MODELLING
LAND USE LAND COVER CHANGES
USING CELLULAR AUTOMATA
IN A GEO-SPATIAL ENVIRONMENT

Anuj Kumar Singh
November, 2003

MODELLING LAND USE LAND COVER CHANGES USING CELLULAR AUTOMATA IN A GEO-SPATIAL ENVIRONMENT

by

Anuj Kumar Singh

Thesis submitted to the International Institute for Geo-information Science and Earth Observation in partial fulfilment of the requirements for the degree of Master of Science in Geoinformatics

Thesis Assessment Board

Chairman: Prof.Dr.A Stein
External examiner: Dr. ir. P.H. Verburg
Supervisor: Dr. T. Bouloucos
IIRS Member: Dr. P.S. Roy

Supervisors

At ITC
Dr, Theo Bouloucos
At IIRS
Mr C.Jeganathan

ITC

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

I certify that although I may have conferred with others in preparing for this assignment, and drawn upon a range of sources cited in this work, the content of this thesis report is my original work.

Signed

Disclaimer

This document describes work undertaken as part of a programme of study at the International Institute for Geo-information Science and Earth Observation. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Acknowledgements

To start with, I am very grateful to Dr P.S.Roy Dean IIRS who provided me this opportunity to do this course and encouraged me all time during my course. I am very grateful to the authorities of “GEONEDIS “ Project-a collaborative program of the ITC, the Netherlands and the Indian Institute of Remote Sensing (National Remote Sensing Agency), Department of Space, Government of India, to provide me this opportunity
.

I am very thankful to the Royal Government of The Netherlands also and I acknowledge their efforts for providing subsistence support for nice and pleasant stay at ITC for the last leg of the program of three months duration. I am extremely thankful to Professor B.S.Rajput then Vice Channcellor Kumaun University, Nainital, Uttranchal for sponsoring me for this course.

I am thankful to all the IIRS family (Staff, Administration, Warden, Library, CMA, and Mess) who made my life very smooth and full of joy during my stay at IIRS.

My special thanks to my guides, Dr Theo Boloucos, and Mr C.Jegannathan for their valuable guidance and encouragement. At this moment I cannot forget to remember Ms Y. Sun and Mr. Luc Boerboom for their discussions and guidance too.

At this moment I would like to remember all of my classmates also, for their cooperation and discussions during all the course.

Last, but not the least, I have deep sense of gratitude to my parents, brothers, sisters, brothers in law for their encouragement too.

.

Anuj Kumar Singh

Date: 27.11.03
Enschede,
The Netherlands

Abstract.

. Model is an abstraction or simplification of real world. From early days of our childhood we learn (subconsciously) from models and play with models. Playing with guns, cars and dolls, making houses with sands, are some of the examples, which we never realize in our daily life. Models are being used as a very powerful learning tool. They are not only static like road maps but they can be dynamic like playing with the car or plane with remote control or in video games Modeling used in the present thesis comes under the category of 'scenario models'.

.

Land use activity is a major issue and challenge for town and country planners (as well as environmentalists) to design the Eco friendly and sustainable economic growth. The human activity for development is forced against the environment, which results in consequences such as soil erosion, global warming, pollution etc. The causes for change in land use type activity may be due to socio-economic development or due to changes in the environment or may be due to both. For example, an increase in total export demand for a particular agriculture product will be translated into increase in demand for land for this particular product whereas increase in tourist influx will result in increased demand of land for constructions. Land can be gained by conversion of agricultural fields or by clearing the forest. In both cases the consequences will be different. Envisioning the consequent effects of land use changes, IGBP (International Geo-sphere and Biosphere Program) and IHDP (International Human Dimension Program) co-organized a working group to set up research agenda and promote research activity for land use and land cover changes (LUCC). The working group suggested three core subjects for LUCC research, such as: situation assessment, modeling and projecting, and conceptual scaling. The ultimate goal of global change study is to assess the impacts under each possible scenario and suggest preventive actions. The modeling and projecting of land use change is essential for scenario analysis and the assessment of LUCC. Consequently, issues related to data, information, and modeling have attracted many research interests ranging from local authorities to global organizations.

Adding the spatial component in the model was evolutionary concept in the Cartographic modeling in 1960s and 1970s. From the 1980s to current the focus has been on the spatial modeling. A cartographic model tends to be more 'static,' meaning it depicts spatial variation in quantitative data, but does not tell about the spatial influence on the variation. The cartographic data model uses points, lines and polygons (topologically encoded) with one, or only a few, attached attributes, such as a land use layer represented as polygons with associated land used code. Spatial models can be seen as the extension of the Cartographic models. It has all those geometric shapes mentioned in cartographic model with set of multiple associated attributes A spatial model depicts spatial processes, or the influence of spatial factors on spatial variation. Finally, spatial models can supply a more 'realistic' view of reality than a cartographic model - as spatial factors can have great influence on the variability of a statistic. A spatial model is the integration of spatial components into mathematical models. With spatial models, spatially dependent factors (e.g. distance or slope) can be incorporated with other statistical data (e.g. population or agricultural production) variables can help to refine model solution and yield better results.

Since the evolution of cellular automata, it is being used in many disciplines ranging from sciences to commercial fields. Because of its capabilities to address the complex patterns with the help of very simple transition rules it was accepted in every corner of the field of research.

In comparison with traditional approaches based on differential or difference equations the CA has advantages. CA can incorporate spatial component. And they address dynamism with simple rules, which increases computational efficiency. Since computational efficiency translates into better handling of dynamism CA becomes favorites to many modelers.

The advantages of CA are many. The construction of model is simple and easy. It has an ability to perform spatial dynamics, and time explicitly. After Analyzing the similarities and capabilities of CA it was proposed by Wagner that CA can be considered as analytical engine of GIS. Raster GIS with map algebra can be integrated with enhanced capabilities.

CA is considered to have a "natural affinity" with raster data. It has similarities with GIS, such as both represents attribute information in a layered fashion, and manipulate that information with operators (Overlay in GIS, Transitional rules in CA). The focalsum or focalmean functions of GIS has direct analogous with neighborhood functions. Having a natural affinity with the GIS it was obvious to have adopted by geographers as a tool for modeling spatial dynamics

Here an attempt was made to integrate non-spatial information with spatial information using GIS and Cellular automata concept integrated with Multi Criteria Technique. It was found that despite the limitations at this stage it can be used to generate the different scenario which can address 'what-if ' issues. This model can become the basis of further improved model.

TABLE OF CONTENTS

1. Chapter one

1.1. Introduction:

The first key word in the title of this thesis is modelling. Model is an abstraction or simplification of real world (Benders 1996). From early days of our childhood we learn (subconsciously) from models and play with models. Playing with guns, cars and dolls, making houses with sands, are some of the examples, which we never realize in our daily life. Models are being used as a very powerful learning tool. They are not only static like road maps but they can be dynamic like playing with the car or plane with remote control or in video games. There is another form of model called as mental model or mental map (Benders 1996). We can find the things even if some times lights are not on in our house. A visit to a restaurant is an example often used; using a mental map we do all the actions (ask for the menu card, order the meal, eat the meal and pay the bill) in the correct order [Johnson-Laird in, Benders (1996)]. The *two* main objectives (Peccei, Saraph in Benders 1996) of the modeling are *Communication* (or to convey the knowledge concerning education, training, negotiation, and gaming) and to *reduce Uncertainty* for decision support for Planning, forecasting, backcasting. The *second* objective (reducing uncertainty) can be further divided into planning, forecasting, and backcasting models (Benders 1996). The forecasting models are characterized by "what if", which illustrates the type of questions to be asked. These models contain a broad range of time scales, from a few years to more than a hundred years (e.g. global change models). Forecasting models are also called scenario models. From a certain starting point and certain expectations about (for the model) crucial developments (a scenario), a situation in the future can be described. Modeling used in this thesis comes under this category.

The second key word in the title of this thesis is *land use land cover change*. Land use activity is a major issue and challenge for town and country planners (as well as environmentalists) to design the Eco friendly and sustainable economic growth. The human activity for development is forced against the environment, which results in consequences such as soil erosion, global warming, pollution etc. The causes for change in land use type activity may be due to socio-economic development or due to changes in the environment or may be due to both. For example, an increase in total export demand for a particular agriculture product will be translated into increase in demand for land for this particular product whereas increase in tourist influx will result in increased demand of land for constructions. Land can be gained by conversion of agricultural fields or by clearing the forest. In both cases the consequences will be different. Envisioning the consequent effects of land use changes, IGBP (International Geo-sphere and Biosphere Program) and IHDP (International Human Dimension Program) co-organized a working group to set up research agenda and promote research activity for land use and land cover changes (LUCC). The working group suggested three core subjects for LUCC research, such as: situation assessment, modeling and projecting, and conceptual scaling. The ultimate goal of global change study is to assess the impacts under each possible scenario and suggest preventive actions. The modeling and projecting of land use change is essential for scenario analysis and the assessment of LUCC. Consequently, issues related to data, information, and modeling have attracted many research interests ranging from local authorities to global organizations.

Third key word in the title of this thesis is *Geo-Spatial environment*. Adding the spatial component in the model was evolutionary concept in the Cartographic modeling in 1960s and 1970s. From the 1980s to current the focus has been on the spatial modeling. A cartographic model tends to be more 'static,' meaning it depicts spatial variation in quantitative data, but does not tell about the spatial influence on the variation. The cartographic data model uses points, lines and polygons (topologically encoded) with one, or only a few, attached attributes, such as a land use layer represented as polygons with associated land used code. Spatial models can be seen as the extension of the Cartographic models. It has all those geometric shapes mentioned in cartographic model with set of multiple associated attributes A spatial model depicts spatial processes, or the influence of spatial factors on spatial variation. Finally, spatial models can supply a more 'realistic' view of reality than a cartographic model - as spatial factors can have great influence on the variability of a statistic. A spatial model is the integration of spatial components into mathematical models. With spatial models, spatially dependent factors (e.g. distance or slope) can be incorporated with other statistical data (e.g. population or agricultural production) variables can help to refine model solution and yield better results.

Fourth key word in the title of the present thesis is Cellular Automata (CA). CA are dynamic, discrete space and time systems. A cellular automaton system consists of a regular grid of cells, each of which can be in one of a finite number of for example) k possible states, updated synchronously in discrete time steps according to a local, identical interaction rule. The state of a cell is determined by the previous states of a surrounding neighborhood of cells. However, it is important to remember that the early work on CA systems was fundamental theoretical work. Using it with powerful computing technology (both hardware and software) and appropriate data did not start until relatively recently.

Since the evolution of cellular automata, it is being used in many disciplines ranging from sciences to commercial fields. Because of its capabilities to address the complex patterns with the help of very simple transition rules it was accepted in every corner of the field of research.

In comparison with traditional approaches based on differential or difference equations(Baker 1989), the CA has advantages. CA can incorporate spatial component. And they address dynamism with simple rules, which increases computational efficiency. Since computational efficiency translates into better handling of dynamism (White *et al.* 1997) CA becomes favorites to many modelers.

The advantages of CA are many. The construction of model is simple and easy. It has an ability to perform spatial dynamics, and time explicitly. After Analyzing the similarities and capabilities of CA it was proposed by Wagner (Wagner 1997).that CA can be considered as analytical engine of GIS. Raster GIS with map algebra can be integrated with enhanced capabilities as discussed by Takeyama *et al* (Takeyama *et al.* 1997).

CA is considered to have a “natural affinity” with raster data. It has similarities with GIS, such as both represents attribute information in a layered fashion, and manipulate that information with operators (Overlay in GIS, Transitional rules in CA). The focalsum or focalmean functions of

GIS has direct analogous with neighborhood functions. Having a natural affinity with the GIS it was obvious to have adopted by geographers as a tool for modeling spatial dynamics

The roots of CA in Geography can be traced in "A Monte Carlo approach to diffusion" a paper by Hagerstrand (1968) (as in(Wagner 1997). Though he did not use the term CA but much of CA can be seen in his approach. Tobler (1979) (as in Wagner 1997) introduced concept of CA in his cellular geography. He discussed how CA could be useful aid in planning. He discussed how planners can get the best transition rule and how they can use CA in estimating the time taken in transition. The property that CA can generate complex spatial patterns from the simple set of rules was realized in many simulation studies (Couclelis 1985). CA has been used in many ecological applications.(as in Wagner 1997)

CA has been used intensely in one particular area, which is modeling urban processes. (Couclelis 1985; Couclelis 1997) (as in Wagner 1997) explored the potential of CA in urban planning and discussed theoretical obstacles in incorporating CA models in Geographical context. White *et al* (White *et al.* 1993; White *et al.* 1994) have used CA model in evolution urban land use. (Clarke *et al.* 1997; Clarke *et al.* 1998) also used CA in modeling urban forms. (Batty *et al.* 1994; Batty *et al.* 1997; Batty *et al.* 1999; Batty 2000) have come up with urban models and urban studies.

1.2. Objective

To develop a CA model that can amalgamate different ground processes in predicting the land use land cover changes.

1.3. Research questions

What kind of process can lead to changes?
What are the causalities?
How can one choose spatial variables?
How can one establish relationship between causalities and changes?
What type of transition rule can be used?
Does different size of neighborhood has an impact of result?

2. Chapter two

2.1. Literature review:

Land use land cover change models can be used for different purposes. These models can be categorized according to amount of information they contain. These are Whole landscape models (Baker 1989), Distributional landscape models (Baker 1989), or spatial landscape models (Baker 1989). Thanks to enormous progress in remote sensing and GIS technology it is easy to capture spatial details and use them in analysis (Takayama *et al.* 1997). Land use land cover change is influence by various natural and human activity processes. Spatial details plays important role in these process (White *et al.* 1997). Therefore spatial modeling has more relevance than other methods of modeling in research. Difference or differential equation based models are dynamic and generate relatively complex results, both temporally and spatially. However solutions that are better than a very crude spatial resolution are hard to achieve computationally (Chen *et al.* 2002).

Different approaches have been attempted in spatial modeling. To name a few models, based on approaches 1) Cellular Automata models, 2) Artificial neural network models 3) Multi Agent models 4) Statistical models 5) Fractal models.

Cellular Automata Models (CA): The approach in this model is 'bottom to top'. The final global structure emerges from purely local interactions among the cells. CA not only offers a new way of thinking for dynamic process modeling it also provides a laboratory for testing the decision making processes. As stated earlier they have natural affinity with GIS and remotely sensed data (Torrens *et al.* 2001). One of the most significant properties of CA is perhaps its simplicity. There are many studies done depending upon different approaches for transition rules and calibration techniques. Some of these will be mentioned later.

Artificial Neural Network Models (ANN):

Artificial Neural network consists of simple processing elements, called neurons, and connection links operating in parallel. Two or more of the neurons can be combined in any layers. A network may contain one or more layers. Typical three kinds of layers in neural network architecture are known as input layer, hidden layer and output layer. The layer between input and output layer is called as hidden layer, because they have no direct relationships with outer real world. However, the number of hidden layers and neurons required to obtain an accurate approximation could not unanimously decided yet. Some found fewer neurons reduce information for processing. On the contrary, others found too many neurons amplify the information amount and increases training hour directly to the proportion of neuron's number. Moreover, the network doesn't converge to stable state. The neuron is a processing element whose output is produced by multiplying its input signals by a weight. Each neuron sums up the information from lower layer, and applies a transfer function to make an output.

ANN model is taught by sample data taken in real area as a human brain learns, thinks and reacts against stimulus. During the training, initial weights that are assigned to interconnection links are modified repeatedly until the ANN can produce acceptable outputs that matches the original target values even though not exactly the same.

Agents based modeling: 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 modeled, often at the individual level. They can be developers, individuals, state policy etc. their influence can be at different scales. 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 behavior to the environment. Multi-agents also have attractive features as discussed by White *et al* (White *et al.* 2000)

Spatial-Statistical models: Although 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, they needed to have spatial component within themselves so that they can used to their full potential in geography. But even after adding the spatial component they are criticized by many authors, as time and spatial domain do not follow standard distribution like normal distribution. Therefore the sampling technique is also questioned.

Fractal based model: fractals are spatial objects having properties of 1) self-similarity (scale independent), 2) fractional dimension. They can be formed by repeating themselves. The natural objects like ferns, coastlines etc has been represented by fractals successfully. Many workers (Batty et al in (Cheng 2003a)have used this approach in urban studies

2.1.1. Modeling with CA

Many CA models have been developed so far especially in growth urban process studies,
Depending upon their transition rules, and calibration methods they differ from each other.
Some of the CA models are mentioned below:

- Macro and micro integrated CA models (RIKS model):

These models (Engelen *et al.* 1997) have two parts, macro model and micro model. Macro model generate different land use consumption based on socio-economic factors, whereas the allocation of the cells to the different land use type is done on the micro scale.

- SLEUTH Model

Prof. Clarke in University of California, Santa Barbara, developed this model (Clarke *et al.* 1997). Since this model is based on influence factors, which are *S*lope, Land use map, Excluded area, Urban area, transportation map, hillside area. Therefore model is called as SLEUTH. Five factors control the behavior of the system. These are: DIFFUSION factor which determines the overall dispersiveness of the distribution both of single grid cells and in the movement of new settlements outward through the road system; a BREED coefficient which determines how likely a newly generated detached settlement is to begin its own growth cycle; a SPREAD coefficient which controls how much normal outward 'organic expansion takes place within the system; a SLOPE-RESISTANCE factor which influences the likelihood of settlement extending up steeper slopes: and Road gravity factor which has the effect of attracting new settlement on to the existing roads if they fall within the given distance from the road

- Fuzzy CA Model

Wu (1996; 1998a) in his study introduced a unique approach ie. heuristically defined transition rule with CA to simulate rural-urban land conversion in a fast growing metropolis. Unlike previous studies in which transition rule is defined by a mathematical equation, this approach introduced the concept of the fuzzy logic control (FLC) into mimicking land conversion process. Preconditions of an action are described by fuzzy sets and state changes are simulated according to these fuzzy sets.

- ANN CA Model

It is not always easy to develop a detailed transition rule in CA model. It is critical to get the correct parameter values in CA to have good and more realistic result. If large number of parameters are involved in the model then calibration becomes difficult. Artificial Neural Network provides a way to get these parameter values automatically ((Li *et al.* 2001a). When simulation is done for multiple land use changes ANN is considered as a promising tool.(Li *et al.* 2002a) The model based on ANN has been successfully applied by Li *et al* (2002) to the simulation of multiple land use changes in a fast growing area in southern China.In this paper, a three-layer neural network with multiple output neurons is designed to calculate conversion probabilities for competing multiple land uses. The model involves iterative looping of the neural network to simulate gradual land use conversion processes. Spatial variables are not deterministic because they are dynamically updated at the end of each loop. A GIS is used to obtain site attributes and training data, and to provide spatial functions for constructing the neural network. The parameter values for modeling are automatically generated by the training procedure of neural networks.

- MCE-CA Model

The transition rules have been defined by various ways. One of the novel way has been proposed by Wu (1998c). The combination of three elements, GIS, CA, and MCE, as claimed by the author has several advantages: visualization of decision-making, easier access to spatial information, and the more realistic definition of transition rules in CA. the present thesis is based on the principle of these model.

- Multi-CA Model

Possibility of using more than one cellular automata for transition rule was explored by (Rinaldi 1998). In reality while addressing the complex problem it may be possible that problem is too complex to be addressed by single cellular automata. Therefore concept of multi cellular automata (MCA) evolved where two cellular automata is linked with each other. Rinaldi (1998) in his study, suggested a schema for more than one cellular automata which can be dependent on each other. The schema can be implemented in GIS environment. It can be implemented manually as well as it can be made automatically. If attempted manually user has to start the sequence every time for each cellular automata.

- Statistic Based CA Model

Some studies based on statistic with CA, has been reported (Wu 2000; Sui *et al.* 2001). Some workers have used Monto Carlo simulation to get the parameter values (Wu 2000) and some have used regression technique. Sui *et al* (2001) had calculated weights by a GIS-based resampling method in three following steps. First all the changed pixels were extracted by overlaying the temporal data. Second then 100 samples of those cells which were converted into urban were collected with the help of moving window of size 10×10 cells. The proportion of converted cells in each window is taken as trasition probabilities for that window cells. Third, the mean values of factors like elevation etc were calculated of those cells in the window. With the help of multiple regression value for the weights were calculated.

- Stochastic CA model:

Some studies has been reported in which the parameter values were obtained stochastically.(Almeida *et al.* 2003) has come up with formal framework based on Bayes theory of probability and weight of evidence. (Wu 2002) had developed a stochastic CA model, which derives its initial probability of simulation from observed sequential land use data. Furthermore this initial probabilities is updated dynamically through local rule based on the strength of neighborhood development

3. Chapter Three

3.1. Origin

The Cellular automata concept is not new. This concept evolved in 1940s in the field of computers. Von Neumann and Ulam are considered the developers of this concept. Later Conway contributed in this field, which is known as "Game of life" People tried to apply this concept in artificial life (robots) but till date they are not completely successful. Although the concept of cellular automata originated in computer science for the development of robots, it is being widely applied to various other disciplines such as Physics, Mathematics, Natural Sciences, and GIS etc. This approach is successfully used in various problems in many disciplines. The beauty of this concept is that it evolved from a very simple pattern to a very complex pattern with the help of simple rules.

3.2. The conventional CA

Cellular Automata (CA) get their name from the fact that they consist of *cells*, like the cells on a checkerboard, and that cell states may evolve according to a simple transition rule, the *automaton.* Cellular Automata are dynamic model that inherently integrates spatial and temporal dimension.

Cellular Automata are composed of five elements as described below:

Cell Space: The cell space is composed of individual cell. Theoretically, these cells may be in any geometric shape. Yet, most CAs adopt regular grids to represent such a space, which make CA very similar to a raster GIS.

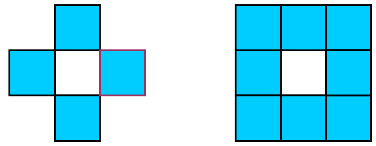
Cell states: The states of each cell may represent any spatial variable, e.g., the various types of land use.

Time steps: A CA will evolve at a sequence of discrete time steps. At each step, the cells will be updated simultaneously based on transition rules.

Transition rules: These rules are the heart of a CA that guides its dynamic evolution. A transition rule normally specifies the states of cell before and after updating based on its neighborhood conditions.

Neighborhood: Each cell has two neighbors in one-dimensional cellular automata, whereas in two-dimensional cellular automata model there are two ways to define it. Von Neumann has considered four neighboring cells as neighbors. Moore considered eight neighboring cells as neighbours. This can be shown as below:

.



The colour cells are neighbours to centre cell which is shown in white colour. The first one is Von Neumann's neighbor and the second one is Moore's neighbors.

3.3. Geoinformatics adopted CA

Models are an abstraction, simplified versions of real objects and processes that can be used as experimental tools to study the real world problems. However the basic CA defined by Ulam, Von Neumann, and Conway etc. is not well suited for the applications in Geoinformatics. Therefore it is necessary to modify it heavily from its formal characteristics, which can be seen as follows:

Lattice (Cell space): In the classic CA two-dimensional lattices are defined as infinite planes or torus (both side edges are connected), which are structured into uniformly spaced, squares or cells. For mimicking the real world, it needs to be modified. The space is limited either by natural constraints or by policies. The spatial units may be different either in size and shape (O' Sullivan proposed a graph cell form, 2000, 3d spatial pattern proposed by Sembolini, 2000).

Cell States: In basic CA, the cell space is closed. Therefore it is assumed that external events cannot influence the dynamics within CA but in the real world for study purposes regions are taken, which are not closed and have many exogenous links. To incorporate the external influence into the CA model, constraints and algorithms are used. To impart greater flexibility into the CA model, states can be divided into two groups, "fixed" and "functional". The states, which are not changed in the process, are treated as fixed such as water bodies and terrain. The states that change with respect to time are considered as functional for example population and land use.

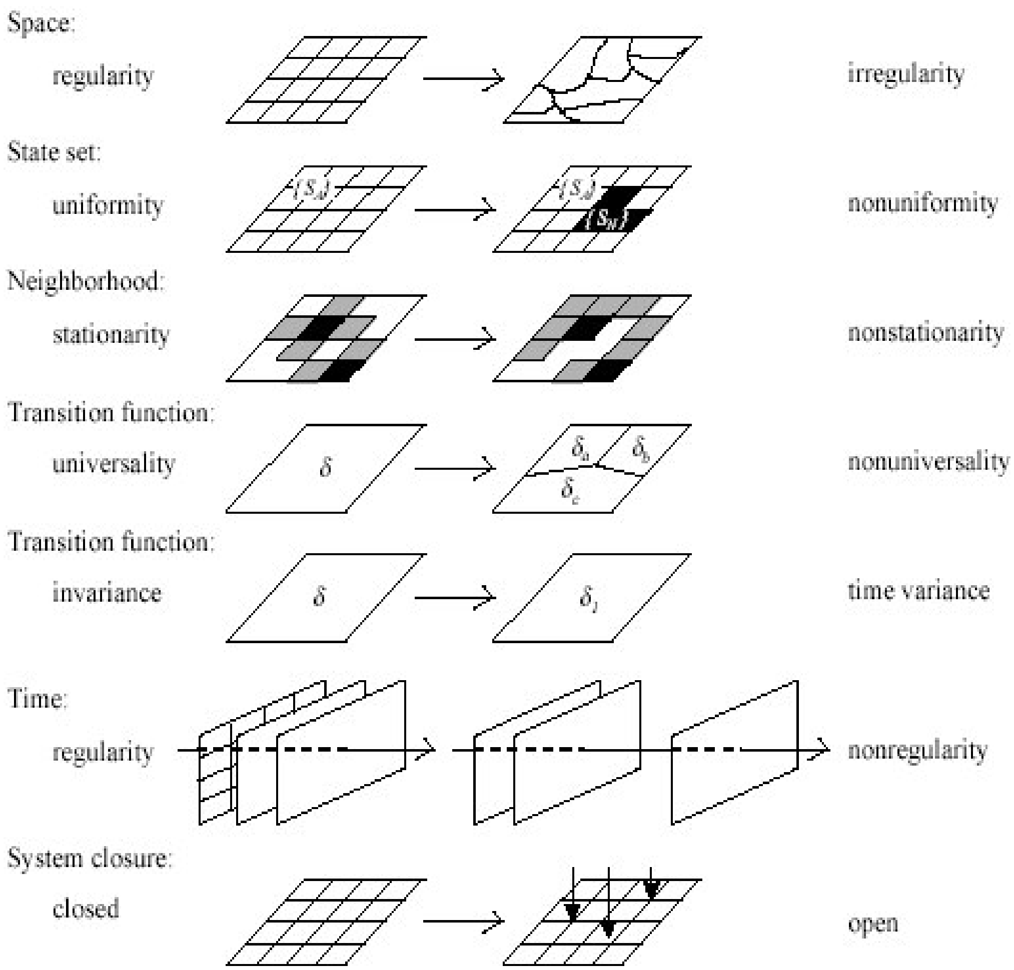
Neighborhoods: There are two types of neighborhoods in conventional CA. They are Von Neumann and Moore. The assumption that just adjacent neighborhood cells have an influence on the cell and not others is not realistic. Here different weights can be used to define distance decay effects (White *et al.* 1997). Different shape and size of neighborhood is taken now a day.

Time: The classic CA takes discrete time interval, which is same for all cells and transition rules are applied simultaneously at every cell. In the real world this does not happen always. Different cells have different rates. Researches are attempting this problem by developing a combined CA_Agent model (Barros and Soberiera, 2002 as sun)

Transition rules: The real world behavior is translated into a CA model by transition rules only. The transition rules make the CA model dynamic. In classic CA transition rules are deterministic and unchanged during evolution. In several studies they are modified into stochastic expression and fuzzy logic controlled methods (Wu, 1998).

In basic cellular automata studies each cell is open for changing its state depending upon the transition rule. In some studies constraints were realized by making some cells unavailable for change (White *et al.* 1997). The unavailability can be based on some global factors like state policy for not cutting the jungles, or to preserve park area within the city etc.

The following figure shows the comparison between the basic cellular automata and cellular automata adopted in GIS. The left side is characteristics of basic cellular automata and right side is modification in these characteristics of cellular automata adopted in GIS.



Note: the figure is taken from (Rinaldi 1998)

4. Chapter four

4.1. Model construction

:

4.1.1. Conceptualization:

As stated in the First section, a model is an abstraction of reality that is used to understand complex relationships. A model is usually the result of the examination of a relationship between two (or more) sets of data. A model can be used to understand how, and even why these data interact together or how does that relationship help to understand more about reality and its systems. The simplifications in the models used make them fit for a certain abstraction level for the persons to teach (cf. the simplified Rutherford-Bohr model for atoms taught in high school).

Land use depends on three types of factors (White *et al.* 1997): the inherent qualities of land itself; the effects of neighboring land use activities; and the aggregate demand of land for particular activity. The problems with traditional cellular automata are 1) cell itself has no characteristics. All the cells same quality and they are characterized by their state only. As explained in chapter three, in transition rule, only neighborhood plays role, whereas in land use change model geographical locations also play role. Particular piece of land can be more suitable to particular land use type/activity. To incorporate these the cell space is no longer homogeneous cell space. 2) In the traditional cellular automata any number of cells can undergo through the transition. Therefore number is unconstrained. In the land use change process this number is constraint by demand of cells to change. Only that number of cells will go under transition, which satisfy the demand.

Since every model is only an abstraction of reality. It has to go through some simplifications. In general we can define change in land use land cover as a function of various factors. Mathematically it can be defined as:

$$\Delta L = \Delta L_1 + \Delta L_2 + \Delta L_3 + \ldots\ldots\ldots \qquad (1)$$

Where Δ L is total change in land use type. $\Delta L_{1,}$ $\Delta L_{2,}$ $\Delta L_{3,}$ etc are changes in different land use types. For an example ΔL_1 may represent change in settlement area, ΔL_2 may represent change in agriculture area, ΔL_3 may represent change in forest area. And so on. This series can be truncated according to application. If only two kinds of changes to be studied then only two components of the series can be taken.

Note: First simplification is to put number of land use type changes.

One important point here is to be mentioned that this change is not scalar only. It has positional value also. Theoretically it can happen that just shifting takes place ie one or more land use type can shift the places without changing the magnitude (area).[Here lies the importance of spatial analysis in GIS.]

Each land use type change can be function of different factors. For an example

$$\Delta L_1 = F(x_1, x_2, x_3, \ldots) \qquad (2)$$

Where ΔL_1 is change in land use type (say) settlement. x_1, x_2, x_3, x_4,…are factors (like population growth, economic growth, policy etc.) responsible for change. F is a function of various factors x_1, x_2, x_3, x_4, etc. these factors can influence the other land use type change also. For an example population growth can force settlement area to increase, and it can increase the demand for agricultural land. Also these factors can influence each other for an example due to population growth, state policy may change etc. not only this, the factors itself may be representing the processes for an example population growth which can have their own drivers. Due to these reasons land use land cover change becomes more complex and challenging phenomenon.

Assuming factors to be independent the function $F(x_1, x_2, x_3, \ldots)$ can be decomposed further into additive series as below:

$$F(x_1, x_2, x_3, \ldots) = f(x_1) + f(x_2) + f(x_3) + \ldots\ldots(3)$$

Where $f(x_1)$ is a function of only factor x1, $f(x_2)$ is a function of factor x_2 only, and $f(x_3)$ is a function of factor x_3 only. It could be decomposed into multiplicative series also but in the case of multiplicative series if any of the process does not show change it will make resultant zero. Assume that a population of one study area is constant therefore there will no change in population factor which will make value of function $F(x_1, x_2, x_3, \ldots)$ as zero. But in that case also a change in land use can take place, for an example, due to state policy (such as protecting forest) there can be shift in settlement area is required. Therefore to avoid this situation the additive series is considered to be better.

Again this series can also be truncated according to level of study. More components will add accuracy in the result but methodology will be the same.

The next step is to transform these factors functions i.e. $f(x_1)$, $f(x_2)$, etc into land requirement. This can be called as quantifying of causalities. The important point to remember here is that the factors, say $f(x_1)$ here can be process. For an example population growth. Now these can have their own nature and growth rate. And these can be very much location and time based. Population growth at one place can be entirely different from the population growth at the other place or this growth can be different at different times even for same region. If we take a region, which comprises of more that one urban area, and more than one rural area, then these processes can vary from area to area. Therefore for the simplification it was felt that for a particular region the processes could be seen at regional scale. The quantification of requirement will be for whole region.

Assuming population growth is the major reason for a change in settlement type in a region, in this model, to quantify the demand for settlement change is calculated from only population growth

process. Theoretically this point can be challenged, as there may be other major reasons also for an example because of tourism demand for settlement can be generated whereas in population growth only permanent residents are considered. Depending on the situation they can be calculated provided data is available. In that case also the basic methodology will be same. Moreover this issue can be addressed by introducing *correction term*, which is explained later.
On the basis of these arguments

$$f(x_1) + f(x_2) + f(x_3) + \ldots\ldots \approx f(x_1) + c \qquad (4)$$

Therefore the equation (3) reduces to

$$F(x_1, x_2, x_3, \ldots) \approx f(x_1) + c \qquad (5)$$

Therefore the equation (2) reduces to

$$\Delta L_1 \approx f(x_1) + c \qquad (6)$$

Where c is correction term, $F(x_1, x_2, x_3, \ldots)$ is change in settlement type only, $f(x_1)$ change in settlement type due to factor x_1 i.e. population growth. From the statistical data the population growth can be studied and population can be extrapolated for the next coming years. The larger the data set the better would be the extrapolated value. How much population requires how much land is very subjective issue. This varies from one region to another. Other important point is population growth can be of two types one is urban and another is rural. The land requirement for each unit of these two classes of the population varies.

$$f(x_1) = f(u) + f(r) \qquad (7)$$

Where u represents urban population growth and r represents rural population growth. Therefore equation (6) can be written as

$$\boxed{\Delta L_1 \approx f(u) + f(r) + c \qquad (8)}$$

The growth rates u and r can be obtained from the statistical data (historical data). These will be in form of equations with the help of which urban and rural population for coming years can be extrapolated. Once getting the populations the area requirement for that population is calculated. This is region specific. This depends on the composition of the population, what percent of population comes under rich class, middle class and poor class. The average land holdings of these classes will be region specific. Multiplying with the average land holding or according to planning body the average size of land for different classes with the size of respective classes will give the land requirement for that population.

Again moving to equation (1). The other component i.e. ΔL_2 the change of other type of land use can be taken as change in agriculture area. On the similar guideline as stated above ΔL_2 can be represented as

$$\Delta L_2 = G(x_1, y_1, y_2, \ldots) \qquad (9)$$

Where $G(x_1, y_1, y_2, \ldots)$ is function of factors $x_1, y_1, y_2, \ldots$etc.

As stated earlier factors can be common also therefore x_1 represents population growth here, others can be for example state policy, labor availability, economic factor such as export-import policy etc. these factors can again influence each other. For an example more population will demand for more food production. On the other hand it will eat up agricultural land near urban and sub-urban areas. More food production can be achieved through more than one way. By increasing productivity through new technologies is one way (India has seen this in late 70s which is termed as green revolution). Other is to increase agricultural land through converting wastelands into agricultural land, or cutting jungles. Since this type of component cannot be broken into simple components easily therefore another assumption is made here. Assuming, $G(x_1, y_1, y_2, \ldots)$ can be related with production trend of last 30 to 40 years of data. Therefore we can write

$$G(x_1, y_1, y_2, \ldots) \approx f(p) + c_1 \qquad (10)$$

Therefore

$$\boxed{\Delta L_2 \approx f(p) + c_1 \qquad (11)}$$

Where p is production trend and c_1 is correction factor. The form of p can be obtained from historical data and on the basis of this form (equation) the production for coming years can be extrapolated. Since new technology is coming in and awareness is more and more per day the productivity is also increasing. Even after the quality of soil is deteriorating (because of excessive use of chemical fertilizers) so far people are managed to get better production in comparison to previous years. Therefore in area calculation increase in productivity should also be considered. Total production divided by productivity will give the area required for that production.

For the model at present only two classes are taken as competitor classes. Therefore final equation can be written as:

$$\boxed{\Delta L \approx f(u) + f(r) + c + f(p) + c_1 \qquad (12)}$$

Second step in the model is to decide where these changes would take place over the time. To solve this problem concept of constrained cellular automata (which is explained in chapter three) has been borrowed.

The region can be seen as an integration of small size of parcels, which have regular and square shape. These can be termed as cells. All these cells can act as systems having attributes (geophysical as well as socio-economic), which can change their state. Therefore the region is considered as collection of *discrete* system. While assuming this we are introducing some approximation as in reality the different classified areas are not of regular shape. And at the boundary they may not follow the square or regular shape. If region under study is large in comparison to unit i.e. individual cell size this approximation can work fairly well. But so far in literature I could not find what this ratio should be? This point will be further touched in discussion part of this thesis.

Since these cells are representing the geographical areas, and areas are undergoing through changes (in terms of land use types) under some 'influences' over the time, they are *dynamic* in nature.

Therefore the region is a collection of discrete and dynamic systems and these systems are capable of changing their states following some rules. This fulfills *partially* the definition of cellular automata.

There are various processes under which the *transitions* from one *state* (land use type) to another take place. Major processes are *Organic growths, Breeding, Slope resistance growths, Road gravity growth, Diffusion*. When the urban area grows around its perimeter it is called as organic growth. Major urban areas have their influence on the nearby areas and influence them to convert into urban area. When the area between the two urban areas starts growing into urban area and soon starts to influence other adjoining cells then it is called as Breeding. It has been observed that better transport network also influences the urban growth. The roads influence the cells near to it for settlements. This is called road gravity growth. When the area may not nearest to urban area but it is flat and it starts growing in to urban area. This is called Diffusion. Since the higher slope is found to resist the growth of urban area that is called slope resistance.

Besides these processes the cells have the geo physical attributes also which influence the land use land cover change. For an example altitude, dem, aspect etc. Therefore ideally transition rule should address all these processes and it should consider all the influencing factors.

On the basis of above discussions the quantisation of the causalities can be done but still how to decide where this changes should take place remains unanswered. For this reason the model look for CA concept.

Cellular automata can be represented as quadruple as follows:

$$(U,S,N,T) \tag{13}$$

Where U is universe (cell space or lattice) S is set of all possible states which a cell can attain, N is neighborhood of a cell, and T is a set of transition rules.

We can represent a state of a cell at time t+1 as a function of its state at time t, its neighborhood and transition rule as follows:

$$S_{t+1} = f(S_t, N, T) \tag{14}$$

Where S_{t+1} is state of a cell at time (t+1), S_t is a state of a cell at time t, and N is neighborhood, and T is a set of transition rules followed by cells.

Cell space (or lattice) is fixed as our region under study is fixed. First approximation is introduced here. Since our region is subset of large geographical set there are large number of cells are left outside of our consideration. There is large number of factors which influences the cells globally and there can be a situation when contribution of a state dominates the neighborhood and make those outside cells more suitable to undergo transition.

Cells representing a geographical area have attributes. Besides geo-physical attributes there can be socio-economic (like population or population density etc) attributes also. In this model for simplicity attributes are geo-physical (e.g. slope, aspect etc). Land use type is represented as state of a cell. Number of states available now is two only. Principally any cell should be able to attain any one of the states, but considering some ground realities the number of states available is restricted. If one cell attains a state 'settlement', it is unlikely to change its state (unless some major event such as earthquake, takes place). Therefore after attaining the settlement state a cell is not available for changing its state. Similarly due to some other reason like state policy to preserve the forest, or water bodies, makes some cells of a particular state unavailable for transitions. Here the concept of constrained cellular automata is applied.

As discussed earlier transition rules should address all the factors responsible for different growth processes and factors, which represents geo-physical characteristics. Factors can be divided into three categories. They are local factors, regional factors, and global factors. Local (neighborhood) factors calculate the number of development cells (state) or suitability in the small set of cells around a location. Regional factors measure the influences away from a distance. They can be measured by the use of some kind of distance decay functions. The values of global factors are invariant spatially, but changeable temporally.

Since macroscopically it has been calculated how many cells should have attained a particular state at a particular time transition rule will only allocate that number of cells to that particular state selecting them from the set of *available* (the cells which can undergo transition) cells spread over the region. To evaluate the candidature of a cell for a particular state model adopts Multi Criteria Evaluation (MCE) Technique. The reason behind this, for one, it is very simple and powerful technique and second it has been used successfully in the past by others (Wu 1998c; Chen *et al.* 2002). With the help of MCE all the candidate cells (those cells which are available for transition) can be given score on the basis of factors. The final score should be the integration of the scores obtained by local factors, regional factors, and global factors. Based on the requirement, the top scorer cells will be allocated to the particular state (land use type). In this model the final score (S) is divided into two parts. First part is suitability and second part is neighborhood influence. In mathematical form it can be represented as:

$$Sc = Su + N \qquad (15)$$

Where Sc is final score of a cell, Su is suitability score and N is neighborhood influence. Suitability again can be divided into two parts. First is due to some inherent geo-physical properties like slope, aspect, etc and other is due to distance from some centers such as distance from city, distance from road etc.
Therefore Su becomes

$$Su = Su\ (\text{physical factors}) + Su\ (\text{proximity factors}) \qquad (16)$$

Therefore equation (15) can be seen as follows:

$$Sc = Su\ (\text{physical factors}) + Su\ (\text{proximity factors}) + N \qquad (17)$$

To calculate the suitability score of a cell, model uses the following equation:

$$\boxed{Su = (\sum wi \times fi) / \sum wi \qquad (18)}$$

Where summation is over i. i take value from 1 to number of factors considered. wi is weight corresponding to factor fi.

To calculate neighborhood influence three things are important one, size of neighborhood second shape of neighborhood and third nature of influence. Size of neighborhood should not be that big that it violates the local factor concept. Principally CA takes either Neuman's neighbor or it takes Moor's neighbor. Here Moor's neighbor (nearest eight) is considered. Shape of the neighborhood is considered as square. Influence can be linear or non linear in nature. For simplicity reason it is considered as linear in nature. The score for neighborhood is calculated as follows. The Euclidian distances of neighbor cells from the cell is calculated in terms of unit. All the Neumann neighbors have 1 unit distance. The other four neighbors (Moor's neighbors excluding Neuman's neighbors) have Euclidean distance 1.41 times unit distance. Taking inverse of these distances are considered as points. Now depending on the number of cells present (*of the same state)* in neighbor these points can be added to get the neighborhood score. Therefore

$$N = \{\sum ni \times (1/d)\} / (\sum ni) \qquad (19)$$

Theoretically there can be a situation where the score of available cells comes equal. To resolve this conflict preference is given to 'settlement state'.
Therefore transition rule will be based on the following equation:

$$\text{If } S_{Set} >= S_{Ag} \text{ then 'state' = 'Settlement'} \qquad (20)$$
$$\text{If } S_{Ag} > S_{Set} \text{ then 'state' = 'Agriculture'} \qquad (21)$$

FLOW DIAGRAM OF THE MODEL

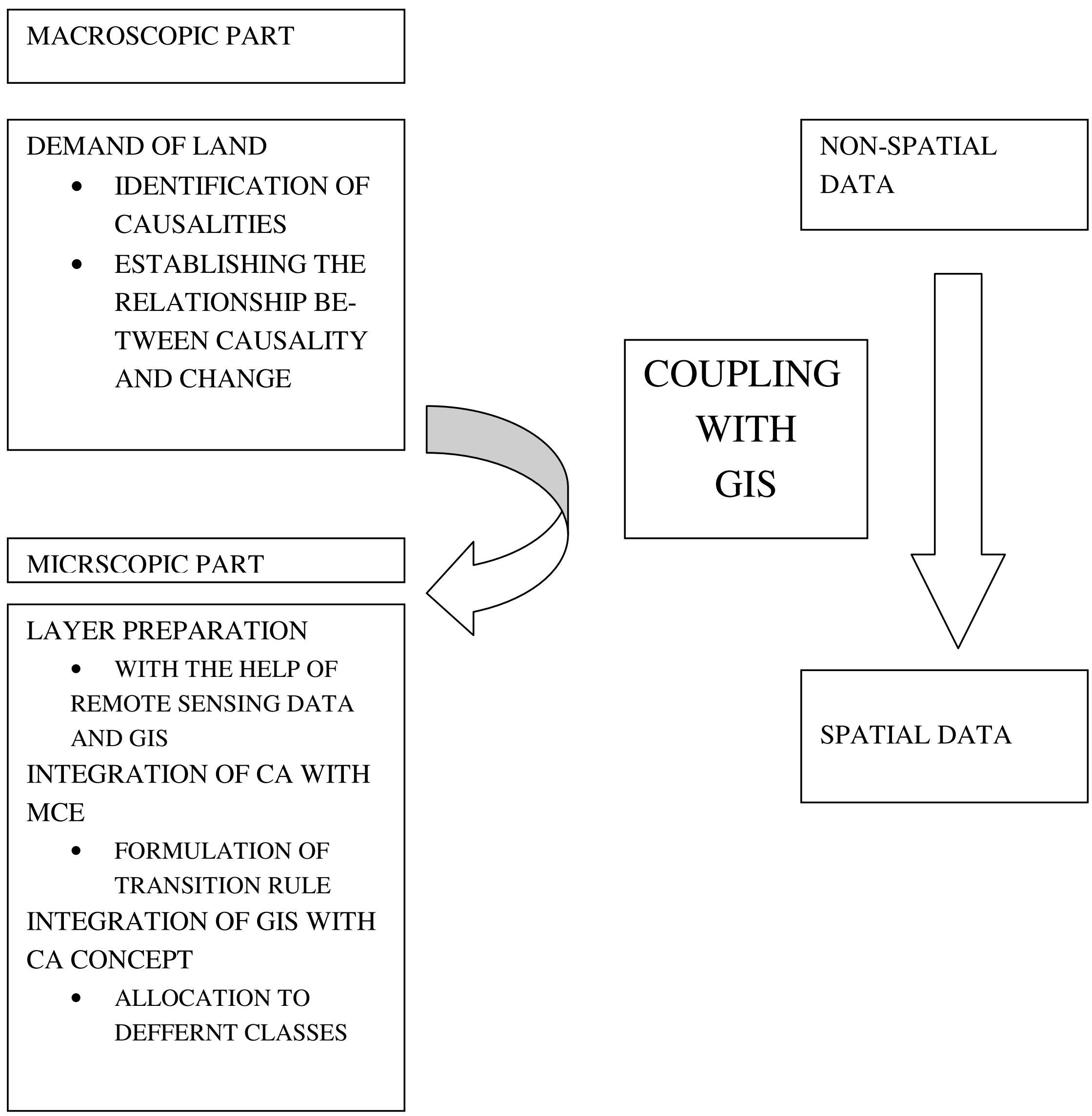


Figure explaining the conceptual flow of model

5. Chapter five

5.1. Case Study - Shimla

Shimla is the state capital of Himachal Pradesh in India (figure 1). It is one of the most important hill stations of north India. The state is also known for its historical values. The earliest known inhabitants of the region were tribal called Dasas. Later, Aryans came and they assimilated in the tribes. In the later centuries, the hill chieftains accepted suzerainty of the Mauryan empire, the Kaushans, the Guptas and Kanuaj rulers. During the Mughal period, the Rajas of the hill states made some mutually agreed arrangements, which governed their relations. In the 19th century, Ranjit Singh annexed/subjugated many of the states. When the British came, they defeated Gorkhas and entered into treaties with some Rajas and annexed the kingdoms of others. The situation more or less remained unchanged till 1947. After Independence, 30 princely states of the area were united and Himachal Pradesh was formed on 15th April 1948. With the recognition of Punjab on 1st November 1966, certain areas belonging to it were also included in Himachal Pradesh. On 25th January 1971, Himachal Pradesh was made a full-fledged State. The State is bordered by Jammu & Kashmir on North, Punjab on West and Southwest, Haryana on South, Uttar Pradesh on Southeast and China on the East.

Shimla (also spelt as 'Simla') derives its name from goddess 'Shayamla Devi', which is another manifestation of Goddess Kali. The capital of Himachal Pradesh came into light when the British discovered it in 1819. Till then, it was a part of the Nepalese kingdom. In 1864 Shimla was declared as the summer capital of India. After Independence, Shimla became the capital of Punjab and was later named the capital of Himachal Pradesh. In 1903 a rail line was constructed between Kalka and Shimla.

Shimla has been blessed with all the natural bounties, one can think of. Dwelling on a panoramic location, the hilly town is surrounded by green pastures and snow-capped peaks. The spectacular cool hills accompanied by the structures made during the colonial era create an aura, which is very different from other hill stations.

Bulging at its seams with unprecedented expansion, Shimla retains its colonial heritage, with grand old buildings; among them are the stately Viceregal Lodge, charming iron lamp posts and Anglo-Saxon names. The Mall, packed with shops and eateries, is the centre of attraction of the town, and Scandal Point, associated with the former Maharaja of Patiala's escapades, offers a view of distant snowclad peaks

Since shimla is state capital, it is natural attraction for the people of the state to move here. Therefore it is under the influence of lot of socio-economic forces. The geographical extent of shimla is as follows:

30 45’ 50.54”N to 31 43’ 11.93”N and 76 59’ 33.92”E to 78 18’ 53.84”E.

Shimla is divided into twelve administrative boundaries. These are termed as ‘Tehsils’. The figure I shows its location in its state ie Himachal Pradesh and in country ie India

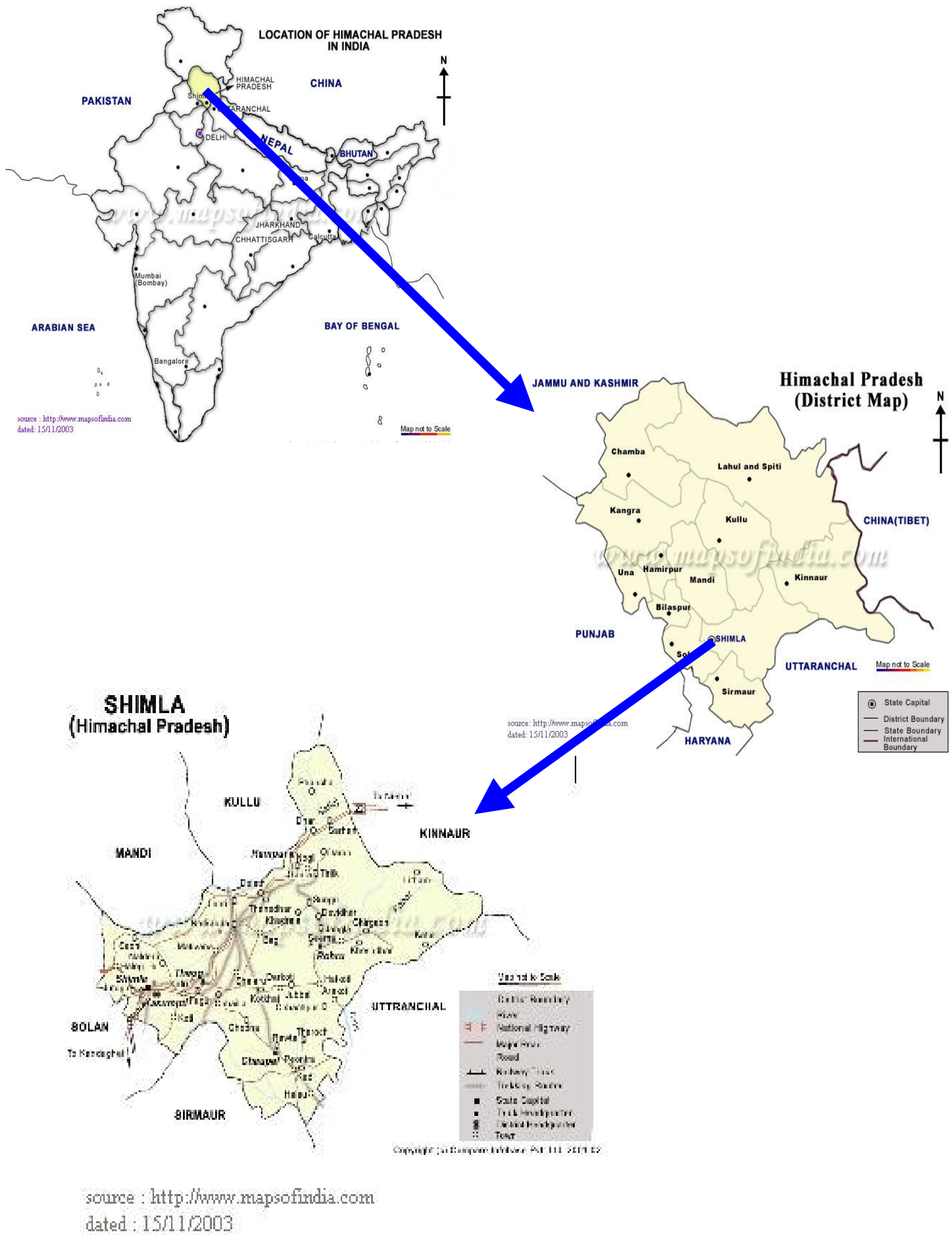


Figure-1 location of Shimla

5.2. Methodology.

The Landsat TM data for 1987 (figure 2a), and IRS 1D for 1999 (figure-2b) were acquired by Geoinformatics Division, Indian Institute of Remote Sensing (IIRS), Dehradun, India from National Remote Sensing Agency (NRSA), Hyderabad, India. The satellite data were georeferenced in Lambert Conical Conformal (LCC) projection system.

Due to cloud, snow and shadow in 1987 satellite data only 53% (approximately) of the total area could be classified. Snow cloud and shadow area were taken out from both the images for analysis. Because of hilly terrain it was very difficult to have classified map with high accuracy. Data were visually interpreted with unsupervised classification technique. Later after intensive ground truthing it was tried to correct it as much as possible. Due to difficulty in classification barren land and grassland were merged, and horticulture and agriculture were also merged. The final classes (figure3a, 3b) are taken as follows: Settlement, Agriculture (including horticulture), Barren (including grassland), Scrub, forest (figure 3a).

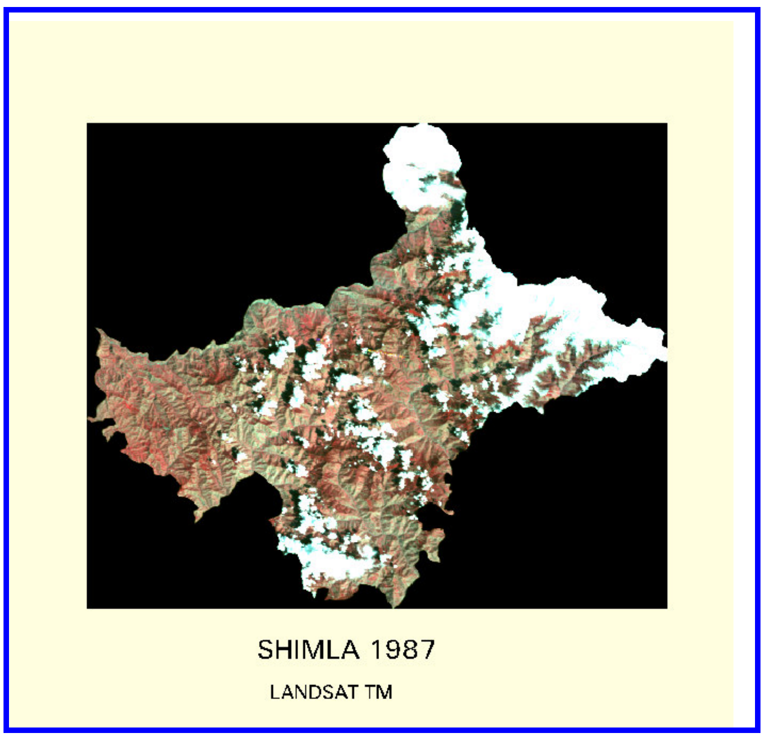


Figure 2a

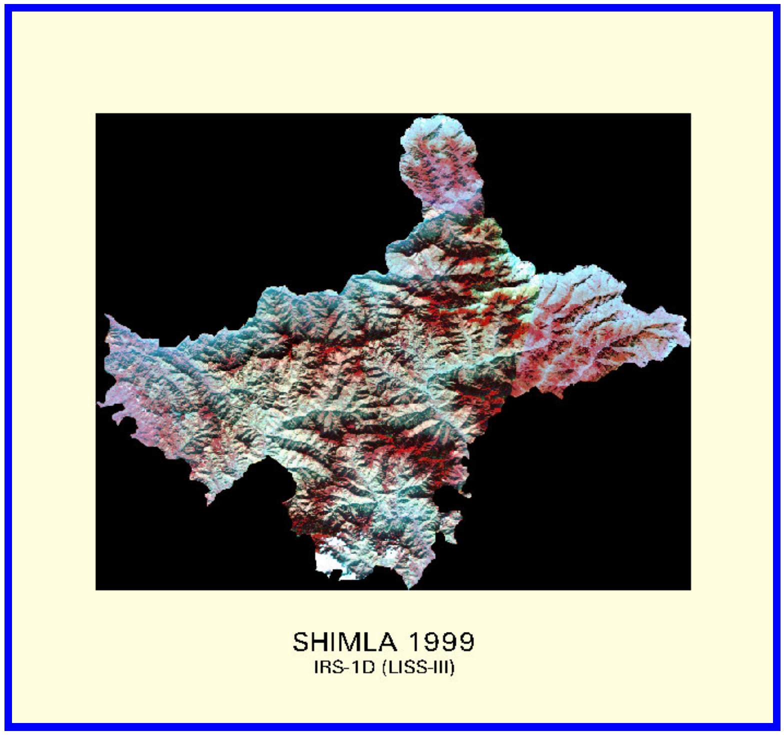


Figure-2b

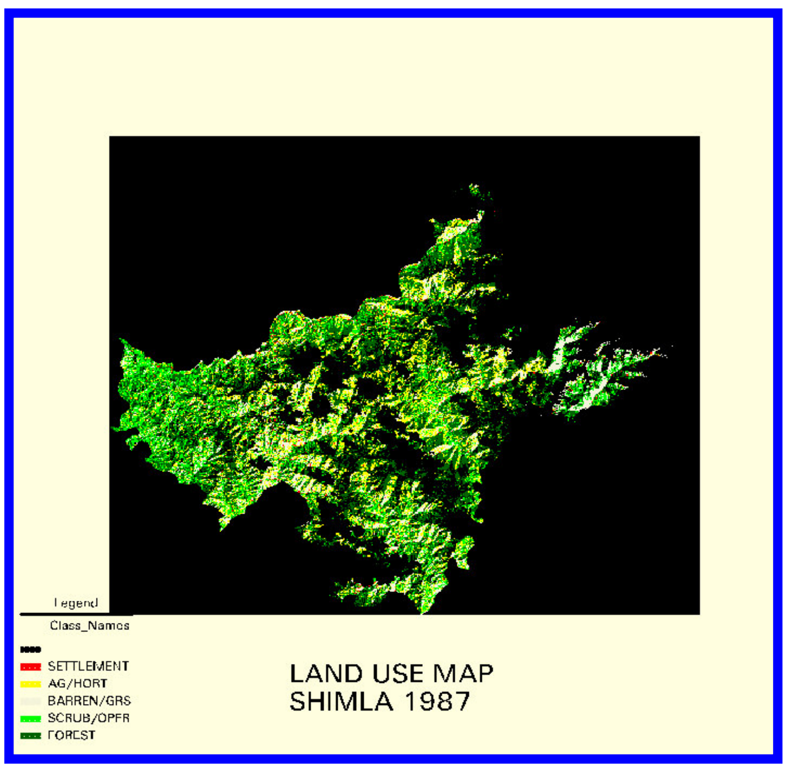


Figure 3a

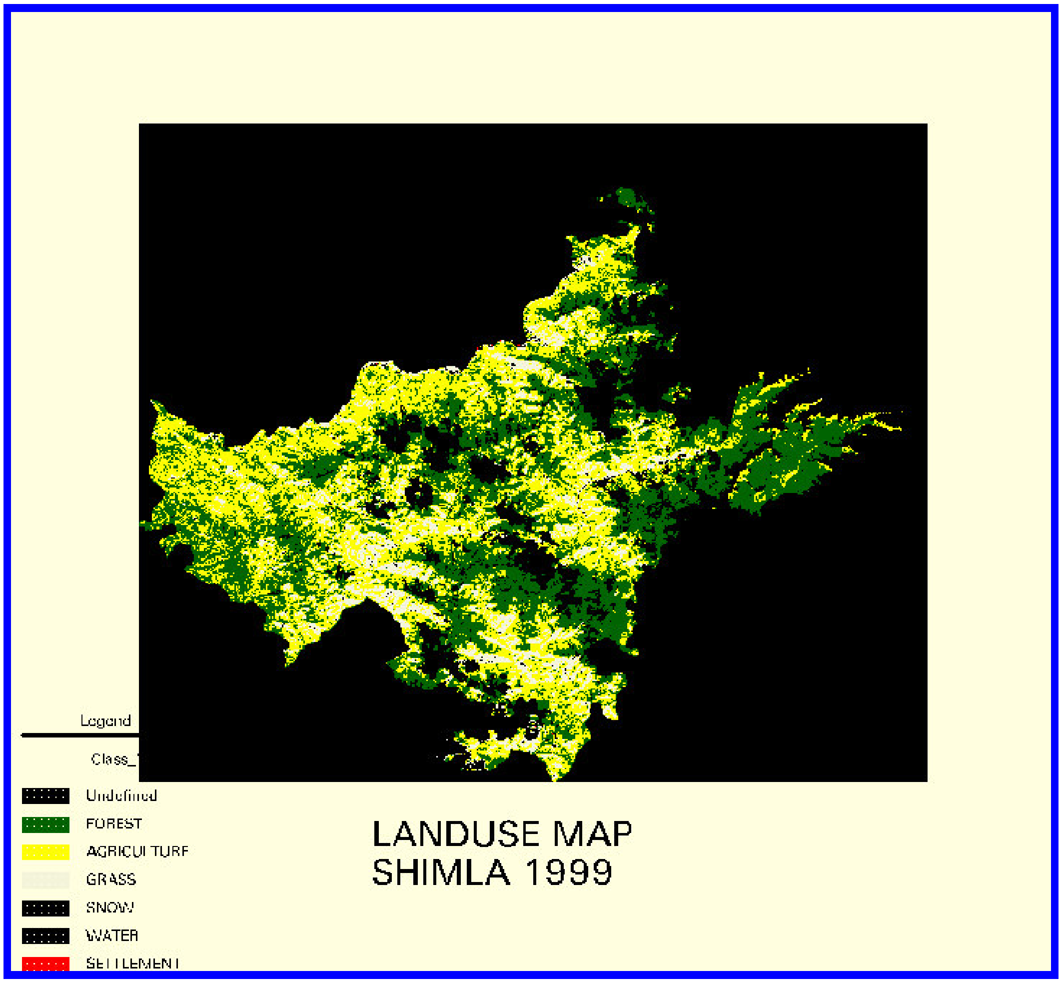

Legend
Undefined
FOREST
AGRICULTURE
GRASS
SNOW
WATER
SETTLEMENT
LANDUSE MAP
SHIMLA 1999

Figure 3b

5.2.1. Selection of Causalities (factors):

It is not possible to incorporate all the variable therefore, optimal variables were chosen with the help of literature (Agarwal *et al.* 2002) expert knowledge and data availability. For expert knowledge Mr A.N.Gautam, senior urban town planner, (Town and Country planning Department, Shimla) and Mr Sharma, Deputy Directorate Agriculture Department, Solan (Himachal Pradesh) has been consulted. If more variables are taken the result can be better, but the methodology here will be the same. The list of factors are shown in the table 3a

5.2.2. Quantifying the causalities (factors):

This is very important part of the model. For a good model it is very important to have accurate demand of the cell for a particular class. This can be achieved only if the relationship can be defined between the change in causalities and change in the land use types. Here quantification has been done on the guidelines set in the chapter four. Settlement and Agriculture were selected as competitor classes in the model.

As formulated in chapter four, first step was to generate demand for land for each class.

1. Demand of land
    - *Settlement*

To use the equation (8) of chapter four, f(u), f(r) and c has to be calculated. The population data (table 1a, and 1b) was collected during field visit. Using Microsoft Excel the best trend line (with highest regression coefficient value ie R^2) with equation (fig 4a, 4b) was observed on total population data. This equation was used to project the future population. Since the requirements for both kind of population (urban and rural) are different, the growth of urban population was observed in similar fashion as in case of total population was observed. The estimated urban population was deducted from the estimated total population each year to get the rural population each year. Then the urban population was divided by the average urban family size (Source: Mr A.N.Gautam) to get the number of urban families. The standard land requirement for a family (provided by Mr. A.N.Gautam), in that region was multiplied by the number of family to get the total land required for total number of fanily. The same exercise was done for calculating the land requirement for the rural population. Then the total was compared with the area shown in the satellite data. Assuming the difference between these two figure as error, mainly due to mis classification this figure was added to each year's land requirement for the settlement. This error is the value of c.

Therefore the form of f(r) and f(u) in equation (8) are (figures 4a,4b) as follows:

$$f(u) = 13.375e^{0.1261x} \qquad (22)$$

Total population growth, f(t), rate is:

$$f(t) = [96917x + 226485] \qquad (23)$$

Therefore,

$$f(r) = [96917x + 226485] - f(u) \qquad (24)$$

and

$$c = \text{area from satellite} - \text{estimated area} \qquad (25)$$

TEHSILWISE POPULATION (DISTRICT SHIMLA) 1961-2001

NAME OF TEHSIL AND SUB-TEHSIL	POPULATION				
	1961	1971	1981	1991	2001
SUNI	18750	22274	27003	29950	31423
SHIMLA RURAL	combined f	combined f	51058	66969	73550
SHIMLA URBAN			78206	102186	142161
KASUMPTI	27114	UN	UN	UN	UN
JUNGA	UN	UN	9630	11444	12420
THEOG	41167	47237	57138	65981	77967
KUMARSAIN	23828	30809	355650	39174	40569
NANKHARI	UN	16157	19443	23326	25152
RAMPUR	43847	40631	45960	52833	71779
ROHARU	51869	62910	32693	33554	44881
TIKKAR	UN	UN	10065	11450	12556
CHAUPAL	40391	18254	21476	24265	28386
CHETA	UN	UN	11313	13871	16829
NERUA	UN	28491	22621	27052	31158
JUBBAL	17776	21857	27383	31918	33609
CHIRGAON	UN	UN	28680	32895	39497
KOT KHAI	20583	24285	28583	32458	34147
DODRA KAUR	UN	UN	4120	4956	5661
TOTAL	329379	419844	510932	604282	721745

UN - Unavailabe

sources:

1. statistical outline 1961 - population for 1968
2. district census hand books 1981 and 1991 - population 1971, 81 and 91
3. physical figures by census deptt - population 2001

Table 1a

years	total	urban	% of total
1961	329379	UN	UN
1971	419844	UN	UN
1981	510932	78206	15.30654
1991	604282	102186	16.91032
2001	721745	142161	19.69685

UN - Unavailable

Table 1b

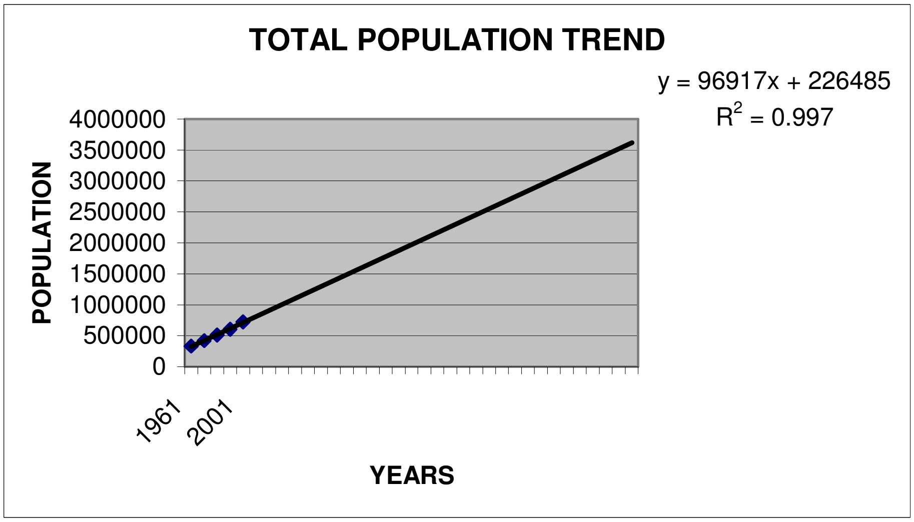


Figure 4a

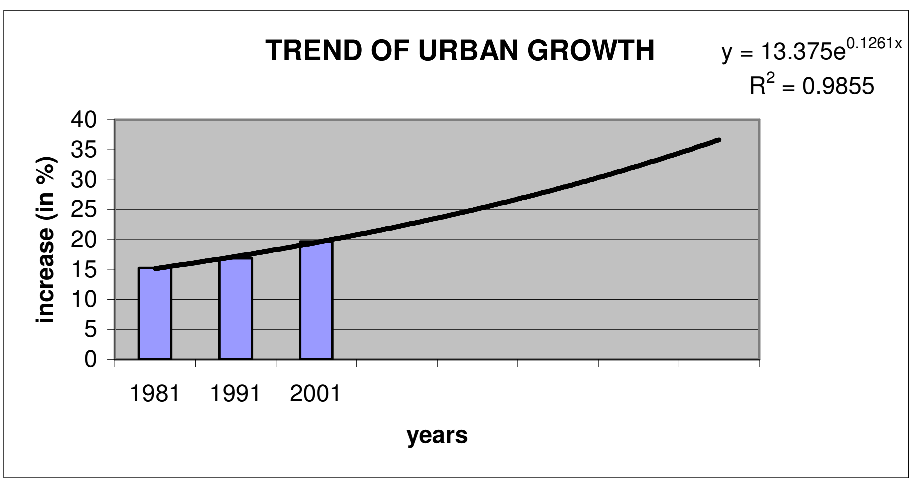


Figure 4b

- *Agriculture*

To get the form of f(p) in equation (11) the total production (table 2) trends (figure 5a, 5b) was observed using Microsoft Excel. The best-fit curve (with maximum regression coefficient (R^2) shown in figures 5a, 5b) was obtained by adding the trend line on the curve. This trend (equation) was used to estimate the production in each coming year. The productivity was calculated by dividing production with the area under agriculture. The productivity trend was also obtained. Here assumption was made that to increase the productivity continuously with same rate is not feasible. Though it can be increased for some time using new technology it is more likely that it will become nearly a constant. Therefore logarithmic curve was thought to be best even though it was not giving maximum regression coefficient. Dividing the estimated total production by estimated productivity the area for each year was estimated. The difference in area obtained from estimated method and calculated from classified data of 1987 is assigned to correction term c_1 in equation (11).
Therefore the form of f(p) and value of c_1 are as below:

$$f(p) = [78514Ln(x) + 154519] / [0.5406Ln(x) + 1.7941] \quad (26)$$

$$c_1 = \text{area from satellite} - \text{estimated area.} \quad (27)$$

years	area(ha) total	poduction(mt) total	productivity prod./area
1976	90714	125852	1.38734925
1981	95388	257757	2.70219524
1983	98001	270492	2.76009428
1986	101999	222135	2.17781547
1987	101215	279260	2.75907721
1991	107465	307286	2.85940539
1993	108063	281816	2.60788614
1996	109793.97	297520	2.70980273
1997	109856	308000	2.80367026
2001	114107.82	380970	3.33868441

Table 2: Total production and productivity

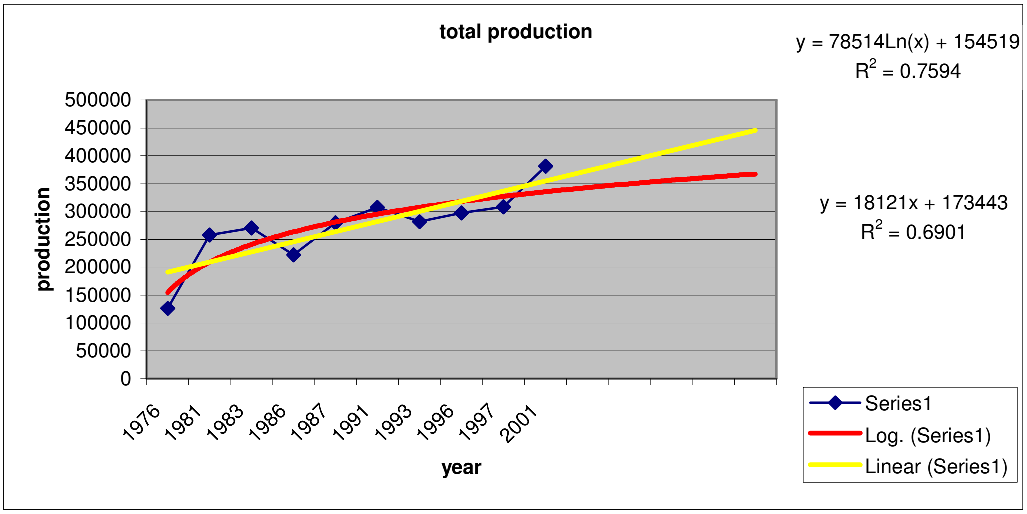


Figure 5a TOTAL PRODUCTION OF FOOD GRAINS (MT)

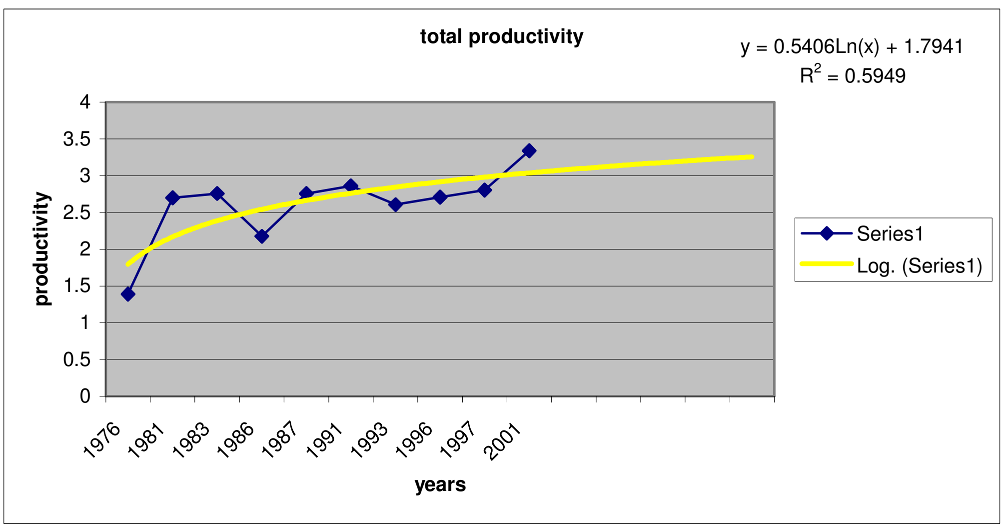


Figure 5b PRODUCTIVITY (PRODUCTION / AREA)

MODEL QUESTIONNAIRE

WEIGHTS

CLASSES → PHYSICAL FACTORS ↓	SETTLEMENT	AGRICULTURE/HORTICULTURE
RAINFALL	4	8
SLOPE	8	2
ASPECT	6	1
ALTITUDE	5	1

CLASSES → PROXIMITY FACTORS ↓	SETTLEMENT	AGRICULTURE/HORTICULTURE
DIST FROM ROAD	**10**	**6**
DIST FROM CITY	**5**	**7**
DIST FROM TOURISTS CENTER	**5**	**1**
DISTANCE FROM INDUSTRY	**3**	**3**

Table 3a

RATINGS

<table>
<tr><th colspan="2">FACTORS</th><th colspan="2">TYPE1</th><th colspan="2">TYPE2</th><th colspan="2">TYPE3</th><th colspan="2">TYPE4</th><th>TYPE5</th></tr>
<tr><td rowspan="2">RAINFALL</td><td>SETTLEMENT</td><td colspan="2">7</td><td colspan="2">8</td><td colspan="2">5</td><td colspan="2"></td><td></td></tr>
<tr><td>AG/HORTICULTURE</td><td colspan="2">1</td><td colspan="2">9</td><td colspan="2">1</td><td colspan="2"></td><td></td></tr>
<tr><td rowspan="2">SLOPE</td><td>SETTLEMENT</td><td colspan="2">10</td><td colspan="2">10</td><td colspan="2">7</td><td colspan="2">5</td><td>1</td></tr>
<tr><td>AG/HORTICULTURE</td><td colspan="2">9</td><td colspan="2">6</td><td colspan="2">1</td><td colspan="2">1</td><td>1</td></tr>
<tr><td rowspan="2">ASPECT</td><td>SETTLEMENT</td><td>2</td><td>7</td><td>8</td><td>10</td><td>8</td><td>7</td><td>4</td><td>2</td><td></td></tr>
<tr><td>AG/HORTICULTURE</td><td colspan="2">4</td><td colspan="2">2</td><td colspan="2">1</td><td colspan="2">2</td><td></td></tr>
<tr><td rowspan="2">ALTITUDE</td><td>SETTLEMENT</td><td colspan="2">10</td><td colspan="2">8</td><td colspan="2">5</td><td colspan="2">2</td><td>1</td></tr>
<tr><td>AG/HORTICULTURE</td><td colspan="2">5</td><td colspan="2">10</td><td colspan="2">3</td><td colspan="2">1</td><td>1</td></tr>
<tr><td rowspan="2">DIST FROM ROAD</td><td>SETTLEMENT</td><td colspan="2">8</td><td colspan="2">5</td><td colspan="2">3</td><td colspan="2">1</td><td></td></tr>
<tr><td>AG/HORTICULTURE</td><td colspan="2">9</td><td colspan="2">7</td><td colspan="2">5</td><td colspan="2">3</td><td>1</td></tr>
<tr><td rowspan="2">DISTANCE FROM MAIN SETTLEMENT</td><td>SETTLEMENT</td><td colspan="2">7</td><td colspan="2">4</td><td colspan="2">3</td><td colspan="2">1</td><td></td></tr>
<tr><td>AG/HORTICULTURE</td><td colspan="2">9</td><td colspan="2">7</td><td colspan="2">5</td><td colspan="2">3</td><td>1</td></tr>
<tr><td rowspan="2">DISTANCE FROM TOURIST CENTER</td><td>SETTLEMENT</td><td colspan="2">5</td><td colspan="2">4</td><td colspan="2">3</td><td colspan="2">2</td><td>1</td></tr>
<tr><td>AG/HORTICULTURE</td><td colspan="2">1</td><td colspan="2">1</td><td colspan="2">1</td><td colspan="2">1</td><td>1</td></tr>
<tr><td rowspan="2">DISTANCE FROM INDUSTRY</td><td>SETTLEMENT</td><td colspan="2">5</td><td colspan="2">4</td><td colspan="2">3</td><td colspan="2">2</td><td>1</td></tr>
<tr><td>AG/HORTICULTURE</td><td colspan="2">8</td><td colspan="2">6</td><td colspan="2">4</td><td colspan="2">1</td><td>1</td></tr>
</table>

DESCRIPTION OF TYPES (IN BRACKETS TYPE1,2,3,4,5 RESPECTIVELY)

RAIN FALL : (<1000 mm, 1000-1400 mm, >1400 mm)
SLOPE : (<15, 15-30, 30-45, 45-60,>60 DEGREE)
ASPECT :(N, NE, E, ES,S, SW, W, WN)
ALTITUDE (<1500, 1500-2000, 2000-2500, 2500-3000, >3000 m)
DIST FROM ROAD (200, 400, 800, 2000 m)
DIST FROM CITY (200, 400, 800, 2000 m)
DIST FROM TOURIST CENTER (200, 400, 800, 2000 m)
DIST FROM INDUSTRY (200, 400, 800, 2000 m)
Table 3b

5.2.3. Calculation of scores

Weight allocation to factors.

To decide the weights to the factors a model questionnaire were prepared and filled with the help of experts (Mr. A.N.Gautam, senior urban town planner, Town and Country planning Department, Shimla, Himachal Pradesh and Mr Sharma deputy director, Directorate of Agriculture, Solan, Himachal Pradesh). This is shown in table[3a]

Weight allocation to neighbors

Weight given to neighbor was according to the distance as discussed in chapter four. The number of neighbors can be maximum eight. The nearest neighbor was given a 2 unit weight. And other is given one unit

Constraints

According to state policy cutting the forest in the Himachal Pradesh is banned. Therefore forest is considered as constraint. Another constraint is considered as settlement as it was assumed while discussion in chapter 4 that once cells have attained the settlement state they cannot change the state.

5.2.4. Transition rule

on the basis of model questionnaire the equation (18) takes the following forms:

for *settlement* the the suitability score [equation (18]

4 * [urb_rain] + 8 * [urb_slop] + 6 * [urb_asp] + 5 * [urb_alt] + 10 * [urb_rd] + 5 * [urb_set] + 5 * [urb_tr] + 3 * [urb_ind] (28)

for *agriculture* the suitability score [equation (18)]

8 * [ag_rain] + 8 * [ag_slop] + 6 * [ag_asp] + 5 * [ag_alt] + 10 * [ag_rd] + 5 * [ag_set] + 5 * [ag_tr] + 3 * [ag_ind] (29)

name of the factor maps are explained in section (Layer preparation)

Substituting the values in equation (18) and (19) the same set of transition rules [equation (20), and(21)] were applied.

5.3. Implementation

5.3.1. Software used:

Microsoft Excel was used to obtain trends as discussed before in this chapter.

ERDAS Imagine-8.5 was used to geo-reference the satellite data, and then classify them to prepare the landuse maps.

ArcView 3.2 was used to digitized the factor maps required such as road map, city centers, tourist centers etc. these maps were obtained from the district plan map.

ArcGIS 8.3 was used for analysis.

AML was used for coding, the Introduction to it given as :

"The ARC Macro Language (AML) is the language used in the ARC environment to program and tailor the application. AML allows one to automate frequently performed actions, create one's own commands, provide startup utilities to help new or inexperienced users perform operations that require specific command settings, and develop menu-driven user interfaces to meet the needs of end users.".Source : ArcDoc

One approximation has been taken in quantity of demand factors that demand factor is linearly dependent area of the region. Since only 53% (approximately) of the area could be classified therefore only 53 % of the magnitude of demand of land was taken.

5.3.2. Layer preparation:

Landsat image had a resolution of 30 meters (m) therefore IRS image was resampled to bring the image on 30 m. the following maps were generated in grid format having the cell size of 30 m.:

- Classified maps for 1987, and 1999.
- Altitude, slope, aspect.
- Major city centers, major tourists centers, major industries, national and state highways.
- Distance maps (from city centers, tourists, industries, and highways).

All these factors were recoded according to their ratings shown in the table (3 b)

[urb_rain] : rainfall data recoded for urban according to rating
[urb_slop] : slope data recoded for urban according to rating
[urb_asp] : aspect data recoded for urban according to rating
[urb_alt] : altitude data recoded for urban according to rating
[urb_rd] : road data recoded for urban according to rating
[urb_set] : settlement data recoded for urban according to rating
[urb_tr] : tourist data recoded for urban according to rating
[urb_ind] : industry data recoded for urban according to rating

[ag_rain] : dem data recoded for agriculture according to rating

[[ag_slop] : slope data recoded for agriculture according to rating
[ag_asp] : aspect data recoded for agriculture according to rating
[ag_alt] : altitude data recoded for agriculture according to rating
[ag_rd] : road data recoded for agriculture according to rating
[ag_set] : settlement data recoded for agriculture according to rating
[ag_tr] : tourist data recoded for agriculture according to rating
[ag_ind] : industry data recoded for agriculture according to rating

for neighborhood influence the 3*3 matrix was taken and with the help of docell operation (grid function available in ArcWorkstation, within AML codes) the weights were allocated as discussed earlier. Here an approximation was taken while calculating the neighborhood influence. Since weights are given on the scale of 10 units and rating was also done on the scale 10 units. So each factor has scale of 100. The score of neighborhood influence [equation (19)], which was in between 0 to 12, was multiplied by 8 before putting the value in equation (15). Therefore, the neighborhood influence on a cell can have the maximum value as 96, whereas the other factors can contribute maximum 100 units to a cell. Assuming this will not have much affect on the result model was run.

The results are shown in chapter 6.

6. Chapter six

6.1. Results

The model was run for different neighborhood size. in the first attempt the moore's neighborhood was taken into account. The weight assigned to them according to their distance with respect to center cell. First simulation was done for the year 1999 and keeping the same neighborhood second simulation was done for year 2011. The results are shown in figure(6.1)

Since the display is not clear therefore the areas of maximum changes in the settlement classes are marked in different colored rectangles. Then these are shown in different figures (figure 6.1.1, figure 6.1.2, figure 6.1.3, figure 6.1.4). In all these mentioned figures, blue color represents additional settlements in year 2011, red color represents the additional settlements in year 1999, and green color represents settlements in 1987. For the year 1999 and year 2011 the settlements are taken from simulation and for 1987 the settlement is taken from the landuse map 1987 prepared from the satellite data of 1987. in both the simulations the maximum number of neighbors are eight (3*3 matrix). The simulated result was compared with the actual land use map of 1999. results were compared visually. The red area shows settlement in year 1999 in landuse map prepared from satellite data IRS-ID (figure 6.4).

The simulation was done for increased neighbors. This time maximum numbers can be 24. (5*5) matrix. Since the display is not clear therefore the areas of maximum changes in the settlement classes are marked in different colored rectangles. Then these are shown in different figures (figure 6.2.1, figure 6.2.2, figure 6.2.3, figure 6.2.4). In all these mentioned figures, blue color represents additional settlements in year 2011, red color represents the additional settlements in year 1999, and green color represents settlements in 1987. For the year 1999 and year 2011 the settlements are taken from simulation and for 1987 the settlement is taken from the landuse map 1987 prepared from the satellite data of 1987

In the next step the effect of different size of neighbors was compared. These results are compared with the result with 3*3 matrix neighbors. For the year 1999 this comparison is shown in figure 6.3.The blue color represents settlement in year 1999 when the neighborhood was taken of 24 cells, but these were not present in the simulated result of the same year 1999 when the neighborhood was taken as eight cells. The red color is represents the opposite.

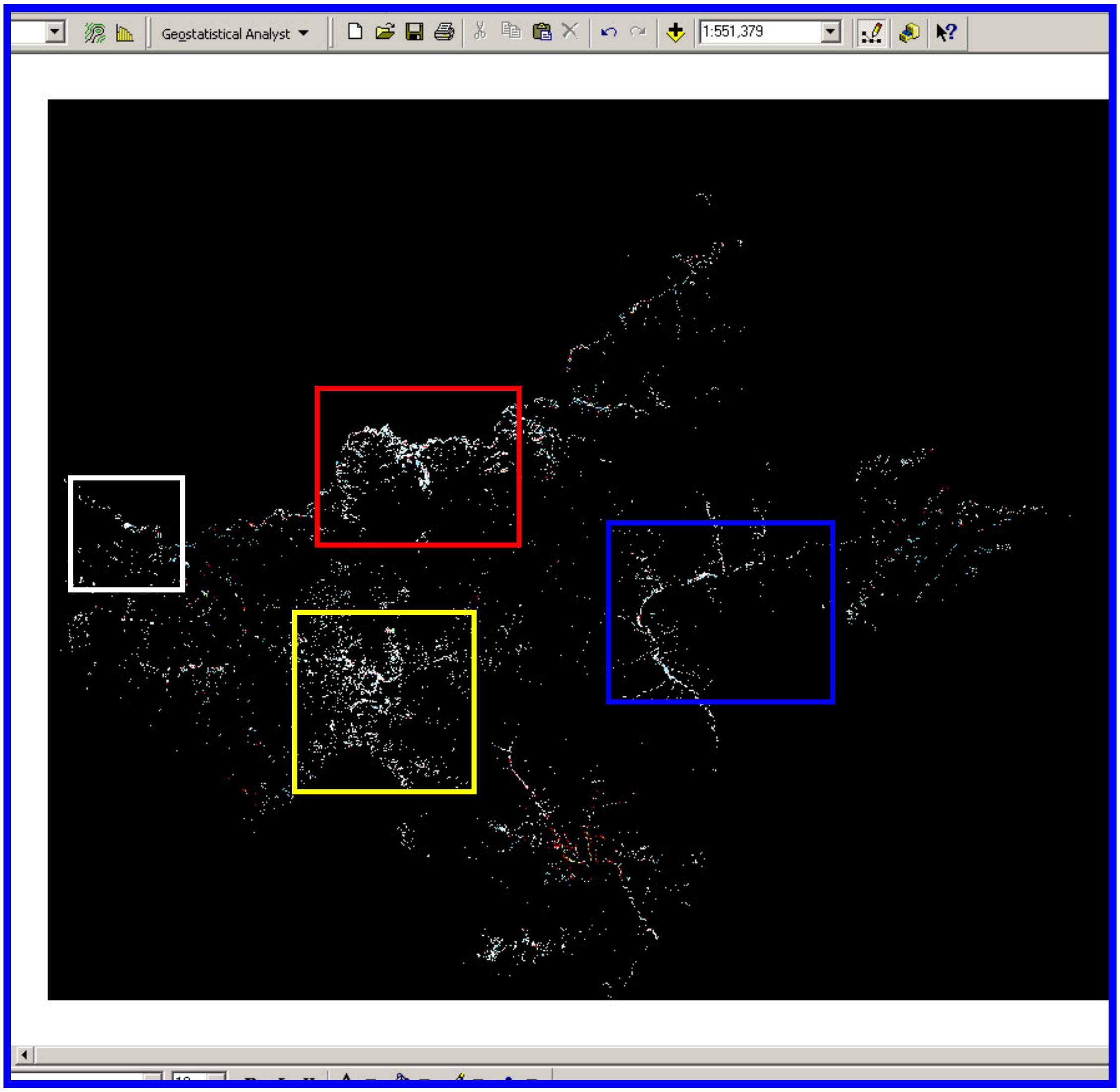


Figure 6.1: Simulated results for the year 1999 and 2011.

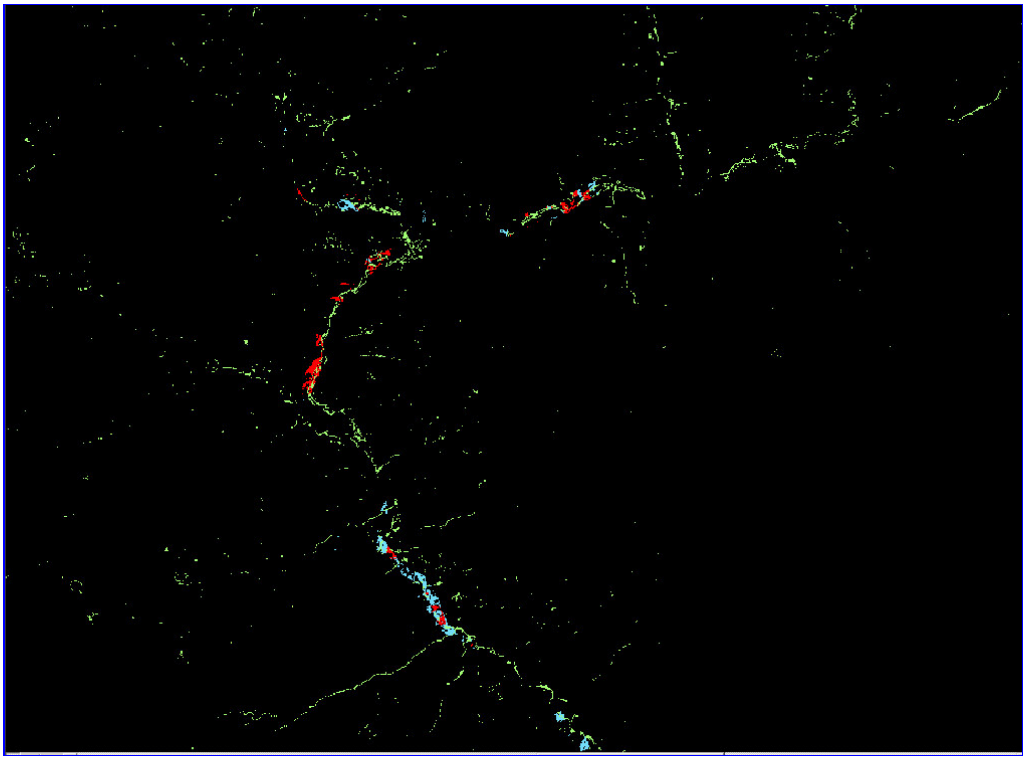


Figure 6.1.1 the area in blue rectangle in figure 6.1

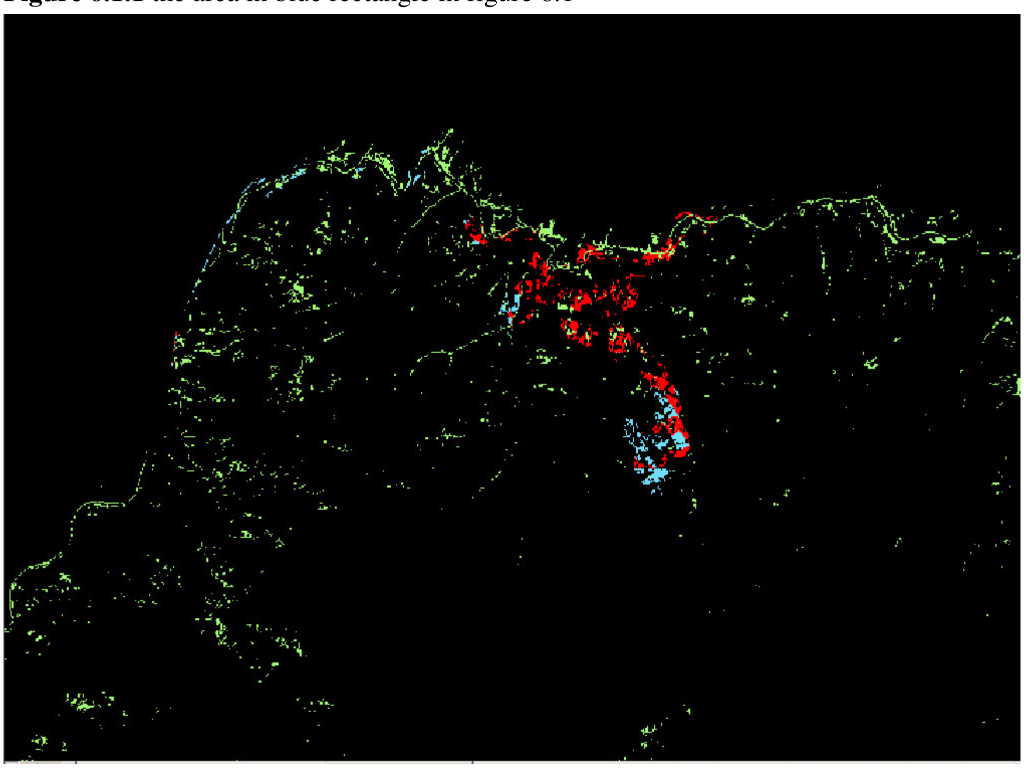


Figure 6.1.2 the area in red rectangle in figure 6.1

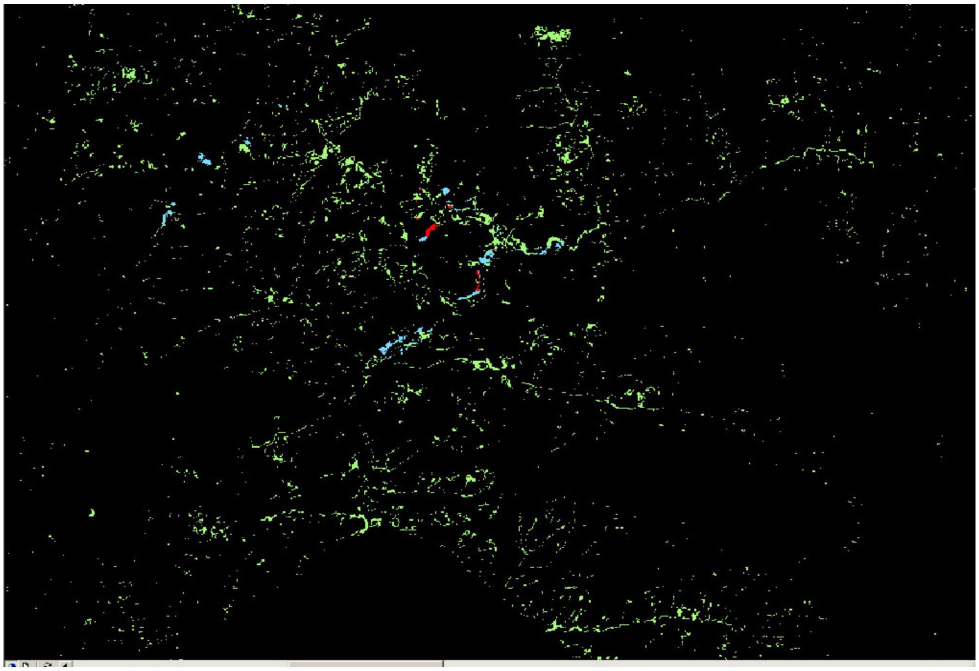


Figure 6.1.3 the area in yellow rectangle in figure 6.1.

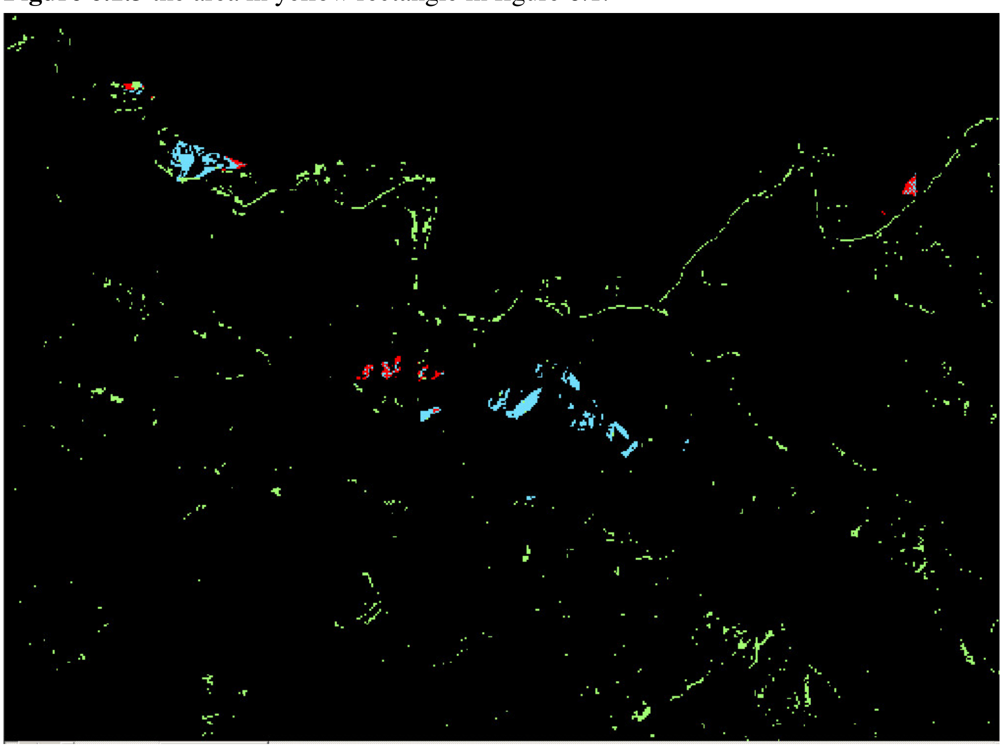


Figure 6.1.4 the area represented in white rectangle in figure 6.1

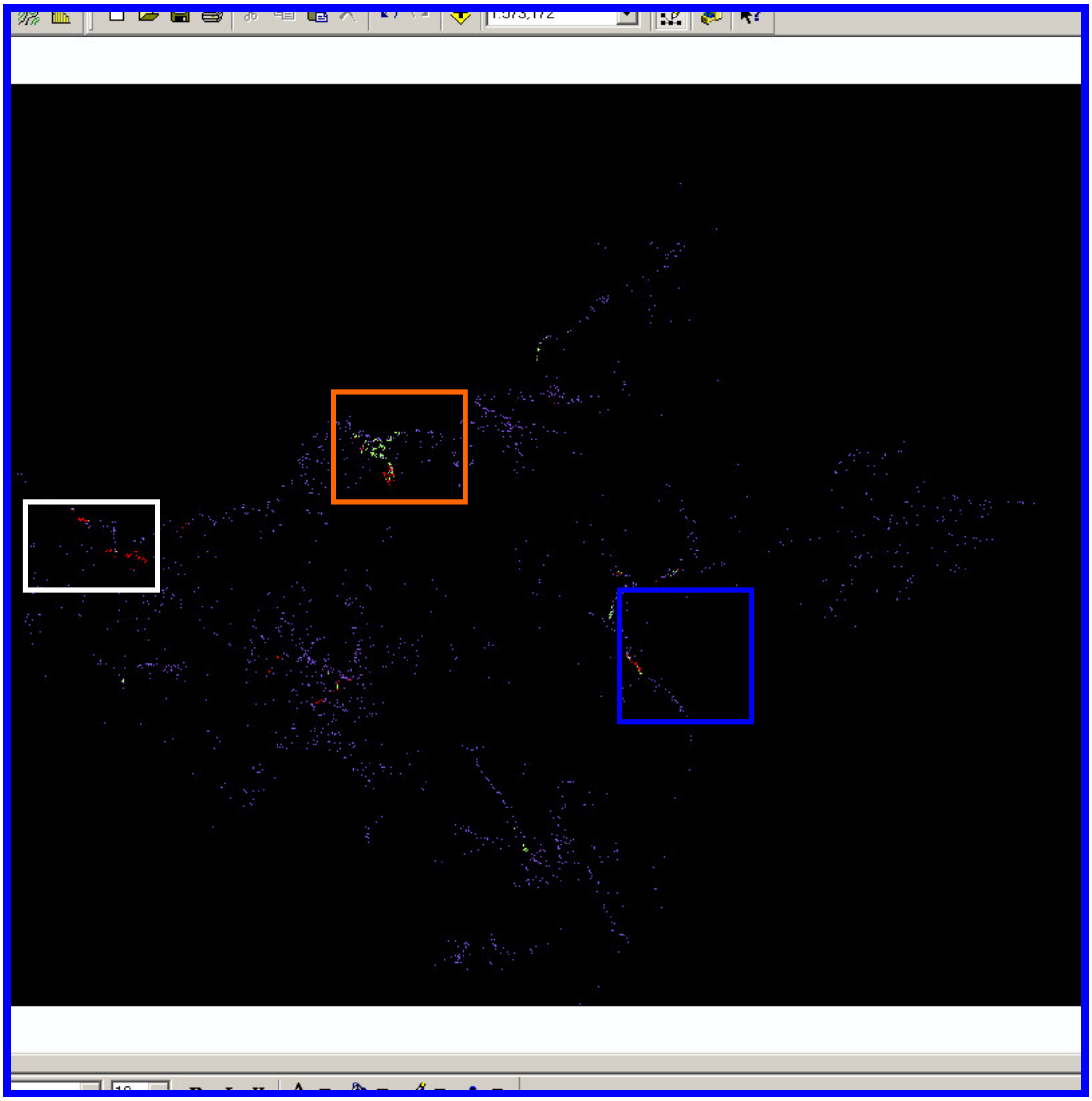


Figure 6.2 simulations for year 1999 while taking 24 cells neighborhood size.

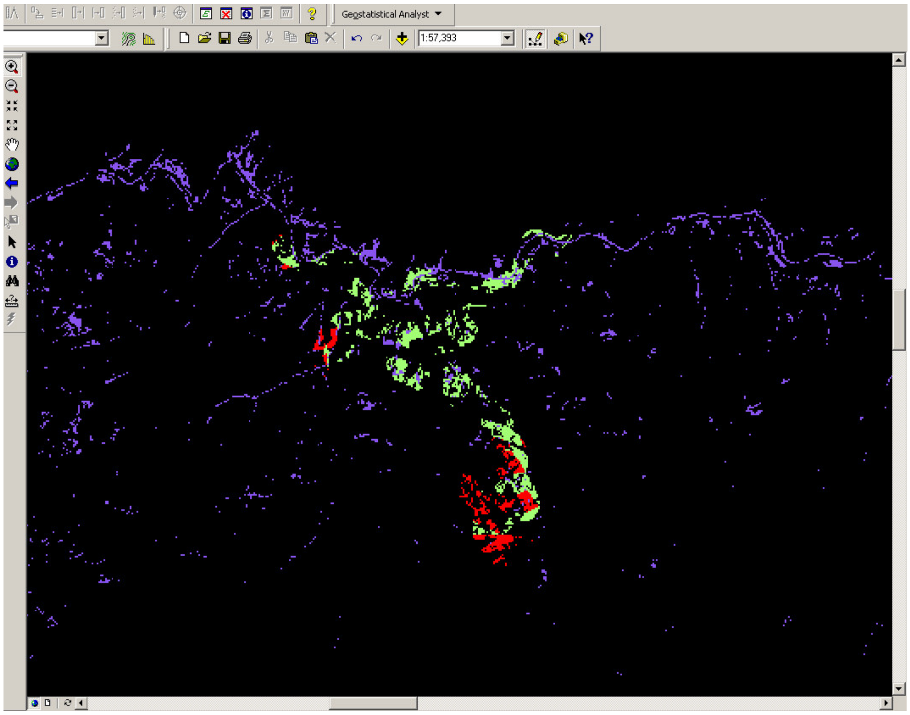


Figure 6.2.1 the are in red rectangle in figure 6.2

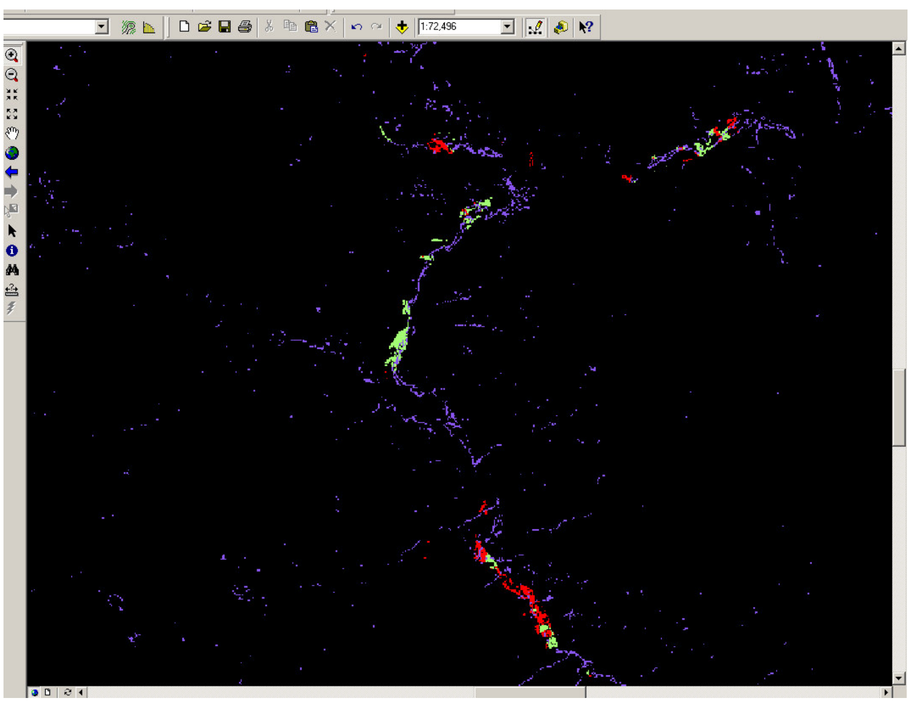


Figure 6.2.2 the area shown in blue rectangle in figure 6.2

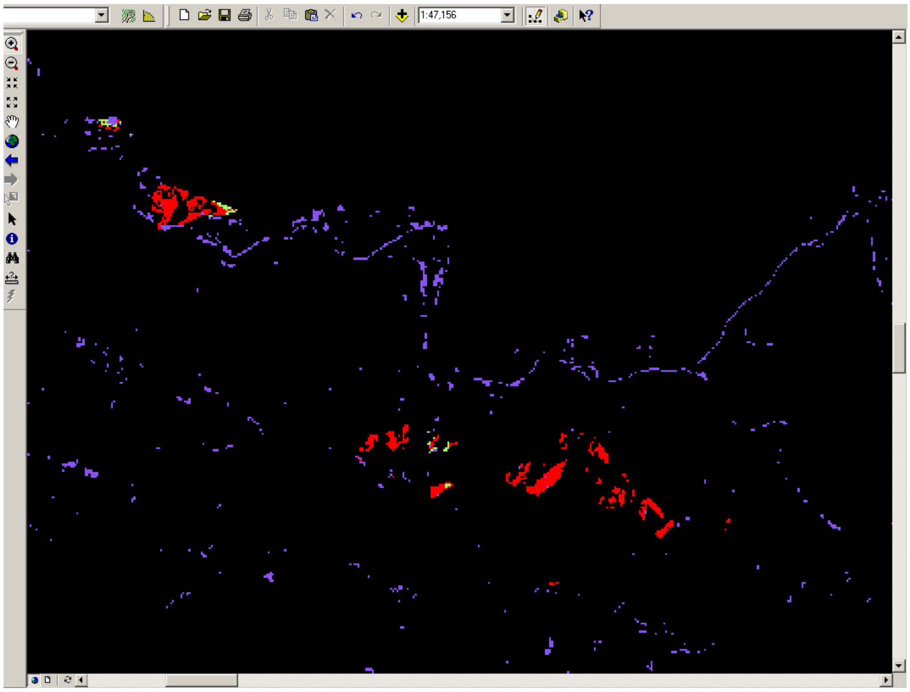


Figure 6.2.3 the area represented in white rectangle in figure 6.2

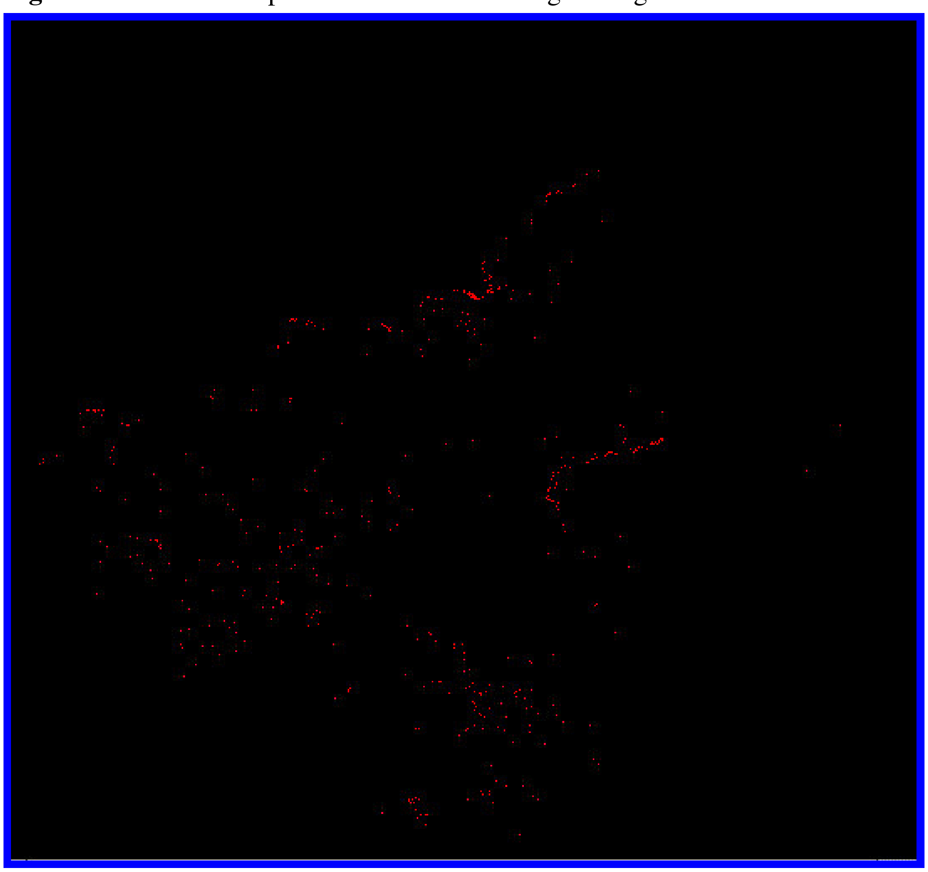


figure 6.4 the red area shows settlement in year 1999 in landuse map prepared fromsatellite data IRS-ID.

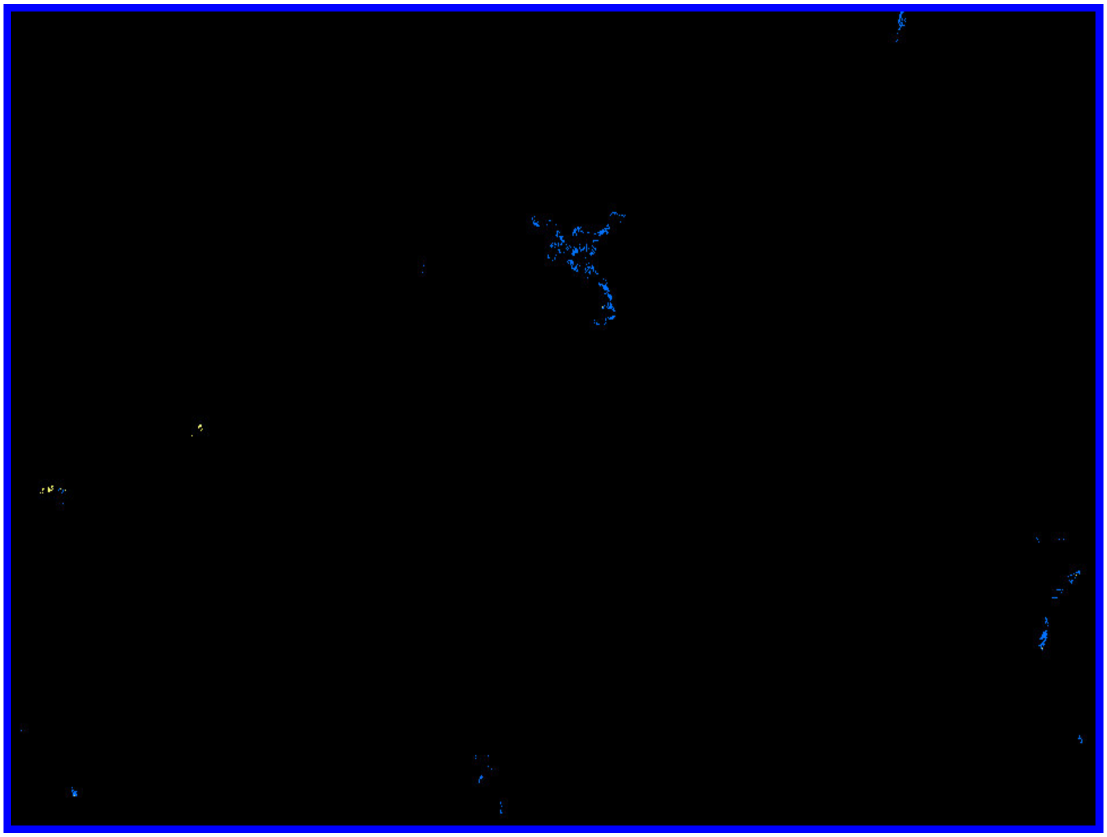


Figure 6.3 the effect of change in the neighbourhood size in the year 1999.

6.2. Discussions and conclusions

There are different kinds of processes, which are responsible for change in land use land cover. such as Road gravity influence, organic growth, breeding etc. In some of the region all of them can be present and in some of the region some may be not present. This is very much dependent on the area taken under case study. The causalities may be socio-economic like population growth, or economic growth, or state policy etc. Shimla is observing population growth and being the capital of the state it is achieving high economic growth. The state policy for not cutting the forests is also playing major role in land use land cover change in the region. Since this is hilly area and there is scarcity of flat lands, therefore changes are occurring vertically more. But this issue is addressed while calculating the demand for land as, for each family the standard amount of land is taken.

The model is highly dependent on the spatial variables taken in to consideration. In this model the spatial and non-spatial variables were taken with the help of literature and experts of that region as discussed in earlier chapters. More variables can increase accuracy of the simulations. Some variables like drainage and river should also have been considered as they may improve the results. Since the river in this hilly region was narrow it was not easy to classify with the satellite data used. It would have taken much time to prepare the drainage layer as they have to be identified according to their importance for giving weights. As time was short drainages were not taken into consideration.

The relationship between the causalities and changes in the land use land cover is established in the chapter four. Model is only abstraction of the reality. Each model has to work under some approximations and assumptions. In this model also, assumptions and approximations were made. As far as possible they were kept in the limits of logic and ground realities. This model is based on the trends, which were derived with the help of 20 to 30 years of the data. Therefore it has a strong basis for generating the land requirements for the future. To minimize the error the correction term was introduced into formulation. This can be used for calibration. Although in this model due to time limitation it could not be explored much but this can be tried for the future. While calculating the demand for the population the inputs can be enrich if this population is divided into further categories such as according to high class, middle class and low class and considering the standard requirement land for each class separately. Similarly for calculating the agriculture land requirement export- import, labour availability, etc terms also can be taken into consideration which will improve the simulated scenario.

The transition rules were based on the score achieved by cells. These cells were calculated using linear weighted summation method of Multi- Criteria Evaluation technique. Although this technique is very useful and simple, it is very much based on the expert's opinion. Experts can have different opinions and to arrive at the unanimous decision is not simple. For refinement and maintaining consistency in the preferences given by the experts AHP method can be more useful. While calculating the scores some time more cells were equally potential than required. MCE could not resolve this situation and therefore some more is to be done to resolve this conflict.

As shown in the results earlier the different size of neighbour hood has an impact of the simulated results. But how much neighbourhood should be considered in the model is unanswered so far. It appears that the neighbourhood influence is very much region and time specific. Depending on the region this can be taken.

Any model's accuracy is highly dependent on the accuracy of input data and assumptions made. While calculating the demand for the land, it was observed that there was difference in the area observed in the satellite and the area calculated by formula which were conceptualised in chapter four and used in case study (chapter five). This difference has been taken as a correction factor and assumed that it will be fixed in all years. This type of correction can be very useful and instead of keeping it as a constant one can use it as varying quantity. As with the time the development will be according to plan more, the correction term while estimating demands will be less with the time The case study has used the satellite data, which has moderate resolution. By using the high-resolution data or other techniques the classified inputs can be improved.

This model is basically for answering 'what –if' scenario. The simulations have been done on the assumptions that 'if these trends are continued then….'. similarly one can ask 'what will happen if….'.

The present model can address some more kinds of processes and influences. For example, it can address the influence of border between two states. On the ground it has been observed that border can influence positively as well as negatively, depending upon the state policies. Therefore a distance map can be prepared for this and this can be considered as one of the factor while calculating the scores of the cells.

The cells near the boundary of the study area could not get full neighborhood influence. Although in basic cellular automata there are two ways with which this is tackled, one to assume the cells are joined with the ends, second assuming the cells has neighbors of same state (class in the present context) at the perimeter. This may introduce some inaccuracy in the results. To avoid this problem an extra area should be included under the study area.

Although a cell size was dependent upon the spatial resolution of the satellite data, but this can be used for higher resolution of data. It can be a case of study if one can establish any standard ratio between the size of study area and cell size. How much error introduces because of the cell size? and at what stages the error comes into the models and how this error propagates in the whole process? are some very interesting issues which can be further investigated.

Can cell have socio economic attributes also?. For example population density. If yes how can it be used in the algorithm is one more issue which can be investigated further.

This model is flexible in the sense that it can accommodate change the forms of the causalities. Depending upon the region under study the different form of causalities can be adopted. For an example in some region the population growth can be exponential and in other region this may be linear equation. Therefore it may be easily adapted to other regions also.

Cells were allocated to diffeent classes according to their scores. The situation was arrived when required cells were less than the total number of cells of equally suitable scores. To resolve this problem those cells were randomly picked up. This introduces some uncertainty in the model. This can be reduced to some extent by putting some more preferences before allocation of the cells to the classes for example, first, near to city centres the allocation can be considered, and then near to road, or if cell is near to more than one such factors those cells should be allowed to undergo transition first.

While working with the grid in ArcWorkstation and developing an algorithm for selecting cells randomly among equally potentially cells the difficulty is faced. When the agriculture class was taken, due to large number of cells were to be selected. The value attribute table (VAT) was not able to generate for large grids. Therefore, to handle large data set more robust programming is needed. One approach can be to divide the whole grid into equal size of grids and then select equal number of cells from each grid randomly.

The breeding process can also be addressed in this model with some modifications and with more programming experience. In the model one iteration is considered as one year. After each year one convolution (matrix of some specified rows and columns) can be run over the simulated result before it could go for second iteration. With this mask it can be checked that at what places it has all the cells of settlement class. If such cells are found then the distance from this cell should be calculated and near to this cell should assign some weightage.

Here neighborhood influence was considered as one among other factors. Therefore the importance (weights) to it given was nearly equal, it would be interesting to see how much sensitive the results are if this is changed (either increased or decreased in comparison to other factors).

Despite the fact that various issues, as discussed earlier, are still to be addressed this model shows promising abilities to incorporate many growth processes with ease and efficiency.

References:

Agarwal, C., G. M. Green, J. M. Grove, T. P. Evans and C. M. Schweik (2002). "A review and assessment of land-use change models: dynamics of space, time, and human choice."

Almeida, C. M. d., M. Batty, A. M. V. Monteiro, G. Camara, B. S. Soares-Filho, G. C. Cerqueira and C. L. Pennachin (2003). "Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation." Computers, Environment and Urban Systems **In Press, Corrected Proof**.

Baker, W. L. (1989). "A review of models of landscape change." Landscape Ecology **2**(2): 111-133.

Batty, M. (2000). "Editorial:Less is more, more is different: complexity, morphology, cities, and emergence." Environment and Planning B **27**: 1-2.

Batty, M. and Y. Xie (1994). "From cells to cities." Environment and Planning B **21**: s41-s48.

Batty, M. and Y. Xie (1997). "Possible urban automata." Environment and Planning B **24**: 175-192.

Batty, M., Y. Xie and Z. Sun (1999). "The Dynamics of Urban Sprawl." CENTRE FOR ADVANCED SPATIAL ANALYSIS Working Paper Series 15.

Benders, R. M. J. (1996). Models and modelling. INTERACTIVE SIMULATION OF ELECTRICITY DEMAND AND PRODUCTION. Groningen.NL, University Groningen.

Chen, J., P. Gong, C. He, W. Luo, M. Tamura and P. Shi (2002). "Assesment of Urban Development Plan of Beijing by Using a CA-Based Urban Growth Model." Photogrammetric Engineering & Remote Sensing **68**(10): 1063-1071.

Cheng, J. (2003a). Understanding Spatial and Temporal Processes of Urban Growth. *Modelling Spatial and Temporal Urban Growth*. J. Cheng. Enschede.NL, ITC.

Clarke, K. C. and L. J. Gaydos (1998). "Loose-coupling a cellular automata model and GIS:long-term urban growth prediction for San Francisco and Washington/Baltimore." int. j. geographical information science **12**(7): 699-714.

Clarke, K. C., S. Hoppen and L.J.Gaydos (1997). "A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area." Environment and Planning B **24**: 247-261.

Couclelis, H. (1985). "Cellular worlds: a framework for modelling micro-macro dynamics." Environment and Planning A **17**: 585-596.

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

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

Li, X. and A. G. O. Yeh (2001a). "Calibration of cellular automata by using neural networks for the simulation of complex urban systems." Environment and Planning A **33**: 1445 -1462.

Li, X. and A. G. O. Yeh (2002a). "Neural-network-based cellular automata for simulating multiple land use changes using GIS." int. j. geographical information science **16**(4): 323-343.

Rinaldi, E. (1998). THE MULTI – CELLULAR AUTOMATON:A TOOL TO BUILD MORE SOPHISTICATED MODELS.

A THEORETICAL FOUNDATION AND A PRACTICAL IMPLEMENTATION. Venezia Italia, Laboratorio sulla Simulazione.

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

Takayama , M. and H. Couclelis (1997). "Map Dynamics Integrating Cellular Automata and GIS Through Geo-Algebra." International Journal of Geographical Information Science **11**(1): 73-91.

Takeyama, M. and H. Couclelis (1997). "Map dynamics: integrating cellular automata and GIS through Geo-Algebra." int. j. geographical information science **11**(1): 73-91.

Torrens, P. M. and D. O'Sullivan (2001). "Editorial:Cellular automata and urban simulation: where do we go from here?" Environment and Planning B **28**: 163-168.

Wagner, D. F. (1997). "Cellular automata and geographic information systems." Environment and Planning B **24**(2): 219-234.

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

White, R. and G. Engelen (1994). "Urban Systems Dynamics and Cellular Automata:Fractal Structures between Order and Chaos." Chaos, Solitons & Fractals **4**(4): 563-583.

White, R. and G. Engelen (1997). "Cellular automata as the basis of integrated dynamic regional modelling." Environment and Planning B **24**(2): 235 - 246.

White, R. and G. Engelen (2000). "High resolution intergated modelling of the spatial dynamics of urban and regional systems." Computers, Environment and Urban Systems **24**: 383-440.

White, R., G. Engelen and I. Uljee (1997). "The use of constrained cellular automata for high-resolution modeling of urban land use dynamics." Environment and Planning B **24**: 323-343.

Wu, F. (1996). "A linguistic cellular automata simulation approach for sustainable land development in a fast growing region." Computers, Environment and Urban Systems **20**(6): 367-387.

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

Wu, F. (1998c). "SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules." int. j. geographical information science **12**(1): 63-82.

Wu, F. (2000). *A Parameterised Urban Cellular Model Combining Spontaneous and Self-Organising Growth*. Geocomputation: Innovation in GIS 7. P. A. a. D. Martin. London, Taylor & Francis: London**:** pp. 73-85.

Wu, F. (2002). "Calibration of stochastic cellular automata: the application to rural-urban land conversions." int. j. geographical information science **16**(8): 795–818.