclelis, 1997; Itami, 1994; Wolfram, 1984), map dynamics (Takeyama and Couclelis, 1997), CA calibration (Li and Yeh, 2001; Wu, 2002), graph-
based CA (O'Sullivan, 2001), Voronoi-based CA (Shi and Pang, 2000), event-based CA (Gronewold and Sonnenschein, 1998) and fuzzy CA (Wu, 1998d). In complexity, many contributions come from other areas such as informatics, biology, physics and ecology. They use abstract models for exploring such general properties of complex systems as emergence, self-organising criticality and chaos. As regards the spatial complexity of the urban systems, as Torrens and O'Sullivan (2001) argue, CA models have been used to explore the self-organising properties of urban systems and experiments with fractal geometry and feedback mechanisms. However, there remains room for connecting that work with studies in other disciplines. Indeed, many aspects of complexity studies remain relatively unexplored by urban CA. In GIS, they attempt to develop more advanced spatial analytical functions based on CA modelling or they try to expand CA from raster data structure to another format. This direction still shows an increasing trend.

The second type links CA to theoretical urban studies, e.g. urban development patterns (Batty, 1998), self-organising competitive location theory (Benati, 1997), polycentric structure (Wu, 1998a), emergent urban form (Xie and Batty, 1997), land use dynamics through their life cycles (Batty et al., 1999b), real estate investment simulation (Wu, 1999), and urban socio-spatial segregation (Portugali et al., 1997). In these studies, transition rules are linked with urban theories to test theoretical hypotheses by using simulated or real data. Published literature has shown that this is a very promising direction, although little explored, which may bring new means for developing and interpreting new urban theories. One of the manifold potential uses of CA in urban theoretical research is the generation of novel city-like phenomena from theoretically informed components (Torrens and O'Sullivan, 2001).

In the third class, CA works as a spatial decision support system for simulation, prediction and planning based on real case study areas. This is a category of practice-oriented research where data availability and quality largely affect the application of CA on various scales (regional, metropolitan and town). Examples include urban land use dynamics (White and Engelen, 1993, 2000), the prediction of future urbanisation patterns (the San Francisco Bay and Washington/Baltimore corridor) (Clarke and Gaydos, 1998) (Gold Coast in Australia) (Ward et al., 2000a), Spanish cities (Silva and Clarke, 2002), long-term simulation of sprawl in the Ann Arbor Region (Batty et al., 1999a), land development process simulation (Guangzhou) (Wu and Webster, 1998), identification of diffused city (central area of Veneto region) (Besussi et al., 1998), urban form planning (a city in Guangdong, China) (Yeh and Li, 2001a), regional-scale urbanisation (Li and Yeh, 2000), urban development density (Yeh and Li, 2002), urban development plan (Chen et al., 2002), landscape dynamic (Soares-Filho et al., 2002), urban expansion based on population density surface (Wu and Martin, 2002), and suburban expansion of a peripheral municipality (town of Amherst, in metropolitan Buffalo, NY) (Batty and Xie, 1994).

In these applications, classic CA have been modified to incorporate urban theories and the understanding of specific practical issues of the study area. These applications span various spatial and temporal scales. They have adequately shown that CA offers a flexible and advanced spatial modelling environment that has not been available before.
However, of the complexity of urban growth, first they principally touch on spatial and decision-making complexity, with little about temporal complexity. The former includes pattern-oriented growth simulation, such as shown by Clarke and Gaydos (1998). The latter aims to aid the decision-making process of land conversion in urban growth (Wu, 1998c) or to simulate the fuzzy behaviour of decision-making in agricultural land encroachment (Wu, 1998d). Second, these applications focus on the simulation of spatial patterns rather than on the interpretation or understanding of the spatio-temporal processes of urban growth. CA models are constrained by their simplicity, and their ability to represent real-world phenomena is often diluted by their abstract characteristics (Torrens and O'Sullivan, 2001). As a consequence, there are many tasks waiting for further exploration of urban growth complexity based on CA.

### 2.3.2 Agent-based modelling

Multi-agent (MA) systems are designed as a collection of interacting autonomous agents, each having their own capacities and goals but related to a common environment. This interaction can involve communication, i.e. the passing of information from one agent and environment to another.

An agent-based model is one in which the basic unit of activity is the agent. Usually, agents explicitly represent actors in the situation being modelled, often at the individual level. Agents are autonomous in that they are capable of effective independent action, and their activity is directed towards the achievement of defined tasks or goals. They share an environment through agent communication and interaction, and they make decisions that tie behaviour to the environment.

From the perspective of modelling, multi-agents also have attractive features (White and Engelen, 2000): (1) as a tool to implement self-organising theory such as a straightforward way of representing spatial entities or actors having relatively complex properties or behaviours; (2) an easy way to capture directly the interactive properties of many natural and human systems, as well as the complex system behaviour that emerges from this interaction. Agent-based simulation is ideally suited to exploring the implications of non-linearity in system behaviour and also lends itself to models that are readily scalable in scope and level. The approach is useful for examining the relationship between micro-level behaviour and macro outcomes. Multi-agent models can locate agents and other resources of the environment in space and thus include the effects of space on the behaviour of the agents and the effects of the agents on the environment (Frank, 2000).

It is important to realise that agents are not necessarily either spatially located or aware. In many models, spatial mobility is not considered at all, although sometimes the term "space" appears as a metaphor for "social distance". The implications of the outcomes of such models for actual, physical spatial outcomes are not generally considered, because in most agent-based models the researchers' main concern is understanding how individual behaviour leads to global outcomes in a generic sense, rather than in the modelling of the real world per se (Haklay et al., 2001).
Agent-based models of this kind have only recently made their appearance in the social sciences (Batty, 2002), largely due to advances in computation and data that enable individual objects or events to be simulated explicitly, and to date most applications have been to theoretical situations (Batty, 2002; Epstein and Axtell, 1996). For the urban system MA are excellent tools for representing mobile entities in urban environments, e.g. people, households, vehicles etc. They have been used in urban contexts to simulate pedestrian movement in dense urban environments (Kerridge et al., 2001) and relocate householders (Benenson, 1998).

Benenson (1998) reported a multi-agent simulation model of the population dynamics in a city, in which inhabitants can change their residential behaviour depending on the properties of their neighbourhood, neighbours and the whole city. The agent in this model is characterised by its economic status and cultural identity and these two properties differ in nature. This model is based on an artificial city, which is used to test some urban theories such as social segregation. The most substantial application of agent-based models in the socio-economic domain is the monumental TranSims. This is a hybrid, lying somewhere between more traditional transport gravitation-interaction models and a full-blown real-time agent-based simulation. It currently models the activities of up to 200,000 individual travellers, which is where the model departs from previous transport planning models (Haklay et al., 2001).

Consequently, current applications of MA mainly focus on abstracted theoretical research or micro-behaviour simulation. There is no report that MA has been applied solely for understanding urban growth on a certain scale. However, it can be inferred that MA are an ideal tool for understanding decision-making complexity of urban growth at a micro scale, such as a single large-scale project.

### 2.3.3 Spatial statistics modelling

Traditional statistical models, e.g. Markov chain analysis, multiple regression analysis, principal component analysis, factor analysis and logistic regression, have been very successful in interpreting socio-economic activities. Markov chain (Lopez et al., 2001), multiple regression (Theobald and Hobbs, 1998) and logistic regression (Wu and Yeh, 1997; Wu, 2000b) have been widely used for modelling urban growth with varied strengths and weaknesses.

Lopez et al. (2001) report a model for predicting land cover and land use change in the urban fringe, a case study in Morelia city, Mexico. They conclude that the most powerful use of the Markov transition matrices seems to be at the descriptive rather than the predictive level. Linear regression between urban and population growth offered a more robust prediction of urban growth in Morelia.

Wu and Yeh (1997) apply logistic regression for modelling land development patterns in Guangzhou between 1978 and 1992, based on a series of aerial photographs. They found that the major determinants of land development have changed: from distance from the city
centre to closeness to the city centre; from proximity to inter-city highways to proximity to city streets; and from more related to less related to the physical condition of the sites etc. This demonstrates that various factors are changing their roles in the process of land development. This research has shown that logistic regression has a stronger capacity for interpreting urban development based on the probability of land conservation.

However, traditional statistics are criticised as being ineffective in modelling spatial and temporal data. The major reason is that spatial and temporal data often violate basic assumptions such as the normal distribution, appropriate error structure of the variables, independence of variables, and model linearity (Olden and Jackson, 2001). Two alternatives are frequently adopted. One is incorporating spatial sampling into traditional analysis (Atkinson and Massari, 1998; Dhakal et al., 2000; Gobin et al., 2001). The other is developing new statistics based on spatial relationships such as spatial dependence and spatial heterogeneity. New methods for analysing spatial (and space-time) data include spatial data analysis (Griffith and Layne, 1999; Haining, 1990), spatial econometrics (Anselin, 1988), local spatial analysis (Ord and Getis, 1995) and geographically weighted regression (GWR) (Fotheringham et al., 2000).

2.3.4 ANN-based modelling

An artificial neural network (ANN) is a system composed of many simple processing elements operating in parallel, whose function is determined by network structure, connection strengths, and the processing performed at computing elements or nodes. The development of a neural network model requires the specification of a "network topology", a learning paradigm and a learning algorithm.

Unlike the more commonly used analytical methods, the ANN is not dependent on particular functional relationships, makes no assumptions regarding the distributional properties of the data, and requires no a priori understanding of variable relationships. This independence makes the ANN a potentially powerful modelling tool for exploring non-linear complex problems (Olden and Jackson, 2001). According to published literature on its various applications, its strength lies in prediction and performing "what-if" types of experiment (Corne et al., 1999).

In the geographical sciences, recent ANN applications include spatial interpolation (Rigol et al., 2001), transport planning (Shmueli, 1998), transport and land use interaction (Rodrigue, 1997), land cover classification (Foody, 2002), image classification (Skidmore et al., 1997), urban change detection (Liu and Lathrop, 2002), and land cover transformation (Pijanowskia et al., 2002).

Shmueli (1998) used an ANN model to test whether or not there is a connection between socio-economic and demographic variables and travel activities. Skidmore et al. (1997) found that the neural network did not accurately classify GIS and remotely sensed data at the forest type level. Kropp (1998) applied a self-organising map (SOM) ANN model to classify 171 cities into four dimensions that represent all relevant features of the system and assess their sensitivity to change. As a form of non-linear dimension reduction, SOM
successfully provided an effective tool to identify cities that are susceptible to perturbations of human-nature interactions. Rodrigue (1997) provided an overview of a parallel transportation/land use modelling environment and concluded that parallel distributed processing offers a new methodology to represent the relational structure between elements of a transportation/land use system and thus helps to model these systems. He also considered that sequential urban modelling does not represent complex urban dynamics well, and he proposed a parallel network (back-propagation algorithm) model to simulate the spatial process and spatial pattern of integrated transport/land use system.

In urban growth, Pijanowskia et al. (2002) integrated ANN and GIS to forecast land use change, where GIS is used to develop the spatial predictor variables. Four phases were followed in their research: (1) design of the network and of inputs from historical data; (2) network training using a subset of inputs; (3) testing the neural network using the full data set of the inputs; and (4) using the information from the neural network to forecast changes.

These applications show that ANN is an ideal method of understanding non-linear spatial patterns, on which short-term prediction may be based. However, the major drawbacks of ANN, including its black-box and static nature, make it of limited value for modelling the urban growth process.

2.3.5 Fractal-based modelling

Benoit Mandelbrot, who coined the term in 1975, defines a fractal as "a set for which the Hausdorff-Besicovitch dimension strictly exceeds the topological dimension (Mandelbrot, 1967; Mandelbrot, 1982). Fractals were originally used for natural objects such as coastlines, plants and clouds or ill-defined mathematical and computer graphics. These are essentially spatial objects whose forms are irregular, scale-independent and self-similar. Recently, however, increasing analytical geographical analysis and analytical urban modelling has shown that planned and designed spatial objects such as urban forms and transportation networks can also be treated as fractals (Batty and Longley, 1994; Frankhauser, 2000; Shen, 1997; Shen, 2002a; Vicsek, 1991). It is considered that fractal dimension is one of the few concepts that are directly relevant to the problem of urban complexity (White and Engelen, 1993; Yeh and Li, 2001a).

The complexity represented by fractals is measured by a fractal dimension of a real number rather than an integer at various spatial dimensions. The fractal dimension may provide a less ambiguous approach to analysing the spatial structure and phenomena than current complexity measures. It is thought that a comparison between conventional density measures and the fractal dimension index gives more insight into the usefulness of fractal dimensions for modelling urban form, growth and development.

Cities are similar in a variety of ways, central place theory being the clearest demonstration of this principle (Batty and Longley, 1994). Fractal models give us a very different perspective on studies of urban density. This book explains how the structure of cities evolves in ways which at first sight may appear irregular, but when understood in terms of fractals reveal a complex and diverse underlying order. Recent studies have shown that the
complex spatial phenomena associated with an actual urban systems are better described using fractal geometry consistent with growth dynamics in disordered media (Makse et al., 1998). Makse et al. (1998) proposed and tested a model that describes the morphology of cities, the scaling of the urban perimeter of individual cities, and the area distribution of city systems. The resulting growth morphology can be understood from the interactions among the constituent units forming the urban region, and can be modelled using a correlated percolation model in the presence of a gradient. Shen (1997) applied a box-counting fractal dimension to calculate the fractal dimension of 30 urban transportation networks and then further tested the relationship between the fractal dimension and the urban population. It is thought that a comparison between conventional density measures and the fractal dimension index would give more insight into the usefulness of fractal dimension in modelling urban form, growth and development. Road network density is closely tied to many other parameters of urban development, such as population, urban growth, land use etc. The fractal dimension of a transportation network may also be used as an indicator of the complexity of the network.

Diffusion limited aggregation (DLA), a physical model used to describe aggregation phenomena, has been applied to describe urban growth (Batty and Longley, 1994). The growth of an urban area simulated through DLA can generate a fractal structure similar to that of real cities. But Makse et al. (1998) criticise the DLA model for generating only one large central place or cluster, whereas a real urban area is formed by a system of central places spatially distributed in a hierarchy of cities. They also propose a correlated percolation model which could predict the global properties (such as scaling behaviour) of urban morphologies. The model is better able to reproduce the observed morphology of cities and the area distribution of sub-clusters and can also describe urban growth dynamics. But this model studied the impact of urban policy on growth only from the perspective of interactions among dependent units of development.

A considerable number of studies report that fractal analysis can be used to measure the similarity between real and simulated spatial patterns created by cellular automata (Yeh and Li, 2001a). But it should also be noted that fractal measures of spatial complexity are difficult to interpret due to the fact that the same value of the fractal dimension may represent different forms or structures. It is also limited in urban process modelling as the temporal dimension is not incorporated in modelling.

### 2.3.6 Chaotic and catastrophe modelling

Catastrophe theory (Clarke and Wilson, 1983) and the theories of bifurcating dissipative structures (Allen and Sanglier, 1981b) attempt to model urban changes. But they have been pitched at the traditionally macro level and thus it has been hard to develop coherent explanations of the kind of changes emerging from the smallest scales which subsequently restructure the macro form of the system (Batty, 1998).

Chaos theory effectively means that unpredictable long-term behaviour arises in deterministic dynamic systems because of their sensitivity to initial conditions. For a dynamic system to be chaotic it must have a "large" set of initial conditions that are highly
unstable. No matter how precisely you measure the initial conditions in these systems, your prediction of its subsequent motion goes radically wrong after a short time. The key to long-term unpredictability is a property known as sensitivity to initial conditions. A chaotic dynamic system indicates that minor changes can cause huge fluctuations. As a result, it is only possible to predict the short-term behaviour of a system, especially for socio-economic systems such as cities. Although, chaos theory is able to explain the complex temporal behaviour of urban growth from a theoretical research viewpoint, the temporal scale of data available from urban growth is too limited to uncover its long-term behaviour.

Self-organised criticality (SOC) is a universal phenomenon occurring across a broad range of disciplines. It is thus a powerful interdisciplinary approach for understanding system complexity in a more general framework. Batty (1998) applied the concept of SOC to explain the temporal urban development pattern by using the cellular automata technique. He suggested that real cities in their evolution over time display this characteristic, which has not yet been tested in his research. Wu (1999) modified a simple sand-pile model from SOC theory to explain the urban development process resulting from real estate investment through cellular automata simulation.

Sprott et al. (2002) tested the phenomena of SOC in the field of landscape ecology, based on a simple cellular automata model. They found that spatial distributions and temporal fluctuations in global quantities show power-law spectra, implying scale-invariance, the characteristic of self-organised criticality when a system evolves into a self-organised system.

2.4 Evaluation of Modelling

2.4.1 Review of urban modelling history

Planning is a future-oriented activity, strongly conditioned by the past and present. Planners have always sought tools to enhance their analytical, problem-solving and decision-making capabilities. Consequently, urban modelling should be able to assist planners in looking to the future. It should facilitate scenario building and provide an important aid to future-directed decision-making.

Urban modelling bloomed in the late 1950s and throughout the 1960s in both the USA and Western European countries, e.g. the Lowry model was designed in 1964 and first introduced into the process of urban planning by using aggregated data. However, with the massive transformation from an industrial to an informational economy, urban modelling gradually faded away as a dominant planning and decision-making paradigm in the late 1970s and through most of the 1980s (Sui, 1998). Modelling techniques from the 1960s to the 1980s were dominated by a-spatial, static, linear, cross-sectional, deterministic approaches, such as regression analysis, mathematical programming, input-output analysis and even system dynamics. They proved inadequate to reflect the complex, dynamic and non-linear factors inherent in urban systems or subsystems (Lloyd-Jones and Erickson, 1997; Sui, 1998), and were of limited value in supporting planning decision-making.
Consequently, the new challenge requires that the focus of modern urban modelling be shifted from macro to micro, from aggregate to disaggregate, from static to dynamic, from linear to non-linear, from top-down to bottom-up, from structure to process, from space to space-time, due to the unpredictability, instability, uncomputability, irreducibility and emergence that exists in the process of urban evolution. Famous examples including TransSim and UrbanSim (Waddel, 2002) indicate the current trends of urban modelling. The time and space dimensions need to be incorporated into the urban modelling process by further integrating with GIS and complexity and non-linearity theories.

2.4.2 Criteria of evaluation

A major distinction among methods can be drawn on the basis of their purpose and the objective of their study. Their purpose can be descriptive, explanatory, predictive, prescriptive. The major criteria for evaluating the operation of various methods are, in terms of data requirement, their linkage with GIS, and their interpretability.

(1) Data requirements

Questions of urban growth have attracted interest among a wide variety of researchers concerned with modelling the spatial and temporal patterns of land conversion and understanding the causes and consequences of these changes. Aided by new spatial data capture technologies such as very high-resolution remote sensing satellites and global positioning systems (GPS), relatively accurate and comprehensive digital data sets of metropolitan areas collected and maintained by public agencies are now becoming widely available (Longley, 1998). Remote sensing potentially provides a strong data-source framework within which to monitor change and understand urban growth, e.g. frequently used Landsat TM, SPOT, IRS and even IKONOS imagery. Nevertheless, it is well known that classified urban land cover does not bear a spectrally identifiable correspondence with urban land use as urban land use is defined by a social purpose and not a set of physical quantities. Remote sensing data are useful for providing outline descriptions of urban form but are less helpful in understanding the functional characteristics of urban growth.

Spatially and temporally explicit models at fine levels of spatial and temporal resolution – the individual parcel level – are increasingly being developed as the required computational and technological infrastructure improves continuously and as data at this level become available. However, in the developing world, poor data infrastructure has been a major barrier in implementing some advanced methods of modelling. Socio-economic attributes based on various levels of spatial statistical units (see chapter 4) and parcel-based land ownership are still not available or accessible to the modelling community. Our inability to monitor land cover changes in a consistent way in the long term also seriously limits our capacity to understand the driving forces and processes controlling these changes (Petit and Lambin, 2001).

As illustrated in figure 2.1, understanding urban growth involves pattern, process and behaviour. However, current data infrastructure only offers pattern and partial process with spatial data at limited spatial and temporal scales. Consequently, urban growth modelling
remains dominated by macro spatial models (pattern and process); and the spatial behaviours linked with micro-scale functional data and temporal complexity based on higher temporal resolution data are still in the state of theoretical research. This situation is even worse in the developing world. Fractal, CA, ANN and logistic regression studies have widely utilised remote sensing imagery as inputs to their modelling.

(2) Linkage with GIS

GIS first came to fame in the early 1980s as a technique for geo-referenced data input, data storage, data processing, data retrieval, and data output, with simple data models and a few spatial analysis functions. The first GIS provided only limited decision-making capacity, due to insufficient spatial modelling functions. The inability to incorporate urban models and to more directly support policy-making processes are two main deficiencies of the current geo-spatial technologies and tools (Nedovic’-Budic’, 2000). The integration of both did not take place until the late 1980s. GIS can provide the urban modeller with new platforms for data management, spatial analysis and visualisation. Loose, close and tight coupling strategies are frequently adopted. At present, ANN and CA have been integrated into GIS such as the ArcView extension (spatial modeller: ANN, fuzzy logic and logistic regression) and IDRISI (CA). Open source software development is becoming popular, such as UrbanSim, which has a free environment for users to develop or modify their own models. Such progress has opened up more opportunities for the applications of these advanced methods of modelling.

(3) Interpretability

Urban growth modelling aims to understand complex dynamic and non-linear processes, and therefore the capacity of interpretation is crucial. Compared with logistic regression, the Markov chain model lacks explanatory power as the causal relationships underlying the transition studies are left unexplored. The transition probabilities are estimated as proportions of cells that have changed state from one point in time to another. This approach remains a useful way of estimating these probabilities despite the development of procedures for estimating transition probabilities on the basis of more complex scientific consideration. ANNs have a greater predictive and non-linear power than traditional approaches. However, their property of "black box" provides little explanatory insight into the relative influence of the independent variables in the prediction process. This lack of explanatory power is a major concern in spatial pattern analysis because the interpretation of statistical models is desirable for gaining knowledge of the causal factors driving spatial phenomena. Traditional statistical approaches can readily identify the influence of the independent variables in the modelling process and also provide some degree of confidence regarding their contribution. Olden and Jackson (2001) concluded that where the underlying data structure and assumptions are met for a particular traditional statistical technique, there is no reason to believe that major differences will exist between traditional approaches and ANNs. However, ANNs were shown to be superior to regression approaches for non-linearly distributed data.
Cellular automata (CA) and multi-agent (MA) approaches overlap to some degree; CA is sometimes considered to be a type of multi-agent system (White and Engelen, 2000). Comparatively, CA focuses on the city level (Wu and Webster, 1998) and the regional level (White and Engelen, 2000). In contrast, MA is only applied on the household (Bernard, 1999) and family (Benenson, 1998). The MA approach deals with decisions posed to people more frequently (Benenson, 1998). CA models focus on landscapes and transitions, agent-based models focus on human actions. CA are most suitable in urban simulation contexts for representing infrastructure. MA are better used to model population dynamics.

MA differ from CA in their spatial mobility: agents can be designed to navigate (virtual) spaces with movement patterns that mimic those of humans, while CA are only capable of exchanging data spatially with their neighbourhoods. Additionally, agents can be given functionality that allows them to evolve over time, altering their attributes and behaviour with the help of artificial intelligence. Comparatively, MA are based more on abstract cellular space as micro data are difficult to access. However, MA applications to urban studies have not been as widespread as those of CA, despite offering the advantages for urban simulation.

2.5 Conclusions

From the literature and the evaluation above, it can be seen that some methods are still in the theoretical stage or applied for artificial city analysis, and need very good data infrastructure. Some methods are more effective on a macro scale than on a micro scale. Each method has its strengths and weaknesses, and respective data requirements and application domains. The selection of methods should depend on the demands of the analysis, the feasibility of the techniques and the availability or limitation of the data framework.

First, as discussed above, urban growth involves three different systems $P$, $N$, $U$. To model their dynamic interactions at varied spatial and temporal scales, current methods of modelling are not adequate to understand all the complexity inherent in urban growth described in the previous sections. Hence, only a limited number of complex phenomena can be modelled.

Second, physical data are becoming more readily available, particularly on the macro-scale now, due to the low price of satellite imagery in recent years. On this macro-scale, socio-economic data are much easier to access as aggregated data are based on annual statistics. This results in the fact that urban growth modelling focuses mainly on spatial complexity understanding such as CA-based dynamic simulation, ANN-based pattern analysis and fractal-based morphology analysis. The difficulty in accessing micro-scale socio-economic data and higher-resolution (spatial and temporal) data limits the understanding of temporal and decision-making complexity in urban growth. Chaos theory and the MA model have not been widely applied for planning practice. The theoretical experiment based on artificial cities is also a feasible modelling means (Batty, 1998; Bura et al., 1996; Wu, 1998a). The poor interpretation capacity of most models (such as CA, fractal and ANN) means that they
are less used for practical applications than traditional or spatial statistics such as logistic regression and geographically weighted regression (GWR).

The conceptual model of the strategy adopted in this research, as illustrated in figure 2.3. Here, the complexity that can be modelled depends on four factors: the demand from urban development planning and growth management, the data that is available from multiple sources, the concepts from other relevant disciplines, and the theories and methods from complexity science. These concepts are based on the theories of complexity and need models to test theories. Methods need data for implementation. The advanced theories and methods discussed above have great potential for understanding urban growth complexity. As the result of the dynamic interactions, urban growth modelling involves numerous variables from three systems $P$, $N$, $U$. This is a basic principle for the models in the later chapters. Their interpretation needs to be linked to the experiences of other disciplines such as agriculture, landscape, ecology and environmental science. Consequently, a multidisciplinary framework is advocated to incorporate the concepts for developing new methodologies for understanding urban growth. In this research, four types of complexity regarding urban growth will be modelled. These are complexity in structure and function (chapter 3), complexity in temporal measure (chapter 4), complexity in pattern (chapter 5) and complexity in process (chapter 6). The specific concepts, methods and data of each model will be elaborated in these four chapters. The major methods for modelling include fractal for structural complexity, landscape metric for functional complexity, data disaggregation and spatial auto-correlation for temporal measures, exploratory data analysis and spatial statistics (logistic regression) for pattern complexity, and cellular automata for process complexity.

Figure 2.3 A conceptual model for the strategy adopted in this research