Masters Program in Geospatial Technologies

URBAN LAND USE AND LAND COVER CHANGE ANALYSIS AND MODELING A CASE STUDY AREA MALATYA, TURKEY

GÜLENDAM BAYSAL

Dissertation submitted in partial fulfillment of the requirements for the Degree of Master of Science in Geospatial Technologies

URBAN LAND USE AND LAND COVER CHANGE ANALYSIS AND MODELING A CASE STUDY AREA MALATYA, TURKEY

Master Thesis

by

Gülendam Baysal

Dissertation Supervised by:

Prof. Dr. Edzer J. Pebesma, Institute for Geoinformatics (IFGI), Westfälische Wilhelms-Universität, Münster, Germany.

Dissertation Co-Supervised by:

Prof. Dr. Jorge Mateu M., Department of Mathematics, Universitat Jaume I (UJI), Castellon, Spain.

Prof. Dr. Pedro Cabral, Instituto Superior de Estatística e Gestão de Informação (ISEGI), Universidade Nova de Lisboa, Lisbon, Portugal.

February 2013

DECLARATION OF ORIGINALITY

I declare that the submitted work is entirely my own and not belongs to any other person. All references, including citation of published and unpublished sources have been appropriately acknowledged in the work. I further declare that the work has not been submitted for the purpose of academic examination, either in its original or similar form, anywhere else.

Münster, 28th February, 2013

Gülendam Baysal

………………...

ABSTRACT

This research was conducted to analyze the land use and land cover changes and to model the changes for the case study area Malatya, Turkey. The first step of the study was acquisition of multi temporal data in order to detect the changes over the time. For this purpose satellite images (Landsat 1990-2000-2010) have been used. In order to acquire data from satellite images object oriented image classification method have been used. To observe the success of the classification accuracy assessment has been done by comparing the control points with the classification results and measured with kappa. According to results of accuracy assessment the overall kappa value found around 75%. The second step was to perform the suitability analysis for the urban category to use in modeling process and it has been done using the Multi Criteria Evaluation method. The third step was to observe the changes between the defined years in the study area. In order to observe the changes land use/cover maps belongs to different years compared with cross tabulation and overlay methods, according to the results it has been observed that the main changes in the study area were the transformation of agricultural lands and orchards to urban areas. Every ten years around 1000ha area of agricultural land and orchards were transformed to urban. After detecting the changes in the study area simulation for the future has been performed. For the simulation two different methods have been used which are; the combination of Cellular Automata and Markov Chain methods and the combination of Multilayer Perceptron and Markov Chain methods with the support of the suitability analysis. In order to validate the models; both of them has been used to simulate the year 2010 land categories using the 1990 and 2000 data. Simulation results compared with the existing 2010 map for the accuracy assessment (validation). For accuracy assessment the quantity and allocation based disagreements and location and quantity based kappa agreements has been calculated. According to the results it has been observed that the combination of Multilayer Perceptron and Markov Chain methods had a higher accuracy in overall, so that this combination used for predicting the year 2020 land categories in the study area. According to the result of simulation it has been found that; the urban area would increase 1575ha in total and ~936ha of agricultural lands and orchards would be transformed to the urban area if the existing trend continued.

Key words: Land Use, Land Cover, Remote sensing, GIS, Cellular Automata, Markov Chain, Segmentation, Supervised Classification, Multilayer Perceptron.

ACKNOWLEDGEMENT

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. First of all I want to thank to my family for their moral support, encouragement and care.

I also want to thank my supervisors Prof. Dr. Edzer J. Pebesma, Prof. Dr. Jorge Mateu M., Prof. Dr. Pedro Cabral for their comments and suggestions. Moreover I would thank to Federal Ministry of Innovation, Science and Research of North Rhine-Westphalia for funding my study.

Special thanks to Firat Development Agency for providing the data and special thanks to my friends Dilşat Kazazoğlu Temiz, expert at Firat Development Agency, and Özge Öztürk, GIS specialist at Ministry of Agriculture, for assisting me while collecting the data for my thesis and for their moral support.

Special thanks to Dr. Christoph Brox, Prof. Dr. Werner Kuhn, Prof. Dr. Christian Kray for their comments during my studies on my thesis. Furthermore I want to thank to Professor Robert Gilmore Pontius Jr., Clark University, for his suggestions and sharing the useful sources.

Finally many thanks to my classmates from all over the world for sharing their knowledge and for their moral support.

ACRONYMS

ANN: Artificial Neural Network

CA: Cellular Automata

ED: European Datum

ETM: Enhanced Thematic Mapper

FDA: Firat Development Agency

LULC: Land Use Land Cover

MARKOV: Markov Chain Model

MCE: Multi Criteria Evaluation

MLP: Multilayer Perceptron

TM: Thematic Mapper

USGS: U.S Geological Survey

UTM: Universal Transverse Mercator

WGS: World Geodetic System

WLC: Weighted Linear Combination

TurkStat: Turkish Statistical Institute

INDEX OF CONTENT

INDEX OF FIGURES

INDEX OF TABLES

1 INTRODUCTION

1.1 Background and Motivation

Land use and land cover changes are dynamic spatial issues and in order to have a sustainable development these changes need to balanced. In many cities this balance spoiled because of rapid population growth in urban areas. The main increase in urban population results by rural to urban migration. In Turkey mass movements from rural to urban has increased after the 1960s and especially after 1990s the increase rate gained a momentum, from figure 1-1 urban and rural population growth rate can be observed.

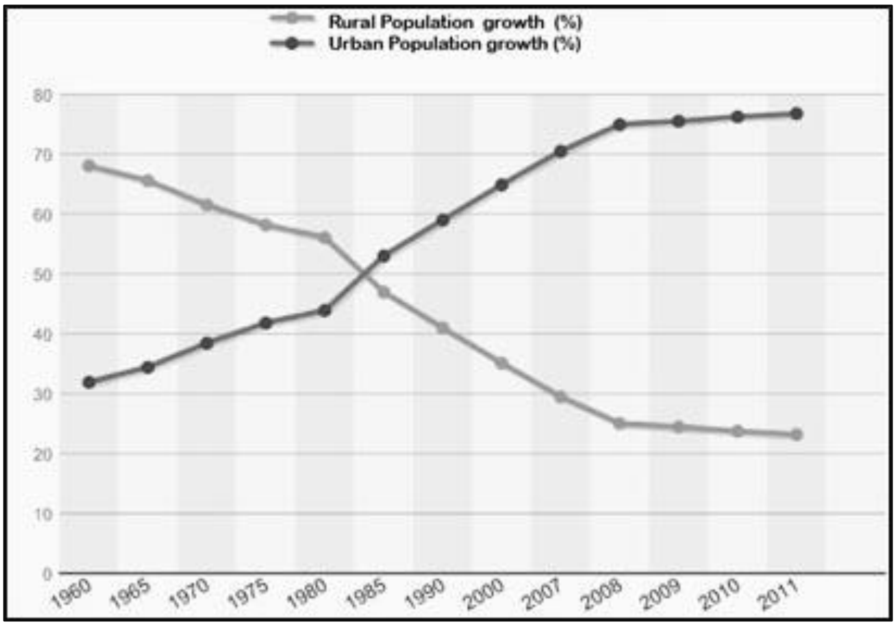

Figure 1-1: Urban and rural population growth rate (%) for Turkey (TurkStat, n.d.)

The growth in the urban population caused the transformation of the agricultural lands on the outskirts of the cities to the urban area. For a sustainable development the transformations need to be balanced, but for many cities in Turkey it was not balanced. Controlling the changes can be managed by planning authorities in order to lead them many researches worked on the land change modeling issues.

1.2 Study Area

Malatya is located at 38°21′N 38°18′E on the East Anatolian Region of Turkey, the city is best known for its apricot orchards. The location of the city is important because it is on the trade ways from east to west (figure 1-2). The population characteristic of Turkey and the Malatya can be observed in the table 1-1 and 1-2. As we can see from the tables the urban population in the study area tripled in 46 years whereas the rural population decreased %50 percent. As a result of this population increase a demand for housing increased.

Table 1-1: Urban and rural populations in 1965 (TurkStat, n.d.)

Year 1965	Total	Urban	Rural
Turkey	31,391,421	10,805,817	20,585,604
Malatya	452,624	147,040	305,584

Table 1-2: Urban and rural populations in 2011 (TurkStat, n.d.)

Year 2011	Total	Urban	Rural
Turkey	74,724,269	57,385,706	17,338,563
Malatya	740,643	480,144	260,499

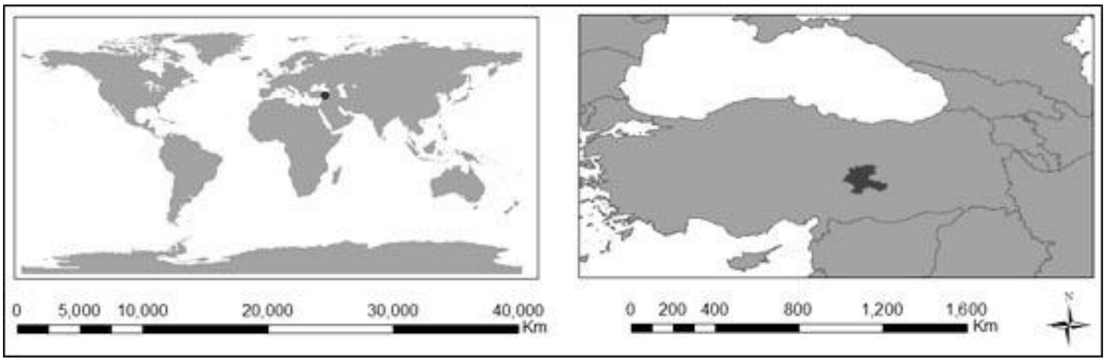

Figure 1-2: Location of study area

1.3 Statement of the Problem

After 1965 the study area started to gain migration because of having more job opportunities and services. Until 1990s this movement was not critical but after that the population of the city increased rapidly, because it became a pole of attraction for the surrounding regions (Ünal, 2010). However the city was not ready for this migration; there was not enough housing for the new population. In the beginning people started to settle down on the outskirts of the city, in farmhouses which are not affecting the agricultural lands. Nevertheless in recent years this has changed and instead of farm houses 7-10 storied apartments raised. The main problem of the study area is unbalanced growth in urban area which means the growth urban area affected the agricultural lands on the outskirts of the city. Since the city has important agricultural sources like apricot orchards in the periphery, this transformation need to be monitored.

1.4 Objectives

The main aim of the study is to understand the urban area changes and the transformations from other land categories to urban area and the main objective of the study using geospatial tools and techniques for monitoring the changes and then modeling the future land classes. According to this aims and objectives we can list main steps of the study as;

- Creating LULC maps from satellite images by image classification techniques (1990-2000-2010).
- Comparing the results of classification with reliable sources (CORINE, STATIP and Google Earth) for the accuracy assessment.
- Using the LULC categories for the years 1990-2000 and 2000-2010 to detect the changes in the study area.
- Using the LULC categories for the years 1990-2000 to simulate the year 2010 in order to validate the model.
- If the model is useful for the study area the simulating the LULC categories for the year 2020.

1.5 Research Questions

- Are the available data adequate for this study?
- Are GIS and remote sensing tools adequate for this study?
- Can the mathematical models help to model LULC changes?
- Where are the main changes and where is the main growth direction of the urban area according to simulated result for the year 2020?

1.6 Theoretical Background and Basic Terminologies

In order to understand the LULC changes it is important to know about the tools and models for change detection and simulation. GIS and remote sensing tools and mathematical models are the main components of the change detection and simulation. In this section we try to summarize the main concepts mentioned in the study.

"***Remote sensing is*** *the acquisition of information about an object or phenomenon without making physical contact with the object.'*"" (Schowengerdt, 2006). Remote sensing is the primary sourcing of multispectral and multi temporal data which will help us to detect the changes.

"***Geographic Information System*** *(GIS) is a system designed to capture, store, manipulate, analyze, manage, and present all types of geographical data*" Geographic information system (n.d.).The power of GIS could help us to process the data and make analysis on the data.

"***A mathematical model*** *is a description of a system using mathematical concepts and languages"* (Mathematical Model, n.d.). The mathematical models help us to understand the system and then by understanding the system behaviors it helps us to predict the future.

Land use is related how the land is used, how the natural environment changes to human built up area. **Land cover** is physical cover of the surface of the earth not related human activities it is natural. (García, Feliú, Esteve, Soba, Hazeu, Rasmussen, Galera-Limdblom & Banski 2010)

1.7 Previous Studies

Many researchers worked on LULC change modeling, with different methodologies and techniques. The table 1-3 prepared in order to summarize some of the works and models which can be implemented in LULC change modeling. Every model has different data requirements, strengths and weaknesses. For this research mainly two approach has been implemented which are combination of Cellular Automata and Markov Chain methods and the combination of Multilayer Perceptron and Markov Chain methods, because of the data availability and the aim of the research.

1.8 Structure of Thesis

The thesis divided into seven chapters in order to explain the each step of the study in detail. In the first chapter the background of the research, study area, the statement of the problem, objectives, research questions, previous studies explained and the main terminologies defined. The second chapter is including the methodology and the data used for the study. Third chapter explains the first phase of the study which is image classification. The fourth chapter deals with the suitability analysis for the urban area which has been used in the modeling part of the study. Fifth chapter explains the changes between the years 1990 -2000 and 2000-2010 in order to investigate the major changes in the study area. Sixth chapter explains the LULC modeling with two different models which are combination of CA and MARKOV methods and the combination of the MLP and MARKOV methods moreover it includes the model validation step and with the accuracy assessment and simulation of the LULC for the year 2020.

Table 1-3: Literature review on the LULC change models

	Model Name	**Builder**	**Model Type**	**What it Explains**	**Variables**	**Strengths**	**Weakness**
(Agarwal et al., 2002)	Markov Model	Wood et al. 1997	Spatial Markov model	Landuse change	Multi Temporal Land Use/ cover maps	Considers both spatial and temporal change	No sense of Geography
(Agarwal et al., 2002)	CA	Clarke et al. 1998; Kirtland et al. 2000	Cellular Automata model	Change in urban areas over time	Extent of urban areas, Elevation, Slope, Roads	Allows each cell to act independently according to rules	Doesn't include human and biological factors
(Adhikari & Southworth, 2012)	Combination of CA and MARKOV methods	Clark Labs	Spatio-Temporal dynamic modeling	Predicts land use/cover in the future	Multi Temporal Land Use/ cover maps, Suitability maps	Creating the Data is easy, CA add spatial dimension to the model, can simulate change among several categories	Socio economic factors are not considered , Calibarting the model with MCE is too much time consuming compared to other methods
(Agarwal et al., 2002)	UrbanSim	Paul Waddell (University of California, Berkeley)	Cellular Automata and individual based model	Spatial maps of housing units by pixel, nonresidential square footage per cell and other economic and demographic characteristics	Parcel files, business establishment files census micro data, Environmental, political, and planning boundaries, location grid control totals from economic regional forecasts, travel access indicators, scenario policy assumptions	Structure allows multiple types of policies to be explored High degree of precision, Employment locations modeled, Designed to provide inputs to the transportation demand model	High data demands, designed for urban areas hard to understand the model, It has rigid model structure, Output must be imported into GIS for viewing
(Li & Yeh, 2002)	Combination of ANN and CA methods	Antony Gar-On Yeh, Xia Li	ANN & Cellular Automata Model	Predicts land selected land class in the future	Multi Temporal Land Use/ cover maps	Calibrating the model with ANN	It can simulate change only in two category
(Pontius & Chen, 2008)	GEOMOD	Clark Labs	Cellular Automata	Predicts land selected land class in the future	Land use/cover map	Need only one time land use map for calibration	It can simulate change only in two category
(Torrens, 2000) (Agarwal et al., 2002)	SLEUTH	Dr. Keith C. Clarke at UCSanta Barbara	Spatially explicit Cellular Automata model	GIS maps of probability (continuous) of urbanization in a specified pixel	Multi Temporal Land Use Map, Impervious surface cover , Road networks (for each time period), Slope (%), Undevelopable land	Relatively easy to transfer among regions, incorporate many different land use classifications systems ,it generates continuous measure of density of development , take into consideration the future developments (such as road)	Designed for urban settings, Data demands are high, Un-calibrated model would produce more error ,Difficult to use
	Land Change Modeler	Clark Labs	Markov Chain , MLP, Logistic Regression , SimWeight	Change Analysis , Predicts land use/cover in the future	Land Use Land Cover data, Road ,DEM, Other Infrastructure	Environmental modeling platform, taking into consideration the future projects, Using the ANN for development of transition potentials, calculating the changes in two time periods	Consideration of one sub model
(Agarwal et al., 2002)	CLUE (Conversion of Land Use and Its Effects)	(Veldkamp and Fresco 1996a)	Discrete, finite state model	Predicts land use/cover in the future	Land suitability for crops, Temperature/Precipitation, Effects of past land use, Impact of pests, weeds, diseases, Human Drivers, Population size and density, Technology level, Level of affluence, Political Structures, Economic conditions, Attitudes and values	Covers a wide range of biophysical and human drivers at differing temporal and spatial scales	Limited consideration of institutional and economic variables
(Agarwal et al., 2002)	LUCAS (Landuse Change Analysis System)	Michael Berry, Richard Flamm, Brett Hazen, Rhonda MacIntyre, and Karen Minser; University of Tennessee	Spatial stochastic model	Transition probability matrix, landscape change. Assesses the, impact on species habitat.	Land cover type, Slope, Aspect, Elevation, Land ownership, Population Density, Distance to nearest road, Distance to nearest economic market, center, Age of trees	Model shows process , output (new land use map), and impact (on species habitat)	*LUCAS tended to fragment the landscape for low proportion land uses, due to the pixel based independent grid method. Patch based simulation would cause less fragmentation, but patch definition requirements often lead to their degeneration into one cell patches*

(Table structure and some model properties adopted from Agarwal, Green, Grove, Evans & Schweik, 2002)

2 METHODOLOGY AND DATA

2.1 Methodology

The methodology of the research composed of three phases.

- Image Classification
- Suitability Analysis
- Change detection and Modeling the LULC

2.1.1 Image Classification

The first phase of the study was acquiring the land use/cover classes from remote sensing sources by classification methods. In figure 2-1 the classification workflow illustrated for summarizing the general outline of this phase. The details have been explained in chapter three.

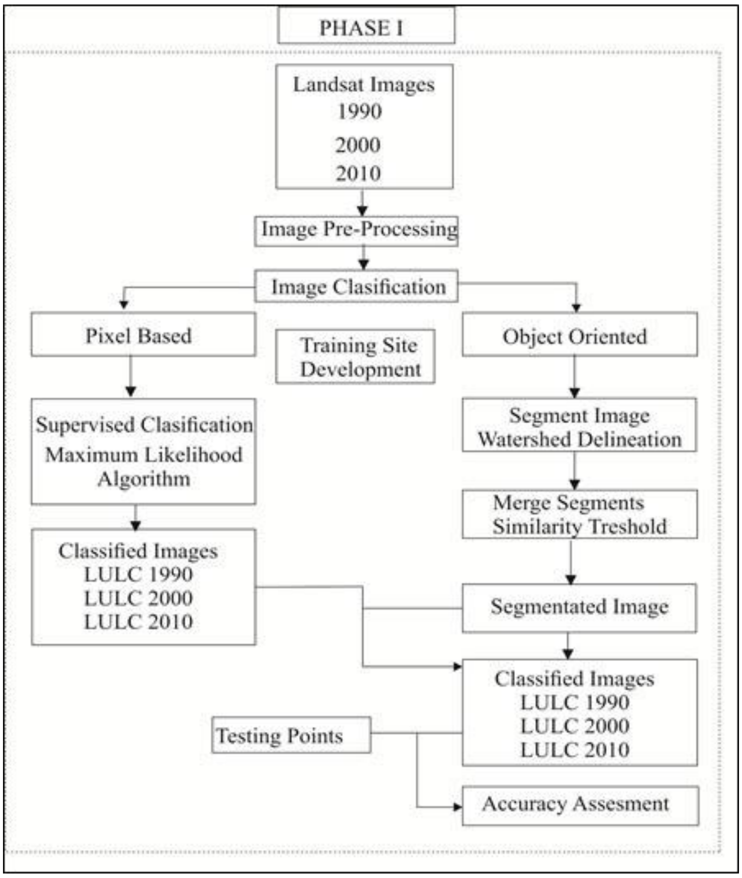

Figure 2-1 : Image classification workflow

2.1.2 Suitability Analysis

Second phase of the study was the suitability analysis for the urban development. In figure 2-2 the schema of the suitability analysis using the multiple criteria evaluation method conceptualized. In chapter four the details of the work can be found.

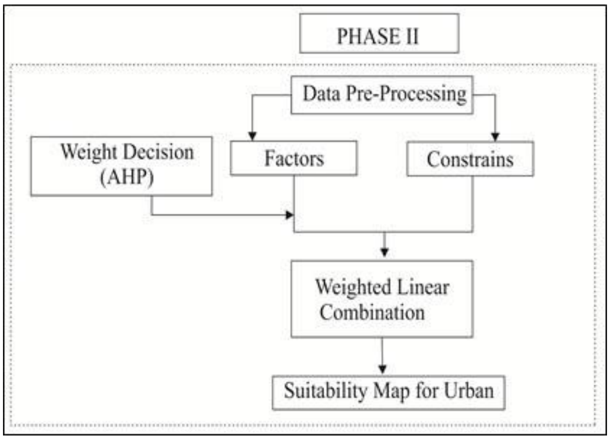

Figure 2-2 : Suitability analysis workflow

2.1.3 Change Detection and Modeling

The third phase of the research was change detection and future prediction. In figure 2-3 this workflow of this phase can be observed. The details for change detection were explained in chapter five and modeling was explained in chapter six.

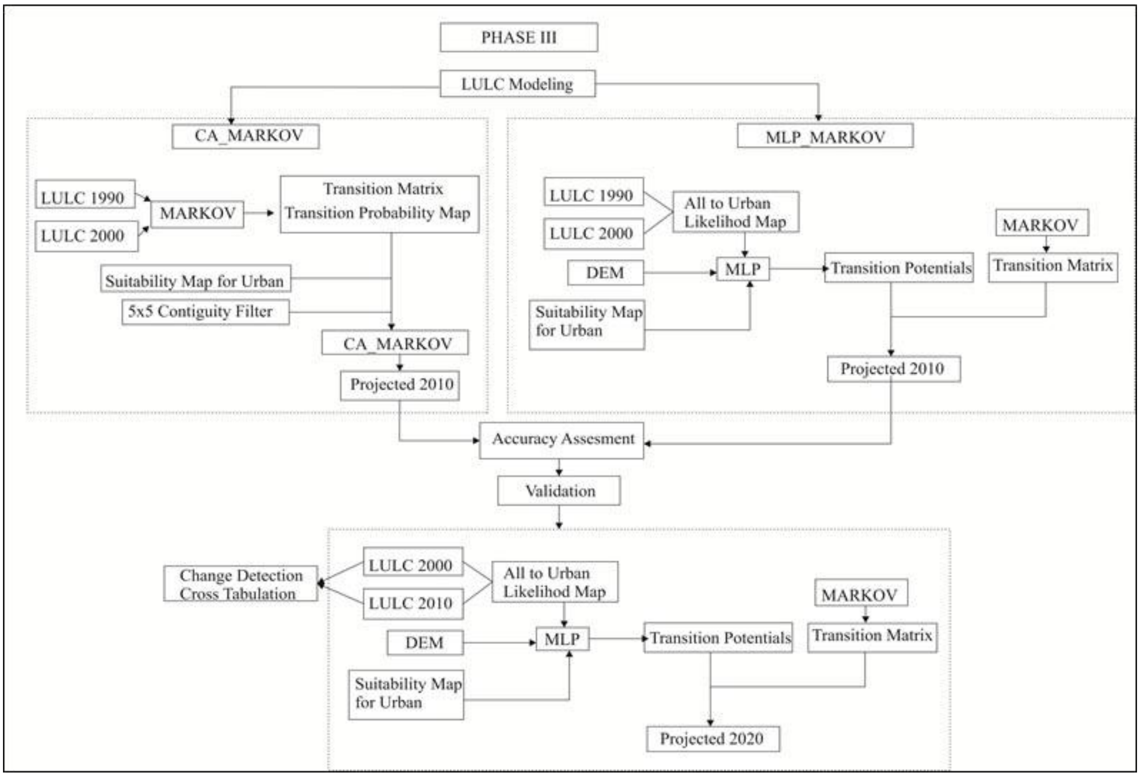

Figure 2-3 : Change detection and simulation workflow

2.2 Multi Temporal and Multispectral Data (Landsat Images)

For the land change detection and modeling the initial step is the data acquisition and the most common sources are the remote sensing sources, because the acquisition of the vector data is expensive and time consuming, finding the multi temporal data in order to detect the changes is not easy, especially in developing countries. For the study area the multi temporal vector data was CORINE which was prepared in 25 ha minimum mapping unit. But this resolution was not detailed enough because of that satellite images have been used. The summer period Landsat images selected for preparation of land use/cover data. The reason of using the summer period was to distinguish the other agriculture from orchards and to observe the transformation from urban area to these classes separately. In order to have better accuracy in classification the cloud free Landsat images used; for this reason in the year 1990 the month for the image is June and the sensor is ETM+ and for 2000 and 2010 the month is July and the sensor is TM. The properties of the images can be found in table 2-1, 2-2 and 2-3.

Table 2-1 : Landsat images basic properties

Date	Date Acquired	Spacecraft ID	Sensor ID
1990	12-Jun-90	LANDSAT_5	"TM"
2000	17-Jul-00	LANDSAT_7	"ETM"
2010	5-Jul-10	LANDSAT_5	"TM"

Table 2-2 : Landsat TM image properties

Landsat 5 (TM sensor)	Wavelength (micrometers)	Resolution (meters)
Band 1	0.45 - 0.52	30
Band 2	0.52 - 0.60	30
Band 3	0.63 - 0.69	30
Band 4	0.76 - 0.90	30
Band 5	1.55 - 1.75	30
Band 6	10.40 - 12.50	120
Band 7	2.08 - 2.35	30

Table 2-3 : Landsat ETM+ image properties

Landsat 7 (ETM+ sensor)	Wavelength (micrometers)	Resolution (meters)
Band 1	0.45 - 0.515	30
Band 2	0.525 - 0.605	30
Band 3	0.63 - 0.69	30
Band 4	0.75 - 0.90	30
Band 5	1.55 - 1.75	30
Band 6	10.40 - 12.5	60
Band 7	2.09 - 2.35	30
Pan Band	.52 - .90	15

2.3 Other Data Sources

The LULC changes can be predicted using the multi temporal land use/cover data but also other data sources needed for the improvement of the study. For this purpose other data related to the study area has been collected from the Firat Development Agency (they have been prepared by different institutions) and processed with GIS software.
The other data sources are;

- **Roads:** Main road network of the study area prepared by Ministry of Transportation in national level.
- **Water bodies**: Lakes and dams in the study area prepared by Turkey General Directorate of State Hydraulic Works.
- **Protection Areas and:** These are including the naturally protected and it has been prepared by Ministry of Culture and Tourism.
- **Archeological Areas**: These areas are approved archeological zones and also prepared by Ministry of Culture and Tourism
- **STATIP** (Problem Identification and Improvement of Agricultural Lands Project): This data is including the main land cover map for the year 2008 which prepared by Ministry of Agriculture of Turkey for identification of problematic agricultural lands, because of this it is mainly concentrated on agricultural lands. The data has been prepared using the SPOT 5 images with 2.5 and 5m resolution.
- **CORINE** (Coordination of Information on the Environment prepared by the European Environmental Agency): This data includes the main land cover maps with 25 ha minimum mapping unit.
- **Approved Development Plan:** This data is including the approved development plan for the study area prepared by municipality of the study area in the year 2009.
- **DEM:** Digital elevation model for the study area (30m resolution).

2.4 Tools

The tools have been used for the study varies, main tools used for the development of the research are;

- ArcGIS 10 used for image and vector data preprocessing.
- IDRISI Selva* used for classification, accuracy assessment of classification, change detection and modeling.
- ENVI** used for testing some other classification methods.
- Map Comparison Kit 3*** used for accuracy assessment of the projected land classes.

**IDRIS Selva is a GIS and remote sensing software.*
***ENVI is remote sensing software.*
****Map Comparison Kit 3 is a tool for map comparisons.*

2.5 Study Area Selection

The study area was defined in order to investigate the changes in the urban area and the outskirts of the urban area. For this purpose bounding box which covers these focus areas covering an area of 40612 ha area has been used (figure 2-4).

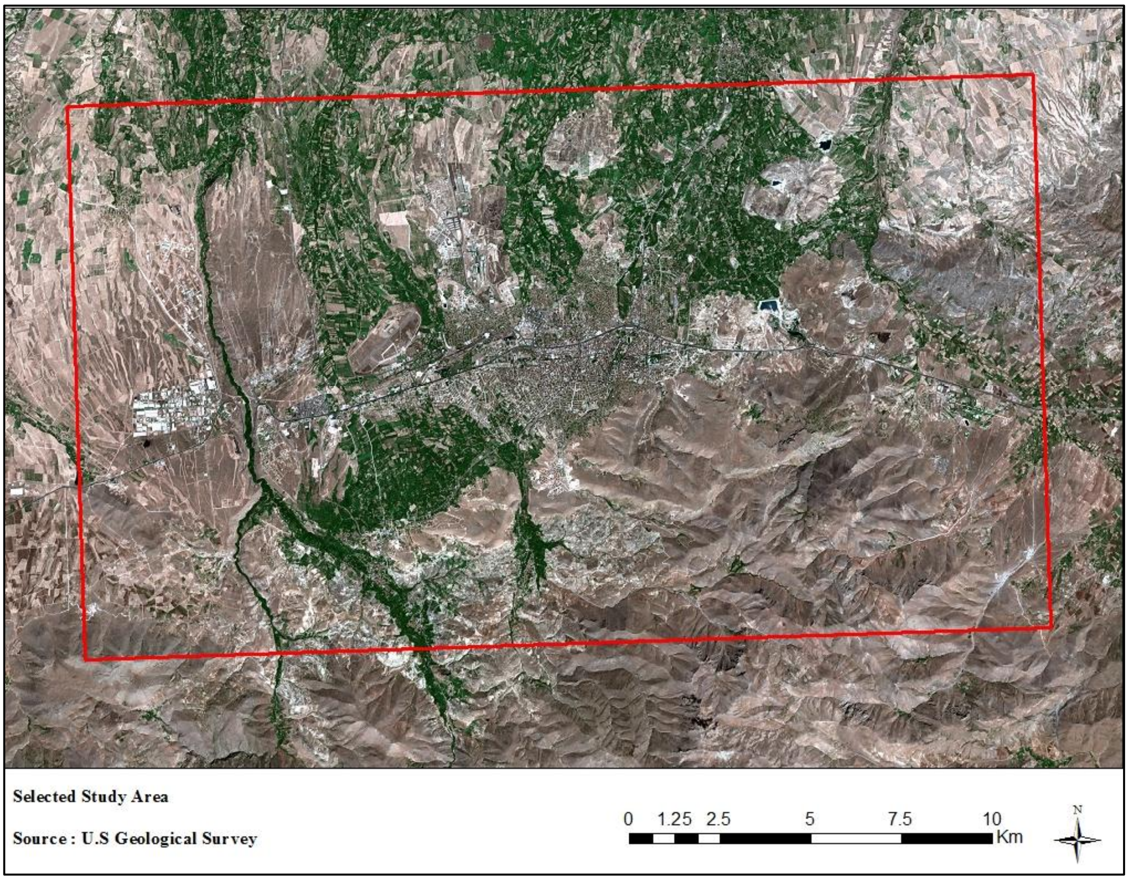

Figure 2-4: Study Area

2.6 Data Preprocessing

The Landsat images were covering a large area (3500000 ha) but for this research the study area was just 40612 ha, because of that Landsat images were clipped according to the bounding box which covers the study area. Moreover Landsat images were in GeoTIFF* format in order to use them in different software especially in IDRISI images have been exported to ERDAS IMAGINE** format. Finally the Vector data received from Firat Development agency was in ED_1950 system, so these data have been transformed to the WGS_84 system in order to be same with Landsat images which are in WGS_84 system and then clipped according to the study area.

**GeoTIFF is a public domain metadata standard which allows georeferencing information to be embedded within a TIFF (image) file* (GeoTIFF, n.d.).
** *ERDAS IMAGINE is remote sensing application.*

3 IMAGE CLASSIFICATION

For land change analysis and modeling the first step is the land use/cover data preparation. Land categories can be acquired by the classifying satellite images. Images are including the color differences, these colors are not the real land classes, in order to get the real category information from the bands the images need to be classified. *Image classification is the process of categorizing image pixels into classes to produce a thematic representation* (Gecena & Sarpb, 2008). There are several methods for classification and each method is specific to the data and the location, because in each location land categories are varies and have different values in the image. For instance the image value (reflectance) of an agricultural land is dependent on the type of crop grows on that land. Even the same crop in different climates can have different colors which change the color on the image. Moreover the seasons also affect the color of land covers.

There are different approaches for classification. According to Caetano (2009) Image classification can be done based on three objectives which are;

- Type of learning (Supervised and Unsupervised)
- Assumptions on data distribution (Parametric, Non-Parametric)
- Number of outputs for each spatial unit (Hard and Soft) (Caetano, 2009)

Moreover there are also objectives regarded levels of classification, which are;

- Pixel based Classification
- Object-oriented Image Segmentation and Classification

3.1 Pixel Based Classification

Pixel based classification is the traditional method of image classification. This is mainly based on the pixel reflectance values of the image (Wang, Sousa & Gong, 2004). According to the type of learning there are mainly two kinds of pixel based classification supervised and unsupervised (Caetano, 2009). For this study the supervised method used for the pixel based classification.

3.1.1 Supervised Classification

"Supervised classification is a procedure for identifying spectrally similar areas on an image by identifying "training" sites of known targets and then extrapolating those spectral signatures to other areas of unknown targets" (Mather & Koch, 2011).

As we can understand from the definition in supervised classification there is a priori knowledge about the image so the image will be classified according to this prior knowledge which is called training sites. Training sites are the areas for which the characteristics are known according to a ground truth or other reliable data. Here the ground truth refers to the data collected from the ground (Ground Truth, n.d.).

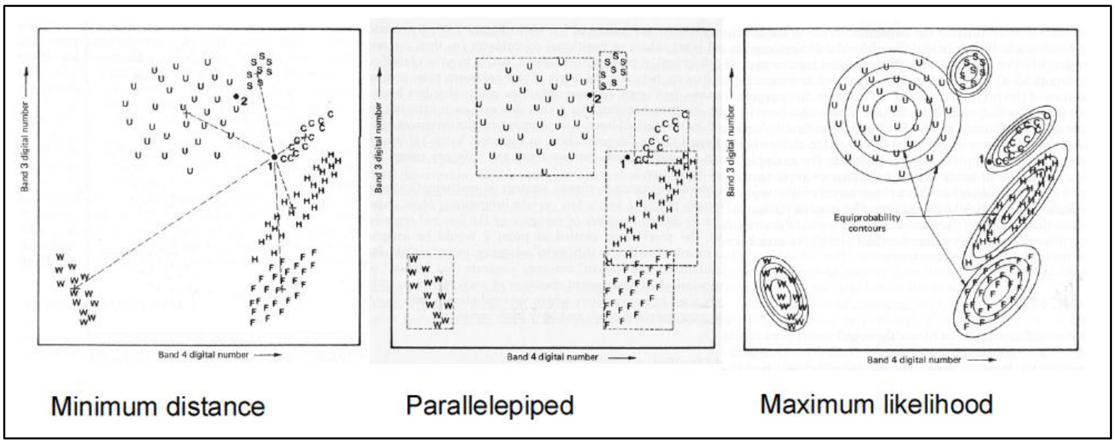

Figure 3-1: Classic supervised classifiers (Caetano, 2009)

There are different algorithms for supervised classification; the classic classifiers are minimum distance, parallel pipelined and maximum likelihood methods. As it can be observed from the figure 3-1 each of these classifiers uses different statistical approaches for the classification, for this research the maximum likelihood used because it gave better results for the study area.

3.1.1.1 Maximum Likelihood Algorithm

The maximum likelihood algorithm uses a maximum likelihood procedure derived from Bayesian probability theory; it applies the probability theory to the classification process. This method is a supervised method which uses the training sites, from these sites it determines the class center and the variability in the raster values in each band for each class. This will help to determine the probability of the cell to be belonging to a particular class defined in training sites. The probability is depending on distance from cell to class center, as it has been illustrated in figure 3-1, class size and the shape of the class in spectral space. The maximum likelihood classifier computes the class probabilities and classifies the cell where the probability is higher (Smith, 2011).

3.1.1.2 Image Enhancement

There are different image enhancements methods are available for images which helps to acquire the category information from the image. Histogram modification, filtering and band compositions are some of the methods. For this research band composition method has been implemented. Each band composition algorithm enhances the different category information from the images. For this study false color composition used which is a combination of VNIR (Visible Near Infra-Red) (4) - red (3) - green (2) in which vegetation seems as red tones, urban areas appear blue towards to gray , water appears blue (Band Combinations, n.d.). Using this composite will help to distinguish between the orchards and other agricultural lands easily and which is one of the objectives of the research. For distinguishing urban area is difficult with any of the composite method, because the urban area of the case study include composed of mixed pixels. Still compared to the other compositions false color composition is the best choice for the study area. The Landsat images for the study area can be observed in figure 3-2, 3-3, 3-4.

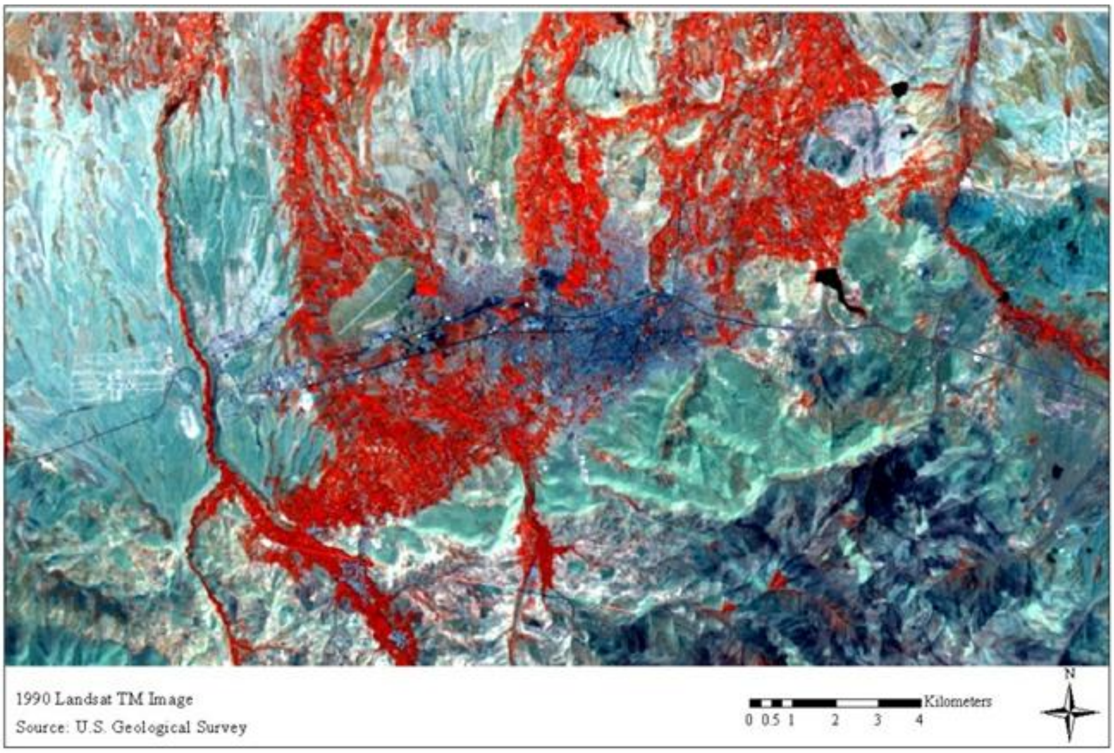

Figure 3-2: Study area profile with Landsat image (1990)

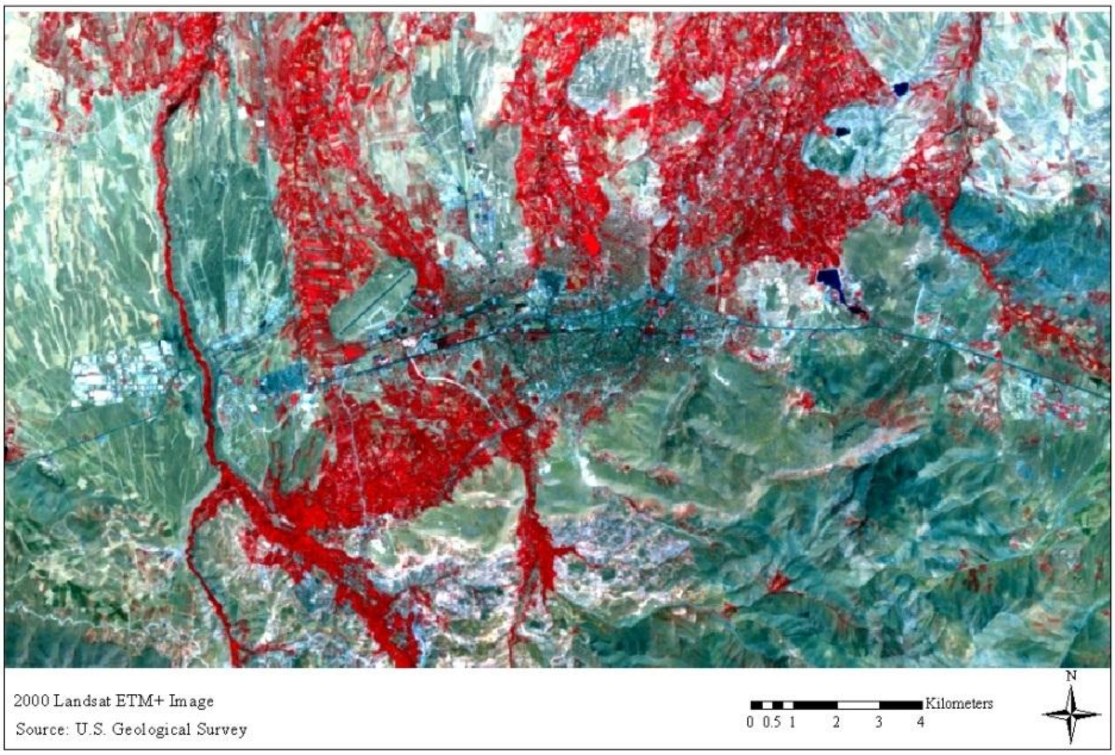

Figure 3-3: Study area profile with Landsat image (2000)

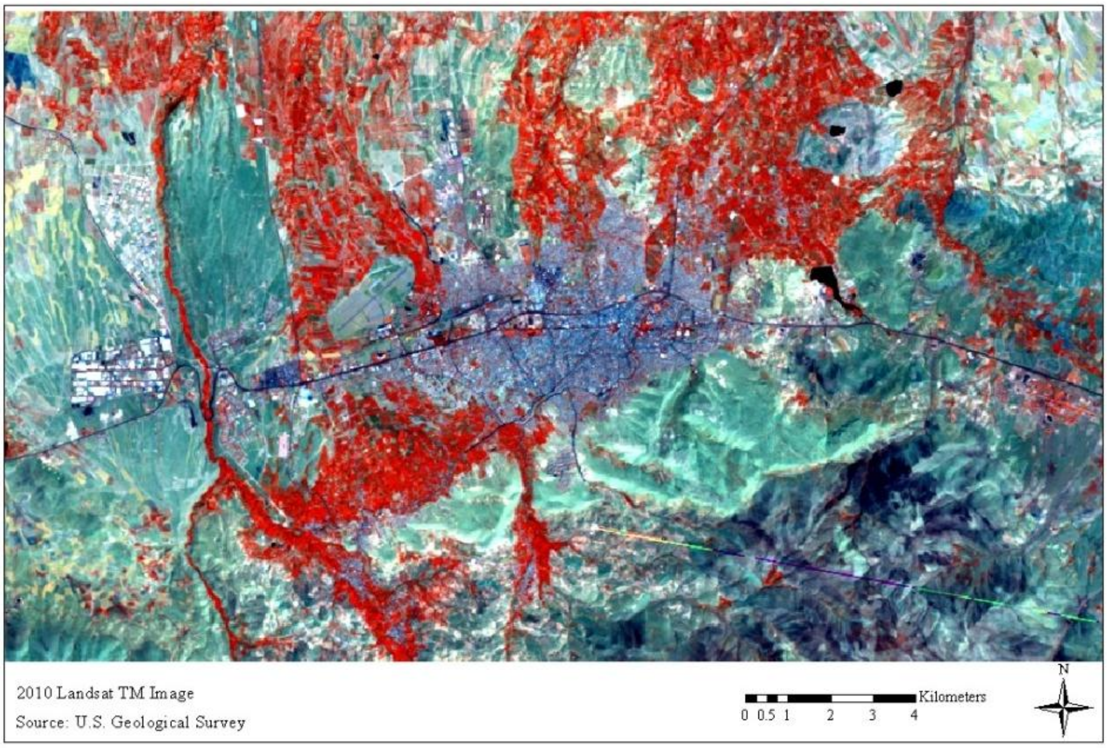

Figure 3-4: Study area profile with Landsat image (2010)

3.1.1.3 Developing the Training Sites

As it has been mentioned before in the supervised classification training sites need to be used as a priori knowledge. For the study are ground truth data is not available for this reason CORINE, STATIP and user interpretation from the satellite image used for development of the training sites. Around 250 training sites selected for the each image the category information for this sites acquired from CORINE and STATIP datasets tested with the user interpretation. For the year 2010 CORINE is not available because of this reason CORINE 2006 and Google earth used for defining the training sites (figure 3-5).

Figure 3-5: Development of training sites

3.1.2 Implementation of Supervised Classification for Study Area

The LULC categories determined according to the objectives of the research in which urban areas, orchards, other agricultural lands, other land cover categories and water body are the main groups. While defining this group's structure in CORINE project used. CORINE has almost 32 categories but for the study area they generalized to 5 main classes as it can be observed from table 3-1. The main research objective of this study was to find the change in urban areas and agricultural usages (agriculture and orchard) and modeling them because of that the classes different than them generalized to other. Water bodies haven't been added to this other class because the reflectance value of the water is not similar to other class.

Table 3-1 : LULC categories (CORINE land cover technical guidelines, 2000)

Generalized Class	**Classes from CORINE**
Agriculture	Irrigated Agriculture , not irrigated agriculture, principally agricultural lands in which some parts covered by natural vegetation, vineyards
Orchard	Irrigated orchards, not irrigated orchards
Urban	Construction sites, Industrial and Commercial Units, Continuous Urban Fabric, Non continuous Urban Fabric, Continuous and non-continuous Rural Fabric, Airports, Roads Railways, Mineral extraction Sites, Green Urban Areas
Other	Sparsely Vegetated Areas and Transitional Woodland Shrub, Natural Grassland , Pastures, Sand dunes, Inland Marshes
Water	Dams , lakes

In order to use the supervised classification the training samples were created based on reliable sources, then the supervised classification based on maximum likelihood algorithm has been used for classification because the other algorithms result was not satisfactory. The results of the pixel based classification can be observed in the figures 3-6, 3-7, 3-8.

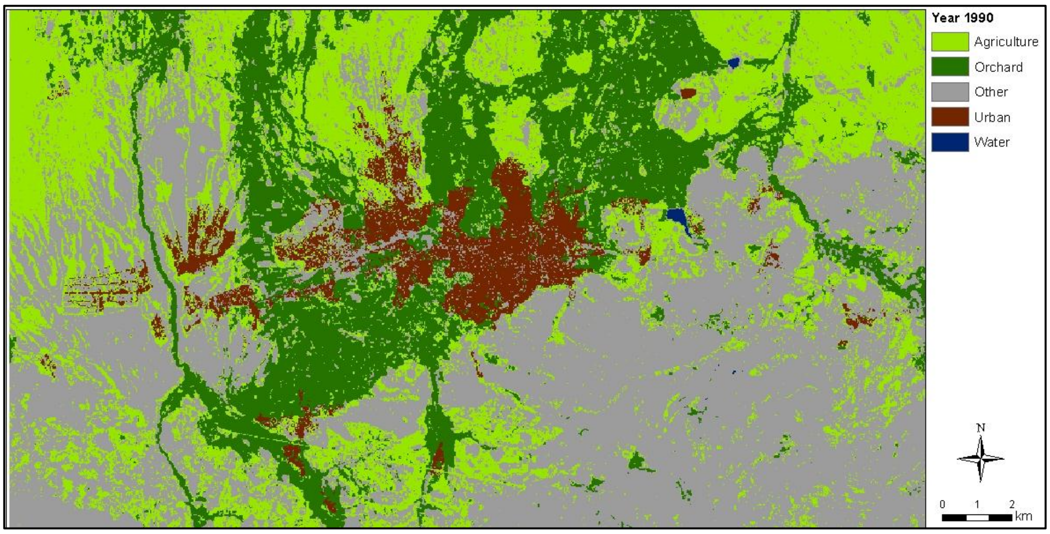

Figure 3-6: Pixel based classification result (1990)

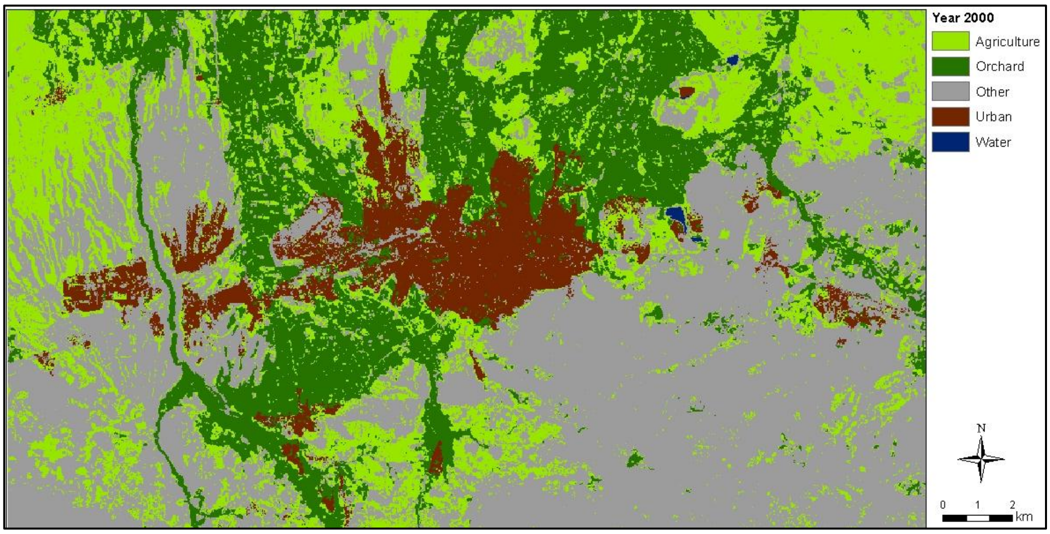

Figure 3-7: Pixel based classification result (2000)

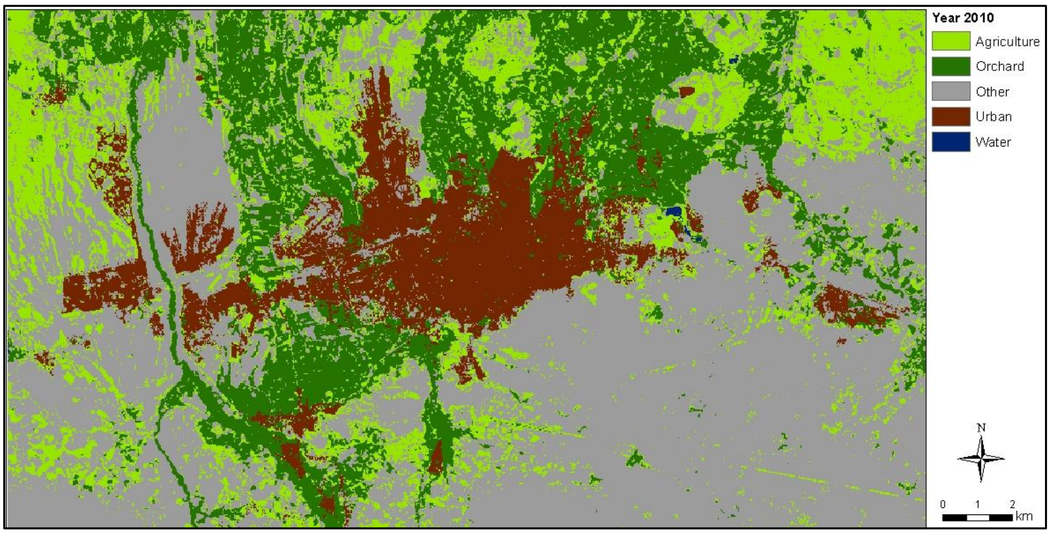

Figure 3-8: Pixel based classification result (2010)

3.2 Object-oriented Image Segmentation and Classification

The pixel based method is very useful for the image classification but the LULC categories can be represented better by objects rather than pixels, so the second step of the classification is using the object oriented method for the classification. Object oriented classification is based on image objects which mean a set of similar pixels (figure 3-9). For acquiring these objects the segmentation is the most common method. *"Image segmentation is the process of partitioning a digital image into multiple segments"* (Shapiro & Stockman, 2001). The main aim in segmentation is dividing the image into more meaningful smaller pieces and then the merging these pixels according to different algorithms.

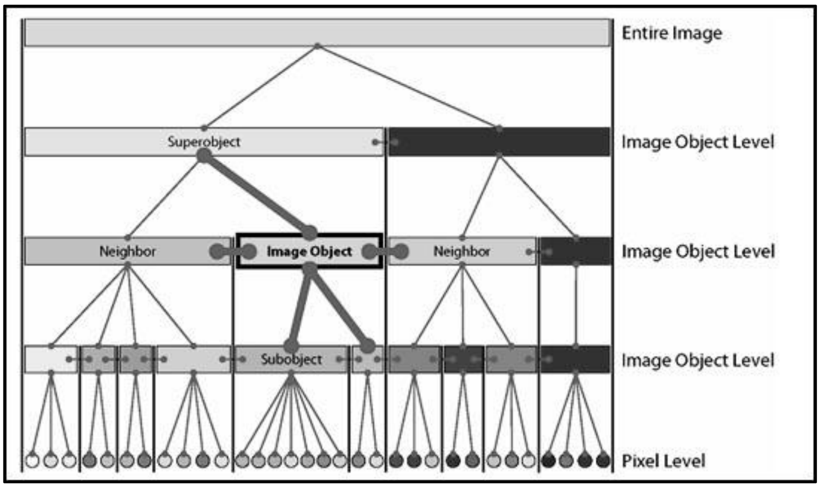

Figure 3-9: Object based Classification (Caetano, 2009)

Common segmentation methods are, thresholding, clustering, region-growing, split-and-merge, watershed transformation, model based segmentation, trainable segmentation (Image Segmentation, n.d.). For this research watershed delineation and & similarity threshold based segmentation methods have been used for classifying the images.

3.2.1 Watershed Delineation Segmentation

The watershed delineation algorithm is using the pixel values within the variance, like elevation values in digital elevation model, then grouping the pixels in the same watershed catchment areas and giving the unique ID to this catchment area and then grouping/merging the pixels which have same watershed ID (Eastman, 2009). The segmentation process is iterative in this method every segment merged with the most similar group, for this purpose user defined similarity threshold, weight variance and weight mean vector were used. For the study threshold value was 5, because the study area composed of mixed pixels for this reason the similarity threshold determined low for having smaller segments which includes fewer amounts of pixels but more similar (Eastman, 2012). After segmentation pixel based classification was used to as a reference image to have the final result. The result of the segmentation based images can be observed from figures 3-10, 3-11 and 3-12.

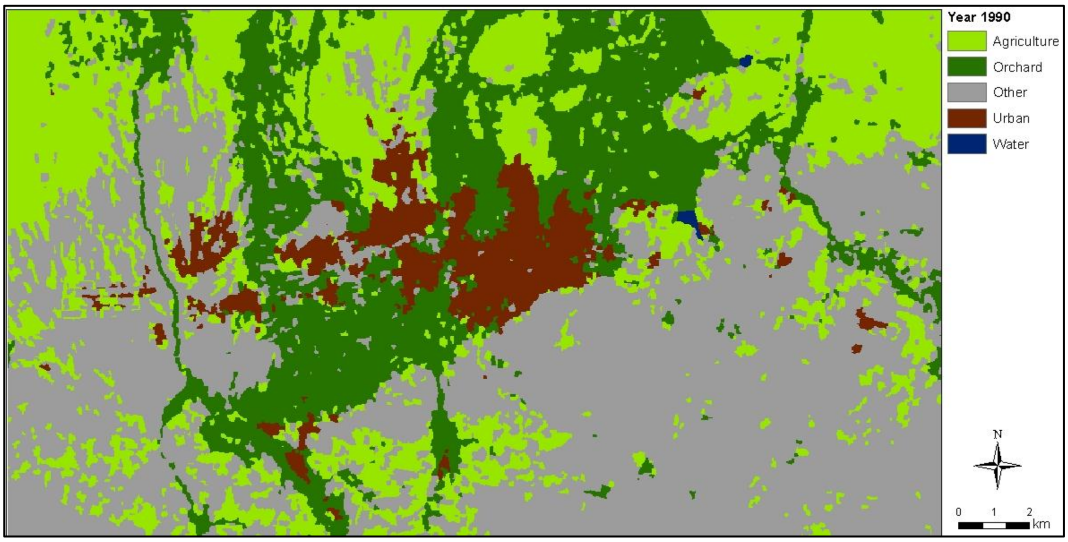

Figure 3-10: Segmentation based classification result (1990)

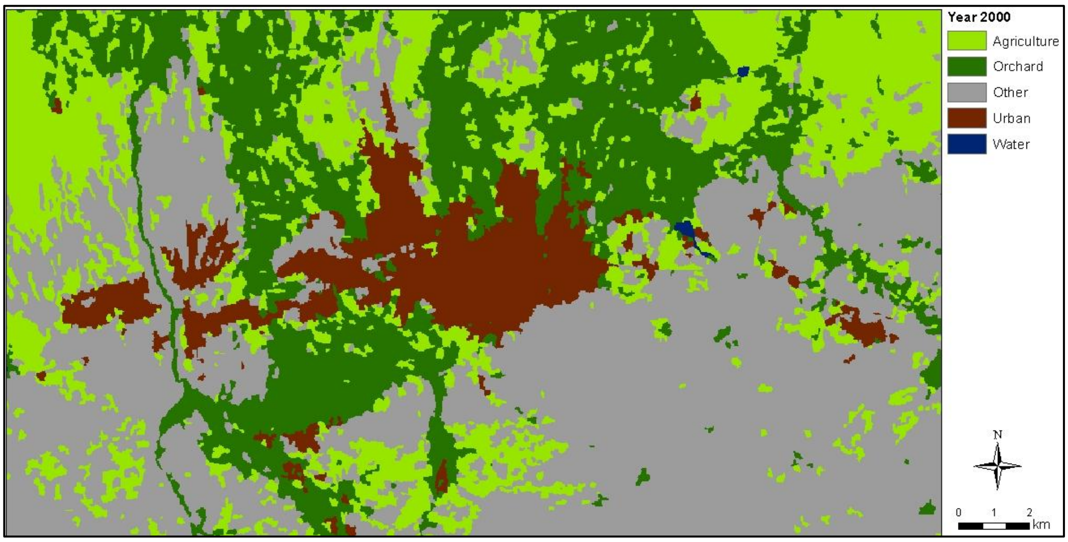

Figure 3-11: Segmentation based classification result (2000)

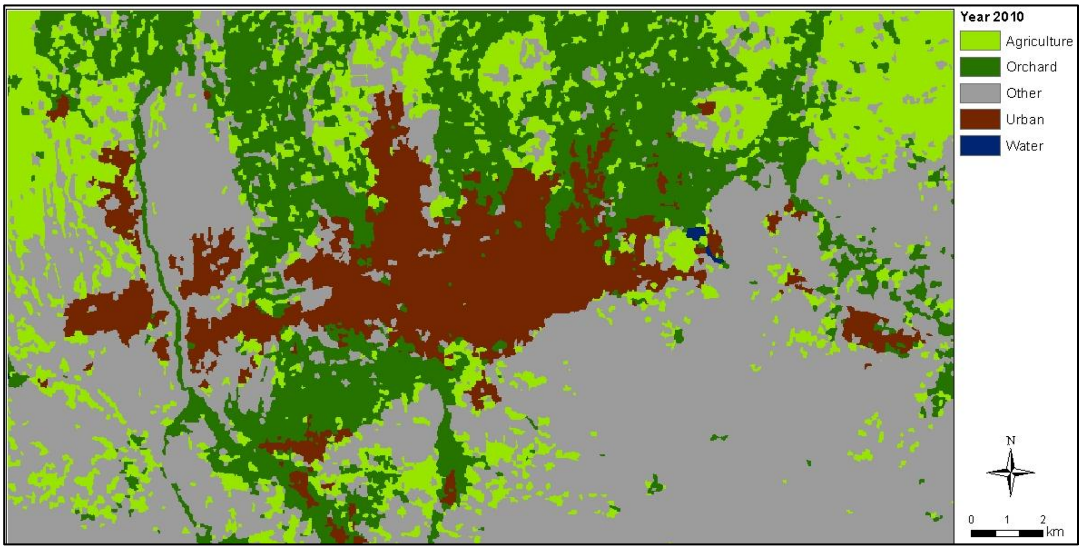

Figure 3-12: Segmentation based classification result (2010)

3.3 Accuracy Assessment

The last part of the image classification process is accuracy assessment. Accuracy assessment is a process to compare the classification with ground truth or reliable sources (Rossiter, 2004). For the accuracy assessment overall kappa which is a statistical measure of overall agreement between two categorical items (Cohen's kappa, n.d.) and conditional kappa which includes, user's accuracy, producer's accuracy and overall accuracy were calculated. The mathematical explanation of the calculation process and the detailed results of the accuracy assessment can be found in appendix A.

For this process we need ground truth data for testing sites, since this is not available the CORINE, STATIP and user interpretation has been used for selecting the testing sites. For each LULC class 19-50 random points were created and then the spatial information for these points acquired from CORINE and STATIP and compared with the satellite images. For the year 2010 the CORINE was not available because of that 2006 CORINE and Google Earth were used for creating the testing points. For each class the general requirement is 50 points (Lillesand & Kiefer, 2004), since the water bodies don't cover a huge area only 19 points were created for water bodies in each LULC time. The method used for accuracy assessment is a comparison technique which is comparing the testing points with the classified image for the each land cover class.

Accuracy assessment results for the segmentation based images:

- 1990: Overall Kappa : 0.7685
- 2000: Overall Kappa : 0.7256
- 2010: Overall Kappa : 0.7511

This level of accuracy found as satisfactory because of the heterogeneity of the study area and the low resolution of the satellite images and according to kappa agreement the accuracy was in substantial agreement level which means it can be used (Pontius, 2000).

- < 0: Less than chance agreement
- 0.01–0.20: Slight agreement
- 0.21– 0.40: Fair agreement
- 0.41–0.60: Moderate agreement
- 0.61–0.80: Substantial agreement
- 0.81–0.99: Almost perfect agreement (Pontius, 2000)

4 SUITABILITY ANALYSIS

Suitability analysis has importance on LULC change modeling process, because with the help of suitability analysis simulation for the future can be grounded with the existing patterns and drivers. The definition of the suitability in general is "*quality of having the properties that are right for a specific purpose*", and for the land-use suitability we can specify the definition suitability as identifying the most appropriate spatial pattern of future land uses according to purpose (Hopkins, 1977). Suitability analysis can be used for different purposes such as, agriculture, ecology and urban development. According to each approach the objectives would be different. The criteria for urban suitability would be different than agricultural suitability because in each case the suitable places have different features (Malczewski, 2004). The features of each category need a different specialization, for instance; the suitable lands for agricultural development are different than the urban development. For this research the focus was the urban area expansion and impacts on other categories, because of that the suitability analysis has been done only for urban category.

Many studies have been done in order to find the urban suitability although the methods and variables dependent on location and time. Each location has its own patterns which have different effects on urban development; this can be depending on the rules and existing trends. For instance the past and present urban development trends are different; in the past the settlements were mainly built near to the industrial zones in order to decrease the travel time, but currently the developments no longer take place near the industrial zones on the contrary the industrial zones are tried to be decentralized. Moreover the urban development patterns are different in developing and developed countries. Therefore for this research the rules in Turkey and patterns in study area taken into consideration for the suitability analysis.

4.1 Multi Criteria Evaluation

For suitability analysis Multi Criteria Evaluation (MCE) is a widely used process. Finding the suitable areas for the urban development requires consideration of different drivers because of that using a system which evaluates multiple criterions are required. This process combines variables with different methodologies and then transforms it into a suitability map output (Drobne & Lisec, 2009). The main criterions in MCE are Factors, Constrains.

Constrains: Constrains are the variables which refer to the restricted areas for development, there are no medium values; the values are either **0** (not suitable) or **1** (suitable) (Eastman, 2012).

Factors: Beside the constrains there are variables which effect the urbanization in a continuous scale, different than the constraints they have medium values and this medium values have different suitability. So the factors can be defined as the variables which have continuous suitability values (Eastman, 2012).

The most common techniques in MCE are;

- Boolean Intersection
- Weighted Linear Combination (WLC)
- Ordered Weighted Averaging (OWA) (Eastman, 2012)

4.1.1 Weighted Linear Combination

WLC is method is simply a weighted overlay operation of the different criterion. Beside the variables the main part of the process is the weight allocation of factors and using the factors and constraints in an overlay analysis in order to find the suitability of the area.

The formula of the Suitability according to WLC is:

$$\mathbf{S=(\Sigma\, W_i\, X_i)\, \Pi C_i}$$

Where, **S:** suitability,

$\mathbf{W_i}$**:** weight of factor i,

$\mathbf{X_i}$**:** criterion score of factor **i**,

$\mathbf{C_i}$: criterion score of constraint **j**,

$\boldsymbol{\Sigma}$ **:** Somme,

$\boldsymbol{\Pi}$ **:** Product (Eastman, 2012)

In this method continuous values (factors) were standardized to a numeric range and then combined according to the weights (Eastman, 2012) and then they were masked (product) by constrains to have the final suitability map. The standardization formula is:

$$\mathbf{X_i = (R_i - R_{min}) / (R_{max} - R_{min}) * standardized\_range}$$

R = raw score

Standardization of the factors can be done according to:

- Fuzzy set membership approach,
- Value/utility approach,
- Function approach,
- The probability approach

For this research fuzzy set membership approach used.

4.1.2 Fuzzy Set Membership

The fuzzy truth represents the membership in vaguely defined sets. This set doesn't have sharp boundaries. For example distance terms like far-close are fuzzy truths because the definition of the proximity is not strict. In suitability analysis, especially if we deal with the proximity issue the fuzzy logic helps us to standardize the variables (Jiang & Eastman 2000). Some of the fuzzy set functions are Sigmoidal, J-shaped and Linear, and for this research linear function has been used (Eastman, 2009). There are different types of the linear function like presented in figure 4-1. These are illustrating how the membership changes. For example in symmetric one the membership (suitability) increases from control point a to b and then decreases from c to d. (For detailed information about fuzzy logic the paper written by Jiang & Eastman (2000) can be referred).

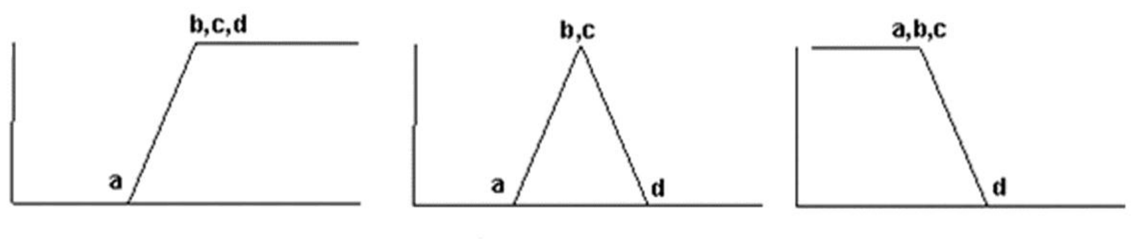

Figure 4-1: Linear membership functions (Eastman, 2009)

4.1.3 Weight Decision

Weights are the values which indicate the impact of each criterion on the process. There are many methods to decide the weights of each factor, in general the weights can be determined by decision makers, but when there are many criterions the process become complex and using the Analytical Hierarchy Process would be a better solution for advanced works. In this system pairwise comparison method has been used, which indicates the relative importance of the each factor (Yu, Chen & Wu, 2009). In the table 4-1 it can be observed that in pairwise comparison table the scale is continuous and according to the importance there are extreme points.

Table 4-1: Continuous Rating Scale

1/9	1/7	1/5	1/3	1	3	5	7	9
extremely	Very strongly	Strongly	Moderately	Equal	Moderately	Strongly	Very strongly	extremely
Less important				More Important				

The pairwise comparison works like this: if one felt that accessibility is important than geological situation in determining suitability for urban development, one would enter a 5 on this scale. If the inverse (geological situation is more important) were the case one would enter 1/5.

4.1.4 Implementation of Suitability Analysis

In order to implement the suitability analysis first of all the variables in other words factors and constrains for the study area were defined.

4.1.4.1 Constrains

Constrains for this project are the locations which are not allowed for urban development by law or existing occupied areas like existing built up area where the development is not possible.

Archeological Lands: According to the regulations in Turkey urban development is forbidden in the approved archeological zones (Protection and Utilization of Archeological Zones, 1999).

Protected Areas: Protected area category is mainly including the naturally protected areas which are also not suitable for urban development (Rule for protection of Culturally and Naturally Important Areas, 1983).

Water Bodies: The urban development is not possible in water bodies.

Existing Built Up: These areas are already occupied so the further urban development is not possible because the models are not considering the vertical growth (Ahmed, 2011). Vertical growth means the increase in the number of floors or replacement of the one storied houses with multiple storied houses which would increase density but not the area.

In figure 4-2 we can observe all constrains to be used in suitability process.

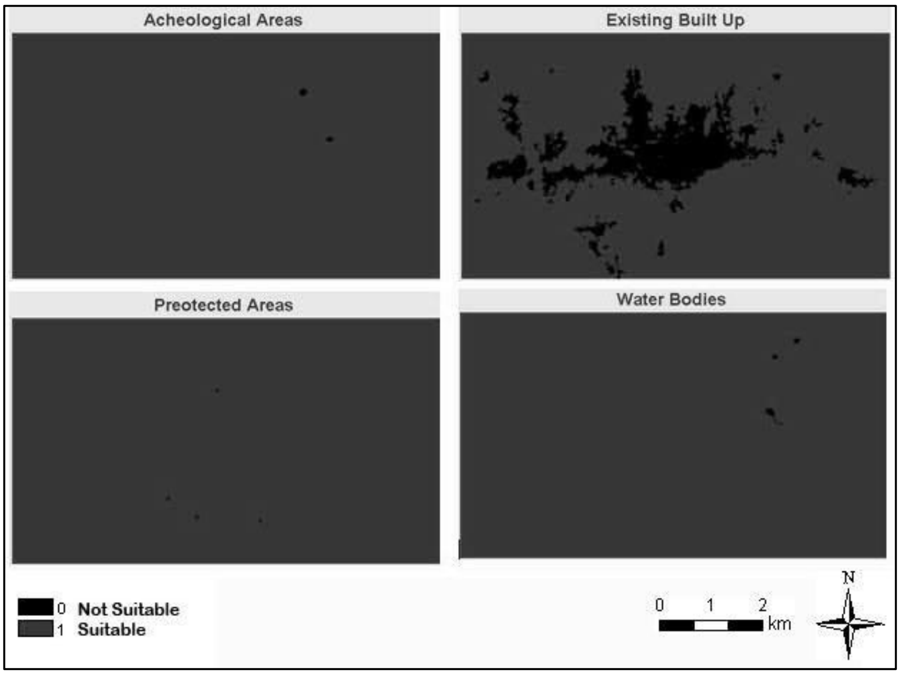

Figure 4-2: Constrains

4.1.4.2 Factors

Factors for this research have been selected according to existing patterns and rules in Turkey. Factors are different than constrains because factors have continuous suitability values different than 0 and 1. The factors used for the study were in different scale so that they need to be standardized to the same scale in order to use in the suitability process. For the standardization of the factors fuzzy logic has been used because the suitability values of factors doesn't have strict values which means fuzzy logic can be helpful. Moreover the factors also were in different level of measurement; some of them were nominal some were numeric. For the nominal factors the reclassification method used and for the numeric ones linear fuzzy set membership used.

Agricultural Lands: The agricultural lands are important for this research as a factor because there are some rules and regulations available for protecting them. However we cannot take them as constrain because even though there are limitations, these areas have been transformed to urban areas. For the agricultural lands Ministry of Agriculture defined different classes, but for this research four of them used and their suitability scale decided according to regulations and then they were standardized to the scale 0-255. Agriculture variable was in nominal scale and this need to be transformed to 0-255 continuous scale. For this factor reclassification method has been used to standardize the values to 0-255 scale. For the class I and II; 0 and for class III; 80-120 and for class IV; 120-150 suitability values assigned and for the other areas 255 values assigned. This values are not strict we just try to show the level of suitability in a continuous scale.

Slope: Slope is an important factor in urban development because it affects the cost of construction. Jantz, Goetz & Shelley (2004) stated that the effect of slope can change probability of urbanization in a specific location. Slope was in continuous scale but not in 0-255 and this need to be standardized. For this variable fuzzy set membership linear function used to standardize the slope values to 0-255 scale. Monotonically decreasing linear function used (Control points c=0, d=20). Which means between the 0-20 the suitability is decreasing and after 20 the suitability is 0 (LaGro, 2001).

Distance from Build Up: Urban development generally takes place near to existing developments because of the existing infrastructure so the proximity to the existing built-up areas has higher suitability compared to the far areas (Araya & Cabral, 2010) (Ahmed & Ahmed, 2012). As it can be observed from the figure 4-3 the expansion of the urban area was mainly covering the previous boundary and it is located mainly in the periphery of the existing built up.

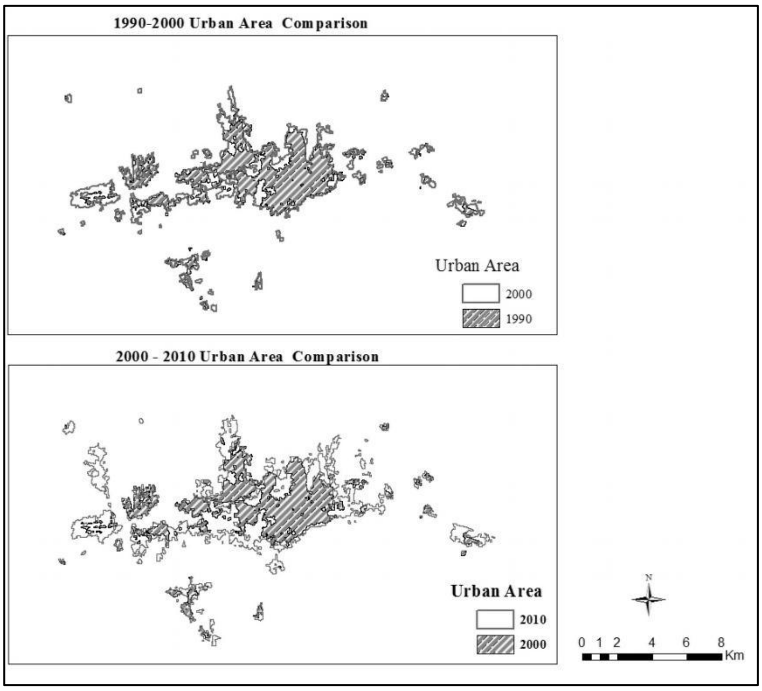

Figure 4-3: Urban area boundary changes

Proximity is fuzzy criterion because there aren't exact boundaries to define the suitability, so we can just say proximity to existing built up increase the suitability. Distance from existing urban area was in continuous scale but not in 0-255 so this standardized by the monotonically decreasing linear function (Control points c=0, d=10000).Which indicated the suitability value decreases when we go far from urban area and after 10 km it is close to 0.

Distance from Main Roads: In developing countries urban development takes place along the road (Kumar & Shaikh, 2012). As it can be observed from the figure 4-4 expansion of the urban area were mainly located along the main roads.

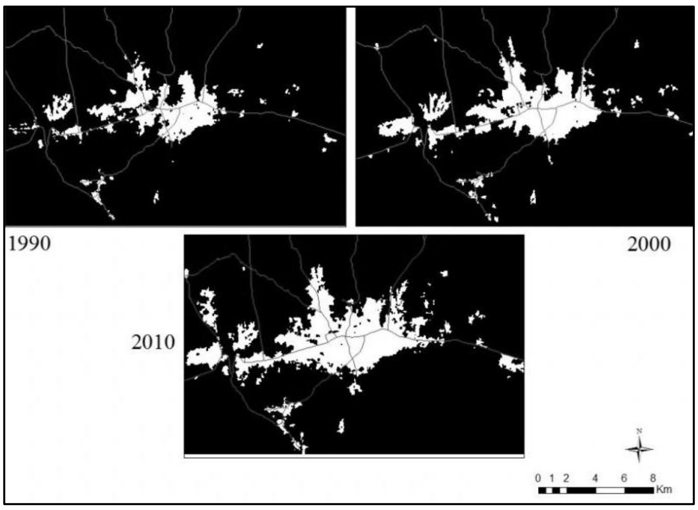

Figure 4-4: Road & urban relation

Moreover as it can be observed from figure 4-5 after 15 km the urban area is decreasing. According to this analysis we can say there is a strong relation between the roads and urban sprawl.

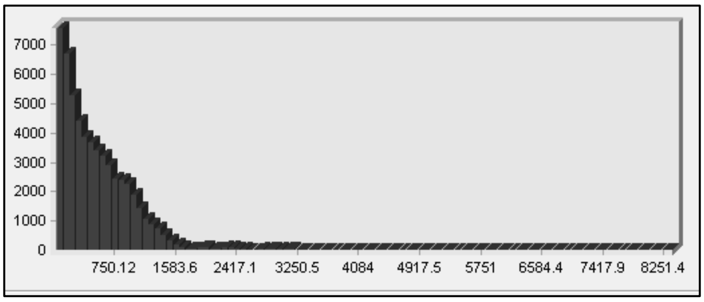

Figure 4-5: Trend chart for road & urban relation

In order to create the suitability related to proximity to road the fuzzy set membership symmetric linear function used. Because within 25m distance the area is not suitable urban development since it is including the road. In this case the control points were a=0 b=25 c=25 d=15000 this values means between the 0-25m the land is not suitable for urban growth, between the 25m to 15000m the suitability is degreasing near 25m it have the highest value after that it is decreasing in a continuous scale and after 15 km it is close to 0.

Geology (Earthquake risk): Geological conditions in other words earthquake risk have an important influence on the urban development, especially after the destructive earthquakes take place in some cities in Turkey these conditions started to be taken into account. Tudes& Yigiter, (2010) used it for the suitability analysis of Adana city in Turkey, and they stated that the earthquake risk could change the suitability of the location; development should take place where there is low risk of earthquake. The earthquake risk map was in nominal scale it had three categories for the study area which are; Suitable area under control 1 (S1), Suitable area under control 2 (S2) and not suitable (NS) area. We order them according to their suitability. These values were in nominal scale in order to convert them to the 0-255 scale reclassification method used and according to the values the suitability degree in 0-255 scale arranged. S1 (230-255) > S2 (160-175)> NS (0). For this factor also we cannot say these values are strict, they have been selected in order to illustrate the level of suitability.

Previous Plans: Previous plans also selected as a factor because in Turkey if there will be a new development this areas are the first places to consider for the expansion so they are more suitable than the other areas. This factor was in nominal level to so this transformed to 0-255 scale by reclassification method. The new development areas categorized as 255 (high suitability) and the other areas in the periphery indicated as 0 (low suitability).

In the figure 4-6 we can observe the standardized suitability maps of each factor defined. From the legend we can observe the pink color represent the highest suitability which is 255 and the black indicates the lowest suitability value which is 0.

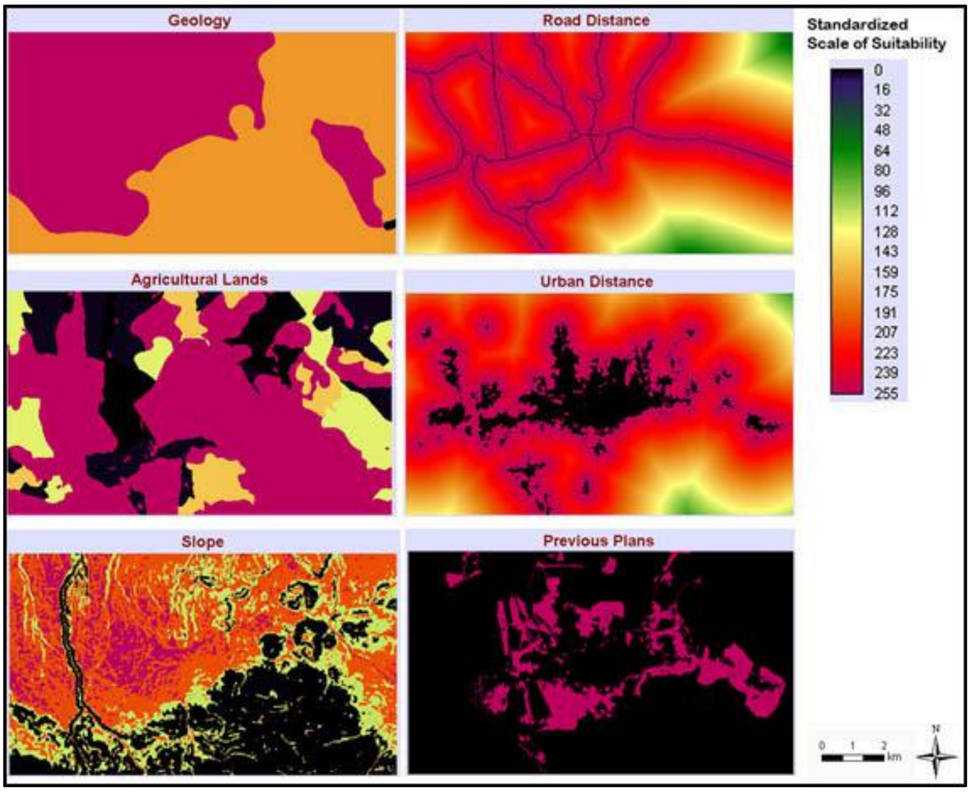

Figure 4-6 : Standardized factors

After defining the factors and constrains and then standardize them the next step is to define the weights of the factors. In order to decide the importance of each factor urban planning expert's opinions were used. According to their opinions the relative importance of the factors determined and the pairwise comparison matrix generated. (Questionnaire form filled by experts can be founded in appendix C).

Table 4-2 : Pairwise comparison matrix

Factors	Agricultural Lands	Geology	Accesibility (Proximity to Roads)	Slope	Proximity to Existing Urban Areas	Previous Plan Boundaries
Agricultural Lands	1					
Geology	5					
Accesibility (Distance to Roads)	3	1/5				
Slope	3	1/3	1			
Proximity to Existing Urban Areas	3	1/3	1	1		
Previous Plan Boundaries	3	1/3	1	1/3	1	1

According to this pairwse comparison table 4-2 the weights were calculated.The weight of the each facto can be observed from table 4-3.

Table 4-3: Weight of Factors

Class	**Weight**
Agricultural Lands	0.0502
Geology	0.4072
Accesibility (Distance to Roads)	0.1244
Slope	0.1701
Distance to Existing Urban Areas	0.1149
Previous Plan Boundaries	0.1332
*Consistency ratio (CR)= 0.04 is acceptable	

* A measure of how far a matrix is from consistency is performed by Consistency Ratio (CR) (Kordi, 2008). If CR > 0.10, then some pairwise values needs to be reconsidered and the process is repeated till the desired value of CR < 0.10 is reached

4.2 Suitability Analysis Results

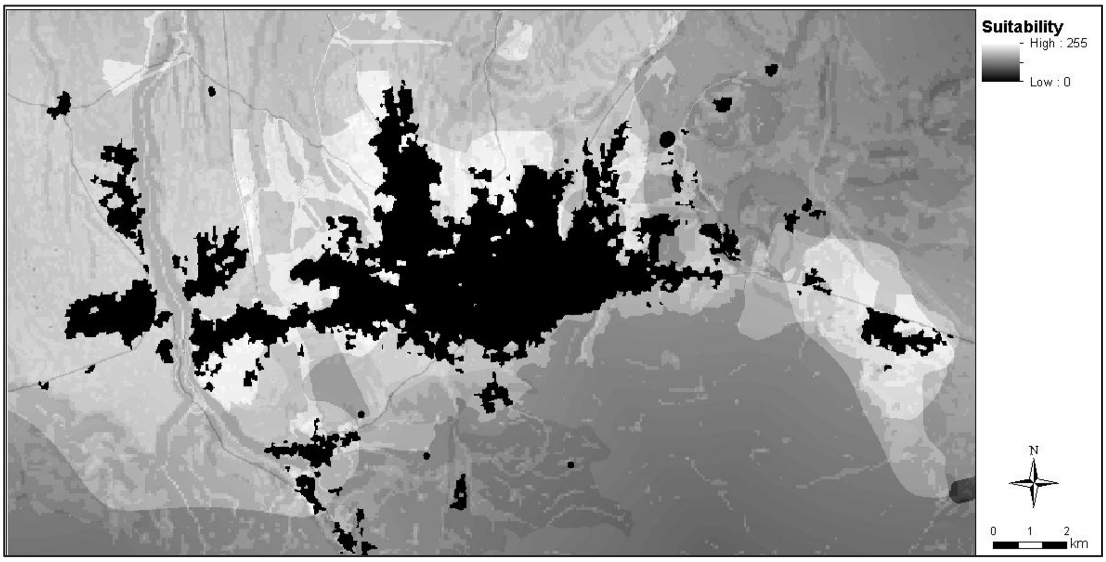

Figure 4-7: Urban suitability

In figure 4-7 we can observe the results of the suitability, here the white color indicates the high suitability and the black indicates the low suitability. In the map as we can observe that constrains was illustrated as black because in these areas urban expansion was not possible.

This suitability map will be used in modeling process in order to ground the model with existing patterns and expert opinions. This might help us to increase the accuracy of the model for the urban class by training the model with existing pattern and expert opinions.

5 CHANGE DETECTION ANALYSIS

Change detection is the process of identifying differences in the state of an object or phenomenon by observing it at different times (Singh, 1989). The land change can be resulted by many factors, in order investigate these factors we need to analyze the changes. By change detection method transitions from one category to other ones can be observed, which can help to understand the interaction between the categories. There are many techniques for change detection. For this research cross tabulation, area calculation, reclassification and the overlay methods were used. Cross tabulation is a statistical process that shows the joint distribution of two or more variables (Cross Tabulation, n.d.). Reclassification is simply classifying the one image according to new categories. Overlay simply means combining information from different layers. By using these methods amount of gain and losses, net change, and contributor to change from each category were calculated (Johnson, 2009). With this method the changes in each category have been observed. As it can be observed from the figure 5-1 in each class there are some gains and loses. Some of them are real changes in the study area and some of them are inaccurate changes resulted from the misclassification of the satellite images.

5.1 Changes between 1990-2000

The figure 5-1 prepared in order to investigate the gains and losses from each category. As we can observe from the figure there are big losses in each class especially in agriculture class and orchard.

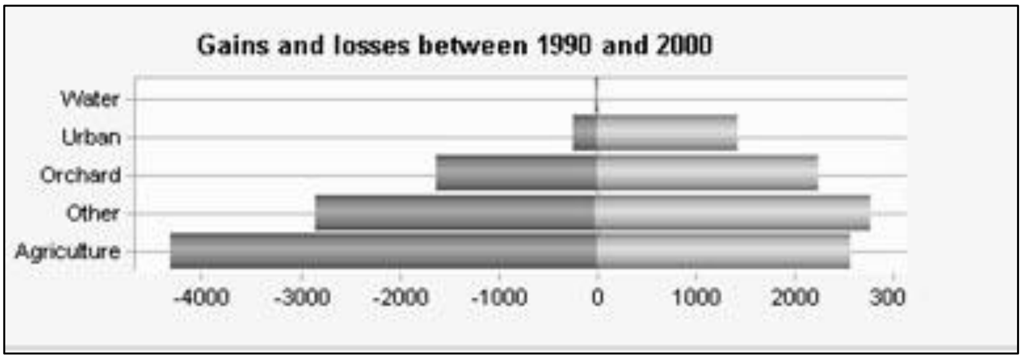

Figure 5-1: Gains and losses in each class (1990 - 2000)

In order to observe the transformation from agriculture and orchard to urban area the figure 5-2 and table 5-1 prepared. The locational contribution to urban area from other classes can be observed from figure 5-2 and the areal changes summarized on the table 5-1. As we can observe from table 5-1 around 900ha of these categories was transformed to urban area.

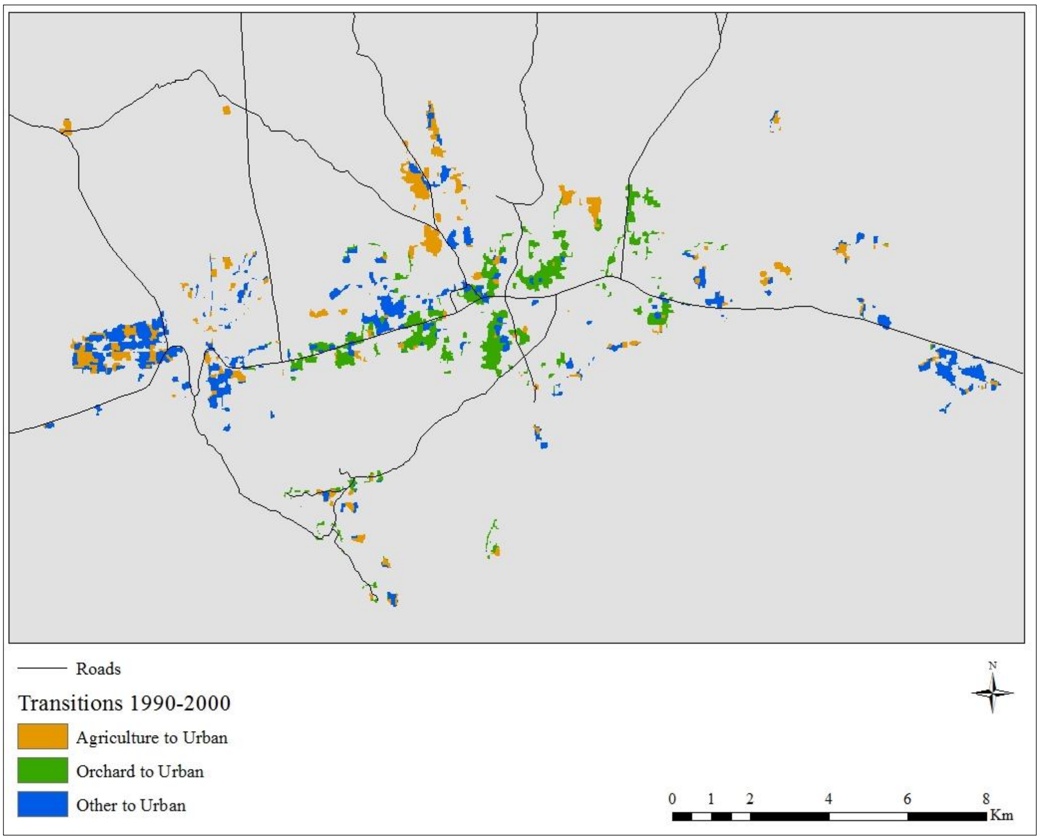

Figure 5-2: Transition from other categories to urban (1990-2000)

Table 5-1: Transition from other categories to urban (1990-2000)

Transition	Area (ha)
Agriculture to Urban	384.31
Other to Urban	575.84
Orchard to Urban	460.69

In this study one of the important objectives was to observe the changes in orchard in order to observe the changes in this category in detail the figure 5-3 prepared, as we can observe the main losses from this category was to urban and the gains were mainly from agriculture and other class.

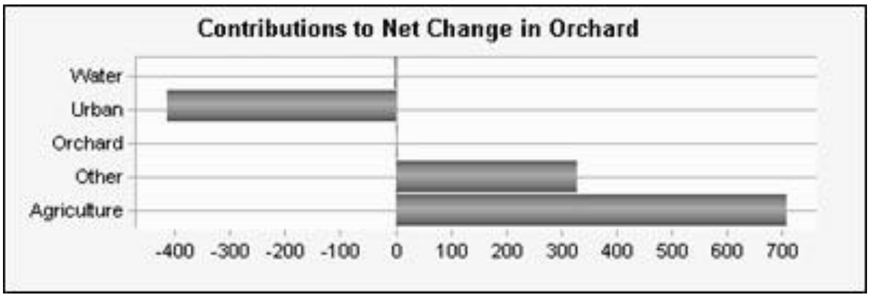

Figure 5-3 : Contribution from other categories to orchard (1990-2000)

From figure 5-4 it can be observed that there was a decrease in agricultural lands due transformation to urban class and other class. The transformation from agriculture to other class might be caused by changes in fertility of the land or resulted from misclasification. In image clasification phase it has been founded that the agriculture class and other usages has a similar reflectance values in some areas, this could mislead the process, because of that here we will mainly concentrate on transformation to urban class.

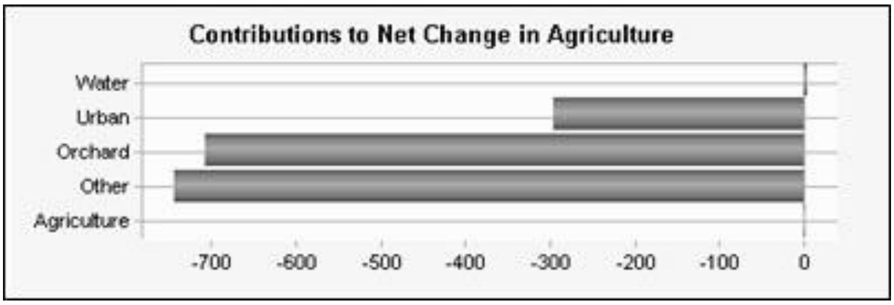

Figure 5-4 : Contribution from other categories to agriculture (1990-2000)

5.2 Changes between 2000 -2010

After observing the changes in the year 1990-2000, the changes in 2000-2010 also investigated in order to understand the change trend in the study area. As it can be observed from figure 5-5 in each class there are some gains and losses. The main transformations are from orchard and agriculture to urban area again.

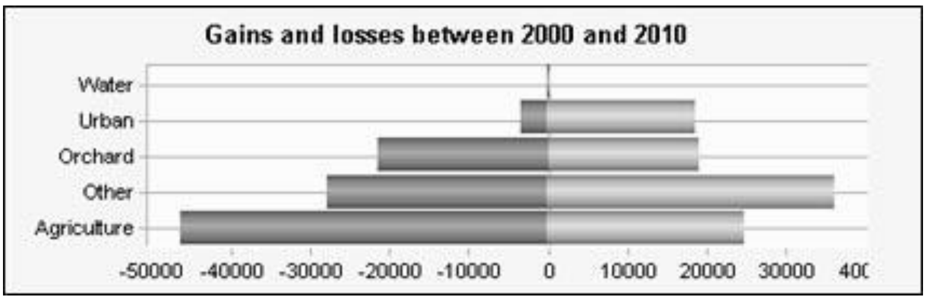

Figure 5-5: Gains and losses in each class (2000- 2010)

In order to understand the locational distribution of the transition to urban area illustrated in figure 5-6 and table 5-2 in order to observe the transformations according to location and quantity.

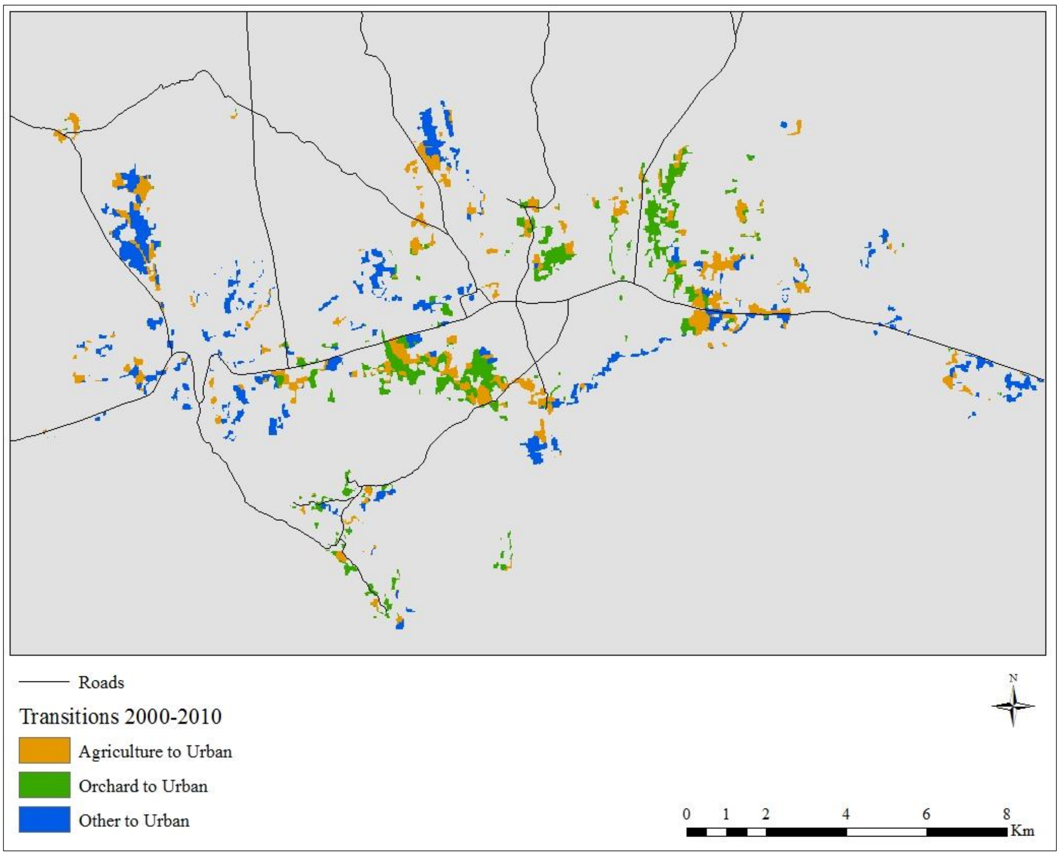

Figure 5-6: Transition from other categories to urban (2000-2010)

As we can observe from figure 5-6 and table 5-2 the urban area increased more than 1500 ha and this mainly observed in outskirts.

Table 5-2: Transition from other categories to urban (2000-2010)

Transition	Area (ha)
Agriculture to Urban	596.12
Other to Urban	614.25
Orchard to Urban	471.55

Furthermore the changes in orchard and agriculture has been observed for this period. As we can see from the figure 5-7 the decrease in orchards mainly resulted from urban area increase.

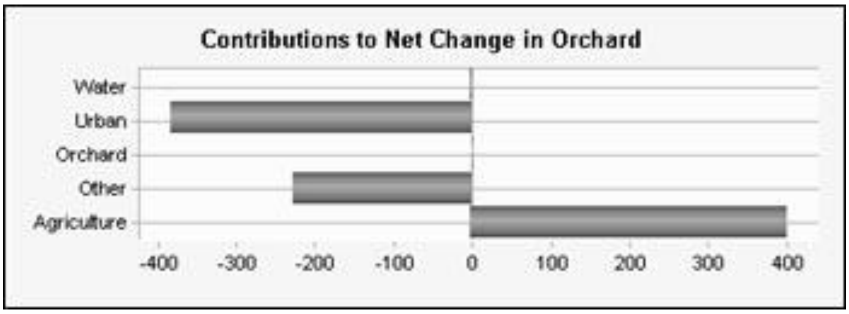

Figure 5-7: Contribution from other categories to orchard (2000-2010)

Moreover the agricultural changes mainly were resulted by other class and urban area increase. As it has been mentioned the transformation from agriculture to other might be resulted because of misclassification so that these changes ignored. As we can observe from the figure 5-8 there is a huge transformation between other class to agriculture which can be a result of misclassification of the classes in different years.

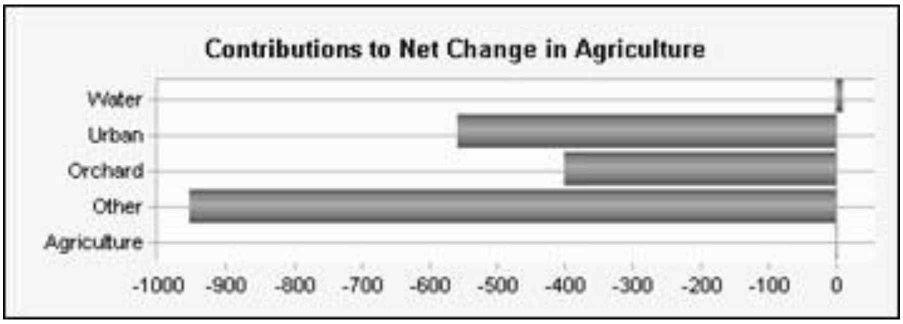

Figure 5-8: Contribution from other categories to agriculture (2000-2010)

6 LAND USE AND LAND COVER CHANGE MODELING

Many studies have been done for LULC change modeling and these studies can be categorized into agent based and pattern based models. Agent based models in other words actor based models are based on actors which are leading the simulation process. And pattern based models are mainly based on the spatial land cover data and the changes over the time (Agarwal et al., 2002).

The most common pattern based models are;

- Cellular Automata (CA)
- Artificial Neural Networks (ANN) and Markov Chain
- Spatial Statistics (Markov Chain)

For this research the combination of CA and MARKOV methods and the combination of MLP and MARKOV methods used , the results of the each model compared and the more successful method used for the simulation of the year 2020 LULC map.

6.1 Cellular Automata and Markov Chain

6.1.1 Markov Chain

The Markov Chain mathematical model invented by Andrei Andreyevich Markov in 1906, it is a Markov process where space is discrete [32, 33] and it is a stochastic process which means it is mainly based on probabilities not certainties. The Markov Chain is a random process based on Markov property which means the next state depends only on current state not the sequence (Markov Chain, n.d.).
The basic assumption in the model is that; the state at some point in the future (t+1) can be determined as a function of the current state (t), in other words the future change will be only depend on the existing change, so the transition between two times can be modeled mathematically (Iacono & Levinson, 2012). Mathematically it can be formulated as;

$$\mathbf{X_{t+1} = f(X_t)}$$

This model firstly has been used by socioeconomic researchers in 1950s but after the 1960s it started to be used in urban studies too, and for LULC modeling it had been used in the late 1970s (Iacono & Levinson et al., 2012). The first usage of this model was on parcel level, later by Bell in 1974 it has been used for remote sensing sources as well (pixel based) (Bell, 1974).

6.1.1.1 Implementation of Markov Chain to Study Area

The variables used for the model summarized in table 6-1. In the study area the t is LULC map of 1990 and t+1 are LULC map of 2000 to predict the 2010.

Table 6-1: Markov Chain variables

Variables	Value
t	1990 LULC
t+1	2000 LULC
Time periods	10

According to MARKOV model the relation between 1990 and 2000 has been modeled to be used in the 2010 simulation. The results of the MARKOV model are;

A transition matrix: This contains the probability of each land use/cover category which can change to every other category (Eastman, 2012). The probability of each class can be observed in table 6-2.

Table 6-2: Markov conditional probability of changing among LULC type

Class	Agriculture	Other	Orchard	Urban	Water
Agriculture	0.6164	0.2107	0.1385	0.0342	0.0001
Other	0.0891	0.8441	0.035	0.0316	0.0002
Orchard	0.098	0.0359	0.8124	0.0533	0.0004
Urban	0.0357	0.0446	0.0197	0.9	0.0001
Water	0.1234	0.0231	0.0283	0.0411	0.7841

A transition areas matrix: This contains the number of pixels which are expected to change from each land use/cover type to each other land use/cover type over the specified time period (Eastman, 2012).The number of cells expected to be transformed to other classes can be observed in table 6-3.

Table 6-3: Cells expected to be transformed to other classes

Class	Agriculture	Other	Orchard	Urban	Water
Agriculture	65159	22275	14642	3617	8
Other	18038	170945	7098	6397	37
Orchard	10125	3707	83906	5502	38
Urban	1449	1811	798	36560	3
Water	48	9	11	16	304

Conditional Probability Images: This reports the probability that each land use/cover type would be found at each pixel after the specified time period (Eastman, 2012). As it can be observed from figure 6-1 the probability is in scale of 0-1 and the pink represent the higher probability which is 1 and the black represent the lowest probability which is 0.

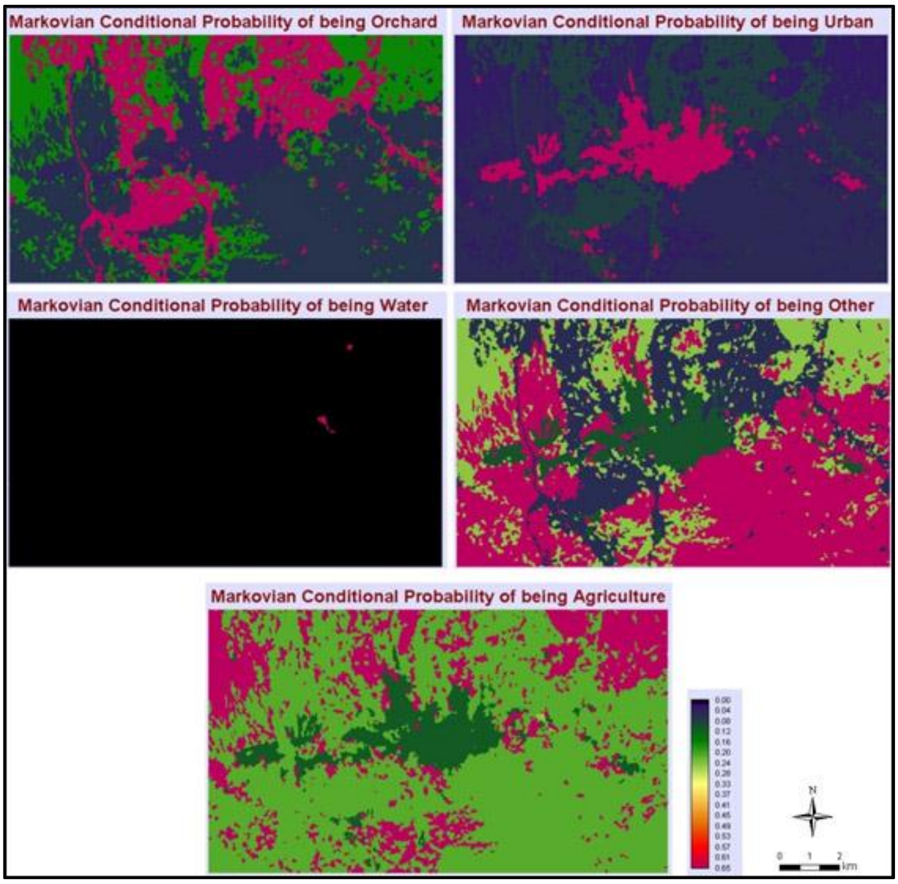

Figure 6-1: Markov conditional probability images

Based on the other studies it has been observed that the Markov Chain doesn't gave successful results because of weaknesses in the spatial side, so it has been used with the combination of different methods (Ye& Bai, 2008).

6.1.2 Cellular Automata

Cellular Automata, invented by John von Neumann, is a discrete dynamical system which models complex behaviors based on simple rules which are animating the cells on a lattice (Zalta, 2012). This method used by many geographers and mathematicians, for instance Tobler tried to explain the world by using cells, interactions with the neighboring cells (Tobler, 1975). Moreover Game of life developed by British mathematician John Horton Conway in 1970 was also using the CA, which was also based on cells and transition rules. In the 1980s the Helen Couclelis, Professor of Geography at the University of California, used CA for urban modeling purposes (Couclelis, 1985). The main components of CA model are "cells", "states", "neighborhood" and "transition rules". It is called discrete dynamic system which means the state of each cell at time t+1 determined by the state of its neighboring cells at the time t which lead to develop the transition rules.

To understand the theory behind the CA an example of linear CA has been used. In a linear CA for the one cell there are three neighborhoods, so the next state of the cell is dependent on these neighborhoods, the next state qi (t+1) of a cell is assumed to be dependent only on itself and on its two neighbors (Maji & Shaw, 2003), mathematical notation of this process is:

$\mathbf{q_i(t+1)= \mathit{f}\,(q_{\,i-1}(t).\; q_i\,(t).q_{\,i+1}(t))}$ **(Maji & Shaw, 2003)**

*Where $q_i(t)$ represents the state of the i^{th} cell at t^{th} instant of time. '**f**' is the next state function and referred to as the rule of the Automata*. (Maji & Shaw, 2003)

Here we try to explain the CA and its components with basic linear model (Shiffman, 2012):

Cells: The CA is composed of cells which are the smallest piece of the system. In this example the array is composed of the cells.

States: It means the value (state) in the each cell (pixel), in this example the states are 0 and 1.

0	1	0	0	1	0	0	0	1

Neighborhood: It explains how the cells connected to other cells. In this figure we can observe the neighborhood in one dimension. Neighborhood for any cell in this system would be the cell itself and its two adjacent neighbors which colored as gray in the figure (Artificial Neural Networks, n.d.).

0	1	0	0	**1**	0	0	0	1

Transition Rules: These rules are generated according to the neighborhoods of the cells. For example in the figure there are sample of transition rules which have been calculated according to the cells and their neighborhoods.

0 0 0	0 1 0	0 1 1	1 0 1	1 1 0	1 1 1
0	0	1	0	1	0

The main aim of the CA is to compute the next state from the current state. Generation 0 in the figure is the current state and we are trying to find the value in generation 1 (next state). In order to find the value we need to consider the transition rules.

Generation 0

0	1	0	0	1	0	0	0	

Generation 1

				?				

The rule set is:

0 0 0	0 1 0	0 1 1	1 0 1	1 1 0	1 1 1
0	0	1	0	1	0

According to the rule set the number with "**?**" in the new generation will be 0.

When we came to the LULC changes the logic is same, the change of the land classes predicted according to the cells and their neighbors.

6.1.3 Combination of Cellular Automata and Markov Chain Methods

The first method for the simulation is the combination of CA and MARKOV methods. The reason of using them together is to give spatial dimension to the MARKOV model which is weak in spatial side (Ye& Bai, 2008). For this method the main variables are; the LULC data in this case it is LULC for the year 2000, Markov transition areas which are found by MARKOV model, a transition suitability image collection which includes the suitability images for the each class and contiguity filter in other words CA filter which is used for *generating a spatial explicit contiguity weighting factor to change the state of cells based on its neighborhoods* (Purves & Pacala, 2008). There are different contiguity filters, which can be 3X3, 5X5 or 7X7. For this research the 5X5 filter used which can be seen in figure 6-2.

0	0	1	0	0
0	1	1	1	0
1	1	1	1	1
0	1	1	1	0
0	0	1	0	0

Figure 6-2: Contiguity filter (5X5)

The suitability image collection includes the suitability image prepared in suitability chapter and the suitability maps for other classes. As it has been mentioned suitability of each class need different specialization, for this research the suitability of the urban class. For other classes the MARKOV conditional probability images, which are based on the probability of each class occurrence in each pixel according to past experiences, used as suitability images (figure 6-3). These images were in the scale of 0-1, but for the suitability image collection they need to be in the scale 0-255, for this reason this images converted to scale of 0-255 (figure 6-3). All the suitability images were combined as a collection and then used in the process. Another important issue in this method is the number of iterations, they can be specified by the user, and they be based on either six months or one year. For this research one year period for iteration has been used which means 10 iterations.

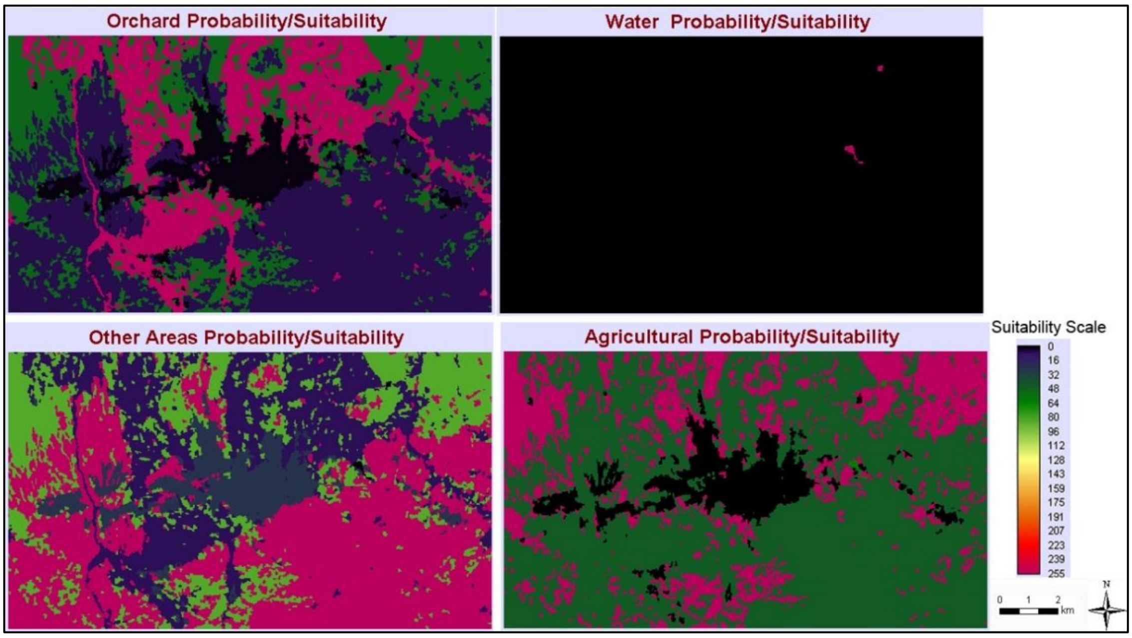

Figure 6-3: Suitability maps for other classes

6.1.3.1 Result of the Combination of Cellular Automata and Markov Chain Methods

After defining the all the variables model has been run and the result of the model can be observed in the figure 6-4 and in order to compare with the existing situation the 2010 LULC map resulted from image classification illustrated in figure 6-5.

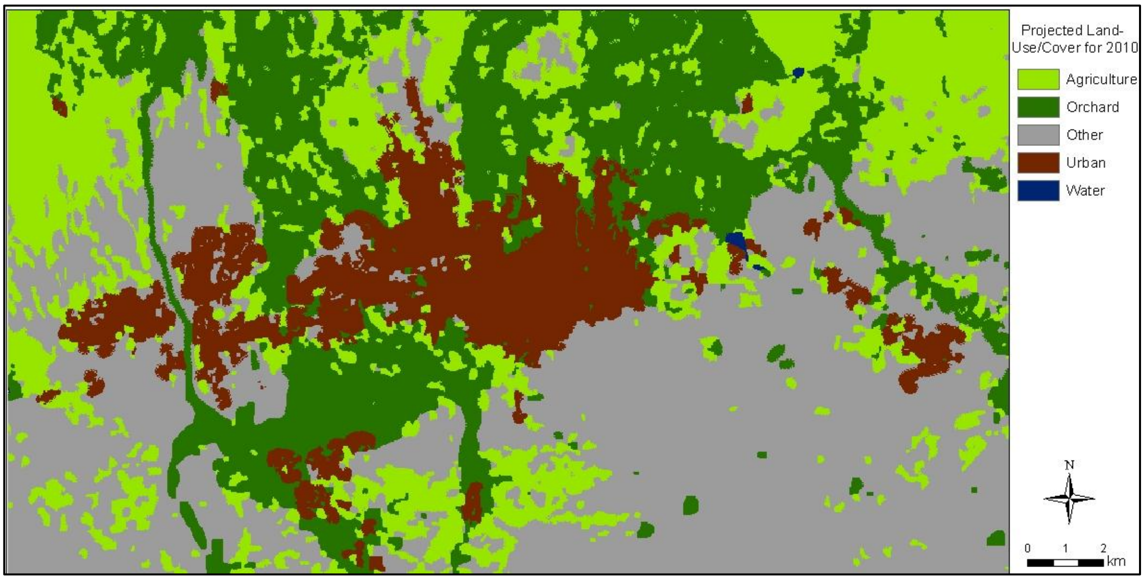

Figure 6-4: Image classification result for 2010 LULC map

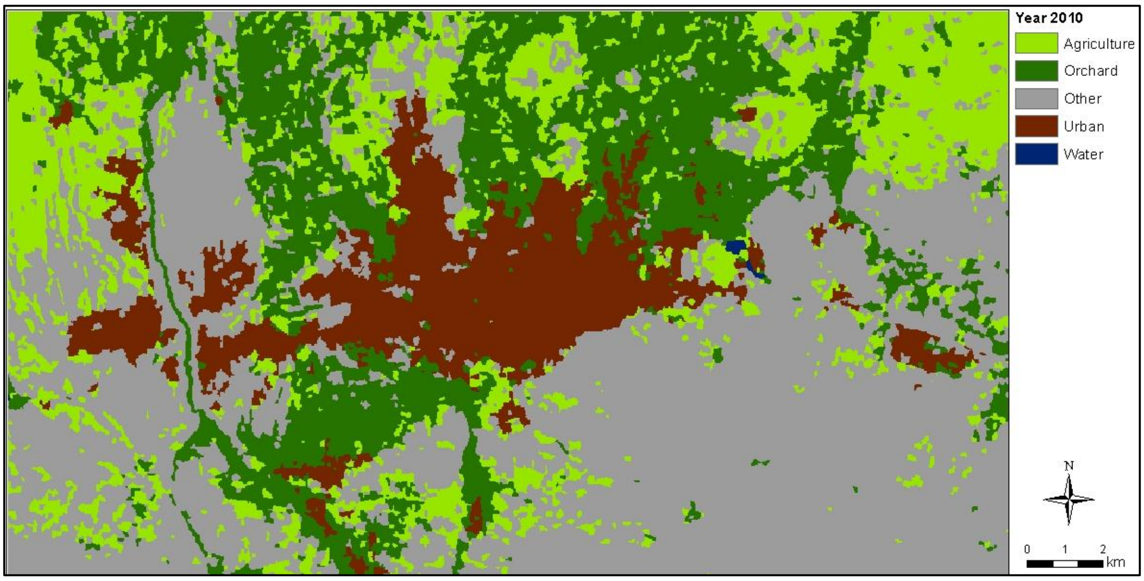

Figure 6-5: Projected LULC for 2010 (CA and MARKOV)

6.2 Multilayer Perceptron and Markov Chain

The second model which was used for modeling the future LULC of the study area was the combination of Markov Chain and Artificial Neural Network based methods. In this model ANN used for defining the spatial allocation of simulated LULC categories, in other works for transition potential development and Markov Chain used for prediction based on transition potentials.

The first stage of the method is detecting the changes; this has been explained in change detection section. The second step of the method is transition modeling which helps to create the transition maps from each LULC category to urban areas, which can be done by using the MLP method. The third step was the change modeling and the validation of the model with the accuracy assessment. Before the implementation of the model the theory of the ANN and MLP need to be understood.

6.2.1 Artificial Neural Networks

Artificial Neural Network is a mathematical model which is inspired from the human nervous system which interconnects the neurons for fulfills the complicated tasks in a short time. This model used for computers in order to solve the nonlinear systems (Artificial Neural Networks, n.d.). The components of the ANN are input, hidden and output layers, as it can be observed in figure 6-6.

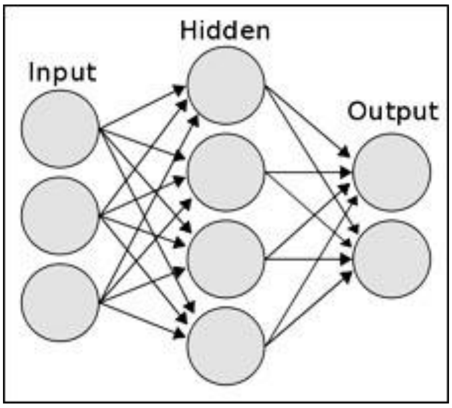

Figure 6-6: ANN structure (Artificial Neural Networks, n.d.)

6.2.2 Multilayer Perceptron

Multilayer Perceptron feed forward ANN is a network of simple neurons called the perceptron (Multilayer Perceptron, n.d.). It is composed of an input layer, output layer and hidden layers between input and output layer. It is a feed forward method which means data flows in one direction from input to output. The main algorithm of this model is computing the linear output from nonlinear inputs according to the weights by using a nonlinear activation function (Multilayer Perceptron, n.d.).
Mathematical notation is:

$$\boldsymbol{y} = \boldsymbol{\varphi}(\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{\omega_i}\, \boldsymbol{x_i} + \mathbf{b}) = \boldsymbol{\varphi}(\boldsymbol{w^T x} + \boldsymbol{b})$$

Where $\boldsymbol{\omega}$ denotes the vector of weights, $\boldsymbol{x}$ is the vector of inputs; **b** is the bias and $\boldsymbol{\varphi}$ is the activation function.
MLP is trained by using the back propagation algorithm. Back-propagation algorithm is composed of two steps which are forward pass and backward pass. In the first step activation transmits from input to output layer. In second step errors propagated from output to hidden layer (Multilayer Perceptron, n.d.) as it has been illustrated in figure 6-7, as a result of these interaction the system would be trained.

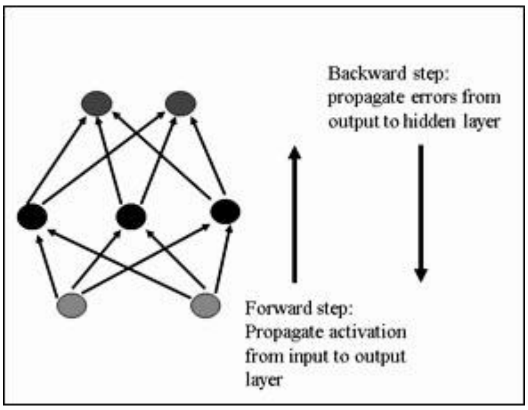

Figure 6-7: Back propagation

6.2.3 Combination of Multilayer Perceptron and Markov Chain Methods

The first stage this combined method is transition potential development with the MLP method. This stage is composed of three sub-stages which are; sub model development, testing the variables and transition potential development. Second stage of the model is simulation with the MARKOV method by using transition potentials.

6.2.3.1 **T*ransition Sub Model Development***

For this research the effect of the urban area was the main indicator, for this reason the changes from other categories to urban used as a transition sub model. In order to model the changes, the main variables which have the effect on the changes need to be detected. These variables are urban suitability which has been explained in suitability section, digital elevation model (DEM) (figure6-9) and the likelihood of transformation from other land uses to the urban. Suitability image and DEM were ready to use for the modeling, although the likelihood image prepared in this stage. For creating this map we need to use the change map (changes from all-to-urban) which has been created on change section and the existing land cover of 1990. By combining these two maps with a variable transformation method we can prepare the likelihood image. There are different methods for variable transformation; which are natural log, exponential, logit, square root, power, and evidence likelihood. For this research evidence likelihood method has been used. Evidence likelihood transformation method is a statistical method for adding the categorical variables into the model. The result of the likelihood image can be observed in figure 6-8.

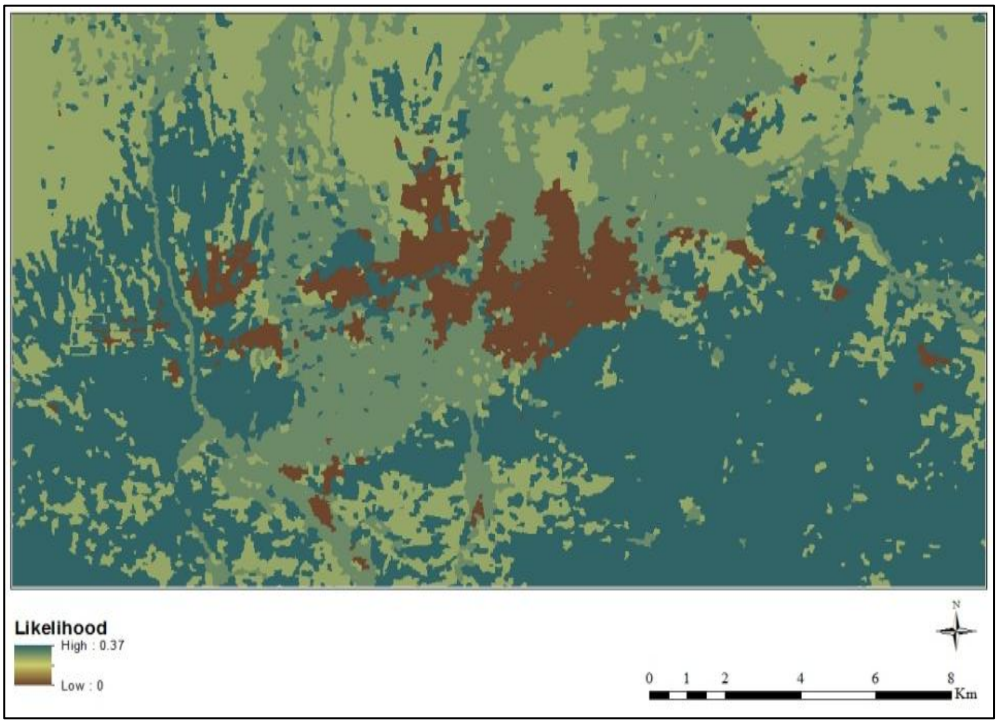

Figure 6-8: Likelihood

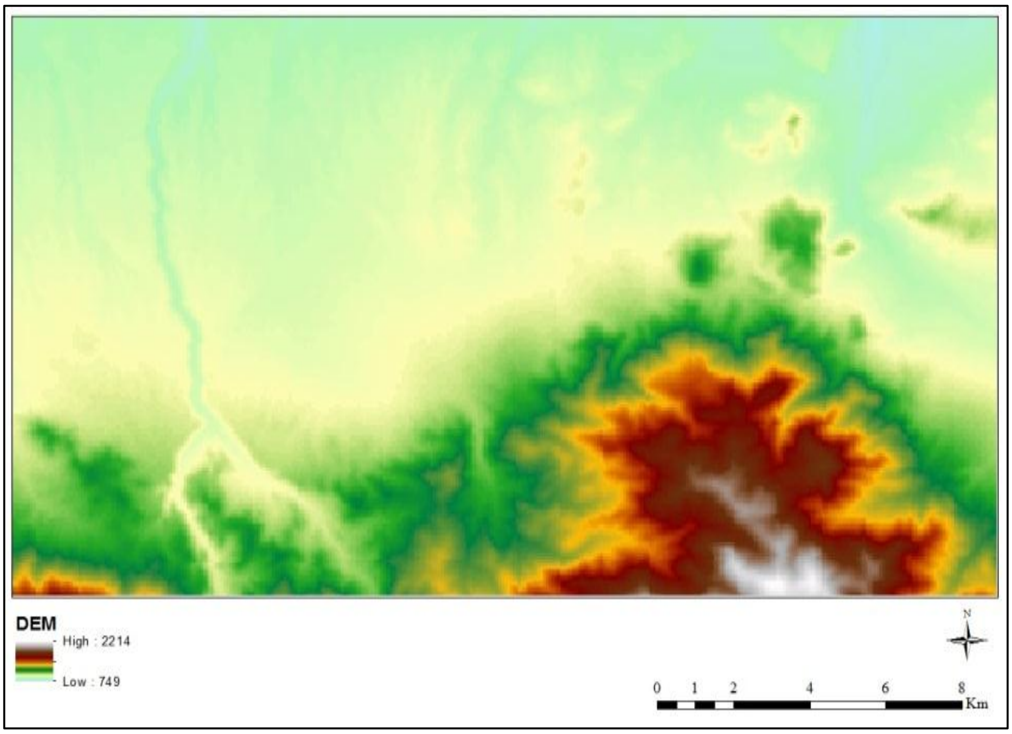

Figure 6-9: Digital elevation model

6.2.3.2 Testing the Selected Variables

Cramer’s V, which is a measure of association between two nominal variables, in the scale of 0-1 (Cramer, 1999), used to determine the association between change and the variables. While deciding the variables the ones which have a Cramer’s V value higher than 0.15 are indicated as useful and the ones which have Cramer’s V higher than 0.4 indicated as good (Eastman, 2009). As we can observe from the table 6-4 all the variables selected for transition development were higher than 0.15, some of them were higher than 0.4 which indicates the selected variables have an association with the changes and we can use them in the process.

Table 6-4: Cramer’s V for each variable (2010)

Variable	Overall Cramer's V	Urban	Other	Orchard	Agriculture	Water
Likelihood	0.5968	0.7318	0.7313	0.7167	0.5603	0.0889
DEM	0.3612	0.5671	0.4009	0.3796	0.324	0.079
Urban Suitability	0.4433	0.7504	0.4599	0.3988	0.19545	0.0689

6.2.3.3 Transition Potential Development

In order to calculate and map these potentials MLP, which is a successful method in solving complex systems, has been used. As it has been explained in MLP definition the one of the basic component of the method is input layer, in this study the input layers are LULC maps, DEM, suitability map and likelihood map.

6.2.3.4 Modeling With the Combination of Multilayer Perceptron and Markov Chain Methods

After calculating the transition potentials the final step was change modeling with the MARKOV method. This time variables in the MARKOV are different than the previous model. Here the model uses potentials not probabilities which have been created by MLP method. Moreover the suitability image used while creating the potentials, so they don't need to be used in the process again. In figure 6-10 we can observe the existing situation in 2010 and in the figure 6-11 the modeled map which has been prepared with this method. As it has been illustrated in the figures the result of the simulation is more or less similar to the existing situation but this need to be tested by accuracy assessment.

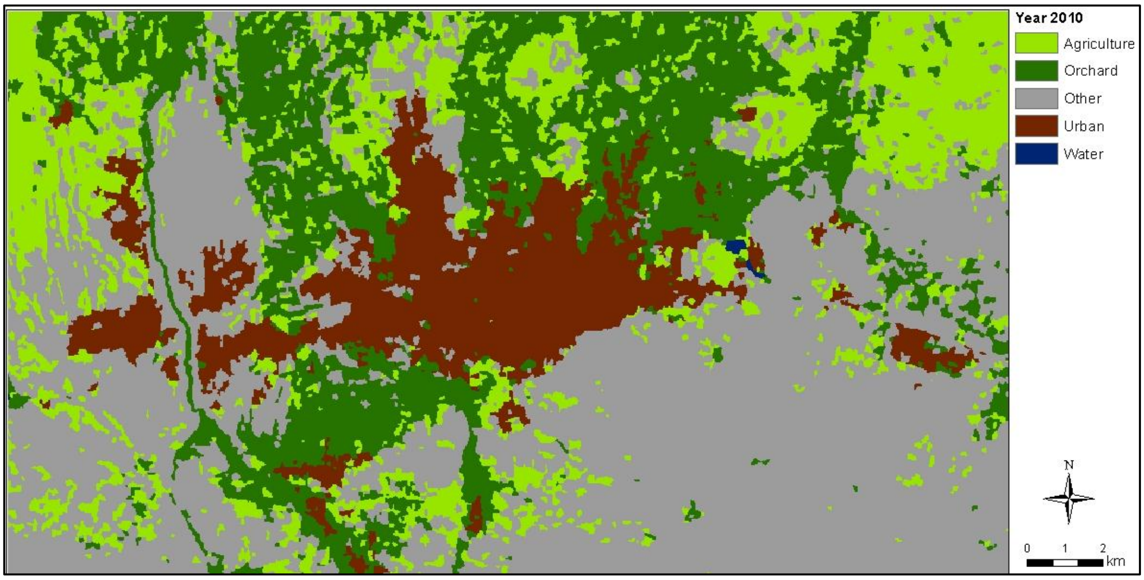

Figure 6-10: Image classification result for 2010 LULC map

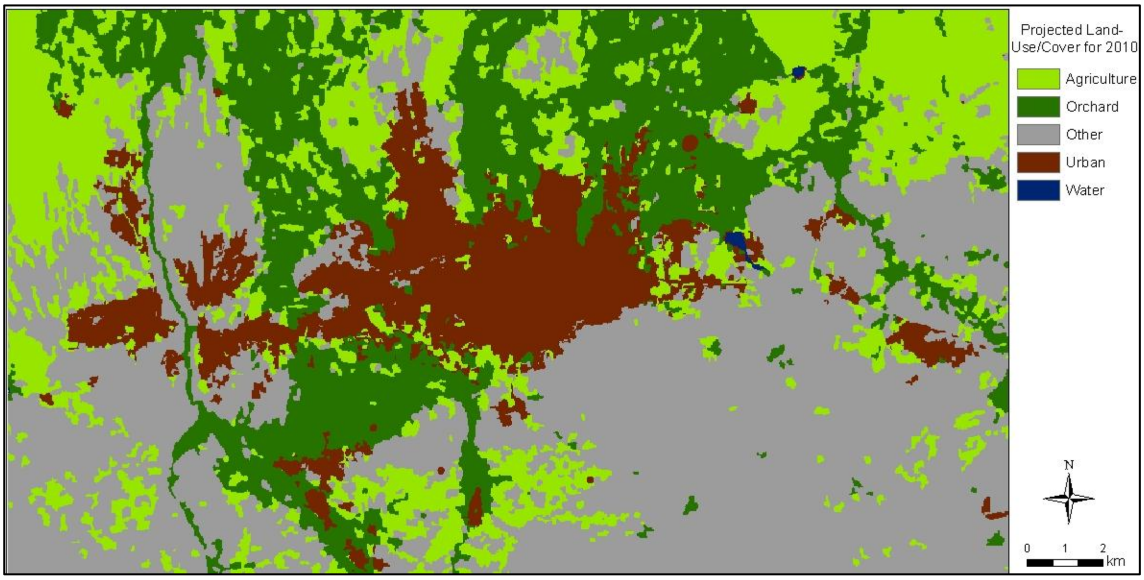

Figure 6-11: Projected LULC for 2010 (MLP and MARKOV)

6.3 Accuracy Assessment

The accuracy assessment of image classification and simulation results is slightly different, because in the latter case the accuracy assessed by using the maps not the specific points which are indicating the real category information (Rossiter, 2004). In many studies kappa statistical measurements has been used for the accuracy assessment but Pontius & Millones (2011) proposed using two simple parameters which are quantity disagreement and allocation disagreement instead of kappa. Quantity disagreement was defined as the difference between two maps due to an unsatisfactory match in the overall proportions of mapped land categories (Pontius & Millones, 2011). Allocation disagreement was defined as the difference between two maps due to an unsatisfactory match between the spatial allocations of the mapped land categories (Pontius & Millones, 2011). For this study kappa agreements and quantity-allocation disagreements have been calculated and according to the results the more successful method has been selected for the simulation. Before performing the calculations for the quantity-allocation disagreement and kappa, comparison between two reference times and the predicted map has been performed for the accuracy assessment. In this step the main components were reference time 1 (2000), reference time 2 (2010) and the predicted map (2010). The aim was to observe the main changes between the years 2000, 2010 and then comparing the results with the predicted map, in order to understand if the overall changes have been predicted correctly or not.

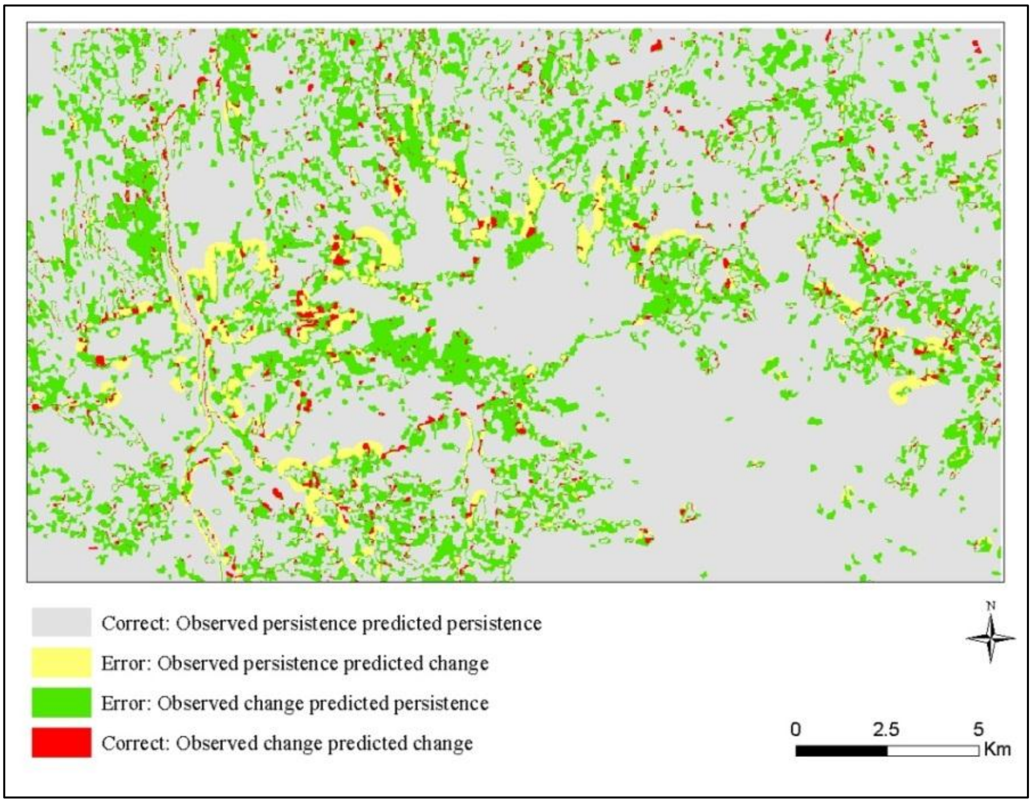

Figure 6-12: Prediction correctness and error based on 2000 (reference), the 2010 (reference) and 2010 (simulation result of CA and MARKOV) LULC maps

The results for the first model can be observed in figure 6-12. It illustrates the comparison of the observed change with the predicted change and shows four types of correctness and error which are; observed persistence predicted as persistence (correct, null successes, color: grey), observed persistence predicted as change (error, color: yellow), observed change predicted as persistence (error, color: green), observed change predicted as change (correct, color: red) (Martins, Silva & Cabral, 2012). As we can observe from figure 6-12 the overall map was covered by yellow and green areas where the simulation failed to predict correctly. As a visual interpretation we can say that the combination of CA and MARKOV methods was not successful in the study area.

Whereas in the figure 6-13 which illustrates the accuracy assessment of the combination of MLP and MARKOV method was including more red areas which means it is more successful method than the previous. And when we look at the overall distribution of the correctly simulated classes we can see that this are mainly concentrated on the periphery of the urban area which means the model is successful in predicting the changes in the urban area. This result is not a coincidence, the variables used in the simulation process was mainly related to urban growth, so that the model predicted the urban growth better than the other classes.

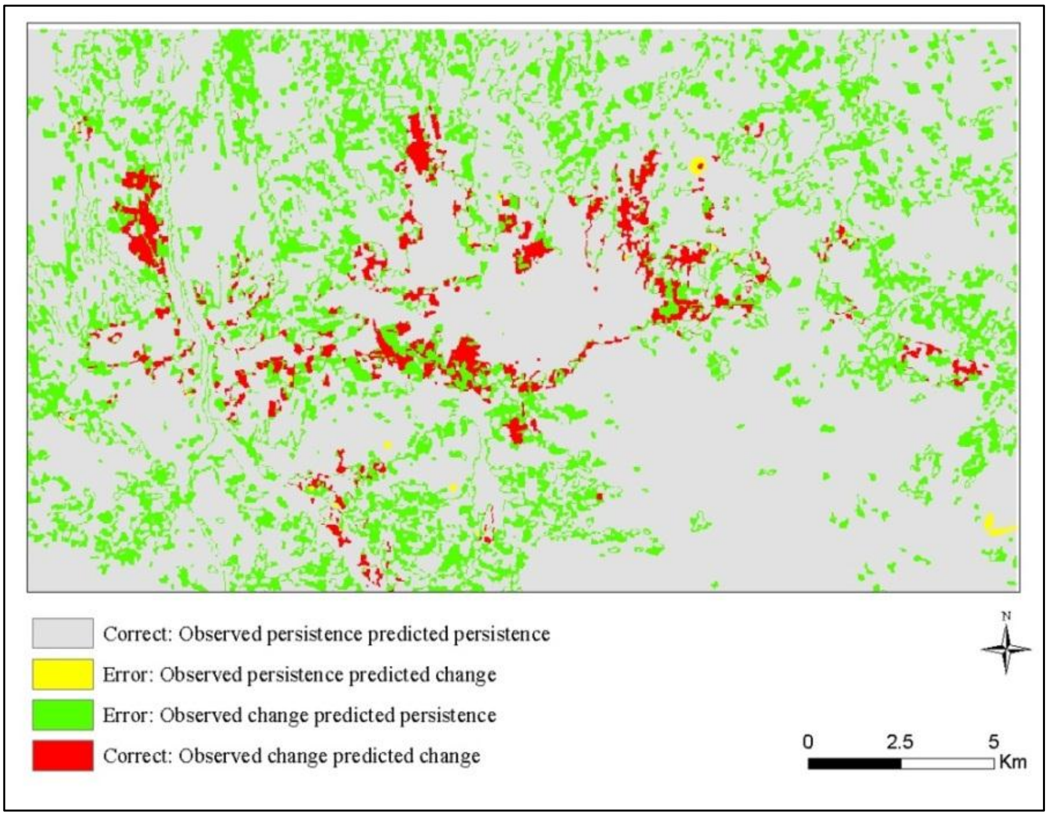

Figure 6-13: Prediction correctness and error based on 2000 (reference), 2010 (reference) and 2010 (simulation result of MLP and MARKOV) LULC maps

6.3.1 Quantity and Allocation Disagreements

In order to find the allocation and quantity disagreement cross tabulation method used to compare the maps and the matrix prepared by Pontius used to calculate the results. For more information about the calculations the paper written by Pontius & Millones (2011) can be referred.

The result of the assessment for the first model can observed in the table 6-5, and from there we can understand that the main disagreement between the maps was due to allocation which is around 19%. The quantity disagreement was only 6 %.

Table 6-5: Components of agreement and disagreement for the combination of CA and MARKOV methods results

Agrement or Disagrement	Value(%)*
Chance agreement	20
Quantity agreement	10
Allocation agreement	45
Allocation disagreement	**19**
Quantity disagreement	**6**

*The values in the table have been computed by entering the cross-tabulation matrix resulted from the comparison of 2010 classification result and the simulation results predicted by two different approaches into a spreadsheet available at http://www.clarku.edu/~rpontius. The crosstab matrixes can be found in appendix D.

The assessment for the second model can be observed in table 6-6. As it can be observed the overall disagreement was lower than the previous model, as a result we can say that the second model was more successful for predicting the changes compared to the first one.

Table 6-6: Components of agreement and disagreement for the combination of MLP and MARKOV methods results

Agrement or Disagrement	Value(%)*
Chance agreement	20
Quantity agreement	11
Allocation agreement	50
Allocation disagreement	**15**
Quantity disagreement	**4**

6.3.2 Kappa

The second method used for the accuracy assessment is calculating kappa values related to the location and the quantity in order to observe the accuracy in each class. These are Klocation and Khisto which demonstrate the source of error, whether it is location or quantity. Klocation shows the similarity in spatial distribution of classes but does not differentiate between classes that are close or distant and it is independent from the total number of cells per class (Serna, 2011) (Pontius, 2000). Khisto measures quantitative similarity between two maps, in other words it checks the similarity in the amount of the cells in testing map and reference map (Serna, 2011).

The accuracy assestment result for the first model can be observed in table 6-7 from this table it can be observed that the overall kappa is ~63 % and for the urban class is ~60 %. Urban area prediction is important for this research so the accuracy of the urban should be high.

Table 6-7: Accuracy assessment (CA and MARKOV)

CA and MARKOV for 2010	**Per Category Values**		
Class	Kappa	klocation	Khisto
Agriculture	0.531	0.629	0.844
Other	0.683	0.779	0.877
Orchard	**0.7**	**0.715**	0.979
Urban	**0.603**	**0.605**	0.996
Water	0.629	0.713	0.883
Overall Kappa	**0.638**	**0.699**	**0.913**

For the second model the result of the accuracy assestment can be observed in table 6-8. In the table we can observe that the overall accuracy is ~73 % which is higher previous model, especially the urban accuracy is ~91 % which is a really good agreement.

Table 6-8: Accuracy assessment (MLP and MARKOV)

MLP_MARKOV for 2010	**Per Category Values**		
Class	**Kappa**	**klocation**	**Khisto**
Agriculture	0.543	0.616	0.882
Other	0.741	0.793	0.934
Orchard	**0.771**	**0.787**	0.979
Urban	**0.912**	**0.914**	0.997
Water	0.657	0.793	0.829
Overall	**0.728**	**0.772**	**0.943**

As a result of both methods we can say that the overall accuracy is better in the second model so that for the future simulation the second model has been used.

6.4 Simulation for the Year 2020

According to the accuracy assessment it has been found that the accuracy value in the combination of MLP and MARKOV methods is higher than the combination of CA and MARKOV methods, because of that for the simulation of the year 2020 the combination of MLP and MARKOV methods has been selected.

For the simulation of the year 2020 the same steps in the simulation of the year 2010 followed. In this process the main inputs also are suitability image, DEM and likelihood images. The likelihood image prepared according to the changes between 2000- 2010 for predicting 2020. We can see the result of the likelihood image in the figure 6-14. After preparation of the variables they have been tested by Cramer's V. As we can observe from table 6-9 the values are higher than 0.15 even likelihood and suitability images are higher than 0.4, which shows a good relation between the changes and the variables. So we can include these variables for transition potential development.

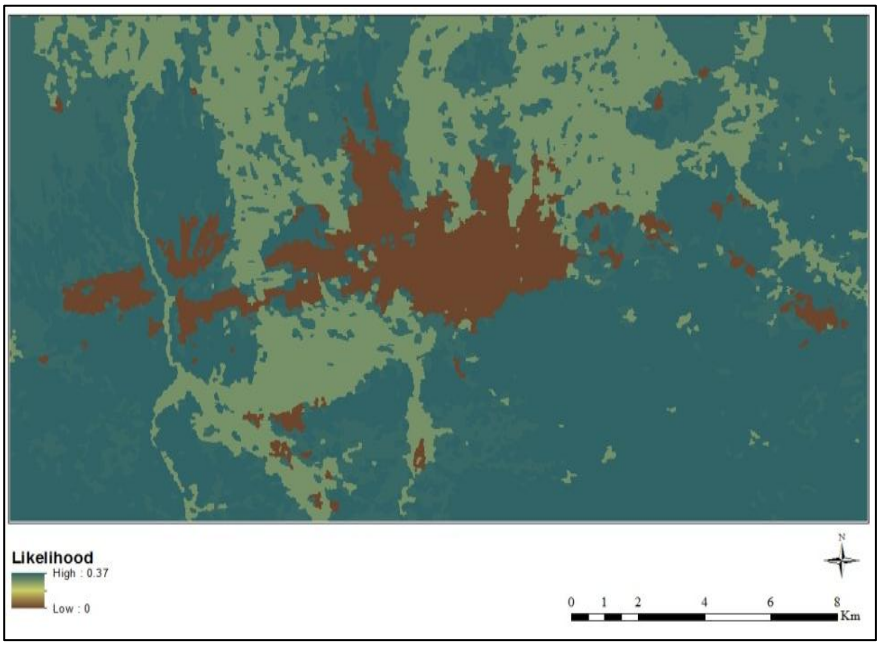

Figure 6-14: Likelihood

Table 6-9: Cramer's V for each variable (2020)

Variable	Overall Cramer's V	Urban	Other	Orchard	Agriculture	Water
Likelihood	0.6028	0.7515	0.7509	0.7284	0.5353	0.0632
DEM	0.3638	0.5491	0.4321	0.3887	0.3146	0.0825
Urban Suitability	0.5472	0.9925	0.5160	0.4295	0.2389	0.0481

After the preparation of potential maps the MARKOV method has been used for modeling. The transition probability matrix and matrix including the cells from each class expected to be transformed to other classes which have been used in the modeling process can be observed in table 6-10 and 6-11.

Table 6-10: Markov conditional probability of changing among LULC types

Class	Agriculture	Other	Orchard	Urban	Water
Agriculture	0.5631	0.2511	0.1227	0.0628	0.0003
Other	0.0787	0.8626	0.0249	0.0338	0
Orchard	0.0822	0.0734	0.7934	0.0509	0.0001
Urban	0.0103	0.0508	0.0244	0.9143	0.0002
Water	0.3582	0.0052	0.0567	0.018	0.5619

Table 6-11: Cells expected to be transformed into other classes

Class	Agriculture	Other	Orchard	Urban	Water
Agriculture	47584	21220	10367	5310	22
Other	16590	181906	5248	7127	7
Orchard	8302	7408	80099	5136	14
Urban	575	2840	1363	51095	12
Water	99	1	16	5	155

In the figure 6-15 we can observe the simulation result for the year 2020.As we can see the main growth direction is west and south-west. These areas are covered by many orchards and agricultural lands.

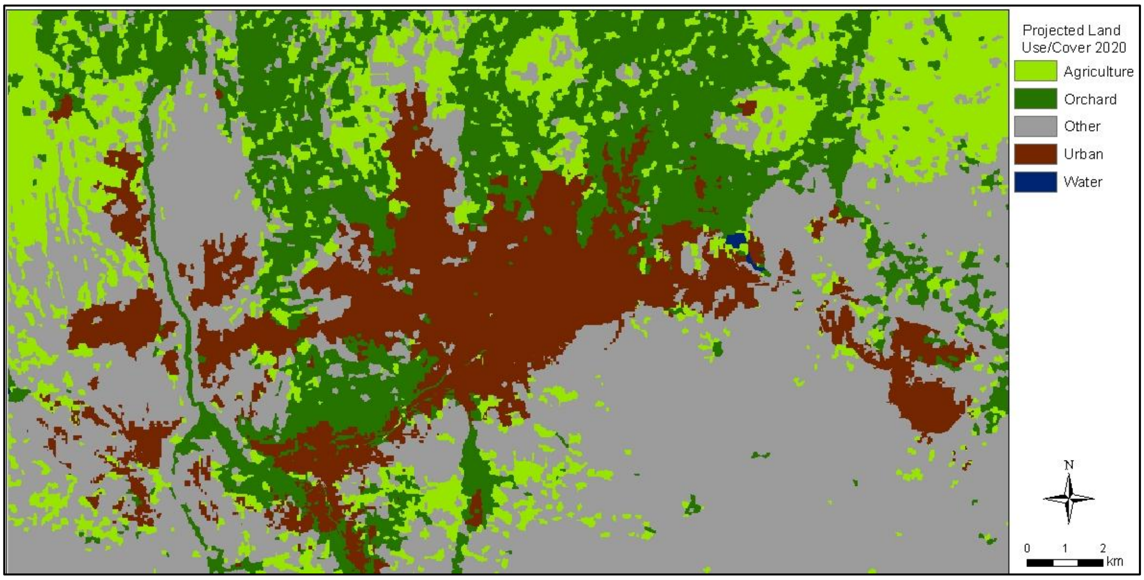

Figure 6-15 : Simulation for the 2020

6.4.1 Changes Based on Simulated 2020 Map (2010-2020)

After modeling the changes we can go further and see the changes in the each category by analyzing changes between 2010 and 2020. As we can observe from the figure 6-16 and table 6-12 the changes in the urban area are in the same trend with the previous years and if this trend followed 936 ha of agricultural land and orchard will be transformed to urban area in 2020.

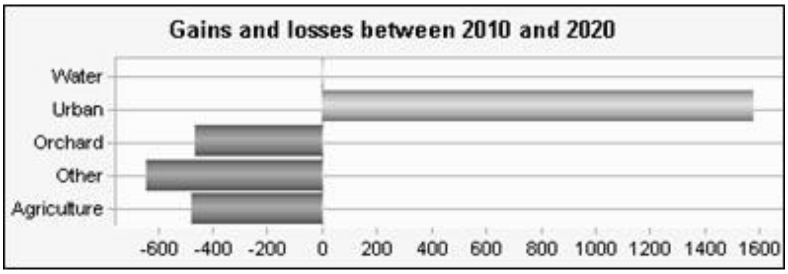

Figure 6-16 : Gains and losses in each category (2010-2020)

Table 6-12 : Transitions from each category to urban

Transition	Area (ha)
Agriculture to Urban	476.306
Other to Urban	639.742
Orchard to Urban	461.228

7 DISCUSSION AND CONCLUSION

This study prepared in order to analyze the existing patterns and to simulate the future LULC maps in order to lead the planning authorities for a sustainable development using the geospatial tools and techniques. While making this research some research questions defined to define the scope of work. After finishing the study the answers of those questions were summarized in this section in order to understand the success of the study both in methods and data used for the study and for the result of the study.

Are the available data adequate for this study?

Spatial problems are complex, it is not easy to simplify them and model them with a mathematical model. Especially LULC changes are depending on many drives which including social and economic drivers. However for this study some of the drivers accepted as constant. The existing trend and the suitability used as the main drivers. Consequently we can observe that in this context the data was more or less adequate. But we observe that with more data the study would give better results.

Are the Landsat images enough for acquiring the multi temporal data from images?

The first problem was about low resolution (30m) property of Landsat images which makes the classification difficult. Moreover beside the properties of the Landsat image the study area which includes many heterogeneous areas make the classification with Landsat images difficult. The overall accuracy for the classification was around 75% and according to Pontius (2000) the minimum requirement of the accuracy in classification should be around 80 % to explain the LULC categories. But for this study because of the resolution of the images and the heterogeneity of the study area we continue the study with the overall 75 % accuracy.

Are GIS and remote sensing tools adequate for this study?

The processing the data with GIS tools help us to solve the problem, but not every tool has the same capacity to solve the problems. Because of that different tools need to be used in the process, but converting the data always results with some lost, for this research this lost was not noticeable so we can say that the tools were useful.

Can the mathematical models help to model LULC changes?

For this study mainly two methods have been implemented namely the combination of CA and MARKOV and the combination of MLP and MARKOV methods. According to the accuracy assessment the second method has a lower disagreement for allocation and quantity which means it was more successful for the study area. But still the accuracy is lower than some other studies which is mainly resulted with the data has been used and the ignored effects of the other drivers. However in the urban area the model had 90 % agreement, because of this we can say the model was helpful to predict change in the urban category, but not for each category.

Where are the main changes and where is the main growth direction of the urban area according to simulated result for the year 2020?

The main growth is to the southwest and west which is the existing trend of the urbanization. According to the result of simulation for 2020 the trend of the urban area is fitting with the existing trend, if this cannot be changed the many orchards will be affected by this change. So the results can lead the planning authorities with a significant accuracy for the urban area.

7.1 Limitation of the Research

The main limitation of the research was finding the high resolution multi temporal images which could help to increase the accuracy of the classification later on the accuracy of the model and simulation. Unfortunately this data couldn't be found and low resolution Landsat images used for the research. The second limitation of the research was finding the multi temporal ground truth data in order to perform the supervised classification of the images. Without ground truth, the classification accuracy always is lower than expected. Third limitation was finding the some other data which could increase the accuracy of the model, such as socioeconomic and climate change data. Moreover the models has been used were limited to include these data to the modeling process. LULC changes are complex problems and there are different drivers which effect these changes. Socioeconomic and climate change data are important components of the changes because land changes are either result of human activities or natural changes.

7.2 Recommendations

For the future works the recommendation can be finding high resolution images and the multi temporal ground truth data in order to increase the accuracy of the classification. Secondly to solve the complex system there are different data requirements without finding the real reason of the changes the model would not be successful for this reason the other drivers which has a significant effect on the changes need to be investigated and the simulation need to be done according to this. For this we also need different approaches which will let the system use different data in the prediction process.

REFERENCES

Ahmed, B., & Ahmed, R. (2012). Modeling Urban Land Cover Growth Dynamics Using Multi‑Temporal Satellite Images: A Case Study of Dhaka, Bangladesh.” ISPRS International Journal of Geo-Information, 1(1), 3-31.

Adhikari, S., & Southworth, J. (2012). Simulating Forest Cover Changes of Bannerghatta National Park Based on a CA-Markov Model: A Remote Sensing Approach. Remote Sensing, 4(10), 3215-3243.

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

Ahmed, B. (2011). Urban Land Cover Change Detection Analysis And Modelling Spatio-Temporal Growth Dynamics Using Remote Sensing And GIS Techniques. Universitat Jaume I (UJI), Castellón, Spain, Universidade Nova de Lisboa (UNL), Lisbon, Portugal, Westfälische Wilhelms-Universität (WWU), Münster, Germany. M.Sc.

Araya, Y. H., & Cabral, P. (2010). Analysis and modeling of urban land cover change in Setubal and Sesimbra, Portugal. Remote Sensing, 2(6), 1549-1563.

Artificial Neural Networks. (n.d.). In Wikipedia. Retrieved November 18, 2012, from http://en.wikipedia.org/wiki/Artificial_neural_network

Band combinations. (n.d.). Retrieved September 30, 2012 from http://gdsc.nlr.nl/gdsc/en/information/earth_observation/band_combinations

Bell, E. (1974). Markov analysis of land use change: an application of stochastic processes to remotely sensed data. Socio-Economic Planning Sciences 8, 311–316.

Caetano M. (2009) ESA Advanced Training Course on Land Remote Sensing 28 june-3 July 2009 (Prague)

Cohen's kappa. (n.d.). In Wikipedia. Retrieved on October 10th, 2012, from http://en.wikipedia.org/wiki/Cohen%27s_kappa

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. Remote sensing of Environment, 37(1), 35-46.

Contingency table. (n.d.). In Wikipedia. Retrieved December 18, 2012, from http://en.wikipedia.org/wiki/Contingency_table

CORINE Land cover technical guidelines, (2000). European Environmental Agency.

Couclelis, H. (1985). Cellular Worlds: a framework for modeling micro-macro dynamics, Environment and Planning A, 17 -585-596.

Cramer, H. (1999). Mathematical Methods of Statistics (PMS-9), Princeton University press.

Cross Tabulation. (n.d.). Retrieved December 18, 2012, from http://www.createsurvey.com/docs/what_is_crosstab.html

Drobne, S. and A. Lisec (2009). Multi-attribute decision analysis in GIS: weighted linear combination and ordered weighted averaging. Informatica 33(4): 459-474.

Eastman JR. (2012). IDRISI Selva Manual. Worcester, MA: Clark Labs, Clark University. 322p. provided as a PDF with the IDRISI Selva software package.

Eastman, J.R. (2009). IDRISI Help System. Accessed in IDRISI Selva, Worcester, MA: Clark University.

Gao, Y., & Mas, J. F. (2008). A Comparison of the Performance of Pixel Based and Object Based Classifications over Images with Various Spatial Resolutions. Online journal of earth sciences, 2(1), 27-35.

García G., Feliú E., Esteve J. F., Soba M.P., Hazeu G., Rasmussen O.R., Galera-Limdblom P.& Banski J.(2010), EU-LUPA European Land Use Patterns, Inception Report, EPSON

Gecena, R., & Sarpb, G. (2008). Road Detection from High and Low Resolution Satellite Images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37.

GeoTIFF. (n.d.). In Wikipedia. Retrieved January 10, 2013 from, http://en.wikipedia.org/wiki/GeoTIFF

Geographic information system (n.d). In Wikipedia. Retrieved December 18, 2012, from http://en.wikipedia.org/wiki/Geographic_information_system

Ground Truth. (n.d.). In Wikipedia. Retrieved December 10, 2012 from, http://en.wikipedia.org/wiki/Ground_truth

Hopkins, L. D. (1977). Methods for generating land suitability maps: a comparative evaluation. Journal of the American Institute of Planners 43(4): 386-400.

Iacono, M., D. Levinson, et al. (2012). A Markov Chain Model of Land Use Change in the Twin Cities, 1958-2005.

Image Segmentation. (n.d.). In Wikipedia. Retrieved September 30, 2012, from http://en.wikipedia.org/wiki/Image_segmentation

Jantz, C. A., Goetz, S. J., & Shelley, M. K. (2004). Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area. Environment and Planning B, 31(2), 251-272.

Jiang, H. and J. R. Eastman (2000). Application of fuzzy measures in multi-criteria evaluation in GIS. International Journal of Geographical Information Science 14(2): 173-184.

Johnson, S. J. (2009). An evaluation of Land Change Modeler for ARCGIS for the ecological analysis of landscape composition. (Doctoral dissertation, Southern Illinois University).

Kordi, M. (2008). Comparison of fuzzy and crisp analytic hierarchy process (AHP) methods for spatial multicriteria decision analysis in GIS (Doctoral dissertation, University of Gävle).

Kumar, M., & Shaikh, V. R. (2012). Site Suitability Analysis for Urban Development Using GIS Based Multicriteria Evaluation Technique. Journal of the Indian Society of Remote Sensing, 1-8.

LaGro, J. A. (2001). Site analysis: linking program and concept in land planning and design. Wiley.

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

Lillesand, T. M., R. W. Kiefer, et al. (2004). Remote sensing and image interpretation, John Wiley & Sons Ltd.

Maji, P., C. Shaw, et al. (2003). Theory and application of cellular Automata for pattern classification. Fundamenta Informaticae 58(3): 321-354.

Malatya Province. (n.d.). In Wikipedia. Retrieved September 10, 2012, from http://en.wikipedia.org/wiki/Malatya_Province

Malczewski, J. (2004). GIS-based land-use suitability analysis: a critical overview. Progress in Planning 62(1): 3-65.

Markov Chain. (n.d.). In Wikipedia. Retrieved October 15, 2012, from http://en.wikipedia.org/wiki/Markov_Chain

Martins, V. N., e Silva, D. S., & Cabral, P. (2012). Social vulnerability assessment to seismic risk using Multicriteria analysis: the case study of Vila Franca do Campo (São Miguel Island, Azores, Portugal).Natural hazards, 1-20.

Mathematical Model. (n.d.). In Wikipedia. Retrieved September 30, 2012, from http://en.wikipedia.org/wiki/Mathematical_model

Mather, P., & Koch, M. (2011). Computer processing of remotely-sensed images: an introduction. Wiley.

Morisette, J. T., & Khorram, S. (2000). Accuracy assessment curves for satellite-based change detection. PE & RS- Photogrammetric Engineering & Remote Sensing, 66(7), 875-880.

Multilayer Perceptron. (n.d.). Retrieved December 30, 2012 from http://www.hiit.fi/u/ahonkela/dippa/node41.html

Multilayer Perceptron. (n.d.). In Wikipedia. Retrieved November 15, 2012, from, http://en.wikipedia.org/wiki/Multilayer_perceptron

Pontius Jr, R. G., & Millones, M. (2011). Death to Kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment. International Journal of Remote Sensing, 32(15), 4407-4429.

Pontius, R. G. (2000). Quantification error versus location error in comparison of categorical maps. Photogrammetric Engineering and Remote Sensing 66(8): 1011-1016.

Pontius, R. G. and M. L. Cheuk (2006). A generalized cross-tabulation matrix to compare soft-classified maps at multiple resolutions. International Journal of Geographical Information Science 20(1): 1-30.

Pontius, R., and Chen, H. (2006). Land change modeling with GEOMOD. Idrisi Kilimanjaro, 2006

Prakash, T. (2003). Land suitability analysis for agricultural crops: A fuzzy Multicriteria Decision Making Approach. MS Theses international institute for geo-information science and earth observation Enschede, the Netherland.

Protection and Utilization of Agricultural Lands (2005), Retrieved December 30, from http://www.mevzuat.adalet.gov.tr/html/23117.html

Protection and Utilization of Archeological Zones (1999). Retrieved February 1, 2013, from http://teftis.kulturturizm.gov.tr/TR,13917/658-nolu-ilke-karari-arkeolojik-sitler-koruma-ve-kullan-.html

Purves, D. and S. Pacala (2008). Predictive models of forest dynamics. Science 320(5882): 1452-1453.

Rossiter, D. G. (2004). Technical Note: Statistical methods for accuracy assessment of classified thematic maps. Enschede (NL): International Institute for Geo-information Science & Earth Observation (ITC)

Rule for protection of culturally and naturally important areas (1983). Retrieved on 1st of February 2013 from, http://www.mevzuat.adalet.gov.tr/html/628.html

Schowengerdt, Robert A. (2007). Remote sensing: models and methods for image processing. Academic Press.

Serna, E. (2011). Quantifying error in vegetation mapping. (Doctoral dissertation, University of Missouri).

Shapiro, L., & Stockman, G. C. (2001). Computer Vision, pp 279-325, New Jersey, Prentice-Hall, ISBN 0-13-030796-3

Shiffman D. (2012). The Nature of Code, California, USA. (Retrieved December 15,from http://natureofcode.com/book/)

Singh, A. (1989). Review Article Digital change detection techniques using remotely-sensed data. International journal of remote sensing, 10(6), 989-1003.

Tobler, W. (1979). Cellular Geography. Philosophy in geography, 9, 379-386.

Torrens, P. M. (2000). How cellular models of urban systems work. WP-28, Centre for Advanced Spatial Analysis, University College London

Tudes, S., & Yigiter, N. D. (2010). Preparation of land use planning model using GIS based on AHP: case study Adana-Turkey. Bulletin of Engineering Geology and the Environment, 69(2), 235-245.

Turkish Statistical Institute (TurkStat). Retrieved October 10, 2012, from http://www.turkstat.gov.tr/Start.do

Ünal, M.R. (2010). Kayisi Araştirma Raporu (Apricot Research Report), Firat Development Agency Reports.

Wang, L., Sousa, W. P., & Gong, P. (2004). Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. International Journal of Remote Sensing, 25(24), 5655-5668.

Ye, B., & Bai, Z. (2008). Simulating land use/cover changes of Nenjiang County based on CA-Markov model. Computer and Computing Technologies in Agriculture, Volume I, 321-329.

Yu, J., Chen, Y., & Wu, J. P. (2009, July). Cellular Automata and GIS based landuse suitability simulation for irrigated agriculture. In 18th World IMACS/MODSIM Congress, Cairns, Australia (p. 7).

Zalta Edward N. (2012). Cellular Automata: The Stanford Encyclopedia of Philosophy.

APPENDIX A

Accuracy Assessment

Accuracy assessment is the way to determine the quality of the information (Congalton, 1991). In order to explain the accuracy assessment there are some terms which need to be defined. **Kappa which is a widely used method for accuracy assessment** is a statistical measure of overall agreement between two categorical items (Cohen's kappa, n.d.). It can be calculated with the following formula:

$$\hat{k} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \cdot x_{+i})}$$

Where:
r = number of rows in the error matrix
x_{ii} = the number of observations in row i and column i
x_{i+} = total of observations in row i
x_{+i} = total of observations in column i
$\mathbf{N^2}$ = total number of observations included in the matrix.
(Congalton, 1991)
Kappa value changes from -1 to +1 (Ahmed, 2011), and the interpretation of the values can be determined according to these values:

- < 0: Less than chance agreement
- 0.01–0.20: Slight agreement
- 0.21– 0.40: Fair agreement
- 0.41–0.60: Moderate agreement
- 0.61–0.80: Substantial agreement
- 0.81–0.99: Almost perfect agreement (Pontius, 2000)

Conditional Kappa

Overall Accuracy is the total Accuracy the ratio of correct plots to the total number of plots. **User's accuracy** corresponds to error of inclusion. Which can be calculated by: the number of samples correctly classified for a given row divided by the total of the row (Morisette & Khorram, 2000). **Producer's accuracy** corresponds to error of exclusion. The number of samples correctly classified for a given column divided by the total for that column (Morisette & Khorram, 2000). Below in the sample example it can be understood more easily, Sample is based on the examples of the paper prepared by Congalton (1991) (Congalton, 1991).

Table A-0-1: Sample error matrix

	Class types determined from a reference source				
Class types determined from classified map	# Plots	A	B	C	Totals
	A	30	4	7	41
	B	20	16	0	36
	C	10	7	6	23
	Totals	60	27	13	100

Total = $Accuracy_{Total} = \frac{30+16+6}{100} * 100 = 51\%$

User = $Accuracy_{User's,A} = \frac{30}{41} * 100 = 73\%$

Producer= $Accuracy_{producers,A} = \frac{30}{60} * 100 = 50\%$

According to This Calculations the Results for the study area images are:

Table A-0-2: 1990 Error matrix and accuracy values

Categories	Agriculture	Other	Orchard	Urban	Water	Total	User Accuracy
Agriculture	34	9	0	1	2	46	73.91
Other	10	45	2	5	0	62	72.58
Orchard	6	0	52	6	0	64	81.25
Urban	0	0	0	38	0	38	100
Water	0	0	0	0	17	17	100
Total	50	54	54	50	19	227	
Producer Accuracy	68	83.33	96.30	76	89.47		
Total Accuracy	81.94						

Overall Kappa	0.7685

Table A-0-3: 2000 Error matrix and accuracy values

Categories	Agriculture	Other	Orchard	Urban	Water	Total	User Accuracy
Agriculture	35	6	5	0	7	53	66.04
Other	7	46	1	3	1	58	79.31
Orchard	6	0	48	4	4	62	77.42
Urban	1	2	0	43	1	47	91.49
Water	0	0	0	0	6	6	100.00
Total	49	54	54	50	19	226	
Producer Accuracy	71.43	85.19	88.89	86.00	31.58		
Total Accuracy	78.76						

Overall Kappa	0.7256

Table A-0-4: 2010 Error Matrix and Accuracy Values

	Agriculture	Other	Orchard	Urban	Water	Total	User Accuracy
Agriculture	26	2	7	2	5	42	61.90
Other	15	49	0	2	1	67	73.13
Orchard	5	0	47	2	0	54	87.04
Urban	2	0	1	48	0	51	94.12
Water	0	0	0	0	13	13	100
Total	48	51	55	54	19	227	
Producer Accuracy	54.17	96.08	85.45	88.89	68.42		
Total Accuracy	80.62						

Overall Kappa	0.7511

APPENDIX B

CORINE Land Cover Maps

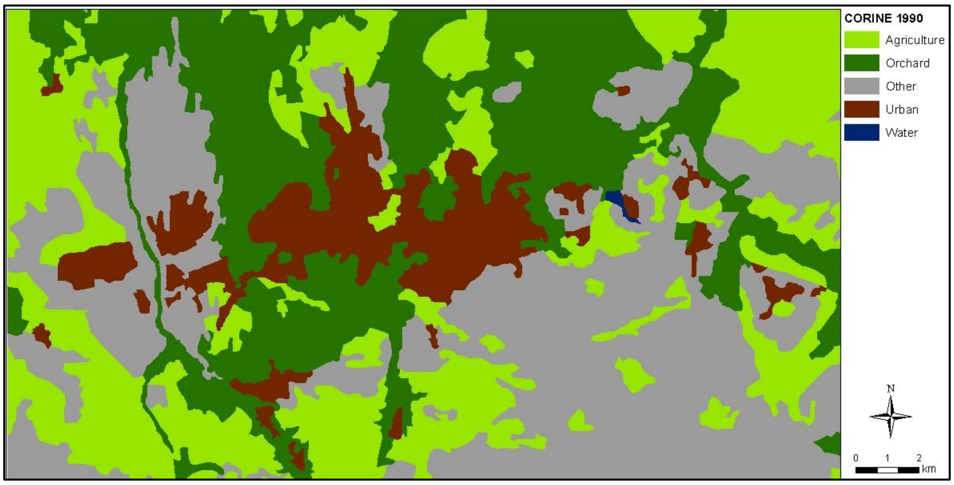

Figure B-0-1: CORINE 1990

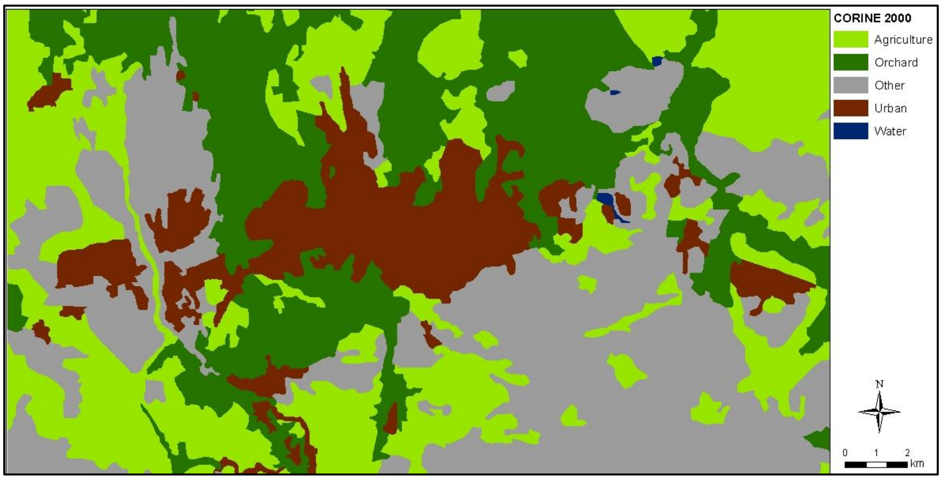

Figure B-0-2: CORINE 2000

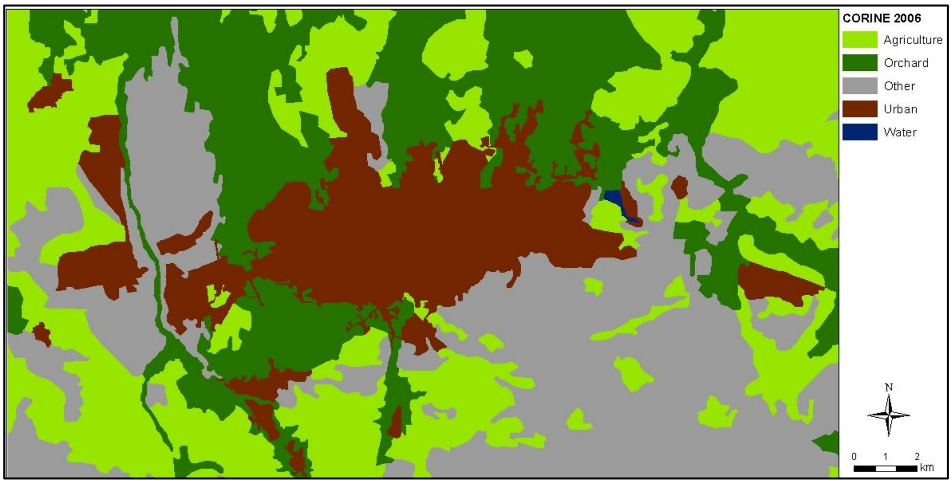

Figure B-0-3: CORINE 2006

APPENDIX C

Table C-0-1: Questionnaire forms filled by experts for suitability analysis

Questionnaire forms for Suitability Analysis of Malatya																
Factors	**Expert 1**	**Expert 2**	**Expert 3**	**Expert 4**	**Expert 5**	**Expert 6**	**Expert 7**	**Expert 8**	**Expert 9**	**Expert 10**	**Expert 11**	**Expert 12**	**Expert 13**	**Expert 14**	**Sum**	**Avarage**
Accesibility (Proximity to Roads)	16	14	18	19	20	18	20	19	20	15	14	10	19	15	237	17
Agriculture	9	11	9	17	9	9	13	11	15	10	16	10	16	10	165	12
Geology	25	20	15	13	19	23	21	20	15	17	18	19	15	24	264	19
Slope	16	17	26	18	15	15	15	16	18	18	19	18	17	20	248	18
Proximity to Existing Built-up	18	16	16	16	20	18	10	10	12	25	18	20	17	16	232	17
Previous Plan Boundaries	16	22	16	17	17	17	21	24	20	15	15	23	16	15	254	18

APPENDIX D

Cross Tabulation

*The cross tabulation is a statistical process that summarizes the categorical data to create a contingency matrix (Cross Tabulation, n.d). A contingency table is a type of table in a matrix format that displays the (multivariate) frequency distribution of the variables (Contingency table, n.d).*Cross tabulation matrix is very important in comparison of two categorical maps. The matrix structure includes the classes of the first map as the rows and the classes of the other map as the columns. (Pontius & Cheuk, 2006).By this method the agreement between the maps can be found. For more information about cross tabulation method with categorical maps the paper written by Pontius & Cheuk, (2006) can be referred.

Table D-0-1: Cross tabulation results of the combination of Cellular Automata and Markov Chain methods

	Reference Map (ClasificationResults)					
		Agriculture	Other	Orchard	Urban	Water
Model Results	Agriculture	60681	28354	12153	6996	39
	Other	12407	161764	2467	6754	7
	Orchard	9145	10769	78800	5566	9
	Urban	2147	9989	7519	36563	24
	Water	123	2	20	7	196

Table D-0-2: Cross tabulation Results of the combination of Multilayer perceptron and Markov Chain methods

	Reference Map					
		Agriculture	Other	Orchard	Urban	Water
Model Results	Agriculture	59381	26408	12963	3328	4
	Other	15925	174449	5036	700	7
	Orchard	8472	7542	81581	172	8
	Urban	586	2477	1357	51679	38
	Water	139	2	22	7	218