

# An Analysis of Land Use Change in the urban fringe using GIS and Logistic regression in Korea

Jung Hoon Kim,

Associate research fellow, Korea Research Institute for Human Settlements

요약: 이 연구는 수도권 외곽지역을 대상으로 도시주변지역관리방안을 마련하기 위해 지리정보시스템(Geographic Information System: GIS)과 공간분석기법을 이용하여 분석적이고 기술적인 적용 가능성을 찾아보고자 한다. 이 연구의 목적은 도시주변지역에 토지이용변화에 영향을 미치는 주요한 토지이용변화 결정변수를 찾아보자는데 있다. 특히 인접필지의 토지이용변수에 중점을 두고 분석을 한다. 하나의 필지가 인접필지의 토지이용현황에 의해 어느정도 영향을 받는지 분석하는 것이다. 그리고 분석결과를 GIS 툴을 이용해서 도면화해보고, 토지이용변화가 일어날 가능성이 높은 지역을 찾아보고자 하는데 있다. 이 연구는 GIS와 공간분석기법을 이용하여 지방자치단체의 업무중 80%이상을 차지하는 지적자료를 기초로 하여 어떻게 도시주변지역의 토지이용변화를 분석할 수 있는가를 보여주고 있다. 분석 결과는 GIS와 공간분석기법을 이용한 향후 응용방향의 한 일단을 제시하고 있다. 이 분석으로 지방자치단체가 도시주변지역의 토지이용을 효과적으로 관리하고 그 지역의 상황을 이해하는데 매우 유익하다. 특히 이 연구에서 분석된 방법론은 실질적이고 경험적인 분석으로 매우 유용하게 다른 지역에도 적용 가능한 것으로 보인다.

Keywords: land use change, urban fringe, GIS, Logistic regression

## 1. Introduction

Large cities in Korea have experienced rapid urbanisation since the 1960s. Population and industry have been concentrated within large cities. As a result, the inner and outer areas of the cities have undergone serious change in land use due to the increase of the demand of land development. In particular, it has a significant impact on the land use of the urban fringe. Rapid land use change occurring in Seoul Metropolitan Regions (SMR) is one of the serious issues at present in Korea. Particularly in the Quasi-Agricultural and Forestry Zone (QAFZ) the change has brought out very serious disorderly development. The phenomenon in the urban fringe in Korea, these days, is the outcome of deregulation since 1994.

The deregulation has contributed to the national economy by increasing the supply of developable land, consequently stabilising the price of land, and encouraging the production activities of enterprises (Sim, 1997). However, there has been bitter controversy over the policy since the deregulation was introduced because it has produced many kinds of negative side effects. Many agricultural and forestry areas in the SMR have changed, for various reasons, to motels, restaurants, and factories without harmony with neighbouring environment.

Land use change is a complex phenomenon differing greatly from place to place. Causes of change include personal choice, legislation, government policies and plans, decisions of developers or transportation entrepreneurs, the nature of the land itself, or the availability of technology to develop the land (Hill, 1989). In addition, it is very difficult to trace the process of land use change because of the complexity of the situation compounded by a lack of comprehension of the range of the information needed. (Ganderton, 1994)

To better understand the rapid land use change occurring in urban fringe, a number of studies have sought to quantify the extent and rate of agriculture and forestry loss. It has lacked a way of displaying the results of analysis because of insufficient development of computer hardware and GIS software. However, planners and policy makers need to know more than the extent and rate of agriculture and forestry loss. They want to know not only the rate and extent of land use change for a given region but also the location of changed land, the pattern changed, and the transition probability of land use.

This research is concerned with analytical and technical capability in using GIS and spatial analysis as relevant to the policy problem of urban fringe management. The purposes of the research are as follows. Firstly, the study will focus on examining the relationship between the nature of the land itself and land use change in order to discover the determinants of the change in the study area. It is imperative that local governments understand the determinants of land use so that future patterns of land use can be projected, planned, and managed under sustainable conditions. Not surprisingly, however, those variables were already investigated by existing studies such as highway proximity, distance to urban centres, elevation, slope, and soil suitability, etc. (Pierce 1981, Firman 1997). Among the site characteristics the variable in which the thesis will be most interested is neighbouring land use. Therefore, secondly, this research will focus more on the impact of neighbouring land use.

Thirdly, this research will develop an analysis of the transition probability of land use, i.e. which areas are most susceptible to land use change. This will be also useful to local governments when formulating a plan to protect the lands from further unplanned development. Finally, the results of analysis will be visualised by using GIS so that local governments can view where a specific change takes place, how the patterns of land use have changed, and what conditions lead to disorderly development. The representation will be useful for local government to assess the impact of policy change.

## 2. Literature Review

There are a number of theories on urban growth and land use conversion. This article will review theories providing important insights related to the purposes of the research such as land use change theories, California Urban Futures (CUF) Model and Cellular Automata (CA) model.

### 2.1 Theories explaining land use change

There are three traditional theories on how to view land use patterns and dynamics: economic theory; sociological theory; and the public interest-oriented approach. Economic theory views land use as a result of economic competition for limited space based on equilibrium theory (Hurd, 1903; Ratcliff, 1949; Wingo, 1961; Alonso, 1964; Muth, 1969). Sociological theory views land use in the context of urban ecology and social structure (Chapin, 1965). The public interest-oriented approach views land use in the context of health, safety, and general welfare (Kaiser & Weiss, 1970).

These traditional land use theories are connected with theories explaining land use changes. Land use change theories generally focus on the conversion of land from non-urban to urban uses. The land use conversion as usual occurs in the urban fringe, along with urban expansion (Fisher, 1954; Bourne, 1969). These land use change theories have been based on three different approaches: factors affecting land use change, leading actor's behaviour, and process oriented approaches (Chang, 1986).

The approach related to factors affecting land use change is concerned with the causal relationships between land use change regarding it as urban growth, and land use change caused by exogenous factors such as proximity between land uses, accessibility to work areas, compatibility between land uses, availability of public facilities, and proximity to ethnic group areas (Chapin, Donnelly, & Weiss, 1962; Chapin and Weiss, 1962a). The studies related to a leading actor's behaviour suggest that the conversion of land use is determined by the decisions of leading actors such as the predevelopment landowner, developer, land speculator, facilitator, local government, and household consumer (Kaiser, 1966; 1968a; 1968b; Clawson, 1971; Rolph, 1973; Baerwald, 1981). The process oriented approach stresses that the conversion process of land use is a sequence of stages brought about by a series of key decisions which include those of both key decision agents and supporting agents in each stage, and which are influenced by decision factors and local public policies (Kaiser & Weiss, 1970).

A number of land use theories mentioned above have contributed to a better understanding of land use and land use change. However, although these theories explain urban expansion and evolving land use patterns, they can hardly serve as tools to simulate, forecast or predict future urban developments. White argues that Land use theories are static models and that they assume rational actors interacting in a market which remains in a stable equilibrium state (White and Engelen, 1993).

## 2.2 The California Urban Futures (CUF) Model

The California Urban Futures model really shows what this study is interesting in such as the application of GIS and spatial analysis and neighbouring land use effects. The Model is a metropolitan simulation model to provide a framework on how realistic growth and development policies at various levels of government might alter the location, pattern, and intensity of urban development in the fourteen-county Northern California Bay Region (Landis, 1994).

The CUF Model was developed as a planning model "to help planners, elected officials, and citizen groups create and compare alternative land use policies" (Landis, 1995; Kramer, 1996). It is a model which forecasts urbanisation at the regional scale (Landis, 1994). The analytical unit of the CUF Model is Developable Land Units (DLUs). DLUs are generated as the spatial union of multiple GIS layers. DLUs are made up of one-hectare (100m x 100m) grid cells (Landis and Zhang, 1998a).

The CUF Model includes spatial database between 1985 and 1995 as follows: i) local population and employment growth; ii) proximity to regional job centres; iii) site slope; iv) whether the site is within or beyond city boundaries or spheres of influence; v) the uses of surrounding sites; vi) the availability of vacant land; vii) site proximity to freeway interchanges and transit stations; viii) site proximity to major commercial, industrial, and public land uses (Landis and Zhang, 1998a).

The Model is one of the first urban models to utilise GIS technology (Wegener, 1994). The Model requires an extensive GIS database because it is developed to make choices based on the knowledge that planners have (Kramer, 1996). In addition, Landis (1995) explains that the CUF model can be applied at the subregional or county level for local growth planning. Landis and Zhang (1998a, 1998b) extended the database by new data categories such as housing prices, employment growth forecasts distributed to individual DLUs, and major new infrastructure investments. However, the model requires a constant and very knowledgeable interaction to be of use for planners (Kramer, 1996).

### 2.3 Cellular Automata (CA) Theory

The cellular automaton approach is to model urban areas and land use temporally and spatially. The Cellular Automata theory is developed to visualise rural to urban transition as a physical process. A cellular automaton is an array of cells, whose states depend on the states of the neighbouring cells (White and Engelen, 1993; Kramer, 1996). Cells of a cellular automaton respond in a mechanical way to surrounding cells. Due to these interactions, cellular automata are considered as discrete dynamic systems with simple construction but complex self-organising behaviour (Wolfram, 1984; Kramer, 1996; Clark et al., 1997). Each point in a regular spatial lattice, called a cell, can have any one of a finite number of states. The states of the cells in the lattice are updated according to a local rule. That is, the state of a cell at a given time depends only on its own state one time step previously, and the states of its nearby neighbours at the previous time step. All cells on the lattice are updated synchronously. Thus the state of the entire lattice advances in discrete time steps (White and Engelen, 1993).

The dynamic urban modelling approach developed by Batty (1992), White (1993), and Clarke et al (1996) has its main roots in the cellular automata theory (Kramer, 1996). They developed new computational dynamic cellular models of urban growth. Their approach has been described as a new school of urban modelling (Clarke et al, 1997). The model on a Caribbean island developed by White is of special interest

because it integrates a cellular automaton with a GIS. In the integration of a cellular automaton model and a GIS, the cellular model can take advantage of the detailed GIS database and at the same time undertake spatial analysis, which is not part of GIS. White sees the ultimate goal in building cellular modelling tools into a GIS in order to create a GIS with dynamic capabilities (White and Engelen, 1993). With this vision White comes close to his objective to develop a theory that offers solutions for the local and specific problems of planners and other practitioners (White and Engelen, 1993; Kramer, 1996).

### 3. Research methods

Urban models in association with urban change processes mentioned above are quite limited in explaining spatial relationships and processes, and their representation. The recent advances in GIS technology have made it possible to represent, collect, store, manage, analyse, and model spatial data, although they can not be used to undertake a wide range of spatial analysis. Most recently, there has been a trend for experts to come together to make the linkage between GIS and methods for the statistical analysis of spatial data as a key issue in research agendas. The linkage of GIS and spatial analysis may help decision making activities for the solving of spatial problems. Thus, this study is studied as a part of this trend and employs two main methods: Logistic Regression and Geographical Information Systems.

Logistic Regression is a form of regression which is used when the dependent variable is a dichotomy and the independents are continuous variables, categorical variables, or both. Logistic regression is used to examine the main determinants of land use change. In this study, logistic regression is easier to use in the SPSS programme than discriminant analysis when it has a mixture of continuous and categorical variables, because it includes procedures for generating the necessary dummy variables automatically (Hinde, 2001). The database will be divided into dependent and independent variables in order to examine spatial patterns of land use changes, the impact of neighbouring land use, and the transition probability of land use.

The dependent variable is the changed or unchanged use of land in a parcel. The independent variables are area, land category, zoning, height, direction, forest cover, neighbouring land use, accessibility to main roads, land price etc. Many items related to land characteristics are excluded as the study area deals with the urban fringe. That is, some urban factor items are unnecessary such as urban planning district and urban planning facilities. Neighbouring land use as an independent variable are added to evaluate the extent to which it is associated with land use change as a dependent variable. Extracted attribute data was entered in Excel spreadsheets and then saved as DBF files.

GIS analysis is used to verify spatially the close statistical relationship between the dependent variable and the independent variables selected by the logistic regression analysis. This study applies vector-based GIS because complete digital parcel maps

are available at the study area. The vector model is favoured for greater manipulation of single parcels. The results of the analysis are visualised using GIS (ARC/INFO and ArcView) to improve the understanding of the land use change in the Seoul Metropolitan Region Perimeter. The GIS also supports storage, analysis, querying, browsing and display of various geographical data available for the study area.

#### 4. Study area and database

Case study in this research is employed to identify the main factors affecting land use change in the urban fringe areas, to analyse particularly the impact of neighbouring land use on the propensity for a land use to change, to view the result of analysis, and to develop an analysis of the transition probability of land use. However, for these analyses, this study could not assemble all the data for the urban fringe around the Seoul Metropolitan Region. Instead of this, it is necessary to examine in detail a small part of the Seoul urban fringe, where land use change pressures are known to be great. Therefore, a single case rather than a multiple case was selected. The capacity of data handled is huge even in a single case and the same case study deals with many independent variables. However, this is not a statistically representative sample of all fringe areas, but was selected to ensure that three main reasons are taken into account: ease of data collection in a digital format, the frequency of land use change, and the adjacency of the Seoul Metropolitan Region. The single case study in this research may offer an insight into the influence of the main variables involved in the land use change process.

Therefore, this study was applied to the urban fringe area of northern Yongin City, Kyonggi Province, as shown in Figure 1, in order to investigate a case study at the parcel level from 1991, 1994 to 1998. The area has an area of about 10 Km<sup>2</sup> and a total of 3,995 parcels. About 64.5% of the total area (2,576 parcels) are Quasi-Agricultural and Forest Areas (QAFA) and 12.7% (508 parcels) are Quasi-Urban Areas as classified in the National Land Use and Management Act. Therefore, the study area will be able to explore in detail the pattern of land use change in the QAFA due to deregulation in 1994.

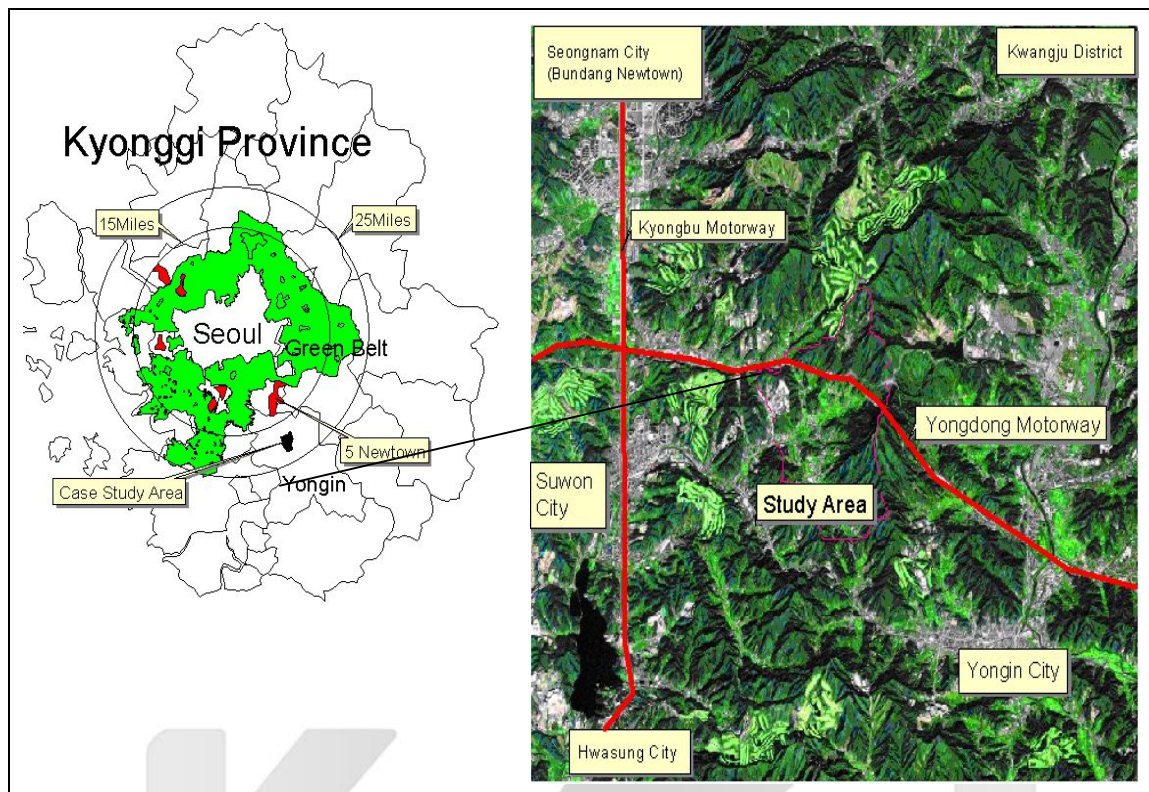


Figure 1 Location of study area

This study requires a digital cadastral map as a basic map. A parcel in a cadastral map is the smallest unit for spatial analysis in this study. The parcel is the ideal unit of analysis for studying land use change because it can have a wide range of information such as parcel numbers, the classification of land category, and the boundaries and ownerships of land parcels. The importance of the parcel approach at the local level is that most administrative affairs in the local authority are associated with cadastral services such as land address information, land characters, land price, the limitation of use, and the resource management of parcels.

The parcel approach can use an administrative code which consists of province code, city code, ward code, block number, parcel number, etc. Local governments maintain many documents for their own local needs such as a land register, an agricultural land register, a forest register, a building register, etc. All of them recorded by each department of local authorities in Korea are based on the same administrative network code. This attribute data can thus be linked easily to the cadastral map by including additional attributes in each record.

There are, however, some technical restrictions on the wide use of the parcel approach. For example, it should deal with very large parcel data at the regional and national level and take much time and expense to analyse, manage, and edit the related database. In addition, the various changes to the cadastral map occur when a new subdivision is created, annexation occurs between parcels, or land development and redevelopment in older areas occurs. In this case, the changes should be updated

by each department of local government but it is not easy to update them promptly. Thus, when analysing large databases, the error of the database may be high.

Therefore the CA approach and CUF model based on the grid cell may be more useful at the regional or national level than the parcel approach. The parcel approach can be used efficiently at the local level to monitor and manage specific local issues. The parcel approach combined with GIS can not only represent the particular local site characteristics that affect land development, but also present a realistic method on very local and specific problems that the local authority must deal with on a daily basis. Both methods will be chosen according to needs and scale.

As attribute data, this study employed the Land Utilisation Character Surveys data (LUCS data) utilised annually by local authority planning departments. They contain locational data as well as place attributes such as perimeter, area of a parcel, urban planning zone, urban planning facilities, land price, state of land use, land category, land height, the position towards which land faces, the condition of adjacent road to land, the straight line from road to boundary of land, the distance from rail to land, etc.

Moreover, this study carried out field investigation to supplement missing data in some parcels and to observe the condition of buildings and work spaces and neighbourhood of parcels changed. This study also referred to documents related with the purpose of the study such as a land register, an agricultural land register, a forest register, etc.

The most important contribution of GIS technology is the logical link between map features and attribute data. By incorporating together map and attribute data, GIS provide a powerful environment within which to measure spatial identities, explore spatial relationships, and ultimately, to model spatial processes. Efforts of link have been employed successfully in other studies (Ludeke et. al. 1990; Pereira and Itami, 1991; Chou et.al., 1993; 1996; Narumalani et. al. 1997; Landis and Zhang, 1998a:1998b).

The method of link in this study is as follows; all data mentioned in previous sections contains information about the same features. That is, both the cadastral map as a base map and the LUCS data as attribute data have the same administrative network code, called *keyid*. These attribute data can thus be linked to the cadastral map by including additional attributes in each record. Once that association is made, many various analyses are possible.

## 5. Empirical Analysis

### 5.1 Extracting variable of neighbouring land use

Among the site characteristics the variable in which the study will be most interested is neighbouring land use. Therefore, this research will focus more on the impact of neighbouring land use. Studies of neighbouring use effects related to a site level have been paid little attention. This is because at site level, it is particularly

difficult to measure and collect data. Common encoding in a form convenient for use with statistical models is particularly difficult. Thus the neighbouring land use variable has been neglected in analysing determinants of land use change. Existing studies on neighbouring use effects are to show the degree of relationship that exists between two or more spatial variables, such that when one changes, the others also change. To do this, the spatial autocorrelation is used. This change can either be in the same direction, which is a positive autocorrelation, or in the opposite direction, which is a negative autocorrelation. Technically speaking, this is to measure the similarity of neighbouring parcels (GIS dictionary).

For measuring spatial autocorrelation, there are many methods suitable for polygons. These methods can be used with two types of data: nominal and interval/ratio data. For nominal data, joint count statistics can be used to measure spatial autocorrelation for polygons. Moran's I Index, the Geary Ratio, Local Indicators for Spatial Association, and the G-statistics can be used for interval/ratio data (Lee and Wong, 2001). These techniques in spatial analysis depend on 'a matrix of proximity measures between zones where each proximity measure represents the degree of influence of one zone upon another. Clearly all these ways of measuring spatial proximity have some validity and there appears to be no one commonly accepted measure' (Fotheringham et.al., 2000: p244).

Apart from the above methods, this study used a new method to capture the neighbouring relationship among parcels. To define the neighbouring relationship, all the surrounding parcels in the study area can be identified as neighbours of the current parcel as long as they touch each other even at a point. To examine neighbouring parcels on all polygons of the study area, this study used ArcView's avenue programming which visited each polygon in turn and produced an outfile containing the parcel record ID and the contents of each of the adjacent parcels. The second thing is to translate the record IDs of neighbouring parcels into land use. For this, this study used the Lookup function in Excel to obtain neighbouring land use. To deal with how to calculate the ratio of each land use to neighbouring land uses, this study used Visual Basic. Finally, this study considered a term of neighbour land use as the ratio of the number of urban use adjacent parcels to the total number of adjacent parcels (Kim, 2002).

## 5.2 The Logistic Regression Model

The specific form of the logistic regression model that this study will use is as follows:

$$P_i = \frac{e^z}{1 + e^z}$$

or equivalently

$$P_i = \frac{1}{1 + e^{-z}}$$

where  $P_i$  is the probability for land use change to occur in the  $i$  th parcel;

e is the base of the natural logarithms, approximately 2.718;

z is the linear combination;

$$Z = B_0 + B_1X_1 + B_2X_2 + \dots + B_pX_p$$

where  $X_i$  is the independent variable associated with each parcel;

$B_0$  is a constant;

$B_i$  is the estimated coefficients associated with each parcel;

The equation notices that when Z is -, the logistic function  $P_i$  equals 0 and when Z is +, the  $P_i$  equals 1. Thus, the range of  $P_i$  is between 0 and 1, regardless of the value of Z. The fact that the logistic function  $P_i$  ranges between 0 and 1 is the primary reason the logistic model is so popular. The model is designed to describe a probability, which is always some number between 0 and 1 (Kleinbaum, 1992). In this study, such a probability gives the potentiality of land use change in the urban fringe.

In this study, a two-scenario modelling approach was adopted to separate the neighbouring land use variable from other independent variables in order to examine the role of neighbouring land uses in the pattern of land use change. In scenario I, this study examined each variable's significance level except the neighbouring land use variable. The model is called a basic model in this study. Only the variables that are statistically significant were employed. In scenario II, such significant variables were incorporated with the neighbouring land use variable. This study called it the modified model (Chou et. al., 1993).

As scenario I, Figure 2 shows coefficients and related statistics from the logistic regression model that predicts land use change from the variables and a constant. In the logistic regression, the criterion for a variable to enter the model is each variable's significance level. Other variables in this analysis such as perimeter, land category 94, land category 98, state of land use 94, state of land use 98, zone 94, land price 98 and height were insignificant. Such variables were excluded in a scenario I. The determinants of land use change in this model thus was reduced to six variables, that is, area of parcel (Area), zone in 1998 (Zone98), land price in 1994 (LndPrc94), forestry (Forst), direction (Drec), and accessibility to main road (Road).

Figure 2 Variables in the Equation based on basic model

	B	S.E.	Wald	df	Sig.	R	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
Step 1									
AREA	.000	.000	7.079	1	.008	.0455	1.000	1.000	1.000
ZONE98	.085	.011	64.906	1	.000	.1600	1.089	1.066	1.112
FORST	-2.163	.435	24.730	1	.000	.0782	.115	.049	.270
DREC	.263	.038	48.948	1	.000	.0962	1.301	1.208	1.400
ROAD	-.131	.022	36.317	1	.000	.1383	.877	.841	.916
LNDPRC94	.000	.000	17.014	1	.000	.1182	1.000	1.000	1.000
Constant	-2.671	.159	282.042	1	.000		.069		

a Variable(s) entered on step 1: AREA, ZONE98, FORST, DREC, ROAD, LNDPRC94.

Given these coefficients, the logistic regression equation for the probability of land use change in the basic model can be written as:

$$P(\text{land use change probability}) = \frac{1}{1 + e^{-z}}$$

$$Z = B_0 + B_1\text{AREA}_i + B_2\text{ZONE98}_i + B_3\text{LNDPRC94}_i + B_4\text{FORST}_i + B_5\text{DREC}_i + B_6\text{ROAD}_i + e$$

where  $\text{AREA}_i$  is the area of the  $i$  th parcel;

$\text{ZONE98}_i$  is one out of the zoning classification of  $i$  th parcel in 1998;

$\text{LNDPRC94}_i$  is the land price of  $i$  th parcel in 1994;

$\text{FORST}_i$  is one out of the forest classification of  $i$  th parcel;

$\text{DREC}_i$  is one out of the direction classification of  $i$  th parcel;

$\text{ROAD}_i$  represents the distance between  $i$  th parcel and the nearest road;

$e$  is a random error term;

To incorporate the neighbouring land use variable in the land use change model, this study defined a term NBRLU as the ratio of the number of urban use adjacent parcels to the total number of adjacent parcels. The estimated coefficients and statistics of the modified model are listed in Figure 3. Compared to that of the basic model in Figure 2, the modified model improved considerably due to the inclusion of the neighbouring land use variable as shown in Figure 3. This means that the neighbouring land use variable can play an important role in determining the pattern of land use change.

Figure 3 Variables in the Equation based on modified model

	B	S.E.	Wald	df	Sig.	R	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
Step 1	AREA	.000	.000	6.277	1	.012	.0417	1.000	1.000
	ZONE98	.082	.011	58.861	1	.000	.1522	1.085	1.063
	FORST	-2.153	.434	24.650	1	.000	.0729	.116	.050
	DREC	.260	.038	47.608	1	.000	.0960	1.297	1.205
	ROAD	-.120	.022	29.926	1	.000	.1363	.887	.850
	LNDPRC94	.000	.000	15.067	1	.000	.1066	1.000	1.000
	NBRLU	.768	.238	10.463	1	.001	.0587	2.156	1.354
	Constant	-2.992	.190	248.330	1	.000		.050	

a Variable(s) entered on step 1: AREA, ZONE98, FORST, DREC, ROAD, LNDPRC94, NBRLU.

Given these coefficients, the logistic regression equation for the probability of land use change in the modified model can be written as:

$$P(\text{land use change probability}) = \frac{1}{1 + e^{-z}}$$

$$Z = B_0 + B_1 \text{AREA}_i + B_2 \text{ZONE98}_i + B_3 \text{LNDPRC94}_i + B_4 \text{FORST}_i + B_5 \text{DREC}_i + B_6 \text{ROAD}_i + B_7 \text{NBRLU}_i + e$$

where  $\text{NBRLU}_i$  represents the ratio of urban neighbouring land use;

The coefficient  $B$  must be interpreted with care. As shown in Figure 2 and Figure 3, two variables (Forest and Road) are negative. This may mean that the probability of land use change will occur with a small change when each variable changes by one unit. This can be affected by the labels of the two variables. In the case of the Forest variable, it is divided into preserved forest land, quasi-preserved forest land and the rest is input as 0. Most of the land used in the analysis were changed in parcels input as 0. Thus the sign of the coefficient may show negative. Even though the Forest variable changes one unit, it may be harder to predict if the event will occur or not. In the case of the Road variable, the label is divided into five general categories according to the nearest distance from each parcel to a road such as: (a) adjacent areas; (b) within 50m; (c) within 100m; (d) within 500m; (e) over 500m. Showing a negative sign for the Road variable may thus be a correct result when taking into account a positive relationship between the accessibility and the change of land use. However, this explanation is not very intuitive. An interpretation of the coefficient which is usually more intuitive is the odds ratio.  $\text{Exp}(B)$  shown in the column in Figure 2 and Figure 3 is the effect of the independent variable on the odds ratio. In the following section, this will be explained in detail.

As shown in the column labelled Wald in Figure 2 and Figure 3, the Wald statistic is commonly used to test the null hypothesis that a coefficient is 0. When a variable has a single degree of freedom, the Wald statistic is just the square of the ratio of the coefficient to its standard error. For example, the coefficient for Zone98 is 0.082, its standard error is 0.011 (under column heading S.E. in Figure 3). The Wald statistic is  $(0.082/0.011)^2$ , or about 58.861. Unfortunately, there is a flaw in the Wald statistic so that very large coefficients may lead to large standard errors and small Wald chi-square values. Whenever one has a large coefficient, one should not rely on the Wald statistic for hypothesis testing. Instead, one should build a model with and without that variable and base its hypothesis test on the change in the log-likelihood (Hosmer and Lemeshow, 1989).

The contribution of individual variables in logistic regression is difficult to determine because the contribution of each variable depends on the other variables in the model. There is a problem, in particular, when independent variables are highly correlated. To estimate the partial correlation between the dependent variable and each of the independent variables, the R statistic is used, shown in Figure 2 and Figure 3. The R statistic can range in value from  $-1$  to  $+1$ . If R is a positive value, the variable increases in value and the likelihood of the event occurring increases as well. If R is negative, the opposite is true. That is, the value of R indicates that the variable has a partial contribution to the model (Kleinbaum, 1992).

In this study, among the values of R, Zone98 variable has the largest partial contribution to the model, shown in Figure 2 and Figure 3. This means that in the study area, zoning has a much higher influence on land use change in the urban fringe than other independent variables. This may be associated with the deregulation in 1994 because rapid land use change in the Quasi-Agricultural and Forest Area occurred after the deregulation.

As mentioned in the above, the logistic model can be rewritten in terms of the odds of an event occurring. The odds of an event occurring are defined as the ratio of the probability (p) that an event occurs to the probability (1-p) that it does not:  $\text{odds} = p/(1-p)$ .

‘Then  $e$  raised to the power  $B_i$  is the factor by which the odds change when the  $i$ th independent variable increases by one unit. If  $B_i$  is positive, the factor will be greater than 1, which means that the odds are increased; if  $B_i$  is negative, the factor will be less than 1, which means that the odds are decreased. When  $B_i$  is 0, the factor equals 1, which leaves the odds unchanged’ (SPSS, 1997: p43). For example, when NBRLU changes one unit, the odds are increased by a factor of 2.156, as is shown in the  $Exp(B)$  column in Figure 3. This implies that the more urban land use parcels there are at the neighbouring location, the higher the transition probability of land use in the study area. In this way the odds ratio is a measure of the strength and direction of relationship between two variables, here between land use change and urban neighbouring land use.

Also, goodness-of-fit tests are available as indicators of success as is the Wald statistic and other tests of the model's significance. Figure 4 and Figure 5 show the goodness-of-fit statistics for the model with all of the independent variables. For the current model, the value of  $-2LL$  is 2253.931, which is smaller than the  $-2LL$  for the model containing only a constant (2455.968), a reduction of 202.036 on 6 degrees of freedom in Figure 4. ‘The *Cox & Snell  $R^2$*  and the *Nagelkerke  $R^2$*  are statistics that attempt to quantify the proportion of explained “variation” in the logistic regression model’ (SPSS, 1997: p47).

There are three additional chi-square entries which are labelled *Model*, *Block*, and *Step* in Figure 4 and Figure 5. ‘The model chi-square is the difference between  $-2LL$

$$\text{Odds} = \frac{e^{B_0 + B_1X_1 + B_2X_2 + \dots + B_pX_p}}{1 - \frac{1}{1 + e^{-z}}} = e^{B_0 + B_1X_1 + B_2X_2 + \dots + B_pX_p}$$

for the model with only a constant and  $-2LL$  for the current model. Thus, the model chi-square tests the null hypothesis that the coefficients for all of the terms in the current model, except the constant, are 0' (SPSS, 1997: p48). In this example,  $-2LL$  for the model containing only the constant is 2455.968, while for the complete model, it is 2253.931. The model chi-square, 202.036, is the difference between these two values (SPSS, 1997).

For seeing whether the independent variables as a whole significantly affect the dependent variable, this is addressed by the Model Chi-square statistic. As shown in Figure 4 and Figure 5, two models in this analysis were analysed to be statistically significant. The comparison between Model Chi-square in the basic mode (Figure 4) and that of the modified model (Figure 5) indicates that the modified model included the neighbouring land use variable increased from 202.036 to 212.444. This means the modified model provides much higher explanation and prediction on land use change, compared to that of the basic model.

The probability of land use change in each parcel will always be between 0 and 1, regardless of the value of  $Z$ . In general, if the estimated probability of the event is less than 0.5, the event will not occur. If the probability is greater than 0.5, the event will occur (Hosmer and Lemeshow, 1989). In order to map land use change probability, it is necessary to construct a probability model of land use change based on variables that are significant (Chou et. al., 1993). Therefore, this study examined the transition probability of land use by the estimated coefficients from the logistic regression model. In this modified model, the independent variables that are statistically significant include the area of a parcel (AREA), Zoning in 1998 (ZONE98), land price in 1994 (LNDPRC94), urban neighbouring land use (NBRLU), direction (DREC), Forest lands (FORST), and accessibility to a main road (ROAD). The identification of such independent variables will be valuable in guiding the implementation of land use management.

Figure 4 The basic model of land use change			Figure 5 The modified model of land use change				
Iteration History <sup>a,b,c</sup>			Iteration History <sup>a,b,c</sup>				
		-2 Log likelihood	Coefficients Constant		-2 Log likelihood	Coefficients Constant	
Step	1	2628.095	-1.631	Step	1	2628.095	-1.631
0	2	2462.185	-2.154	0	2	2462.185	-2.154
	3	2455.984	-2.281		3	2455.984	-2.281
	4	2455.968	-2.288		4	2455.968	-2.288
a. Constant is included in the model. b. Initial -2 Log Likelihood: 2455.968 c. Estimation terminated at iteration number 4 because log-likelihood decreased by less than .010 percent				a. Constant is included in the model. b. Initial -2 Log Likelihood: 2455.968 c. Estimation terminated at iteration number 4 because log-likelihood decreased by less than .010 percent			

Omnibus Tests of Model Coefficients					Omnibus Tests of Model Coefficients				
Step		Chi-square	df	Sig.	Step		Chi-square	df	Sig.
Step 1	Step	202.036	6	.000	Step 1	Step	212.444	7	.000
	Block	202.036	6	.000		Block	212.444	7	.000
	Model	202.036	6	.000		Model	212.444	7	.000

Model Summary				Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square	Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2253.931	.049	.107	1	2243.523	.052	.113

## 6. Results of analysis

The results of analysis were again converted to GIS format files from SPSS files through several steps. The findings thus can be displayed in a GIS environment. These results have three important implications for land use management in the urban fringe.

First, Figure 6 indicate spatial locations where land uses have changed from rural use to urban use from three time periods. When comparing the locations of the original land use in 1991 and the same locations in 1994 parcel by parcel using GIS, Figure 6 shows that there was change in only a few parcels. It also represents quite a substantial increase in building sites and factory sites in the agricultural areas and forest areas between 1994 and 1998.

It can be inferred from the results of analysis that the government's land policy change directly affects land use change in the study area. Before deregulation in 1994, most parcels in the study area were still in the same land use. After deregulation in 1994, the significant changes on the study area have occurred. Many parcels were changed from rural use to urban use. Agricultural areas particularly decreased dramatically and most of the changes involved conversion to building sites and factory sites in 1998.

The problems with this pattern are that the developments are too small in scale and too scattered to provide appropriate local roads. Unplanned and uncontrolled land development like this may lead to many negative side effects in the study area. Unless an appropriate development strategy is formulated to prevent these kinds of problems, the study area is likely to continue the current development patterns. This requires local governments to consider how their own land use regulation and land development policies affect their spatial locations.

The second result shows the distribution of parcels which neighbored to urban land use as displayed in Figure 7. An interesting feature of this pattern is that the distribution map was adjacent to many parcels changed in the land use between 1994 and 1998. This distribution map shows a quite similar pattern with the pattern of land use change.

This distribution map has indicated that any parcel with over 50% urban neighbouring land use as an agricultural area or forestry is likely to be taken over by a building site or a factory site sooner or later. That is, this map has indicated that the transition probability of the parcels adjacent to urban use is higher than that of the parcels adjacent to rural use.

This study has shed some light on the role of neighbouring land use as a determinant of land use change. Among all of the independent variables analysed in the previous chapter, neighbouring land use stands out as the leading element in the evolution of land use change in the urban fringe. Neighbouring land use might be one of the most important factors for land use management in the urban fringe.

The policy implication which could be drawn from this distribution map is that the distribution map of the neighbouring land use may be taken into consideration in spatial processes for explaining and predicting land use change in the urban fringe. Although an appropriate development strategy can help to prevent disorderly development in the urban fringe, land development is carried out according to various causes such as personal choice, legislation, government policies and plans, decisions of developers or transportation entrepreneurs, the nature of the land itself, or the availability of technology to develop the land. Thus the intervention of local government can be limited in this circumstance. Nonetheless, the local government is required still to formulate regulatory policy to encourage a particular type of land conversion or to give priority to a specific type of development (Wu, 1998). For this purpose, local government may be able to utilise the distribution map of neighbouring land use.

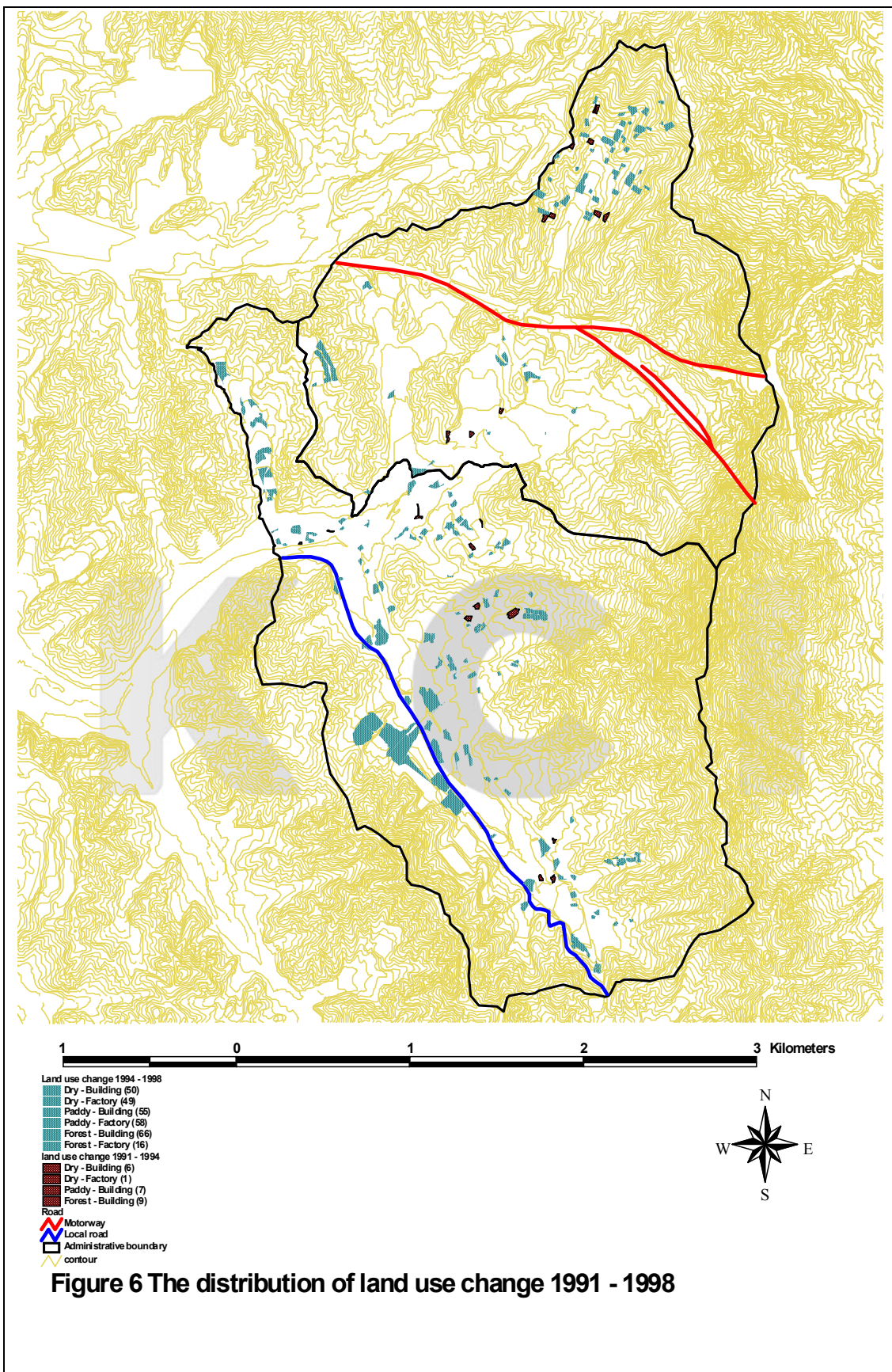
One of the objectives of this study focuses on predicting the future distribution of land use, based on specific locational characteristics. The final result thus shows that a map of the transition probability of land use can be displayed using GIS. Figure 8 pinpoints the spatial location where the transition probability of land use is distributed. The transition probability map was derived from the logistic regression model based on variables that are statistically significant.

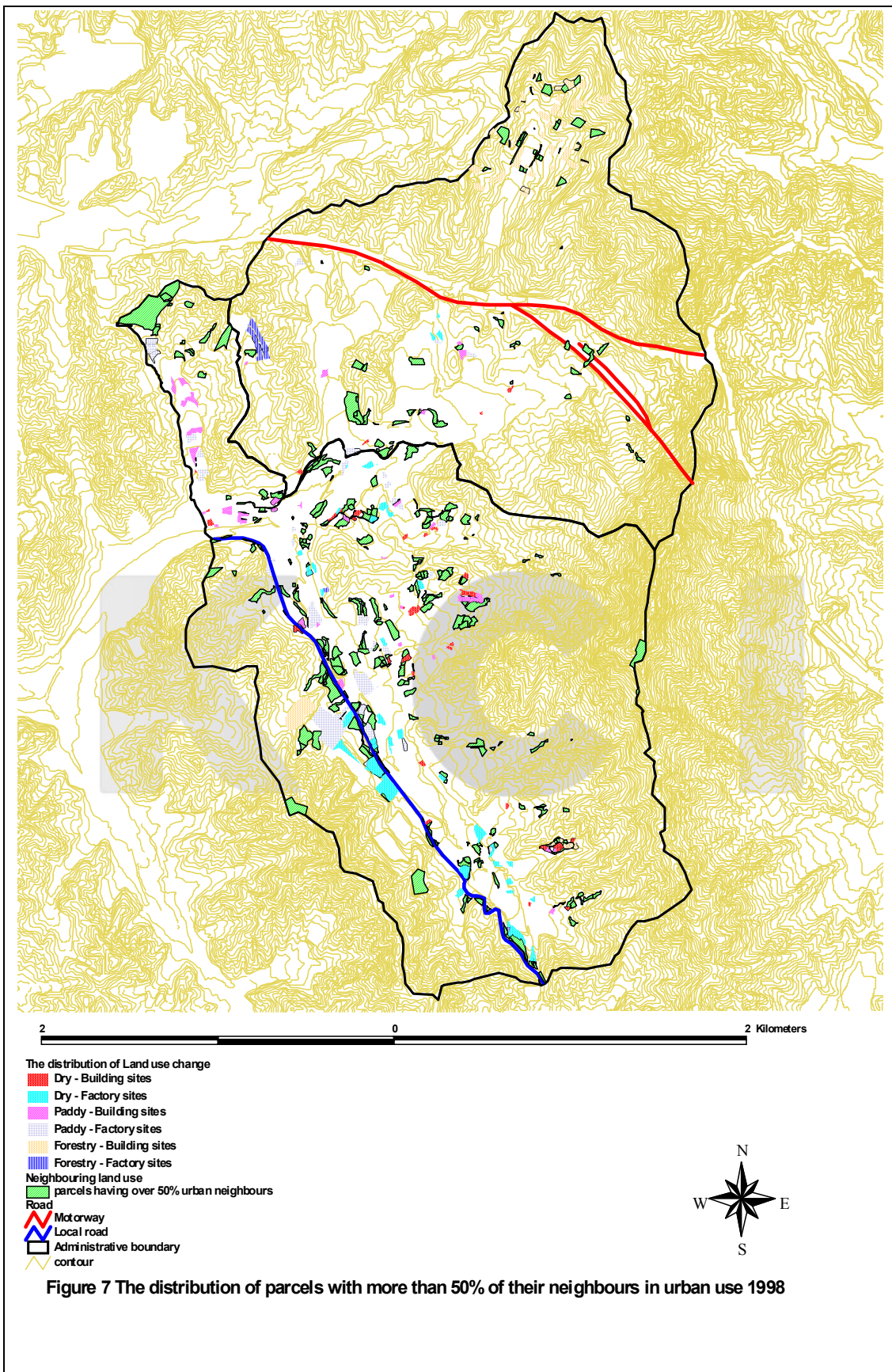
The map is classified into three levels of change potential: less than 10%, between 10% and 50%, and greater than 50%. Even though the values of the transition probability are not high, the distribution of the transition probability of land use is quite interesting. Most of the parcels with a high transition probability are distributed near main roads. That is, accessibility to main roads tends to appear as a very important fact regarding the transition probability of land use in the study area.

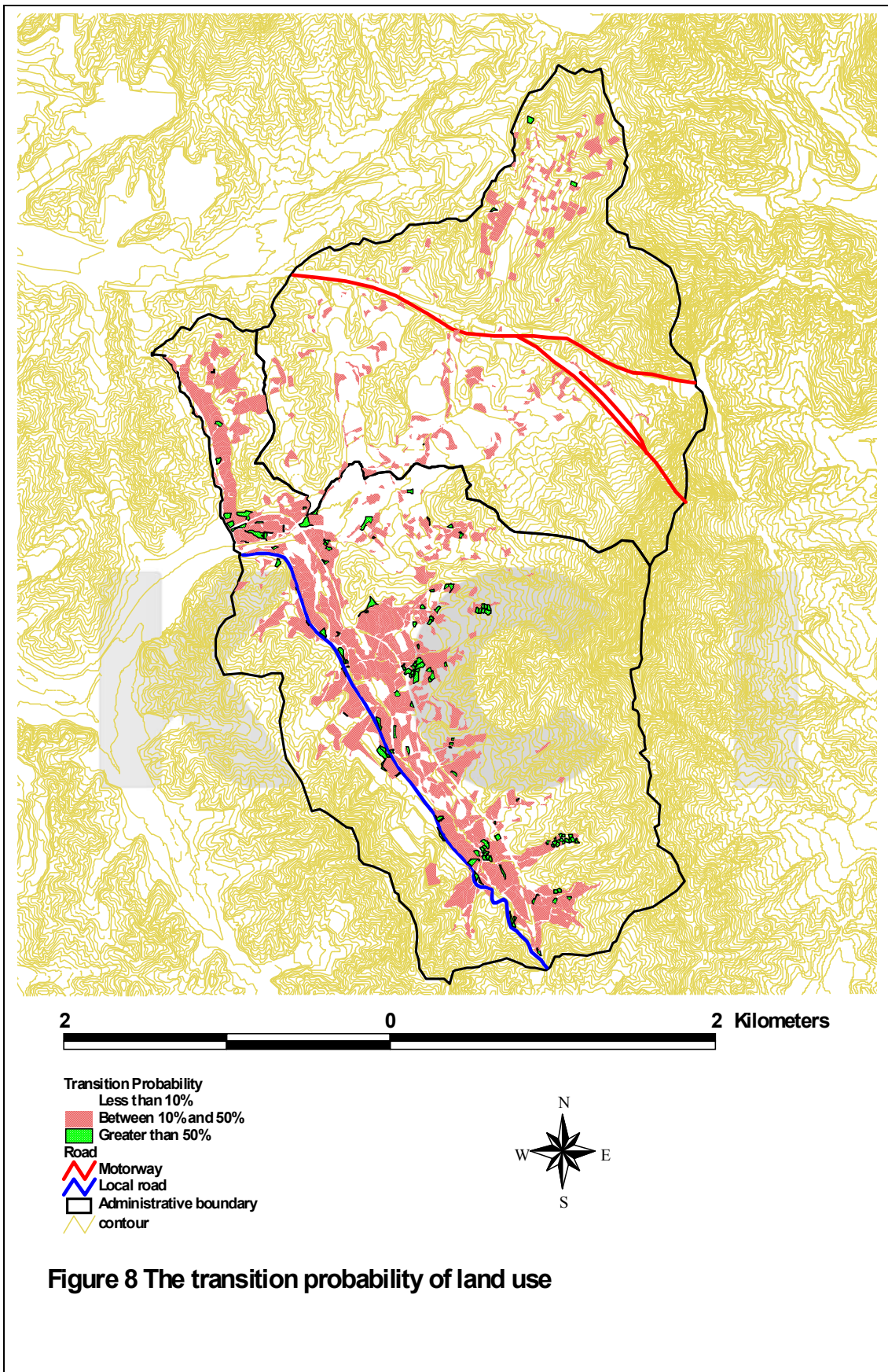
This map is useful for local government to understand target areas when evaluating and deciding on the management strategy of the urban fringe. This map using GIS can be overlaid with another map showing features of major concern. For example, if the protection of paddy fields has a high priority for food production, decision makers can identify protective areas according to overlaying the paddy field map onto the land use transition probability map.

This transition probability map may be able to give not only an important insight in the pattern of future development but also a guideline when evaluating and deciding on the management strategy of the urban fringe. Local government can know which areas are susceptible to land use change by utilising this transition probability map. This map thus can be used as a basic map when formulating an appropriate development strategy.

K C I







## 7. Conclusion

The research results represented important implications for land use management in the urban fringe. Firstly, the distribution of land use change has indicated quite a substantial increase in building sites and factory sites in the agricultural areas and forest areas since 1994 indicating that the government's land policy change directly affects land use change in the study area. Therefore, local governments are required to consider how their land policies affect their spatial locations. Secondly, this study found the determinants of land use change in the study area using a logistic regression model. The result shows that the neighbouring land use variable among the determinants plays an important role in the pattern of land use change. The neighbouring land use map may be taken into consideration in spatial processes for explaining and predicting land use change in the urban fringe. Thirdly, the transition probability map of land use was examined by the estimated coefficients from the logistic regression model. This map is useful for local government to understand target areas when evaluating and deciding on the management strategy of the urban fringe. Finally, the results analysed were visualised by using GIS such as the pattern of land use change from three time periods, the distribution of neighbouring land use and the transition probability of land use.

The study contributes to significant advances on the integration of GIS and spatial analysis. In addition, it is the first study analysing the impact of neighbouring land use at a site level by using vector GIS and spatial analysis. It is also the first analysis of land use change regarding every parcel in the study area using cadastral maps manipulated by local government. The results of this study will assist planners in determining appropriate land use for Korean urban growth areas. These should help local governments understand the effects of land use change in the urban fringe in Korea.

## References

- Alonso, W.**, (1964) *A Theory of the Urban Land Market*, *Papers of the Regional Science Association*, 6, Philadelphia: Wharton School of Finance and Commerce, University of Pennsylvania
- Baerwald, T.J.**, (1981) The site selection process of suburban residential builder, *Urban Geography*, 2 (4) pp339-357
- Batty, M. and Longley, P.**, (1994) *Fractal Cities*. London: Academic Press.
- Batty, M.**, (1992) Urban Modelling in Computer-Graphic and Geographic Information System Environments, *Environment and Planning B* 19: 689-708.
- Bourne, L.S.**, (1969) A spatial allocation-land use conversion model of urban growth, *Journal of Regional Science*, 9(2) pp 261-272
- Burrough, P.A. and McDonnell, R.A.**, (1998) *Principles of Geographical Information Systems*, Oxford University Press
- Chang, B.M.** (1986) *Land Use Transition Model: the case of Korea*, unpublished PhD thesis, Department of City and Regional Planning, University of North Carolina
- Chapin, S.F.Jr. and Weiss, S.F.**, (1962a) *Urban Growth Dynamics in a Regional Cluster of Cities*, New York and London: John Wiley and Sons
- Chapin, S.F.Jr.**, (1965) *Urban land use planning*, Urbana, Illinois: University of Illinois press

- Chapin, S.F.Jr., Donnelly, T.G., and Weiss, S.F.,** (1962) *Factors Influencing Land Development*, Institute for Research in Social Science, University of North Carolina, Chapel Hill
- Chou, Y.H.,** (1992) Management of wildfires with a geographical information system, *International Journal of Geographical Information System*, 6:123-140
- Chou, Y.H., and Minnich, R.A.,** (1993) Mapping Probability of Fire Occurrence in San Jacinto Mountains, California, USA, *Environmental Management* Vol. 17, No. 1, pp.129-140
- Chou, Y.H., and Soret, S.,** (1996) Neighbourhood Effects in Bird Distributions, Navarre, Spain, *Environmental Management* Vol. 20, No. 5, pp.675-687
- Clarke, K.C., Gaydos, L., Hoppen, S.,** (1997) "A self-modifying cellular automaton model of historical urbanisation in the San Francisco Bay area," *Environment and Planning B*, V24. Pp247-261
- Clarke, K.C., Hoppen, S., and Gaydos, L.J.,** (1996) Methods and Techniques for Rigorous Calibration of a Cellular Automaton Model of Urban Growth, *Third International Conference/Workshop Integrating GIS and Environmental Modeling*, Santa Fe, NM, January 21-25, Santa Barbara: National Centre for Geographic Information and Analysis. WWW and CD.
- Clawson, M.,** (1971) *Suburban Land Conversion in the United States: An Economic and Governmental Process*, Baltimore: John Hopkins University Press
- Firman, T.** (1997) Land Conversion and Urban Development in the Northern Region of West Java, Indonesia, *Urban Studies*, Vol. 34, No. 7, 1027-1046
- Fisher, E.M. and Fisher, R.M.,** (1954) *Urban Real Estate*, New York: Henry Holt and Co.
- Fotheringham, A.S., Brunsdon, C., and Charlton, M.,** (2000) *Quantitative Geography: Perspectives on Spatial Data Analysis* SAGE
- Ganderton, P.** (1994) Modelling the land conversion process: a realist perspective, *Environment and Planning A*, volume 26, pp 803-819
- GIS dictionary,** <http://www.agi.org.uk/public/gis-resources/index.htm>
- Hill, R. D.** (1989) *Land-use Change: Proceeding of the Asahikawa-Sapporo International Symposium*, Hong Kong University Press
- Hinde, J.,** (2001) John Hinde's teaching page, School of Mathematical Sciences, University of Exeter, <http://www.maths.ex.ac.uk/staff/JPHinde/psy6010/disclogi.html>
- Hosmer, D.W. and Lemeshow, S.,** (1989) *Applied logistic regression*, New York: Wiley
- Hurd, R. M.,** (1903) *Principles of City Land Value*, New York: The Record and Guide
- Kaiser, E.J.,** (1966) *Toward a Model of Residential Developer Locational Behavior*, Unpublished Ph. D. Dissertation, University of North Carolina
- Kaiser, E.J.,** (1968a) Locational decision factors in a producer model of residential development, *Land Economics*, 44
- Kaiser, E.J.,** (1968b) *A Producer Model for Residential Growth: Analysing and Predicting the Location of Residential Subdivision*, An Urban Studies Research Monograph, Chapel Hill: University of North Carolina
- Kaiser, E.J., and Weiss, S.F.** (1970) Public policy and the residential development process, *Journal of the American Institute of Planners*, 36 (1) pp30-37
- Kim, B.G., Kwon, K.S., and Jeong, D.H.** (2001) Items of cadastral computations in the PBLIS of KCSC, *New Technology for a New Century* international conference, FIG, pp70-90
- Kim, J.H. (2001)** *An Analysis of Land Use Change using GIS and Spatial Analysis: A Case Study of the Seoul Metropolitan Region Perimeter*, Ph. D., Thesis, University of Newcastle
- Kim, J.H. (2002)** A Study on the patterns of land use change based on the digital parcels, *The Journal of Geographic Information System Association of Korea*, Vol. 10, No.1
- Kleinbaum, D.G.,** (1992) *Logistic Regression: A self-Learning Text*, Springer
- Kramer, J.,** (1996) Integration of a GIS with a Local Scale Self-Modifying Cellular Automaton Urban Growth Model in Southeastern Orange County, New York, M.A. thesis, Hunter College of the City University of New York

**Landis, J. and Zhang, M.**, (1998a) The second generation of the California urban futures model. Part 1: Model logic and theory, *Environment and Planning B: Planning and Design* : Vol 30, pp 657-666

**Landis, J. and Zhang, M.**, (1998b) The second generation of the California urban futures model, Part 2: Specification and calibration results of the land-use change submodel, *Environment and Planning B: Planning and Design*, Vol 25, pp 795-824

**Landis, John.**, (1994) The California Urban Futures Model: A New Generation of Metropolitan Simulation Models, *Environment and Planning B: Planning and Design*. 21: 399-420.

**Landis, John.**, (1995) Imagining Land Use Futures: Applying the California Urban Futures Model, *Journal of the American Planning Association*. 61:1. 438-457.

**Lee, J and Wong, D.W.S.**, (2001) *Statistical Analysis with ArcView GIS*, John Wiley & sons

**Ludeke, A.K., Maggio, R.C., and Reid, L.M.**, (1990) An Analysis of Anthropogenic Deforestation Using Logistic Regression and GIS, *Journal of Environmental Management*, Vol. 31, pp. 247-259

**Muth, R. F.**, (1969) *Cities and Housing: The Spatial Pattern of Urban Residential Land Use*, Chicago and London: The University of Chicago Press

**Narumalani, S., Jensen, J.R., Althausen, J.D., Burkhalter, S. and Mackey, H.E.** (1997) Aquatic Macrophyte Modelling Using GIS and Logistic Multiple Regression, *Photogrammetric Engineering & Remote Sensing*, Vol. 63, No. 1, pp. 41-49

**Pereira, J.M.C. and Itami, R.M.**, (1991) GIS-Based Habitat Modelling Using Logistic Multiple Regression: A Study of the Mt. Graham Red Squirrel, *Photogrammetric Engineering & Remote Sensing*, Vol. 57, No. 11, pp. 1475-1486

**Pierce, J. T.**(1981) Conversion of rural land to urban: a Canadian Profile, *Professional Geographer*, 33, pp. 163-173

**Ratcliff, R. U.**, (1949) *Urban Land Economics*, New York: McGraw Hill

**Rolph, E.S.**, (1973) *Decision-making by Residential Developers in Santa Clara County*, Santa Monica, California, RAND Corporation for the National Science Foundation, R-1225

**Sim S. U.**, (1997) The problems of Development in Rural Regions Around Cities, *International Symposium on City Planning*, Nagoya, Japan, pp109-120

**Wegener, Michael.**, (1994) Operational Urban Models: State of the Art, *Journal of the American Planning Association* 60:1. 17-30.

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

**Wingo, L. Jr.**, (1961) *Transportation and Urban Land*, Washington, D.C.: Resources For the Future

**Wolfram S.**, (1984) Universality and complexity in cellular automata, *Physica D* 10 1-35

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