

APPLICATION OF SLEUTH MODEL IN ANTALYA

A THESIS SUBMITTED TO  
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES  
OF  
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ÖZLEM ŞEVİK

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR  
THE DEGREE OF MASTER OF SCIENCE  
IN  
GEODETIC AND GEOGRAPHIC INFORMATION TECHNOLOGIES

APRIL 2006

Approval of the Graduate School of Natural and Applied Sciences.

\_\_\_\_\_  
Prof.Dr. Canan Özgen  
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

\_\_\_\_\_  
Assist. Prof. Dr. Zuhale Akyürek  
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

\_\_\_\_\_  
Assist. Prof. Dr. Zuhale Akyürek  
Supervisor

Examining Committee Members

Assoc Prof. Dr. Oğuz Işık (METU, CP) \_\_\_\_\_

Assist. Prof. Dr. Zuhale Akyürek (METU, CE) \_\_\_\_\_

Prof. Dr.Vedat Toprak (METU, GEOE) \_\_\_\_\_

Assoc. Prof. Dr. Nurünnisa Usul (METU, CE) \_\_\_\_\_

Assist. Prof. Dr. Ayşegül AKSOY (METU, ENVE) \_\_\_\_\_

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Last name: Özlem Şevik

Signature :

## ABSTRACT

### APPLICATION OF SLEUTH MODEL IN ANTALYA

ŞEVİK, ÖZLEM

M.Sc., Department of Geodetic and Geographic Information Technologies

Supervisor: Assist. Prof. Dr. Zuhal Akyürek

April 2006, 112 pages

In this study, an urban growth model is used to simulate the urban growth in 2025 in the Antalya Metropolitan Area. It is the fastest growing metropolis in Turkey with a population growth of 41,79‰, although Turkey's growth is 18,28‰ for the last decade.

An Urban Growth Model (SLEUTH, Version 3.0) is calibrated with cartographic data. The prediction is based on the archived data trends of the years of the 1987, 1996, and 2002 images, which are extracted from Landsat Thematic Mapper and Enhanced Thematic Mapper satellite images and the aerial photographs acquired in 1992 and the data are prepared to insert them as input into the model. The urban extent is obtained through supervised classification of the satellite images and visual interpretation of aerial photographs.

The model calibration, where a predetermined order of stepping through the "coefficient space" is used is performed in order to determine the best fit values for the five growth control parameters including the coefficients of diffusion, breed and spread, slope and road gravity with the historical urban extent data. The development trend in Antalya is simulated by slowing down growth by taking into consideration the road development and environmental protection. After the simulation for a period of 23 years, 9824 ha increased in urban areas is obtained for 2025.

**Keywords:** Urban growth, simulation, prediction, supervised classification, Antalya.

## ÖZ

### SLEUTH MODELİNİN ANTALYADA UYGULANMASI

ŞEVİK, ÖZLEM

Yüksek Lisans, Jeodezi ve Coğrafi Bilgi Teknolojileri Ana Bilim Dalı  
Tez Yöneticisi: Yrd. Doç. Dr. Zuhal Akyürek

Nisan 2005, 112 sayfa

Bu çalışmada, Antalya metropoliten alanının 2025 yılındaki kentsel büyüme tahmini için bir kentsel büyüme modeli kullanılmıştır. Türkiye'nin son on yıldaki toplam nüfus artış hızı %18,28 iken Antalya, %41,78'lik nüfus artış hızı ile Türkiye'de en hızlı nüfus artış hızına sahip metropoldür.

Bir kentsel büyüme modeli (Sleuth V 3.0) kartografik veriler ile kalibre edilmiştir. Kentsel büyüme tahmini, 1987, 1996 ve 2002 yıllarının Landsat TM ve ETM görüntülerinden ve 1992 hava fotoğraflarından elde edilen arşiv veri eğilimlerine dayandırılmaktadır. Kentsel büyüme alanı uydu görüntülerinin sınıflandırılması ve hava fotoğraflarının görsel değerlendirilmesi ile elde edilmiş ve veriler modele girdi olarak sokulmak üzere hazırlanmıştır.

Model kalibrasyonu, difüzyon, breed ve spread, eğim ve yol katsayılarını içeren, beş büyüme kontrol parametreleri için en uygun olan değerleri elde etmek amacıyla gerçekleştirilmiştir. Antalya'daki büyüme eğilimi, yol gelişmesini ve çevresel korumayı göz önünde bulundurarak, büyümeyi yavaşlatarak simüle edildi. Simülasyon periyodu olan 23 yıl boyunca, 2025 yılına kadar kentsel alanlarda 9824 hektarlık bir artış elde edildi.

**Anahtar Kelimeler:** Kentsel büyüme, simülasyon, tahmin, kontrollü sınıflandırma, Antalya.

To my parents Birsen and Hüseyin Şevik  
And to my sister Mihran and to my brother Halil Şevik

Thank you for your love and support

## ACKNOWLEDGEMENTS

I would like to express my sincere thanks especially to my parents for their unaccountable tolerance, support and lifelong love in every day of my life. You provided me everything I need and everything I want. I love you.

I am greatly indebted to my supervisor Assist. Prof. Zuhâl Ayürek for her guidance, continuous encouragement, and interest, constructive discussions, critical reviewing, supervision, and invaluable support throughout every stage of this study. Without her, this thesis could have never been completed. I am very grateful and very lucky to have worked with you. Thank you for not leaving me alone in those dark tunnels and showed me the right way when I really got stuck.

One of the deepest thanks goes to Şafak Burak Çevikbaş who helped me during the model run. Thank you very much for your peerless computer knowledge and patience towards my stressful times.

I would like to express my thanks to Hakan Oğuz who studied SLEUTH before in Texas, USA for his patience and very helpful personality. His helps to me was unforgettable. Thank you for sharing your invaluable experiment with me.

I also appreciate the support of Mark Feller who is a SLEUTH expert for his helps from USA. His invaluable experiment and support was unforgettable.

I like to thank to my dear friend Z.Gözde Kaya for her unforgettable technical and moral support. You are very precious for me.

Thank you Orkun for your technical support. It was unforgettable.

Thanks go to INTA SpaceTurk and its members Ebru, Tuna, Nilhan, Köksal and M.Emin Halitligil for their technical and moral support.

I would like to thank to GGIT Department Assistants, Kıvanç, Ali Özgün, Dilek, Reşat, Serkan, Pınar, Gülcan, Arzu, Aslı for their support throughout my graduate studies and during the thesis.

I ought to thank to all examining committee members whose comments, critics, and advices improved the text a lot.

Finally, I would like to thank my family for their encouragement and patience during my study.

## TABLE OF CONTENTS

PLAGIARISM .....	iii
ABSTRACT .....	iv
ÖZ .....	v
DEDICATION .....	vi
ACKNOWLEDGMENTS .....	vii
TABLE OF CONTENTS .....	ix
LIST OF TABLES .....	xii
LIST OF FIGURES .....	xiii
CHAPTER	
1. INTRODUCTION .....	1
2. THEORETICAL AND PRACTICAL VIEW ON URBAN GROWTH DETECTION AND PREDICTION .....	5
2.1. Urban Area Detection By Means of Remote Sensing .....	5
2.2. Cellular Automaton (CA) Urban Growth Models .....	9
2.3. SLEUTH Model Definition and Urban Growth Predictions with SLEUTH Model .....	19
3. MATERIALS AND METHODOLOGY .....	31
3.1. Definition of the Study Area .....	31
3.2. Data Used in the Study .....	35
3.3. Analyses of the Satellite Images .....	43
3.3.1. Classification Analyses of the Satellite Image of the Year 2002 .....	45
3.3.1.2. Recoding of the Raster Layer of the Year 2002 .....	47

3.3.1.3. Accuracy Assessment of the Classification Result of the Year 2002. ....	50
3.3.2. Classification of the Satellite Images of the Year 1987 and 1996. ....	51
3.3.2.1. Post Classification Analyses of the Years 1987 and 1996. ....	54
3.3.2.2. Recoding of the Raster Layers of the Years 1987 and 1996. ....	55
3.3.2.3. Generation of the Model Function for the Classification Results of the Years 1987 and 1996. ....	56
3.4. Analyses of the 1992 Aerial Photographs. ....	60
3.4.1. The Generation of the Urban Extent Layer of the Year 1992. ....	60
3.4.2. Vector to Raster Conversion and the Recoding of the Urban Extent Layer of the Year 1992. ....	61
3.5. Analyses of the Roads Layer. ....	61
3.5.1. The Generation of the Vector Layers of the Roads. ....	62
3.5.2. Vector to Raster Conversion and the Recoding of the Vector Layers of the Roads 1995 and 2003. ....	62
3.6. Analyses of the Excluded Area Layer. ....	64
3.6.1 The Generation of the Excluded Area Layer. ....	65
3.6.2. Vector to Raster Conversion and the Recoding of the Vector Layer of the Excluded Areas of the Year 2003. ....	65
3.7. Slope and the Hillshade Layers. ....	66
3.8. The Resampling and the Subsetting of the Entire Layers. ....	66
3.9. The Conversion of the Subsetted Layers to the 8 Bit Unsigned GIF Format. ....	66
3.10. Methodology of the Study. ....	67
4. URBAN GROWTH ANALYSES. ....	71
4.1. Model Run. ....	71
4.1.1. Testing the Input Data. ....	72

4.1.2. Calibration Analyses. . . . .	72
4.1.2.1. Coarse Calibration for the Antalya Input Data. . . . .	73
4.1.2.2. Fine Calibration for the Antalya Input Data. . . . .	79
4.1.2.3. Final Calibration for the Antalya Input Data . . . . .	82
4.1.2.4. Forecast Coefficients Run for the Antalya Input Data .	85
4.1.3. Urban Growth Prediction Run. . . . .	87
5. RESULTS AND DISCUSSIONS. . . . .	89
6. CONCLUSION AND RECOMMENDATIONS. . . . .	100
REFERENCES. . . . .	103
ONLINE REFERENCES. . . . .	107
APPENDIX A. . . . .	108

## LIST OF TABLES

### TABLE

2.1. Conway's Game of Life; birth, death, and survival rule (URL 2). . . . .	11
2.2. General Meanings of the Calibration Parameters (Candau, 2002; Silva. and Clarke, 2002). . . . .	26
3.1. Density comparison of Antalya and the Centrum district, (DİE, 2000). . . .	34
3.2. List of the Model Input Data, Naming Format. . . . .	37
4.1. The Naming Format of the Data Stipulated by the Model for the Coarse Calibration. . . . .	74
4.2. Parameters of the Test Run and the Calibration Runs. . . . .	75
4.3. Control_Stats file from the Coarse (120m) Calibration Run. . . . .	77
4.4. Result Parameters of the Coarse Calibration. . . . .	78
4.5. Start, Stop and Step Values Derived from the Coarse Calibration	78
4.6. Control_Stats file from the Fine (60m) Calibration Run. . . . .	80
4.7. Result Parameters of the Fine Calibration . . . . .	81
4.8. Start, Stop, and Step Values Derived from the Fine Calibration . . .	82
4.9. Control_Stats file from the Final (30m) Calibration Run. . . . .	83
4.10. Result Parameters of the Final Calibration. . . . .	84
4.11. Start, Stop, and Step Values Derived from the Fine Calibration. . .	85
4.12. Average .Log file from the Forecast Run. . . . .	86
4.13. Coefficient Values Derived from the Forecast Run. . . . .	87
5.1. Comparison of the Selected Statistical Measures of the Years 2002 and 2025. . . . .	89
5.2. The Output Average File with the Statistical Measures from the Year. 2003 through the Year 2025. . . . .	92

## LIST OF FIGURES

### FIGURE

2.1. Rectangle and Circular Neighborhood for the CA Model: (a) Rectangle and Circular Neighborhood. (b) Simulation Based on Rectangle and Circular Neighborhood (Li and Yeh, 1998). . . . .	16
2.2. The Behavior of Each Metropolitan Area to the Different Coefficients, AML: Lisbon Metropolitan Area, AMP: Porto Metropolitan Area (Silva and Clarke, 2001). . . . .	30
3.1. Location of the Study Area. . . . .	32
3.2. The Study Area in Lat/Long.Projection System with the Municipality Borders. . . . .	33
3.3. Comparison of the Growth Rate of Population of Antalya to Turkey for. the last decade (‰), (DİE, 2000). . . . .	34
3.4. 1987 Landsat TM satellite imagery of the study area with 30 meter spatial. . . . .	38
3.5. 1992 Aerial photograph mosaic of the study area with a scale of 1/40.000. . . . .	38
3.6. 1996 Spot satellite imagery of the study area with 20-meter spatial resolution. . . . .	39
3.7. 2002 Landsat ETM satellite imagery of the study area with 30-meter spatial resolution . . . . .	39
3.8. Road Data for the Year 1995. . . . .	40
3.9. Road Data for the Year 2003. . . . .	40
3.10. Excluded Areas (unresidential areas). . . . .	42
3.11. The Layer of Terrain Slope. . . . .	42
3.12. Layer of the Hillshade (URL 5, 2004). . . . .	43
3.13. Excluded Areas (unresidential areas) . . . . .	44
3.14. The Result Thematic Map of the Supervised Maximum Likelihood Classification Analysis of the Year 2002. . . . .	48

3.15. Aggregation of the Maximum Likelihood Classification for the year 2002. ....	49
3.16. Mode 7 x 7 Filter Application to the Aggregation Result of the year 2002. ....	50
3.17. 2002 Urban Extent Layer after Recoding the Mode Filtered 2002 Image (white is urban, black is non-urban pixels). ....	50
3.18. The Flowchart of the Methodology of the Classification Analysis from the Data Acquisition Step to the Accuracy Assessment Step. . .	52
3.19. The Resultant Thematic Map of the Supervised Maximum Likelihood Classification Analysis of the Year 1987. ....	53
3.20. The Resultant Thematic Map of the Supervised Maximum Likelihood Classification Analysis of the Year 1996. ....	53
3.21. Aggregation of the Maximum Likelihood Classification for the year 1987. ....	54
3.22. Aggregation of the Maximum Likelihood Classification for the year 1996. ....	55
3.23. Mode 7 x 7 Filter Application to the Aggregation Result of the year 1987. ....	56
3.24. Mode 7 x 7 Filter Application to the Aggregation Result of the year 1996. ....	57
3.25. 1987 Urban Extent Layer after Recoding the Mode Filtered 1987 Image (white is urban, black is non-urban pixels). ....	57
3.26. 1996 Urban Extent Layer after Recoding the Mode Filtered 1996 Image (white is urban, black is non-urban pixels). ....	58
3.27. The Model Applied and Corrected 1996 Urban Extent Layer (white is urban, black is non-urban pixels). ....	59
3.28. The Model Applied and Corrected 1987 Urban Extent Layer (white is urban, black is non-urban pixels). ....	60
3.29. The Layer of the Urban Extent of the Year 1992 (Red Polygons are the urban extents) . . . . .	61
3.30. Recoded Urban Extent Layer of the Year 1992 (white is urban, black is non-urban pixels). ....	62
3.31. The Digitized Road Layers of the Years 1995 (a) and 2003 (b) . . . . .	63

3.32. The Recoded Road Layers of the Years 1995 (a) and 2003 (b). . . . .	64
3.33. The Recoded Excluded Area Layer of the Year 2003 (white is urban, black is non-urban pixels). . . . .	65
3.34. General outline of a SLEUTH model run examined in this study (Goldstein, 2004). . . . .	70
5.1. Urban Growth Prediction of the year 2025 with the Probability Colors.	89
5.2. Urban Growth Rates through the Year 2025. . . . .	91
5.3. Simulation of the Spatial Consequences of Future Urban Growth in Antalya. ( Period of the Years shown in the graphic Is 3. The Final simulation year of 2025 is also shown at the end. . . . .	94
5.4. The Urban Growth Prediction of the Year 2025 within the Municipalities of Antalya. . . . .	95
5.5. The Road Influenced Urban Growth in 2025 (a) and 2003 (b). . . . .	96
5.6. An example of the Decrease of sng (spontaneous growth) Measure in Undeveloped Areas. . . . .	97
5.7. An Example of the Increase of og (organic growth) Growth Measure. .	98
5.8. The Urban Growth of 2025 with the 2003 Road Data. . . . .	99

## **CHAPTER 1**

### **INTRODUCTION**

The human population is increasing gradually from day to day. As a result, the requirement for the settlement areas is increasing in all excessively populated world cities. Over the last century, the increase and the intensification of urbanization has become one of the most cited and discussed subjects all over the world.

Clarke et al. (1997) describes urbanization as the conversion of natural to artificial land cover characterized by human settlements and workplaces. This single transformation involves a wholesale modification of natural processes such as runoff and transpiration, and the short-term and long-term impacts touch every member of the human race every day. In a longer timescale, 200 years, total global population has increased six times and the earth's urban population has increased over 100 times. Driven by the Industrial Revolution, cities have gone from being a minor feature to a major one. Nowadays, human beings are trying to control the urban growth because the population is increasing rapidly during the last 50 years. It is an important problem for the world and it should be examined for a spacious life in the future. If the future of urban growth can be seen, this problem can be solved easily.

According to United Nations Population Division of the Department of Economic and Social Affairs (URL 1) official estimates and projections of urban, rural and city populations, all the population growth expected at the world level during the next 30 years will be concentrated in urban areas. Also, for the first time in the world's history, the number of urban dwellers will equal the number of rural dwellers in 2007. The world's urban population reached 2.9 billion in 2000 and is expected rise to 5 billion by 2030. Whereas 30 per cent of the world population lived in urban areas in 1950, the proportion of urban dwellers rose to 47 percent by 2000 and is projected to attain 60 percent by 2030.

With 26.5 million inhabitants, Tokyo is the most populous urban agglomeration in the world, followed by Sao Paulo (18.3), Mexico City (18.3), New York (16.8) and Mumbai (16.5). By 2015, Tokyo will remain the largest urban agglomeration with 27.2 million inhabitants, followed by Dhaka, Mumbai, Sao Paulo, Delhi and Mexico City, all of which are expected to have more than 20 million inhabitants. In Turkey, İstanbul will be the 19<sup>th</sup> in the most populous urban list with the 11,4 million inhabitants in the year 2015 (URL 1). However, Antalya is one of the fastest growing metropolises in Turkey with a population growth of 41,79‰ although İstanbul's growth is 33,09‰ and Turkey's is 18,28‰ for the last decade.

Substantial growth in cities first occurred in Western Europe, America, and Japan, but spread in the latter part of this century throughout Asia, South America, and Africa (Candau, 2002). It is still spreading continuously. Until today, urban growth has not shown any sign of slowing, even in nations where population growth has stabilized.

Understanding the dynamics of complex urban systems and evaluating the impact of urban growth on the environment involve procedures of modeling and simulation, which require innovative methodology and robust techniques. A number of analytical and static urban models have been developed to explain urban expansion and evolving patterns rather than to predict future urban development. Among all the documented dynamic models, those based on Cellular Automata (CA) are probably the most impressive in terms of their technological evolution in connection to urban applications.

The study focus here is not the creation of a new model but understanding and exploration of an existing dynamic model for problem solving in applied urban studies. This is built upon a self-modifying CA urban growth model namely SLEUTH originally developed by Keith Clarke at the University of California at Santa Barbara.

Goldstein (2004) examines SLEUTH's main properties, which cause to be chosen by researchers. It runs with a series of growth rules that form modified

CA. During SLEUTH run time, the growth rules of the CA are calibrated to archive the extent of the urban spatial data. Then it can be used to forecast and predict urban extent under different scenarios. It is a scale independent, transportable, and transparent model. It can also be used as a planning tool by incorporating different human perceptions into the data used for predicting the future footprint of a city

This study is the first application of SLEUTH model in Antalya in Turkey. The primary objective of this study is to simulate the spatial consequences of future urban growth in Antalya. Furthermore, the model's effectiveness when applied to the Antalya area is analyzed and future research directions for more accurate simulations are suggested. The study is presented in six chapters. In the first chapter the importance of the urban growth and the main properties of the SLEUTH, are identified.

Theoretical and practical view on urban growth detection and prediction is defined in the second chapter under three themes. First, the urban area detection by means of remote sensing is determined. Then the other CA urban growth models are presented. Finally, the SLEUTH model definition and urban growth predictions with SLEUTH Model are examined.

In the third chapter, materials and the used methodology, the case study area, preparation of the input data within GIS and Remote Sensing (RS) analysis required for the model implementation, including the classification analysis of the satellite images, post classification methods, accuracy assessment of the classification analyses are described.

In the fourth chapter, urban growth analyses are described. The structure of the model run and the required coefficients, the modes of the model (test, calibration, and prediction) adapted to Antalya are explained. In addition, running of the modes and selection of the coefficients from each run, finally, the outputs of the prediction mode are examined clearly in this chapter.

The results obtained in this study are examined in the fifth chapter. Growth probabilities of each year from the prediction start date through the prediction stop date are described with figures and tables. In addition to these the urban growth quantities, situation of the urban growth rates and the reasons for this situation are examined.

In the conclusion chapter, an evaluation of the study and the significant findings and some impediments of the study are explained. Finally, the recommendations for the further studies are given.

## CHAPTER 2

### THEORETICAL AND PRACTICAL VIEW ON URBAN GROWTH DETECTION AND PREDICTION

In this chapter, the previous studies about urban area detection by means of remote sensing, urban growth models, and the studies on urban growth prediction by using SLEUTH Model (Clarke, 1996) are presented. There are numerous references of urban growth contention, urban growth detection from satellite images to predict the urban growth by means of the models.

#### 2.1. Urban Area Detection by Means of Remote Sensing

Urban Growth Modeling and prediction history essentially started in 1950s, showed less activity in 1970s and 80s but revived vigorously in the 1990s by the help of spatial data availability and GIS.

Urban growth is an essential environmental issue to be monitored and forecasted in order to think about alternatives that could lead to a more sustainable future urban development. Hung (2002) explained that satellite imagery has been a useful tool for monitoring environments since early 70's when MSS (Multispectral Scanner) provided the first commercial satellite image.

In the past three decades, the remote sensing field and other disciplines witnessed remarkable improvements in satellite image quality and quantity, in terms of spectral resolution and spatial resolution. These improvements in satellite images and digital image processing algorithms, exposed some opportunities to do environmental quantitative analysis, rather than just land classification or object identification (Hung, 2002). Unfortunately, some problems reported decades ago remain unsolved, especially in urban areas (Forster, 1985).

The most significant problem from remote sensing of urban areas is mixed pixels. Mixed pixels refer to pixels having more than one land cover material. They are caused by various factors. It is inevitable to have mixed pixels. How to handle them may be a key factor to the success of any urban remote sensing research. Fuzzy representation may be a solution (Zhang and Foody, 1998). In addition, there is a need for a model to simplify heterogeneous urban environments so that they could be represented by a limited number of end members, and quantitative analysis could be performed.

Hung (2002) modeled urban environments from TM satellite images in Salt Lake City. He declares that urban environments are very heterogeneous. The model he used is a conceptual model to simplify urban environments as combination of three basic ground components that are vegetation, impervious surface, and soil. Hung (2002) studied six ground components (two for vegetation, three for impervious surface, and one for soil) are selected as basic components of urban environments. The data used in this study are 1990 TM image of partial Salt Lake City area. Percentages of the six ground components are extracted from a previously developed supervised classifier. In detail these six classes are; healthy green grass vegetation (V<sub>gr</sub>); tree and/or shrub vegetation (V<sub>tr</sub>); bright impervious surface (I<sub>br</sub>), such as rooftop, metal, and tile; medium impervious surface (I<sub>md</sub>), such as concrete and weathered asphalt; dark impervious surface (I<sub>dr</sub>), such as asphalt and darkened concrete; and soil and/or dry vegetation (S<sub>dv</sub>). Soil and dry vegetation are put together as one component, because these two land cover types are quite similar in their spectral reflectance characteristics, as well as the brightness values from satellite images. A supervised classifier was applied then various charts and plots of mean brightness values are generated to demonstrate the capacity of V-I-S composition on urban land cover analysis. As a result, a six-channel image with each channels indicating the percentages of one predefined land cover type are acquired. Hung (2002) adds that it is not very easy to identify all the classes one by one, or to distinguish one from the other, without prior information of the study area. The resultant image provides sub-pixel information about V-I-S components of urban areas. By this representation, heterogeneous urban areas are simplified to combinations of basic ground

components. Various urban land cover types can be displayed on a V-I-S diagram. The resultant image contains considerable biophysical information that is not usually extracted from satellite images with per-pixel classifier.

Esch et al. (2004) use radar data (TerraSAR-X data), which are hardly employed for the built-up area detection by using an object-oriented classification. Their study aims to develop a concept for an automated extraction of built-up areas based on very high resolution, single polarized Xband imagery. They mentioned that for a robust object oriented analysis accurate and reliable image segmentation is required. In their study, they used three image segmentation levels in different spatial scales for the identification of built-up areas. It is realized from the previous studies (Esch et al., 2004; Karakiş, et al., 2005) that image segmentation gives productive results when applied on the high-resolution images as QuickBird, Ikonos and TerraSAR-X data. Moreover, according to them image segmentation is an appropriate way before an object-oriented classification. On the other hand, it is realized that the image segmentation involves significant difficulties, as determination of the optimum number of levels for segmentation and being the corresponding segmentation parameters very complex and therefore time-consuming. In addition, the segmentation of the individual structures in the scene is strongly affected by local characteristics (Esch et al., 2004).

Bauer et al., (2003) use Landsat TM and ETM+ images. They focus on the classification of Landsat data to acquire the land cover changes in Minnesota with an acceptable accuracy. They describe the methods and results of classifications of multi-temporal Landsat TM / ETM+ data of the seven-county of Twin Cities Metropolitan Area of Minnesota for 1986, 1991, and 1998.

The advantages of satellite imagery are discussed in Bauer et al. (2003). Historically, remote sensing in the form of aerial photography has been an important source of land cover and land use information. However, they say that the cost of aerial photography acquisition and interpretation of cover types is prohibitively expensive for large geographic areas. They all say in their study in 2003 that digital satellite imagery such as Landsat TM and ETM+ are

alternatives to acquire the needed information. They touch on the advantages of satellite imagery in their study; (1) the sensor provides coverage of large geographic areas, (2) the digital form of the data lends itself to more efficient analysis and the classified data are compatible with geographic information systems, eliminating the need to digitize interpreted information, and (3) land cover maps can be generated at considerably less cost than by other methods (albeit at 30-meter spatial resolution). The area used in Bauer et al., (2003) study includes a diversity of land cover classes as core urban areas, rural land uses, including agricultural fields, grasslands, wetlands, and forests. The combination of early summer (late May or early June) with mid to late summer (August or early September) images provides the highest classification accuracy (Bauer et al., 2003). Yang and Lo (2002) also use this most common approach to detect land changes in Atlanta, Georgia Area. Then a comparison is done between the area estimates from Landsat classifications, change maps, and U.S. Dept. of Agriculture's Natural Resources Inventory (NRI). The change detection maps are also compared to high resolution, IKONOS, satellite imagery acquired in 2000 and to 1991. In addition to this, another comparison is done to 1998 parcel maps in a GIS database.

As a result they obtained the accuracy statistics from the comparison of parcel maps in a GIS database; the overall accuracies are 95.2% for 1986, 94,6% for 1991 and 95.9% for 1998, respectively, user's accuracy of individual classes that they acquired range from 85 and 98%, producer's accuracy range from 85 to 98%. A Comparison of Landsat and NRI (Natural Resources Inventory) is also done to evaluate the classification accuracy. The two evaluation results are similar. As a result, they obtained a map displaying the urban growth from 1991 to 1998.

The aim of examining these examples is to stiffen that the urban area detection can be done by using high or low-resolution satellite images, and some vector data.

## 2.2. CA Urban Growth Models

Modeling cities with cellular automata is a popular and spreading approach, and it is virtually impossible without the data management capabilities of GIS and powerful workstation technology (Wolfram, 1994; Clarke and Gaydos, 1998).

Cellular automata (CA) were developed by Ulam in 1940s and soon used by Von Neumann to investigate the logical nature of self-reproducible systems (Li and Yeh, 1998). As Torrens (2000) cited, the CA system's basic elements are cells, states, neighborhoods, and rules. Torrens (2000), and Li and Yeh (1998) explain the cells as the smallest units which expose some adjacency or proximity. The state of a cell can change according to transition rules, which are defined in terms of neighborhood functions. The notion of neighborhood is central to the CA paradigm (Torrens, 2000).

The rapid development of GIS helps to foster the application of CA in urban simulation. Current GIS are not designed for fast iterative computation, but cellular automata can be used by creating batch files that contain iterative command sequences. By linking cellular automata to GIS some of the limitations of current GIS can be overcome, in addition to this CA can benefit from the useful information provided by GIS in defining transition rules. The data requirement of CA can be best satisfied with the aid of GIS. In integration of GIS with CA, CA serves as an analytical engine to provide a flexible framework for the programming and running of dynamic spatial models. Also it is possible to put and arrange some constraints in the transition rules of cellular automata so that urban growth can be rationalized according to a set of predefined sustainable criteria (Li and Yeh, 1998).

Cellular automata have been used for simulating urban development (Clarke et al., 1997), as well as for other applications such as simulating change in land cover, freeway traffic, or the spread of wildfires.

The most famous and simply described instance of a CA is the mathematician John Conway's Game of Life (URL 2). The Life CA was developed by Conway to explore the simplest possible configuration for a universal computer. Its

specifications are very simple. Only two possible states are permitted in the game of life: alive and dead. The lattice of the CA is a square grid of infinite dimensions. In other words in its classic form, a cellular automaton consists of a regular array of cells, each of which has a finite number of states. Each state change must be local, depending only on the states of neighboring cells.

URL 2 describes the neighborhoods of the Life CA that are consisting of nine cells and says that the transition rules are straightforward. According to Conway, there are three rules that govern dynamics ('life') in the game: birth, death, and survival. The birth rule specifies that a cell will be born (i.e., that it will transition from a state of 'dead' to 'alive') if it has three 'alive' cells in its nine-cell neighborhood. Cells die (they transition from a state of 'alive' to one of 'dead') from overcrowding between time steps if they have more than three live neighbors. Cells die by exposure if there are fewer than two live neighbors. The survivor rule specifies that a live cell should remain alive in the next time step if it has either two or three live cells in its neighborhood. Conway's game of life is given in the Table 2.1.

In his game of life, he also built a configuration by the help of R. Wilson Gosper at the Massachusetts Institute of Technology and his team, within the game that could generate moving configurations of stable patterns called 'glider guns' that were capable of firing a steady stream of wandering gliders. MIT team had demonstrated that the Life CA was capable of generating a machine that could, in turn, reproduce copies of itself that were as complicated in their structure.

For some time now, cellular automata (CA) have been in popular use for urban simulation. Based on the Torrens's paper (2000), it is relatively easy to generalize the basic specification of CA to represent urban systems. Cellular automata are simple models for the simulation of complex systems (Wolfram, 1994) and cellular model assumes only an action space (usually a grid), a set of initial conditions, and a set of behavior rules.

**Table 2.1.** Conway’s Game of Life; birth, death, and survival rule, (URL 2).

<b>Birth rule</b>
if [cell state is “dead” in time t ]
and if [the number of cells with state “alive” in neighborhood $\geq 3$ ]
then [set state of cell to “alive” in time (t+1)]
end
<b>Death rule</b>
if [cell state is “alive” in time t ]
and if [the number of cells with state “alive” in neighborhood $> 3$ or [the number of cells with state “dead” in neighborhood $< 2$ ]
then [set state of cell to “dead” in time (t+1) ]
end
<b>Survival rule</b>
if [cell state is “alive” in time t ]
and if [the number of cells with state “alive” in neighborhood $> 2 < 3$ ]
then [set state of cell to “alive” in time (t+1) ]
end

Torrens (2000) likens the cell space, on which a cellular automaton operates, to an urban sense in an environment, a landscape, or a territory. The CA lattice can also be generalized to represent urban spatial structures, networks of accessibility, or the physical infrastructure of the city (particularly when the lattice is specified as an irregular tessellation). “CA cells operate just like the pixels that comprise a television screen, except that each cell is capable of processing information, as well as visualizing its state. Cells can correspond to any zonal geography within a city: parcels of land, administrative boundaries, traffic analysis zones, etc. The cell state offers a flexible framework for encoding attributes of a city into the simulation model (O’Sullivan and Torrens, 2000).

By the help of cell state, the attributes of a city can be encoded for the simulation model. In an urban context a cell state can be the representer of any attribute of the urban environment as land use (residential or commercial), density (high density or low density), land cover (forested or concrete), etc. Neighborhoods in urban CA represent the area of influence or activity within the city, e.g., market catchment areas, the walking radius of individual pedestrians, the commuting watershed, etc. The rules of a CA drive the dynamics of change in the model. CA rules can be devised to mirror how phenomena in the real world operate, and can then be coded as algorithms within the simulation.

### **i. The Complex Characteristics of Cellular Automata (CA)**

For a successful modeling process, the identification of key variables or components, and their interrelations that truthfully represent the urban reality are needed.

According to O'Sullivan and Torrens (2000), CA is a good mechanism for exploring emergence in complex adaptive systems as urban areas. They are dynamic and fine-scaled in resolution. In addition, the use of neighborhoods is a good encapsulation of interaction among system elements. In addition to these, CA also exhibit many of the signature trademarks of complex adaptive systems, such as phase shifts, power laws, self-organization, self-similarity, and fractal dimensions.

Rank-size rules or power laws (as they pertain to complexity and CA) link the frequency of occurrence of phenomena to their unit size with linear, consistent relationships across scales. There are many small-sized cities in the world, but only a few large cities. Of course, this makes intuitive sense, but what is remarkable is that the relationship between the population size of a city and the frequency of occurrence of cities of certain sizes is linear (Torrens, 2000).

Things tend to disintegrate over time according to second law of thermodynamics in physics. Many natural systems tend towards disorder, but

other, often open systems (and particularly biological systems) show the reverse tendency. They generate structure rather than disorder as they develop over time, even when starting from disordered or even structureless initial states. Such systems may be regarded as self-organizing. Self-organization is one of the characteristics of complex adaptive systems of CA (Wolfram, 1994; Torrens, 2000).

The patterns that CA generates often exhibit a degree of regularity in structure. Often these regularities are self-similar portions of the evolved pattern of a structure are indistinguishable from the whole, because the CA has a self-similarity characteristic. Cities often exhibit a bi-fractal structure, characterized by two or more zones. Inner zones are the well-developed core of cities and in these parts of the cities urbanization process are essentially completed. Outer fringe zones sprawling and urbanization is still underway (Torrens, 2000).

## **ii. Advantages of a CA**

Torrens (2000) and Wolfram (1994) listed the advantages of the CA as follows:

- Flexible
- Connection of form with function and pattern with process
- Applicable with remotely sensed data and GIS
- Advantages of using cellular automata for urban simulation
- Weaknesses of traditional models
- Spatiality
- Decentralized approach
- Affinity with geographic information systems and remote sensing
- Attention to detail
- Function and form
- Dynamics
- Infusion of complexity theory
- Simplicity
- Linking macro- to micro-approaches

- Visualization (Torrens, 2000; Wolfram, 1994)

The Role of GIS in Urban Cellular Modeling can be listed as follows (Torrens, 2000):

- Storing and Managing Input Data and Results
- Pre-processing of input data (editing, transformation, interpolation, derivation of parameters)
- Analysis and visualization of results
- Providing computational environment and tools for simulations

For better urban forms for sustainable development and for helping planners, CA can be extended and integrated with GIS (Li and Yeh, 1998). The objective of their research is to develop an operational CA model for sustainable urban development. They generalized the standard cellular automata as follows:

$$S^{t+1} = f(S^t, N) \quad (2.1)$$

S is a set of all possible states of the cellular automata, f is a transition function that defines the change of the state from t to t+1 time, and N is a neighborhood of all cells providing input values for the function f (Li and Yeh, 1998).

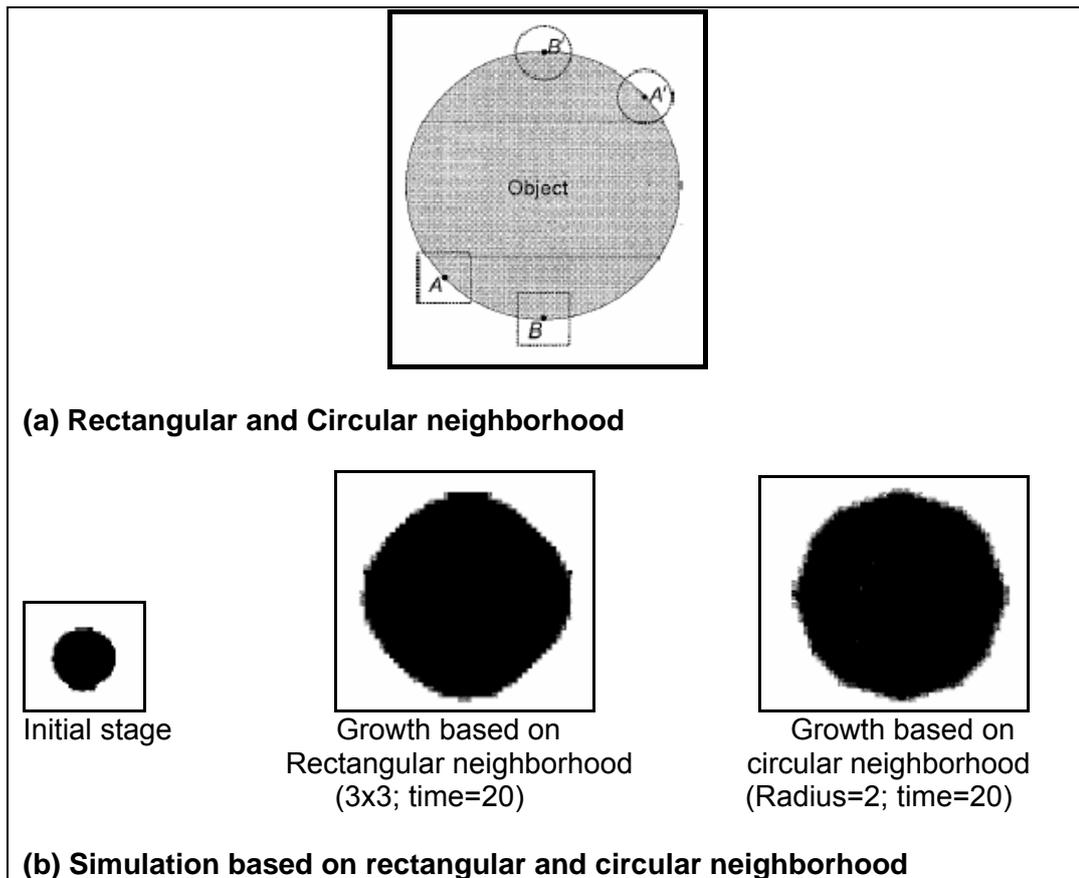
According to Li and Yeh (1998), cellular automata can be explained as three kinds; state based cellular automata, suitability based cellular automata and constrained cellular automata. In a standard (state-based) CA model, the state is the main attribute to describe the development of a cell. A cell cannot take on more than one state simultaneously. However, a state can change from one to another in different periods. In urban simulation through such a CA, the most general state for a cell is developed (alive) or not developed (dead). In the state based cellular automata model, the state of a central cell uses a 3 x 3 window to count the distribution of states in its neighboring cells. These models indicate that cells developed in the neighborhood cells can add some probability for development in the central cell. A suitability based cellular automata is a more sophisticated one. This kind of system simulate urban growth through the concepts of “development probability” and “development suitability” and it

assumes a relation between the states (developed or not), development probability and development suitability. The constraints are used to make more reliable and reproducible predictions of actual urban land use patterns. They are mainly related to land suitability according to accessibility that affects land development probability, such as cost distance to city centers, roads and railways. Dependent on the study area more constraints as local, regional and global can be added to cellular automata for a satisfactory sustainable urban growth form (Li and Yeh, 1998). Li and Yeh used a constrained CA in their study. The model is implemented in the Arc/Info Grid environment. Therefore, the CA model is developed within a GIS to facilitate the convenient access to the land use information in the GIS database. The GIS database consist of land use maps, soil maps, economic data and the monitoring results of land use change detected from remote sensing for the Pearl River Delta. The model is applied to Dongguan City in this delta, the fastest growing region in China. The main point of constraint CA model is its being straightforward in allowing constraints to shape urban growth. Thus, a better urban form can be obtained by reducing urban encroachment on the restricted areas according to the constraint scores. Li and Yeh (1998) cite that circular neighborhood is better than a rectangular neighborhood (the Moore neighborhood) because no bias exists in all directions (Figure 2.1).

In Figure 2.1 (a), the points of A, B, A', and B' should have the same neighborhood for a circular object. Nevertheless, the configuration of neighborhood by rectangles produces a discrepancy of neighborhood between A and B. As it is seen from the Figure 2.1 (b), the simulation from the rectangle neighborhood can produce significant distortions, compared with that from the circular neighborhood (Li and Yeh, 1998).

The basic data used in the study of Li and Yeh (1998) are Landsat TM with 30-meter ground resolution for the years 1988 and 1993. Agricultural suitability maps are produced from soil and slope maps. All the data layers are converted to ground resolution of 50 m with 619 pixels x 889 pixels and their CA model used a circular neighborhood with a radius of two pixels. The simulation start year is 1988 and urban areas of this year are acquired from the classification of 1988 Landsat TM image. Three scenarios are thought by means of three

different agricultural land constraint values. First, the constraint score  $k=0$  is used. In the output of this scenario, a lot of agricultural land loss takes place in the west part of the city where best agricultural lands exist. Then iteratively constraint scores  $k=1$  as a normal constraint and  $k=2$  with a stricter constraint are used. So the minimum agricultural land consumption is seen in the third value scenario.



**Figure 2.1.** Rectangle and Circular Neighborhood for the CA Model: (a) Rectangle and Circular Neighborhood. (b) Simulation Based on Rectangle and Circular Neighborhood (Li and Yeh, 1998).

Li and Yeh (1998) state that more complicated factors besides agricultural suitability can be embedded into the constrained CA model to reflect other environmental settings for sustainable urban forms. They give examples of land resources and economic factors, which vary regionally and globally. Per-capita

agricultural land resources are not same among towns and this affect land supply regionally. Land supply may also change globally because of government policies and intervention. Li and Yeh (1998) cite that development suitability and constraints can be defined with reference to these changeable regional and global factors. Therefore, it can be understood from this study that the integration of CA and GIS provides a useful tool to explore sustainable urban forms under different development scenarios. Moreover, research on geographic modeling with CA is still exploring and building upon modeling capabilities (Clarke et al., 1996; Clarke and Gaydos, 1998; Li and Yeh, 1998). In their study, Li and Yeh (1998) focus on the agricultural land loss but it is considered that local, regional, and global constraints embedded in this model are most important in generating sustainable urban growth in many developing countries where the loss of agricultural land to urban development is serious. Li and Yeh (1998) conclude their study with the thought of that the constrained CA for modeling the sustainable urban growth can be defined using GIS and Remote sensing data. They also added that remote sensing could be used to obtain land use information, which is then transformed into GIS for analysis and modeling.

A GIS based model is used to predict urban growth in terms of land use change and calibrated in the Charleston region of South Carolina by Allen and Lu in 2003. As Wolfram (1994), Clarke et al. (1996), Torrens (2000), Allen and Lu (2003) state that urbans are complex systems and this complexity makes it difficult to figure out their changes using a model based on a single approach. The model that Allen and Lu used in 2003 is coupled with a rule based suitability module and is designed to predict land transition probabilities and simulate urban growth under different scenarios. The prediction year of their study is 2030. In Charleston region from 1960 to 1990, the population increased 41% (Allen and Lu, 2003). To develop an operational model, to simulate future urban growth based on different scenarios and to predict future urban spatial expansions through to the year 2030 is aimed in Allen and Lu's study (2003). The model used in their study (2003) is the integration of three different modeling schemes:

a) Logistic regression models: It is used to identify the significant variables and rules that differentiate urban or city from rural and forest environments

b) Relative probability model (Piyanowski et al., 1997): It uses spatial interactions of neighborhood, distance, patch size (parcels) and site-specific characteristics.

c) The focus group model: It is used to create a human input layer, set the growth scenarios, evaluate predictions, and disseminate the information.

The modeling system is developed as an extension of ArcView and integrated with SPSS statistical software package.

The outputs of these three models are linearly combined in an integrated model to generate a hybrid transition probability grid for the final prediction. As the input data, urban areas are derived from the Landsat MSS and Landsat TM according to Anderson et al. (1976) land use (manmade) and land cover (natural or semi-natural) classification system. This land category includes residential, commercial and services, industrial, transportation, communication, utilities, industrial and commercial complexes, and mixed urban or built-up land. Other remaining land uses are grouped into a single category as non-urban. The reason for using the binary land use classification as urban and non-urban is to emphasize the urban growth similar to the data used for Antalya. Three different resolutions are used. First, all variable grids are prepared at 30 x 30 m, equivalent to the spatial resolution of the Landsat TM imagery used to derive urban land use data. Then they are resampled at 100 and 200 m to create table data sets for conducting statistical analysis (model calibration) at the county and regional levels, respectively. The main purpose of their study is to maintain the data sets at manageable sizes, while keeping the resolution as high as possible. Higher resolution grids (30 x 30 m) are used for the final prediction and mapping in order to obtain a better visual effect. Allen and Lu (2003) state that some constraint areas as environmentally important areas, on which the urban

development is not allowed, can be introduced to the model by the rule-based model. Due to the complexity of land use systems, many factors, or variables cannot be defined or measured, or their relationships cannot be modeled, but their overall intangible effects may well be perceived by people, particularly local planners, developers, or experts with years of experience. In this case Allen and Lu (2003) create a focus group consisted of local experts, planners, developers, landowners, conservationists, and community leaders who have a profound knowledge of the region and urban growth factors. Finally, the outputs of these three models are linearly combined in an integrated model to generate a hybrid transition probability grid for the final prediction. The final prediction result is the year 2030. According to the prediction, the total population of the region will increase from 532,600 in 1994 to 795,800 in 2030. The net growth is about 263,000 people, or 49.41% within a 36-year period, about 7500 people per year. The population growth rate is 1.41% annually.

As a result, Allan and Lu (2003) achieved that the logistic model is useful for identifying significant predictors and it can obtain high prediction success rates. In addition to this, they also acquired that the results of temporal validations indicate the logistic model is statistically reliable for short-term prediction, but becomes less reliable once the time-span becomes longer.

### **2.3. SLEUTH Model Definition and Urban Growth Predictions with SLEUTH Model**

In recent years, dynamic modeling has become a primary research field in Geographical Information Science. It is rapidly gaining popularity among urban planners and geographers for simulating urban and landscape. (Turner, 1987; Pijanowski et al., 1997; Clarke and Gaydos, 1998; Bell et al., 1999; Çelikoyan et al., 2003; Xie, 2003).

SLEUTH is one of the popular urban growth models (Clarke et al., 1996). There are a number of reasons for choosing this model for the current research. It is composed of a series of growth rules and formed modified Cellular Automata

(CA). It accommodates CA rules, which are also growth rules calibrated to historical urban spatial data. Due to its scale independence, dynamic and future oriented structure, transportability, use of under different conditions by modifying some initial conditions and changing input data layers and application of all regions with different data sets, SLEUTH has become a popular tool in modeling urban spreading extent over time or forecasting growth into the future (Yang and Lo, 2003; Goldstein, 2004). Yang and Lo cite the ability of SLEUTH's growth rules for the future prediction, deviating from ordinary line-fitting urban models. Thus, SLEUTH can be used as a powerful planning tool by incorporating different human perceptions into the data used for predicting the future of a city. SLEUTH has been used to model a growing number of geographical regions; Chester County (Arthur, 2001), Washington-Baltimore metropolitan region (Clarke and Gaydos, 1998), Porto and Lisbon, Portugal (Silva and Clarke, 2002), San Francisco (Clarke et al., 1997) are some of the applications of SLEUTH. For understanding the model precisely, the SLEUTH applications in different areas and continents are examined.

Urban and regional models are usually supported by a set of variables and parameters that feed system dynamics and process interactions built into the models. Depending on which variables are required by the model and for policy manipulation, common elements can be defined and assigned behavior and significance, such as the importance of roadways, urban extent, topographic slope, parks, and reserves. Most urban and regional models incorporate these general characteristics of urban settlement (Silva and Clarke, 2002).

The model is written in the C language computer program and it operates as a set of nested loops. The outer loop repeatedly executes each growth "history", retaining cumulative statistical data and the inner loop executes the growth rules for a single 'year'. The starting point for urban growth is an input layer of 'seed' cells from the urban extent identified from historical maps or other sources. The rules apply to a cell at a time and the whole grid is updated as the 'annual' iterations are completed. The basis for urban expansion in each succeeding year is formed by the modified array. Potential cells for urbanization are selected and the growth rules evaluate the properties of the cell and its

neighbors as if they are already urban or not, what their topographic slope is how close they are to a road (Clarke et al., 1997; Clarke and Gaydos, 1998). In the model there are five factors controlling the behavior of the system:

- Breed Factor (Coefficient) determines how likely a newly generated, detached or road-influenced settlement is to begin its own growth cycle.
- Diffusion Factor (Coefficient) controls the overall dispersiveness of growth.
- Spread Coefficient controls how much diffusion expansion occurs from existing settlements.
- Slope Resistance influences the likelihood of settlement extending up steeper slopes
- Road Gravity encourages new settlements to develop near the transportation network (Clarke and Gaydos, 1998; URL3, 2006).

According to Clarke (1996), urbanization is the sum of the four types of the growth:

- Spontaneous Growth models the development of urban settlements in undeveloped areas.
- Diffusive Growth permits the urbanization of isolated cells, which are flat enough to be desirable locations for new urban spreading centers.
- Organic Growth promotes the expansion of established urban cells to their surroundings.
- Road Influenced Growth promotes the urbanization along the transportation network because of increased accessibility.

Newly urbanized cells must pass the random tests of breed and diffusion coefficient, slope resistance and road gravity. During the urban growth computation, a second level of growth rules called "self-modification" prompted by an unusually high or low growth rate above or below a threshold or critical number (Clarke et al., 1996). In that case, the model modifies certain parameters to emphasize trend. Therefore, self-modification is quite important to ensure reasonable results (Yang and Lo, 2003). The finishing values of all the

coefficients (located in a file called param.log) were used to find the final best values that describe the boom and bust periods in the system. This utility averages the finishing coefficient values stored in the param.log file, and returns a set of five integers that represent the best coefficient values resulting from the entire process of calibration, reflecting both the growth rules, and the self-modification rules (Silva and Clarke, 2002).

The growth rate is computed by comparing the number of new pixels urbanized in any period to the total existing urban area. It can be summarized from the Silva and Clarke's study in 2002 that the self-modification of the rules changes the control parameters when modeled growth rates are exceeded, so that the model's behavior includes feedback. The limits of "critical high" and "critical low" initiate an increase or decrease in three of the growth control parameters: diffusion, breed, and spread. The increase to the parameters is by a multiplier greater than one, "boom," imitating the tendency of an expanding system to grow ever more rapidly, while the decrease is by a multiplier less than one, "bust", causing growth to taper off as it does in a depressed or saturated system. By summarizing from Silva and Clarke's study (2002), each time the model records rapid growth, or little or no growth, the model adapts itself to this new set of conditions. In the case of rapid growth, the model multiplies the growth control parameters by a multiplier greater than one. Little or no growth causes the control parameters to be multiplied by values less than one. The parameter values increase most rapidly at the beginning of the growth cycle when there are many cells available to urbanization, and then, with time, the parameters are decreased as expansion levels off and the growth rate falls below the critical low. However, to prevent uncontrolled exponential growth as the system increases in overall size, the multiplier applied to the factors is slightly decreased or lagged in every subsequent growth year. Self-Modification can also increase the road-gravity factor as the road network enlarges, prompting a wider band of urbanization around the roads, and decrease the slope resistance factor as the percentage of land available for development decreases, permitting expansion onto steeper slopes. By means of self-modification, the parameter values increase most rapidly in the beginning of the growth cycle when many cells are available for urbanization, and decrease as

urban density increases in the region. Therefore, without self-modification the model produces linear or exponential growth (Clarke et al, 1996).

The model runs in three modes; test mode, calibration mode and the prediction mode. By means of the test mode, you can test your data whether ready for the calibration then the prediction. Prior to calibration, the first step in the application of the SLEUTH model is the verification of the data sets and their initial reaction to the input data, called test mode, including assuring that they conform to data input specifications. A minimum of four urban years, two road years and at least one excluded layer, one hill shade image and one slope layer are required, and the code verifies the correct input of each of these data sets (Clarke et al., 1996; URL 3). The importance of this step can be understood from Silva and Clarke's study (2002) in the Lisbon Metropolitan Area. The water bodies and land outside the Porto Metropolitan Area was initially not defined correctly in the excluded layer, and consequently the model was seen expanding urbanization to these areas. It was also observed during this test phase that the slope layer was not contributing to the model calibration, for the test mode statistics did not seem to be sensitive to changes in slope. It was found that the percent slope image had been altered during its conversion from TIF to GIF format. Without this initial test, the model could run for days during calibration, and the time would be unnecessarily lost.

Once the test mode is completed, the next phase is the calibration mode that is the most important step for the success of model prediction. The purpose of the model calibration phase is to determine the best-fit values for the five growth control parameters including coefficients of diffusion, breed and spread, slope resistance and road gravity with historical urban extent data. Calibration relies on statistical measures of historical fit. It is the key component of the modeling process by which numerical values are assigned to the model parameters in such a way that the model accurately reproduces the real patterns (Clarke et al., 1996; Clarke et al., 1997; Yang and Lo, 2003). Calibration mode has three phases; coarse, fine and final calibrations as told in Silva and Clarke's study in 2001. Yang and Lo (2003) also have the calibration results comprising the three phases. In the "coarse calibration" input data are resampled to four times of their

initial resolution (100 m resolution data is resampled to 400 m) and the model attempts to simulate the historical growth patterns for a wide range of parameter values across the entire parameter space. “In the fine calibration” the input data are resampled to twice of their original resolution (100 m resolution data is now resampled to 200 m). Using this half resolution data, the narrowed range of parameters from the previous step are used to simulate the historical growth patterns. Results of these simulations are evaluated using spatial metrics of fit, and the range of parameters is narrowed. Finally, in the “final calibration” the input data are used at their full resolution. The first year provides a seed for the set of parameters tested, which then simulate urban growth and then evaluate it compared to the actual control data. The set of parameters that best recreates the urban growth is then used in model forecasting (Dietzel and Clarke, 2004).

In these calibration results for Lisbon the set of initial control parameter values are ranging from 1 (in the case of the diffusion coefficient) to 100 as the maximum values for each of diffusion, breed, spread, slope resistance and road gravity. In the coarse calibration, the resulting values were narrowed to 1, 100, 50, 25, and 20. With the final calibration, the values became more sensitive, respectively, values of 16, 57, 50, 25, and 30.

The comparison of the model final “population” (number of urban pixels) and the urbanization for the control years gives a high summary correlation of 0.90 (compare\_score). This means the prediction of the model based on the initial seed year of the present urban pattern using those refined values is very similar to what happened in reality. The shape and form of urbanization seems also to confirm that calibration adjusts the values to reflect local characteristics. The final calibration correlations are 0.78 in the case of the score  $r^2_{edges}$  (modeled urban edges against the urban edges of control years), and 0.87 in the case of the cluster\_ $r^2$  score (modeled urban clustering against known urban clustering). For Lisbon, leesalle (degree of shape match between the modeled growth and the known urban extent) is 0.35 in the final calibration but it is very hard to obtain high values of shape match (Clarke and Gaydos, 1998). Therefore, a value of 0.35 is very good for the Lisbon Metropolitan Area (Silva and Clarke, 2002). General meanings of these parameters are given in Table 2.2.

The comparison of the five coefficients that control the system is discussed very clearly in Silva and Clarke (2002), Candau (2002), and Goldstein (2004) that if a diffusion value is high it is a sign of the scattering urbanization of the area where SLEUTH applied, the breed value shows the amount of vacant land suitable for development and it is high when this kind of lands are abundant in a case. If the spread coefficient is high, with a high probability it shows us that the transportation infrastructure is recently built in the region with a high capacity, which is an invitation for the spread urbanization. In that case, the road value is also high. Finally, the slope coefficient is low if the area has fewer constraints to urbanization due to slope. The result of these three phases comprise of the optimum values for the diffusion, spread, slope and road gravity parameters.

When the calibration mode is completed, the results are used for forecasting studies in the prediction mode. Prediction Mode is used to project future scenarios of urban growth.

Sleuth is a CA based model. A typical CA consists of four primary components:

1. Cells
2. States
3. Neighborhoods
4. Transition rules

The state of a cell can change in relation to its neighboring cells when a set of transition rules are applied uniformly (Candau, 2002; O'Sullivan and Torrens, 2000).

**Table 2.2.** General Meanings of the Calibration Parameters (Candau, 2002; Silva and Clarke, 2002).

composite score	Acquired by multiplying all the scores together, then by multiplying this value by a ratio, which is the comparison of model final urban areas to the actual urban areas.
Compare value	Comparison of the modeled final population to real data final population.
$r^2$ Population	It is the least squares regression score for modeled urbanization compared with actual urbanization for the control years
$r^2$ Urban	Regression score for urbanization
$r^2$ Urban Edges	Regression score for Urban edges
$r^2$ Urban Clusters	Regression score for Urban Clusters
Mean cluster size $r^2$	Least squares regression score for modeled average urban cluster size compared with known mean urban cluster size for the control years
Leesallee Score	Measures the degree of shape match between the modeled growth and the known urban extent for the control years
Average Slope $r^2$	The least squares regression of average slope for modeled urbanized cells compared with average slope of known urban cells for the control years
pct Urban_ $r^2$	The least squares regression of percent of available pixels urbanized compared with the urbanized pixels for the control years
xmu and $r^2$	These values are the (center of gravity [x] and [y]) least squares regression of average x and y values for modeled urbanized cells compared with average x and y values of known urban cells.
Sdist_ $r^2$	Standard deviation averaged over (XY)

#### Advantages of the Model (SLEUTH):

- Scale independent
- Dynamic and future oriented
- Can be used under different conditions by modifying some initial conditions and changing input data layers
- Can be applied to all regions with different datasets (Candau 2002; Yang and Lo, 2003, URL 3, 2005)

In the SLEUTH model, the growth rules are uniform throughout a gridded representation of geographical space and are applied on a cell-by-cell basis. A single time span is one iteration of the CA, and all changes are applied synchronously at the end of each period (Clarke et al., 1997).

The current version of the urban growth model can drive a land use/cover transition model, which is called land cover deltatron model, but the urban growth model can run independently.

Yang and Lo (2003) simulated the urban growth in Atlanta metropolitan area, one of the fastest growing metropolises in the United States during the past three decades, by using SLEUTH model coupled with a land transition model. Yang and Lo (2003) calibrated the model with historical data that are extracted from a time series of satellite images. Three specific scenarios to simulate the spatial consequences of urban growth under different environmental conditions are applied. Their first scenario is to simulate the continued growth trend by maintaining the unchanged current conditions. The second scenario is to project the growth trend by taking into consideration the road development and the environmental protection. In these two scenarios, the unchecked urban growth would result in the displacement of almost all the natural vegetation and all the open space in Atlanta. The third scenario is to simulate the development trend by slowing down growth and changing growth pattern. In contrast to first two scenarios, the result from the third scenario displays much more greenness and open space, including buffer zones of large streams and lakes could be

preserved. Consequently, the last scenario is the most desirable for the future urban growth of Atlanta.

Clarke and Gaydos (1998) applied a CA model, which is applied to two rapidly growing, but remarkably different urban areas: the San Francisco Bay region in California and the Washington / Baltimore corridor in the Eastern United States. The calibration and prediction results for both regions are presented, and their data requirements are reviewed, the differences in the initial configurations and control parameters for the model in the two settings are compared, and the role of GIS in the applications are discussed. As a result, the model has generated some long-term predictions that appear useful for urban planning.

Silva and Clarke (2002) examined differences in the model's behavior when the model obviously applied to different environments of two European cities, which are captured in the data and modeled. They interpreted and evaluated the model's portability and universality of application.

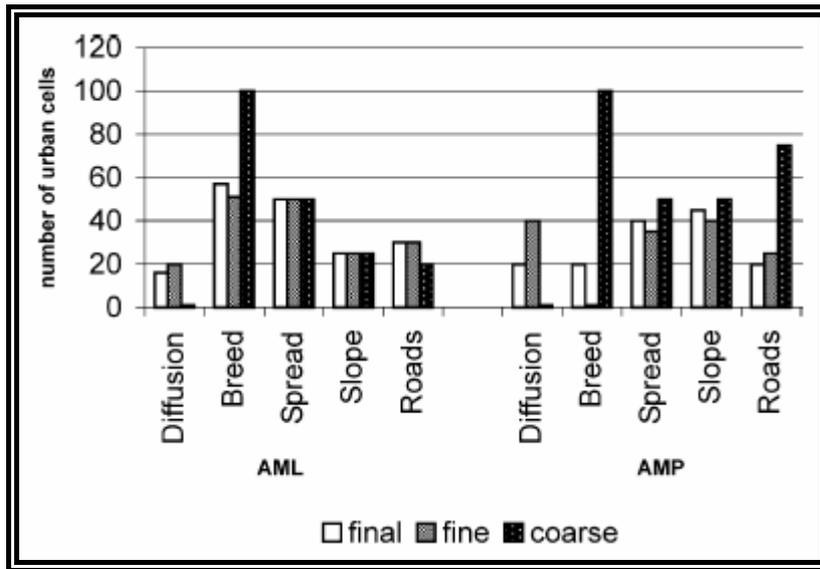
Data assembly problems are revealed as in other SLEUTH applications. The problems that Clarke and Gaydos (1998) encountered with are inconsistent feature definitions, especially for urban areas and major roads, extensive manual generalization in historical maps, integration of multiple image and map sources from different projections, datums, and coordinate systems. After assembly of the various data sets and their conversion to the input format for the model in Clarke and Gaydos (1998), the model calibration is done for the two study areas to find the best fit of observed historical data.

Although the calibration phase is slow and highly dependent on the size of the input data set and on the quality and quantity of historical data, the model can produce useful results. Clarke and Gaydos (1998) think that the model should be applied to new different areas and at different map scales and they plan to test the model at about 1 km resolution for the entire lower 48 United States for the full Anderson Level I classification (Anderson et al., 1976), and for applications in New York, Chicago, Philadelphia, and Portland.

For Lisbon and Porto, as done in the other SLEUTH applications (Clarke and Gaydos, 1998; Yang and Lo, 2003) all data layers for two metropolitan areas required for SLEUTH model are collected (slope, excluded areas, urban extent for the control years of 1984, 1995, 1997, and the seed year of 1975 transportation and hill shade). These data are prepared in ArcInfo GIS format and converted into the 8-bit GIF format. First, the calibration result metrics are compared in Silva and Clarke's study (2002). The metrics acquired from the calibration that best describe each system are explained in terms of their behavior according to the landscape characteristics and history. The scores and coefficients of both metropolitan areas are compared to understand what extent the model reflects different realities, and which metrics are more sensitive. Therefore, some events including different political, socio-economic, and cultural changes during recent times are crisscrossed for both metropolitan area and the country. The first event period is chosen as the period before the 1974 revolution. The second period comprises the years between 1974 and the end of the 1980s (Silva and Clarke, 2002).

Until now, all the existed important periods are determined to use in the crisscross comparison. In addition to these political effects, housing market and rental laws are also taken into consideration. As a result, they acquired the tables of three calibration phases for both cities. Then they compared these two cities by interpreting the parameters in these tables. The result comparison is shown in Figure 2.2 (Silva and Clarke, 2002).

The numbers acquired from the detailed and exhaustive calibration can be used to predict future growth in the SLEUTH model prediction mode. From Silva and Clarke's (2002) study it can be observed that, cities may show a higher degree of influence of infill from a relatively modest number of existing centers (as in Porto) or stronger impact of transportation on growth (Lisbon). Throughout calibration, these different characteristics can be captured in the set of final coefficients that best describe the specific system/reality under study at the same time, and so can predict future developments.



**Figure 2.2.** The Behavior of Each Metropolitan Area to the Different Coefficients, AML: Lisbon Metropolitan Area, AMP: Porto Metropolitan Area (Silva and Clarke, 2001)

## CHAPTER 3

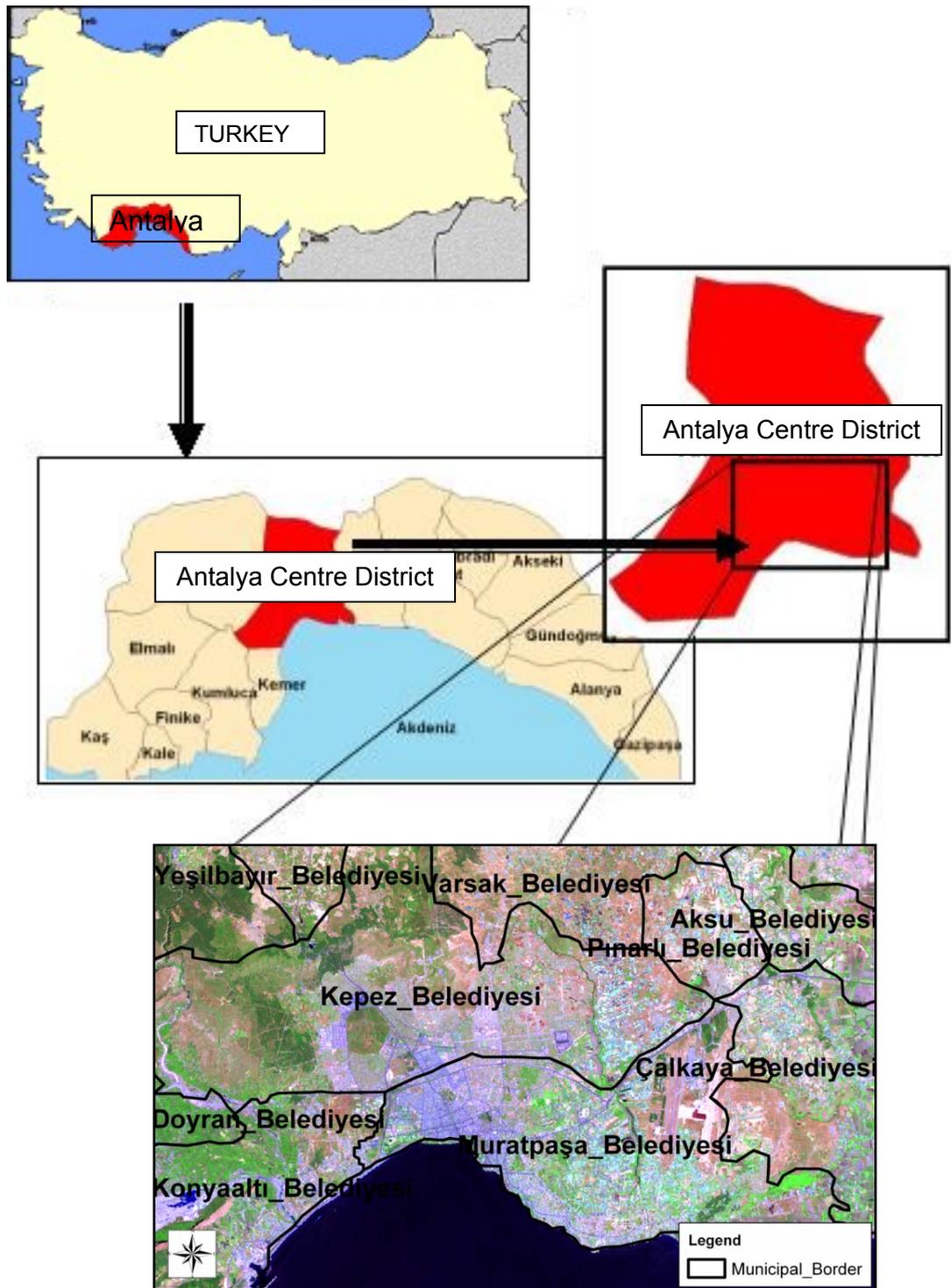
### MATERIALS AND METHODOLOGY

In this chapter, the case study area, the preparation of the input data required for the model SLEUTH and the methodology used in this study are described.

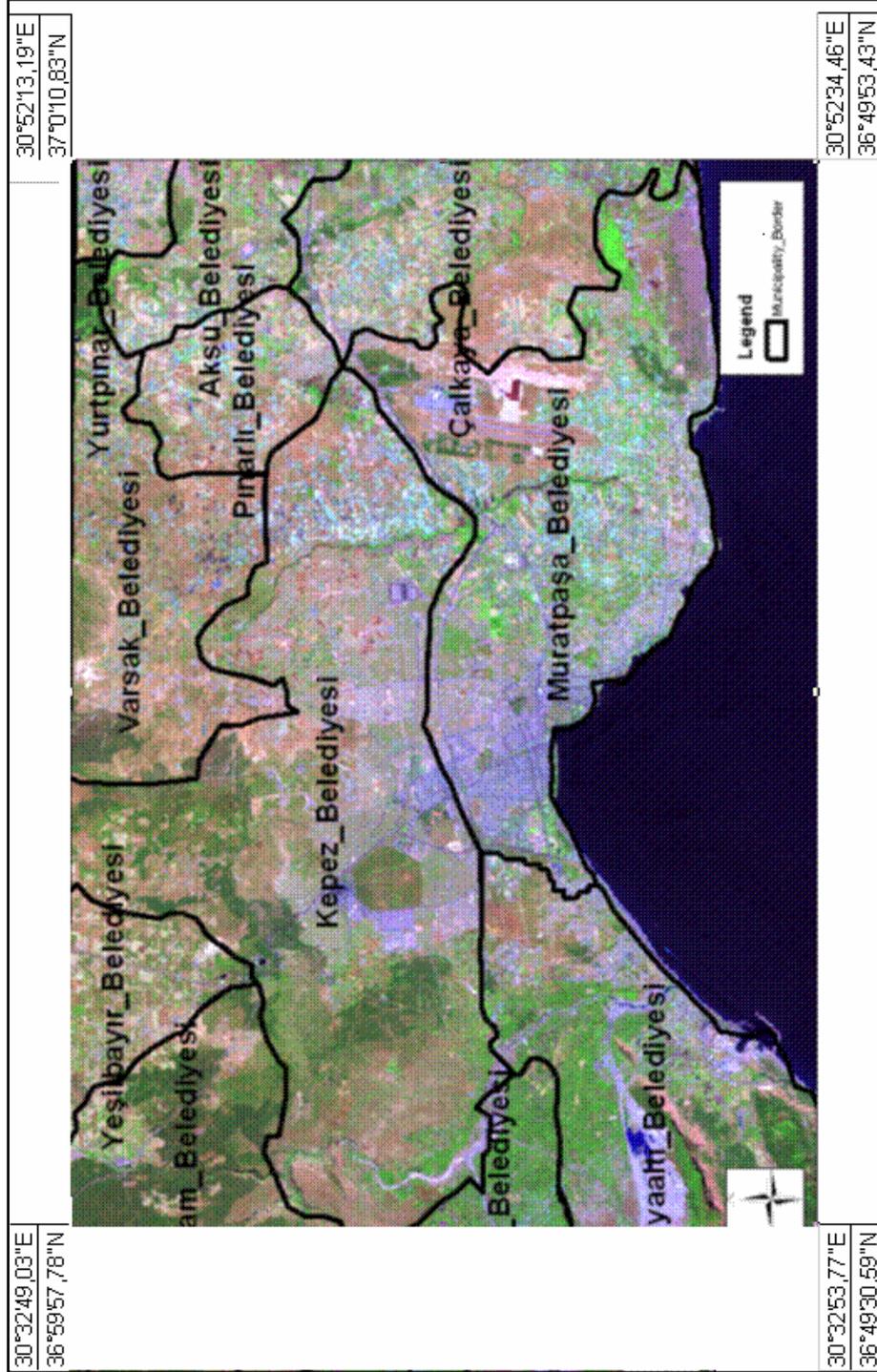
#### 3.1. Definition of the Study Area

The study area is located in Antalya. It is in the Mediterranean Region of Turkey, between the 30°32' E-36°59' N, 30°52'E-37°0'N, 30°32'E-36°49'N, 30°52'E-36°49'N lat/long units. Except centre district, Antalya has fourteen districts, which are Akseki, Alanya, Elmalı, Finike, Gazipaşa, Gündoğmuş, İbradi, Kale, Kaş, Kemer, Korkuteli, Kumluca, Manavgat, and Serik. In addition to this, it has a centre district in which the case study area of this study takes part (Figure 3.1). The modeled area is a rectangle including west part of the Municipality of Çalkaya, whole Muratpaşa Municipality, northeast part of the Municipality of Konyaaltı, and the south part of the Municipality of Kepez (Figure 3.2)

For the last decade, Antalya has been the fastest growing metropolitan city in Turkey as it emerged the premier tourism and industrial urban centre of the Mediterranean Region. The total population of Antalya was 1.132.211 in the year of 1990 and the most crowded district was the Central district with its total population of 448.773 (DİE, 1990).



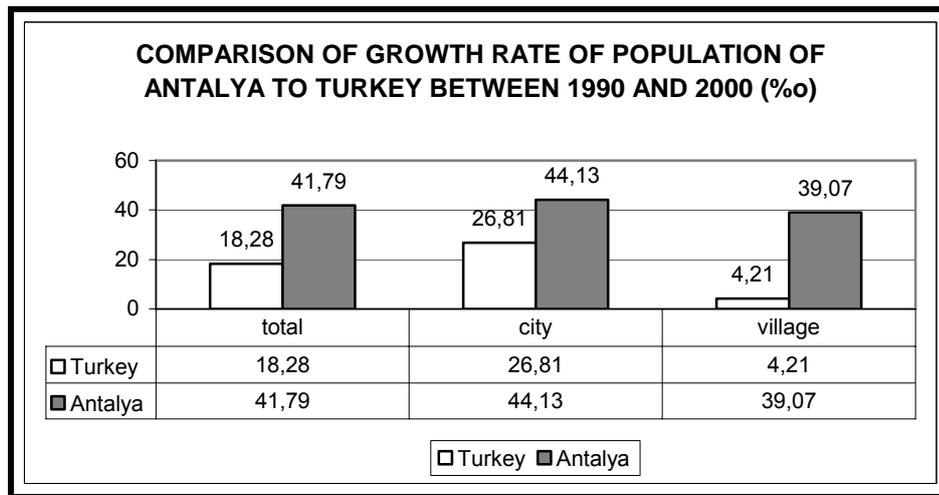
**Figure 3.1.** Location of the Study Area



**Figure 3.2.** The Study Area in Lat/Long. Projection System with the Municipality Borders

By the year 2000, the total population of Antalya rose to 1.719.751 and the population of the central district increased to 714.129. While the total growth rate of Antalya for the last decade is 41, 79% (Figure 3.3), it is 46, 44% for the central district. Beside this, Turkey's growth rate for the last decade is 18, 78%.

(DİE, 2000) This uncontrolled population increase because of the development of the tourism, commerce and industry activities in Antalya has pulled the migration to the city and has caused the city's expanding greatly as urbanization consumes large areas of agricultural and forest land (UTTA Planning, 1995). As it is given in the Table 3.1 while the density of population in Antalya in one square kilometer is 83 (population/km<sup>2</sup>), in the central district it is 354 (population/km<sup>2</sup>) (DİE, 2000). The main element that affects the environmental and physical structure of Antalya is the migration fact. Under the pressure of social-economic factors of migration, the urban structure of Antalya has an unhealthy and unstable growing tendency. Consequently, the quality of life in this city is degraded. By the effect of the population pulled by the tourism activity to the area, urban environment, natural resources, cultural worth, and the ecosystem has begun to loose quality (URL 4, 2005). Thus, Antalya is chosen as the study area due to these effects and its fast growing trend.



**Figure 3.3.** Comparison of the Growth Rate of Population of Antalya to Turkey for the last decade (‰), (DİE, 2000).

**Table 3.1.** Density comparison of Antalya and the Centre district, (DİE, 2000)

	Density (population/square kilometer)	Surface area (square kilometer)
Antalya	83	20.723
Centre District	354	2020

### 3.2. Data Used in the Study

Similar to other predictive models (Hung, 2002; Yang and Lo, 2003; Allen and Lo, 2003; Silva and Clarke, 2002; Dietzel and Clarke, 2004; Li and Yeh, 1998; Clarke and Gaydos, 1998), SLEUTH requires some input data in order to initiate the simulation. The 3.0-beta version of SLEUTH requires an input of at least five types of unsigned 8-bit GIF format data; urban extent, road, slope, excluded areas, and hill shade. In addition to these data if land use/cover is being analyzed for land use prediction of the study area in the future, the land use/cover datum can also be used as input for the model. Data used in the study are collected from governmental organizations, private companies, and internet. The list and the sources of the input data themes are given in Table 3.2.

At least four years of urban extent data are required for the model calibration (URL 3, 2005). For Antalya, four years of Antalya data (three satellite images and air photos) are obtained. The urban extent themes are produced from satellite images and air photos. The urban extents of 1987, 1996, and 2003 are produced from satellite image classifications. All the satellite images are obtained during the summer time namely June, July, or August. The year of 1987 satellite image (Thematic Mapper (TM) Landsat 30 meter image with 7 bands) is downloaded from internet address (URL 5) (Figure 3.4). For the urban extent of the year 1992, air photos obtained from General Command of Mapping are used (Figure 3.5). They are purchased as 18 pieces air photos. Then they are all scanned and georeferenced individually in respect of the coordinate system of the satellite images which have the projection system of UTM WGS 84, Zone 36. Not only air photos and satellite images, but also all the other data sets are projected into the same projection system. The 20-meter SPOT of 1996 image with 3 bands is acquired from INTA Spaceturk (Figure 3.6). The year of 2002 satellite image (Enhanced Thematic Mapper (ETM) Landsat 30 meter image with 8 bands) is also downloaded from the same address (URL 5) (Figure 3.7).

For the road theme at least two road layers are required and the dates of the road themes do not need to match exactly with the dates of urban extent

themes (Yang and Lo, 2003). In this study, the theme of roads contains two layers. These road layers are produced for 1995 and 2003. They contain 4 grade road types which are 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> degree of roads in the case study area. Generally these grades represent the importance, occupancy and consequently, relative urban attractiveness of roads (Yang and Lo, 2003).

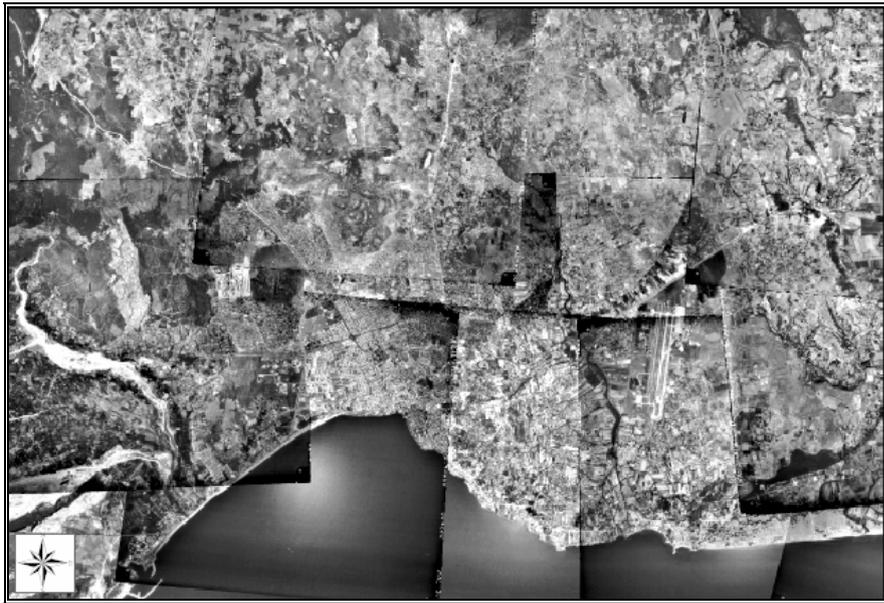
In Antalya, the first-degree roads connect the city's main development directions to the Centre, in which the other degree roads are included. While the first-degree roads represent the most important and occupied roads, the fourth degree ones represent the least occupied roads. However, for Antalya in 1995 the roles of these graded roads are not represented as well. Elker (1995) studied the planning at transportation of Antalya. The road data of the year 1995 is digitized (Figure 3.8) from the existing transportation map of the year 1995 which is prepared by Elker (1995), and the 1996 satellite image is also used in this stage. For the year 2003, transportation plan of Antalya for 2003 prepared by DAMPO Planning is used. The transportation plan, which is in CAD format is converted to shape file (Figure 3.9). The degrees of roads are queried and the unnecessary data are eliminated.

**Table 3.2.** List of the Model Input Data, Naming Format

Theme	Year	Source	Format	Naming Format for the model
Urban Extent	1987	Classified from 1987 Satellite Image (Landsat TM30m)	Raster	antalya.urban.1987.gif
	1992	Digitized from 1992 Air photos	Rasterized	antalya.urban.1992.gif
	1996	Classified from 1996 Satellite Image (Spot 20m)	Raster	antalya.urban.1996.gif
	2002	Classified from 2002 Satellite Image (Landsat ETM 30m)	Raster	antalya.urban.2002.gif
Roads	1995	Digitized from Antalya Transportation Map and Satellite Image of 1996	Rasterized (originally cad data)	antalya.road.1995.gif
	2003	Digitized from DAMPO Planning, 1/5000 Development Plan and Satellite Image of 2002	Rasterized (Originally cad data)	antalya.road.2003.gif
Slope	1996	Computed from SRTM 90m DEM	Raster	antalya.slope.gif
Excluded Areas	2003	Digitized from 1/5000 Development Plan	Rasterized (Originally cad data)	antalya.excluded.gif
Hill shade	1996	Computed from SRTM 90m DEM	Raster	antalya.hillshade.gif



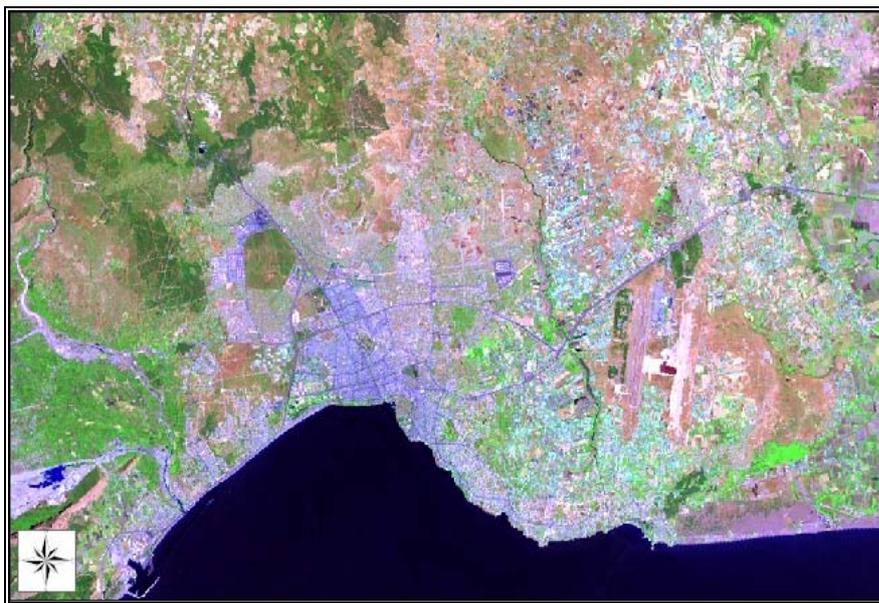
**Figure 3.4.** 1987 Landsat TM satellite imagery of the study area with 30 meter spatial resolution.



**Figure 3.5.** 1992 Aerial photograph mosaic of the study area with a scale of 1/40,000



**Figure 3.6.** 1996 Spot satellite imagery of the study area with 20-meter spatial resolution.



**Figure 3.7.** 2002 Landsat ETM satellite imagery of the study area with 30-meter spatial resolution.



**Figure 3.8.** Road Data for the Year 1995



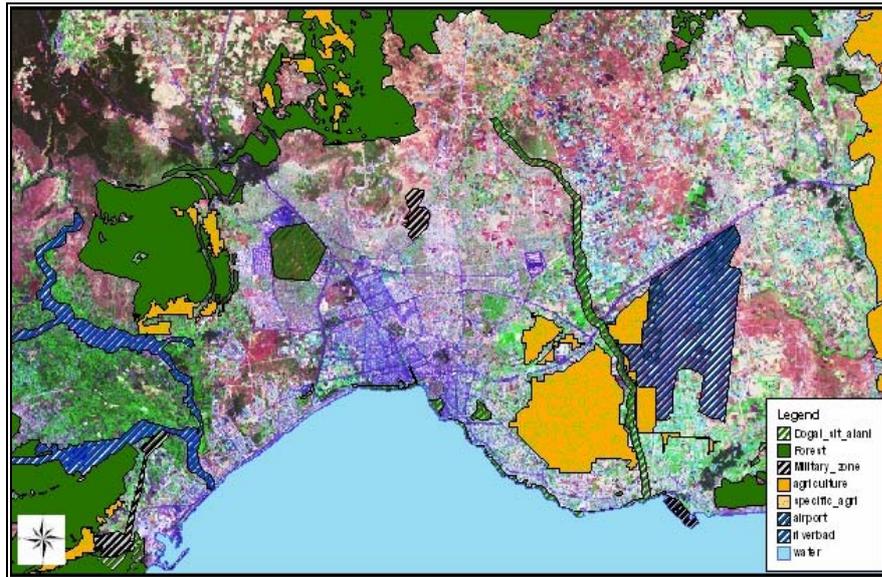
**Figure 3.9.** Road Data for the Year 2003

The roads can be weighted according to their relative urban attractiveness. This can be done by adding a width dimension to highways that are the centre line of roads, and nodes, which are points. Thus, weighting buffers can be applied to the highways and the nodes. Then the buffered features can be assigned

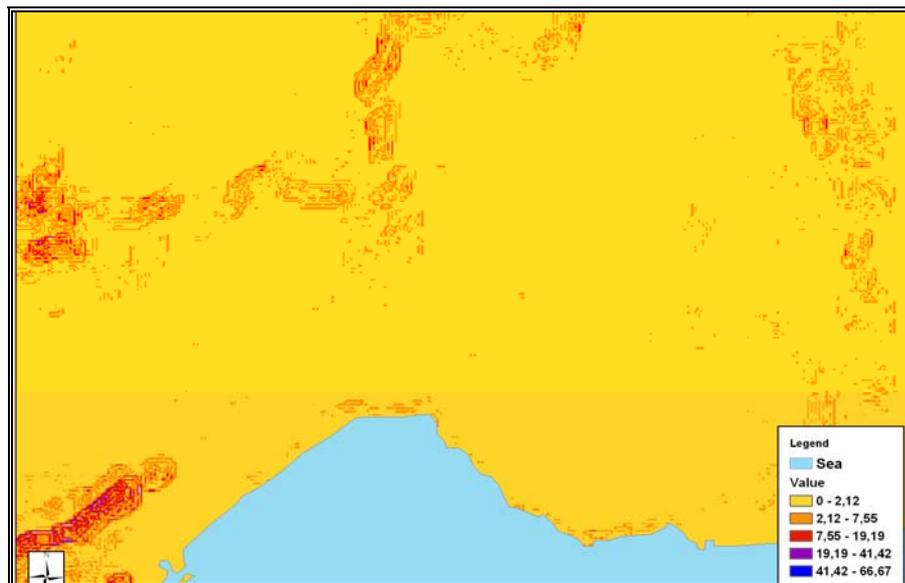
different values ( $0 < n < 255$ ) according to their relative importance before being converted into raster format (Yang and Lo, 2003). This method was not applied in this study.

For the excluded areas, one layer is produced from the Antalya 1/5000 Development Plan (DAMPO Planning, 2003) which is approved by the government. In respect of this plan, there are unresidential areas in Antalya and those must be preserved. Thus, natural conservation areas, forest covered areas, military zones, first and second-degree productive agricultural areas, airport, and riverbeds are digitized (Figure 3.10). The Mediterranean Sea by the Antalya also digitized in this layer as the attribute name with water. Not all these areas in this layer are allowed for urban development. In addition, a second layer of the excluded areas that contains three levels of buffer zones around rivers, forests, riverbeds, and conservation areas could be put into the model. The buffer zone within 50 meter could be assigned a value of 100, meaning that this area is not allowed for urban development at all; the buffer between 50 and 100 meter could be assigned a value of 60, indicating a 60% probability of exclusion; and the buffer zone between 100 and 200 meter could be assigned a value of 20, indicating a 20% probability of exclusion.

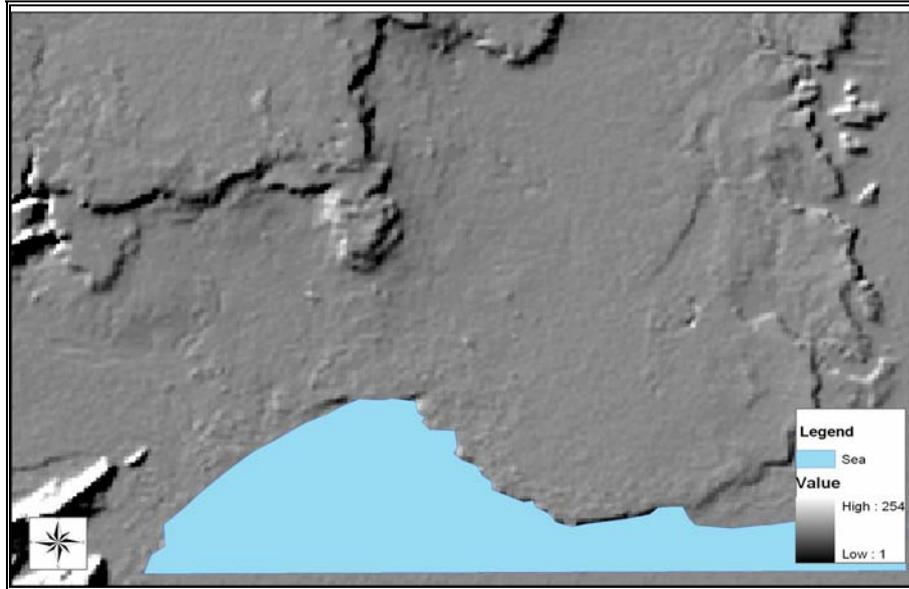
The slope should be derived from a digital elevation model (DEM). A layer of terrain slope is computed from the SRTM 90 meter DEM data that is obtained from URL 5. Cell values must be in percent slope, not in degree (Figure 3.11). Moreover, the pixel value range must be between 0 and 100. A layer of the hill-shaded image is computed from SRTM 90 meter DEM (Figure 3.12). This image shows the topographic relief in the study area and it is used in the model in order to give spatial context to the urban extent data as a background image for visualization purposes. This must be a grayscale image as it is given in the following figure (Figure 3.12).



**Figure 3.10.** Excluded Areas (unresidential areas)



**Figure 3.11.** The Layer of Terrain Slope.

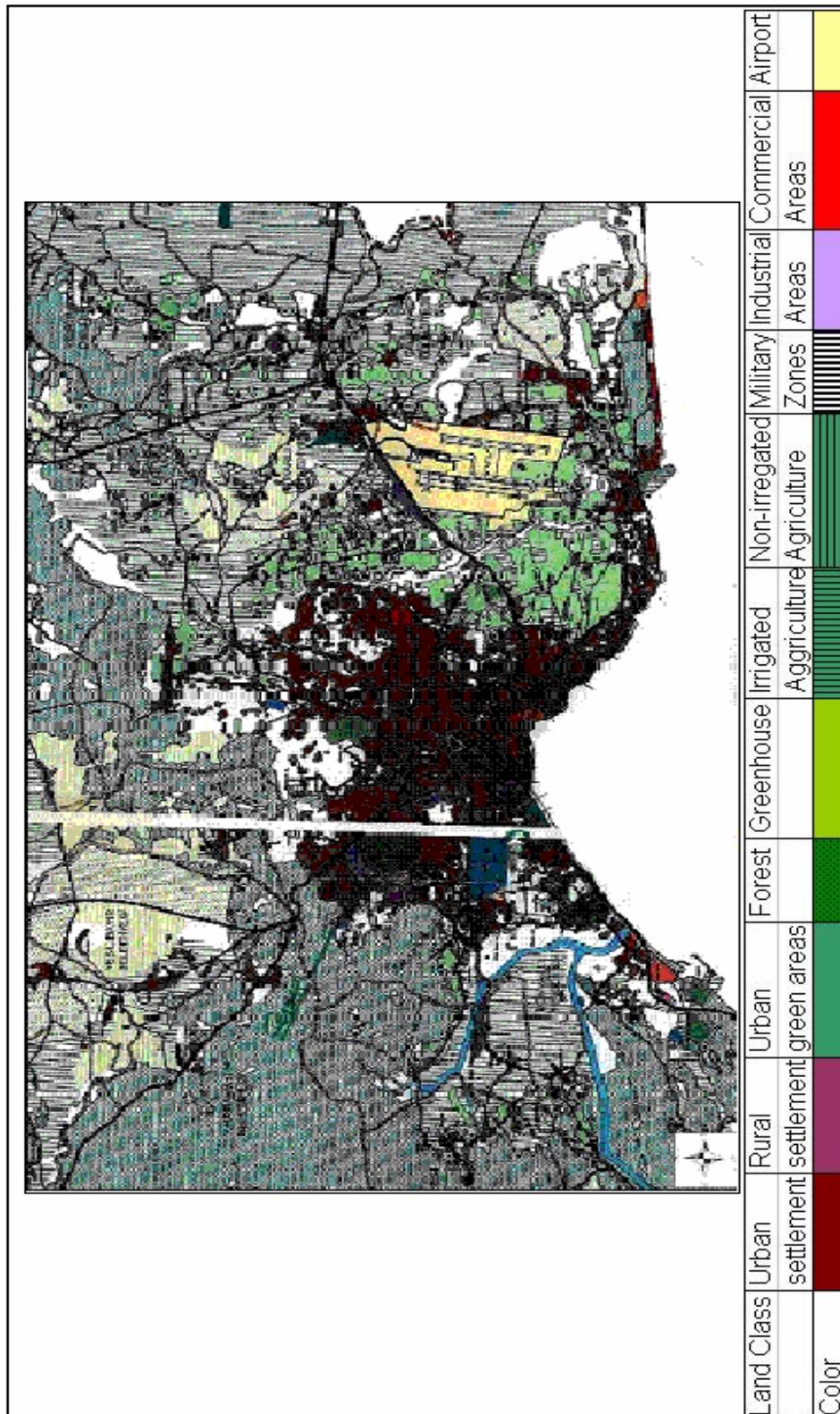


**Figure 3.12.** Layer of the Hillshade (URL 5, 2004).

Addition to all these data, a land use data of the year 2003 (Figure 3.13) is used to assess the accuracy of the classification results. It is obtained from DAMPO Planning Firm, which is the latest Antalya Development Plan of Antalya. It contains all the existing land use and land cover features of the study area as settlement areas, active green areas as parks in the city and the passive green areas as forest, watery agriculture areas, dry agriculture areas, empty spaces, and airport.

### **3.3. Analyses of the Satellite Images**

Anderson et al. (1976) stated that there is no one ideal classification of land use and land cover, and it is unlikely that one could ever be developed. He notices that there are different perspectives in the classification process and says that in almost any classification process, it is rare to find the clearly defined classes that one would like. For the SLEUTH model as it is told in the literature survey (chapter two), generally the Landsat satellite images are used because of their medium resolution (Yang and Lo, 2003). For the model, exposing the urban extent and non-urban extent classes are sufficient. There is no need to expose the details of the urban features.



**Figure 3.13.** Land Use Data

The Landsat satellite images with 30 meter resolution for the years of 1987 and 2002 and the SPOT satellite image with 20 meter resolution for the year of 1996

are classified to obtain the urban extent input data. For the urban extent of the year 1992, there is no need for the classification because aerial photographs with a scale of 1/40.000 are used in order to extract the built up areas as urban extent.

### **3.3.1. Classification Analyses of the Satellite Image of the Year 2002**

To obtain the urban and non-urban classes Landsat 2002 satellite image is classified into five classes, which are urban, green areas together with the active (used in daily life) and passive ones (not in used in daily life), agricultural areas, empty spaces, and water. The classification is performed with PCI Geomatica V9.1. The training areas were collected according to these five classes. However, the result was not acceptable because of the mixed pixels. The water pixels were not identified precisely and were mixed to the agriculture and urban pixels. In addition to this, the empty spaces were mixed with agricultural areas. Therefore, a rule for water, which is valid for Landsat, is applied to the 2002 satellite image. In Landsat Images, water is the unique land cover type such that the band reflectance values always decrease as the band number increases, ignoring the thermal band (band 6). In other words, if pixel (i, j) lies entirely within a water region then  $b_{ij}^1 > b_{ij}^2 > b_{ij}^3 > b_{ij}^4 > b_{ij}^5 > b_{ij}^7$  (Avci, 2000). There were many water pixels in the output so a low pass mode filter with 3 x 3 window is applied. The filter output was not satisfactory. Following, a 5 x 5 window low pass mode filter is applied but the result was not as expected. So it is decided to put water into the excluded area layer. It is understood that to expose the water class from a 30 meter Landsat image is complicated. However, an accuracy assessment by using 2003 land use map is realized to test the classification result. In PCI, 510 random points are appointed. Before the random appointment of the points, Mediterranean Sea is masked from the satellite image. The random points are not allowed to be assigned on this water mask in order not to affect the accuracy of the result. The numbers of samples chosen randomly from each class are proportional to the percentage of the image occupied by each class. As a result, the overall accuracy is determined as 61.644%. It was very low for an acceptable accuracy level according to Anderson et. al., (1976). As Hung (2002) clarified, it

is not very easy to identify all the classes one by one, or to distinguish from each other, without prior information of the study area.

The study area comprises of many land use features. As a stipulation of a supervised classification, the training areas are collected. After collecting the training areas for five classes, it is decided that increasing the number of classes would be useful to identify the features clearly. If the satellite image and the land use map are examined the abundance of the land use classes can be seen frankly. Therefore, for the classification of the satellite image of the year 2002, the numbers of classes are extended from 5 to 13 (Bauer et al., 2003). Therefore, the confusion of the land classes would be minimum. In the study area, four different urban textures are determined. Areas with label of urban\_1 represent the obvious urban areas particularly in the Centre of the city and the areas with the most density of the population. Urban\_2 areas are the areas near the urban\_1 areas with the less population of density. Urban\_3 includes the industrial areas as small industrial complexes as a part of urban. Urban\_4 represents the least population density areas. In addition to these urban areas, the airport in the east part of the Centre is completely labeled as urban 5 because it is also a built up area which reflects as urban texture. Following, three different green area textures are determined from the satellite image and the land use map as green\_1, green\_2, and green\_3. Here, the green 3 represents the darkest green lands which belongs to forest reflectance, green\_2 represents the green areas with less darker reflectance than the green\_3 and green\_1 represents the green areas with the least green reflectance and which particularly belongs to the agricultural green and the green areas inside the urban areas. From the satellite image, two different empty spaces or barelands are determined. One is the obvious empty spaces represented as bareland\_1 and the other is the bareland\_2.

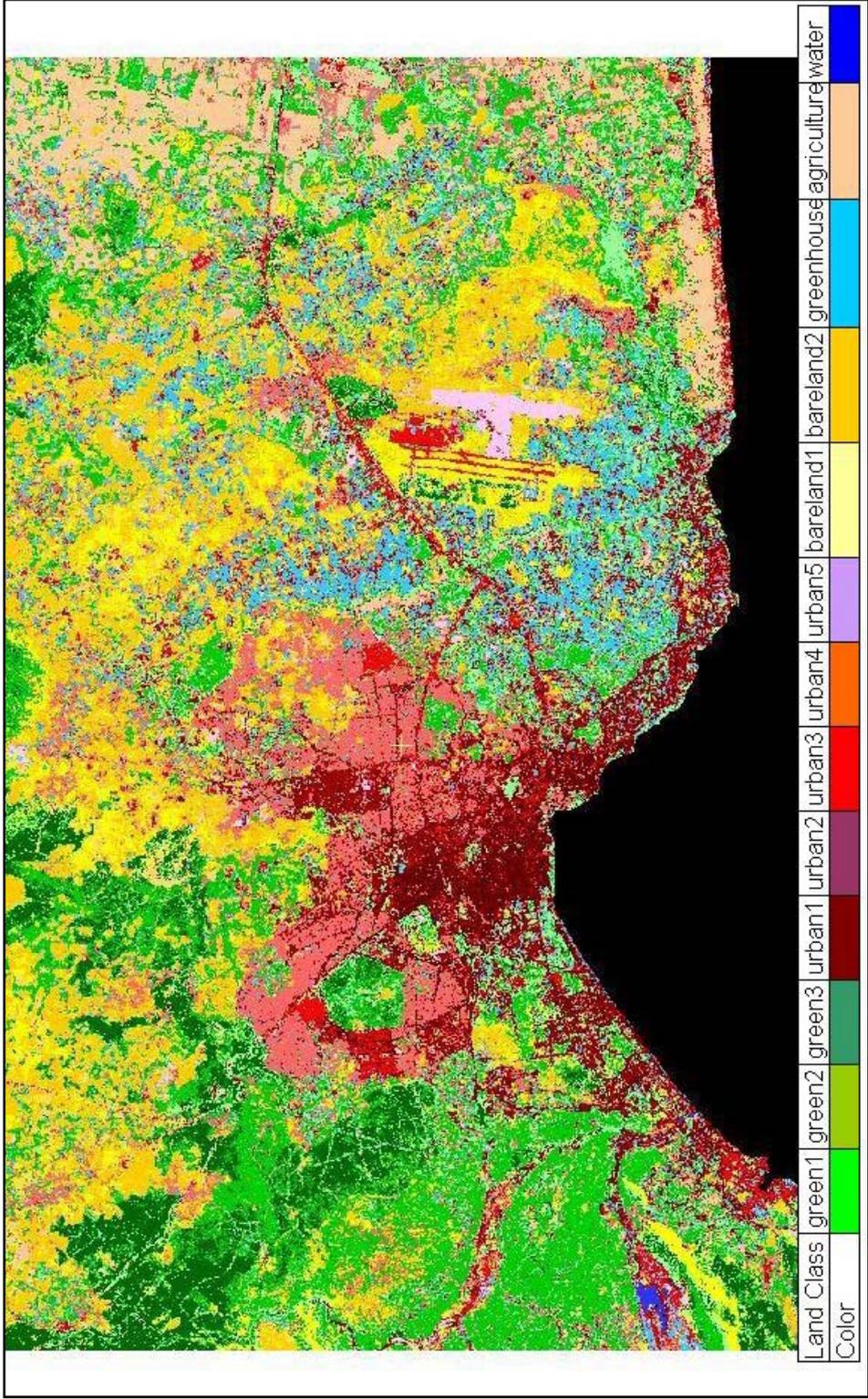
The training areas for all the classes are collected precisely. It was noticed that the minimum number of training area pixels are collected as the 10% of the total number of the pixels of the involved class. For example for the class urban\_1, the collected training area pixels were the 10% of the total urban\_1 reflected pixels in the satellite image.

The classification method used in this thesis is the Supervised Maximum Likelihood Classification, which is the most common method as Bauer et al. (2003) applied in their study in Minnesota. The resultant thematic map of the classification is displayed in Figure 3.14. The classes were aggregated to 5 land use classes and the resultant map is shown in Figure 3.15. The three types of green areas were aggregated to 1 green class, the five types of urban areas were aggregated to one urban class, and the three types of agricultural areas were aggregated to one agriculture class. Water and bareland were the other two classes.

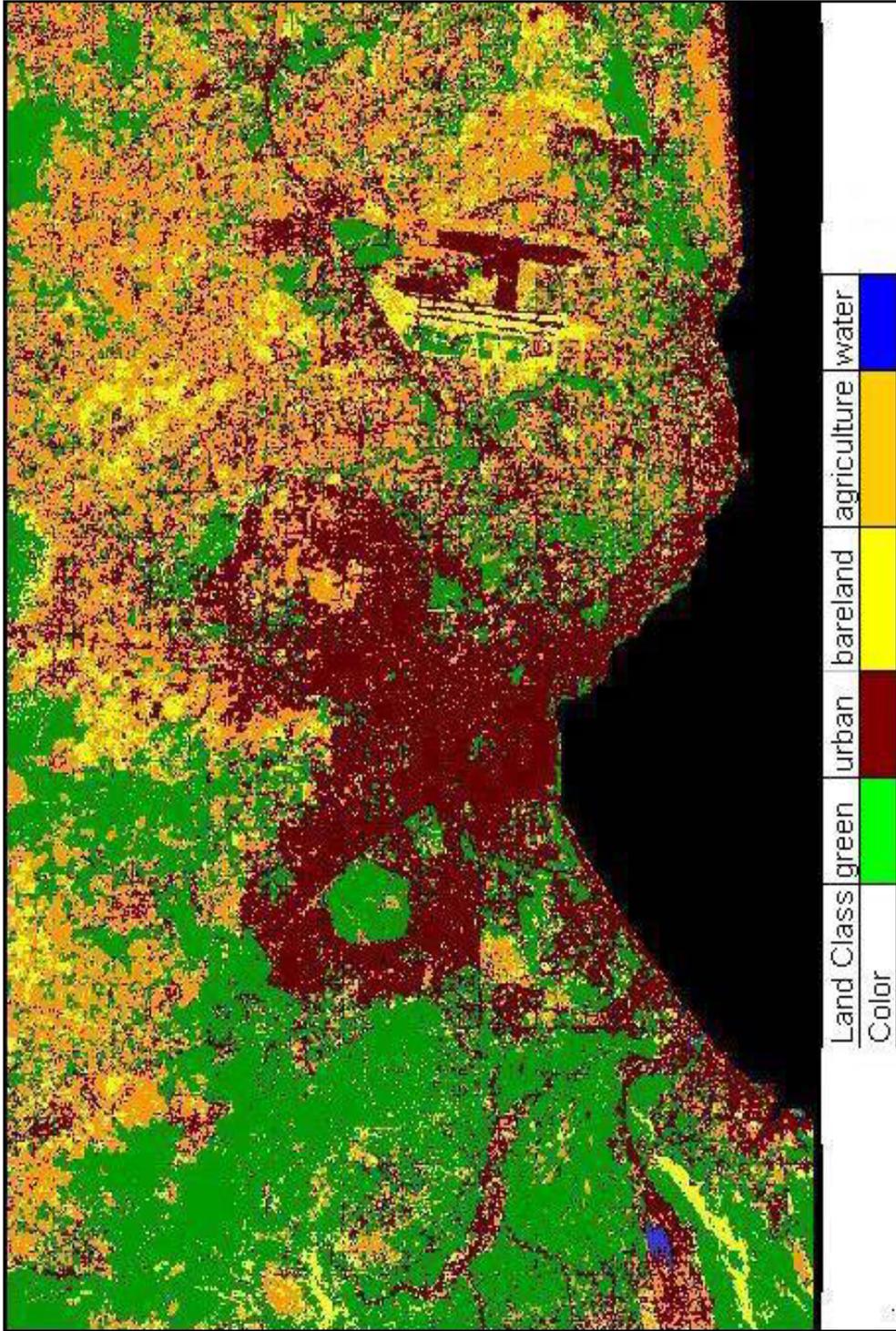
If the aggregation result (Figure 3.15) is compared to land use map (Figure 3.13) of the year 2003 by the human view, the strong similarity can be realized easily. However, to maintain this similarity the accuracy assessment of this classification was also realized by means of the land use map. Before the accuracy assessment, a low pass mode filter with a window of 7 x 7 is applied to the aggregation result to obtain the integrity of the classes (Figure 3.16). A mode filter computes the mode, most occurring value, of the pixel values and replaces the least occurring value with the most occurring value.

### **3.3.1.2. Recoding of the Raster Layer of the Year 2002**

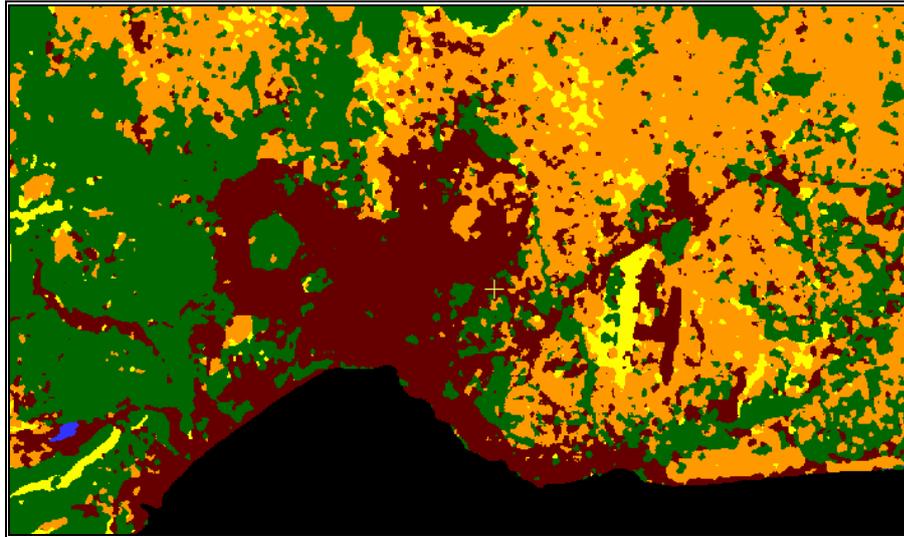
The model requires the urban extent layer as urban and non-urban binary layer. Therefore, in ArcGIS 8.3 and ERDAS 8.6, mode7 x 7 filtered layer is converted to raster imagine format then the pixel values of all classes except urban are recoded as zero and the urban layer is recoded as one because of the urban extent requirement of the model. The recoded new image is shown in Figure 3.17. In the figure, the black pixels are recoded as the pixel value zero and they represent the non-urban areas, and the white pixels are recoded as the pixel value 1 and they represent the urban pixels.



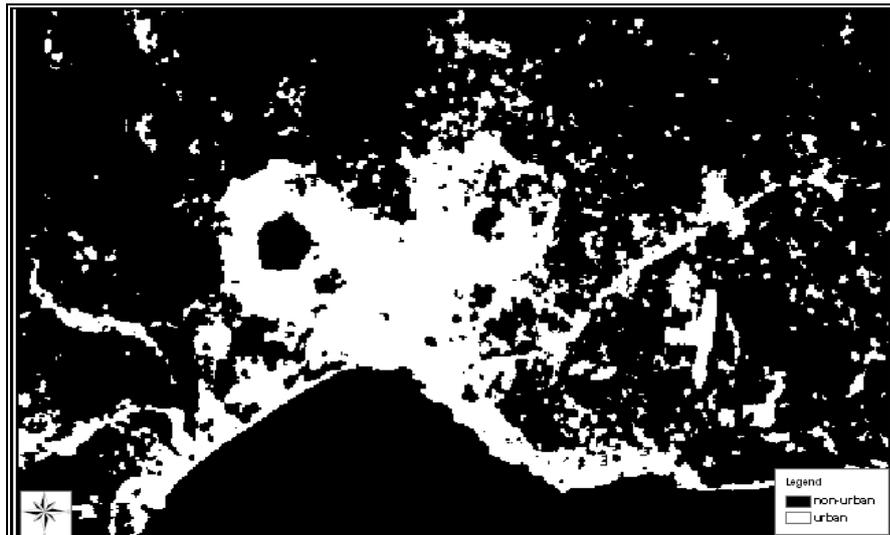
**Figure 3.14.** The Result Thematic Map of the Supervised Maximum Likelihood Classification Analysis of the Year 2002.



**Figure 3.15.** Aggregation of the Maximum Likelihood Classification for the year 2002.



**Figure 3.16.** Mode 7 x 7 Filter Application to the Aggregation Result of the year 2002.



**Figure 3.17.** 2002 Urban Extent Layer after Recoding the Mode Filtered 2002 Image (white is urban, black is non-urban pixels).

### 3.3.1.3. Accuracy Assessment of the Classification Result of the Year 2002

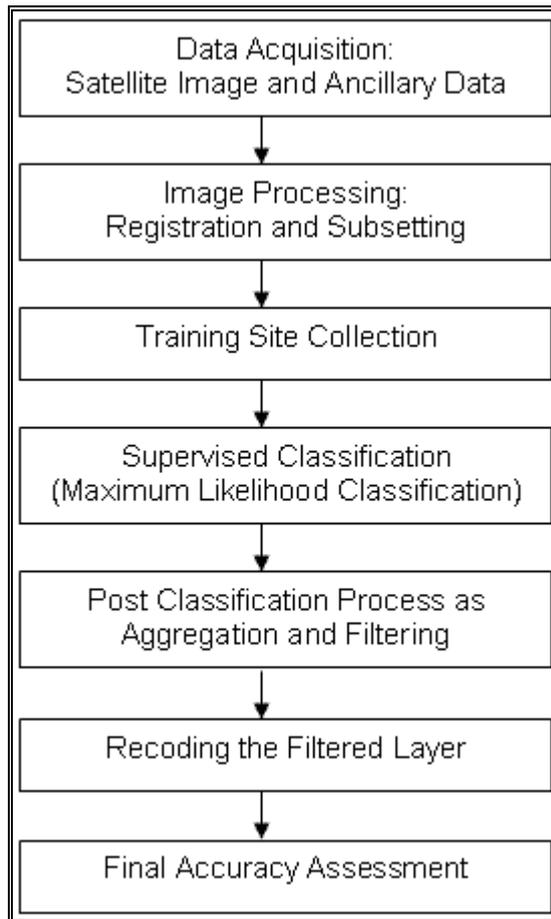
Assessing the accuracy of a remote sensing application is one of the most important steps in any classification exercise. Without an accuracy assessment the output or results is of little value (URL 6, 2005). Therefore, 510 random points are

assigned to this recoded binary layer in order to obtain the accuracy assessment statistics. As a result, overall accuracy is obtained as 91.720% at the 95% confidence interval. The producer's accuracy is obtained as 92.4% for non-urban areas, 88.8% for the urban areas. According to Anderson (1976), accuracy statistics over 85% are adequate. Here, the overall accuracy is the sum of the diagonal elements divided by the total number in the sample. The producer's accuracy represents the percentage of a given class that is correctly identified on the map (the percentage correct for a given row divided by the total for that row) (Lecture Notes of the Integration of RS and GIS, 2004).

On the base of these statistics, this classification methodology steps were decided to be used for the year of 1987, 30 meter of resolution Landsat image and for the year of 1996, 20 meter of resolution SPOT image. The flowchart of the methodology from the step of data acquisition to the step of accuracy assessment of the recoded image is shown in the Figure 3.18.

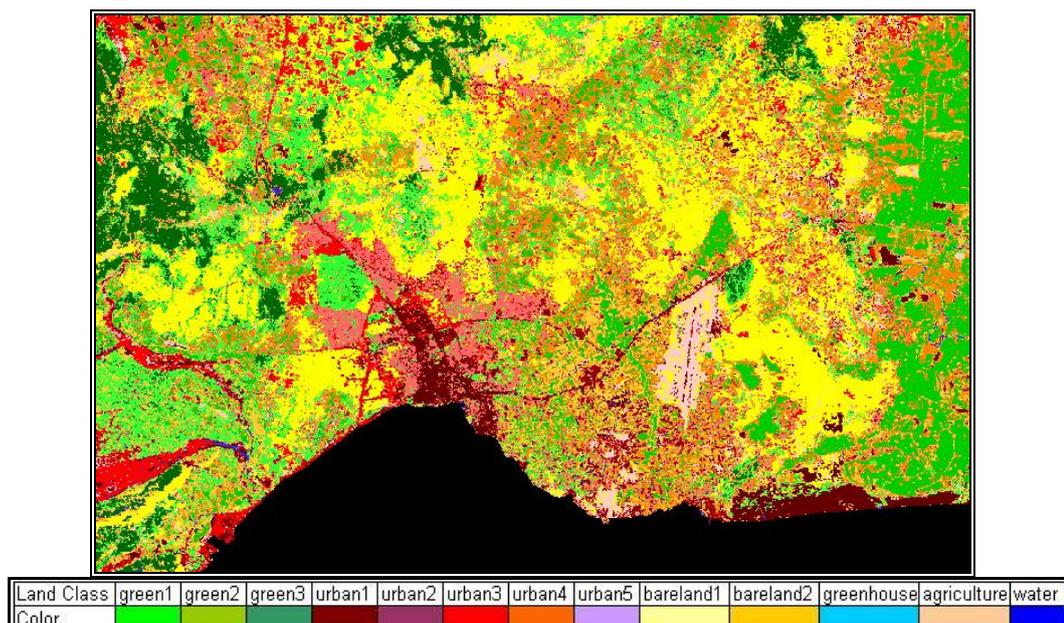
### **3.3.2. Classification of the Satellite Images of the Year 1987 and 1996**

For the classification of the satellite images of the years 1987 and 1996, the above flowchart is used as a guide. The ancillary data as land use or vector data could not be found for these years so it is thought that the methodology used for the year of 2002 could guide these years classification analysis because the accuracy assessment result for this year was very high as 91.720%. After the registration and subsetting the images of the years 1987 and 1996, the training areas for these years are collected precisely for the same thirteen classes. Then the maximum likelihood classification method is run. The results of the classifications are shown in Figure 3.19 and Figure 3.20.

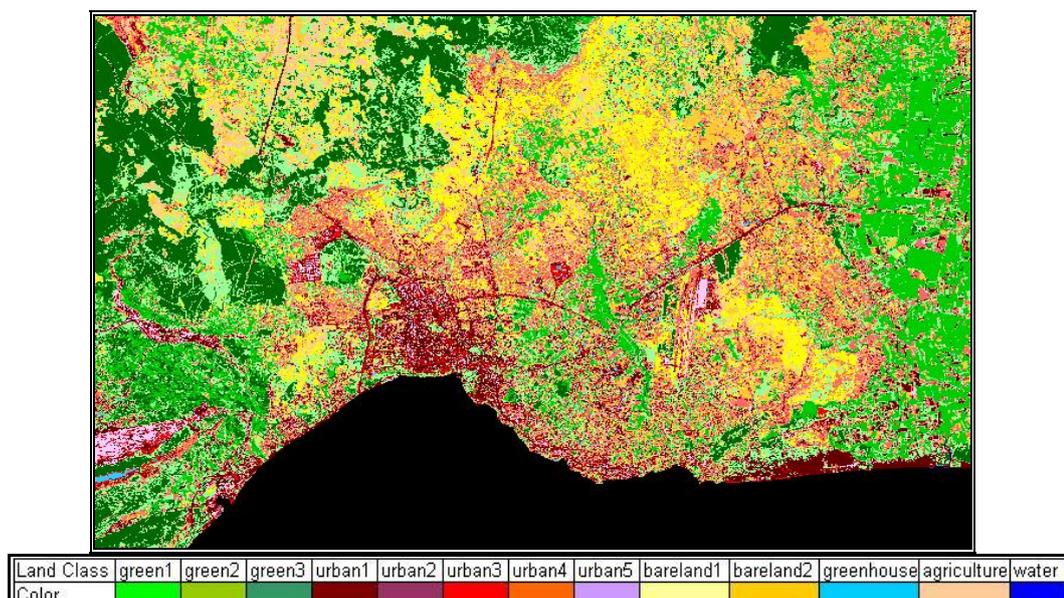


**Figure 3.18.** The Flowchart of the Methodology of the Classification Analysis from the Data Acquisition Step to the Accuracy Assessment Step

From the classification result of the year 1987, it can be figured out that the empty spaces with yellow color are much dominant than the ones in 2002. When the classification results are examined, it was found that the mixed pixels appeared a lot in the satellite image of the year 1996 because of the better pixel resolution of 20 meter (Figure 3.20).



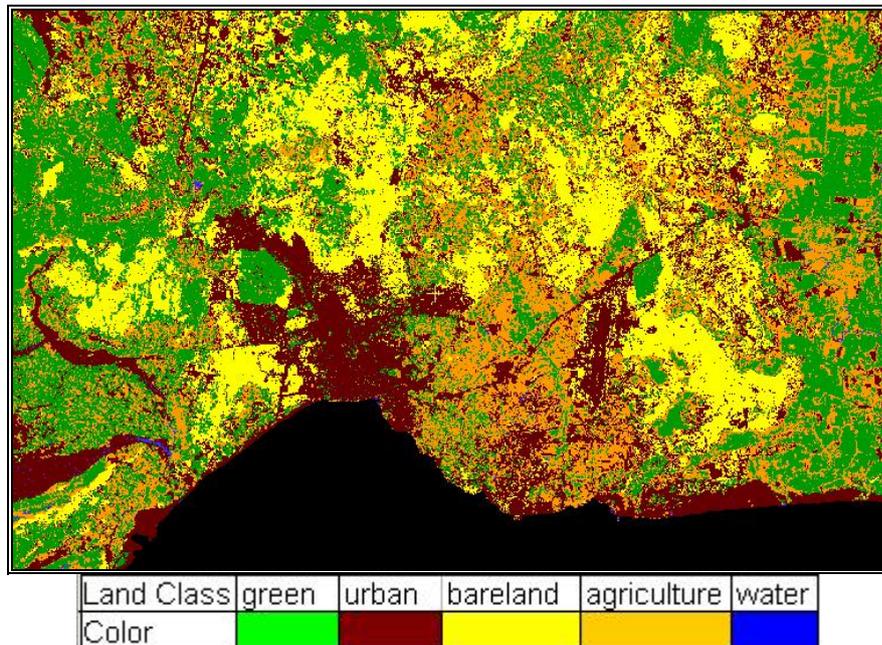
**Figure 3.19.** The Resultant Thematic Map of the Supervised Maximum Likelihood Classification Analysis of the Year 1987



**Figure 3.20.** The Resultant Thematic Map of the Supervised Maximum Likelihood Classification Analysis of the Year 1996

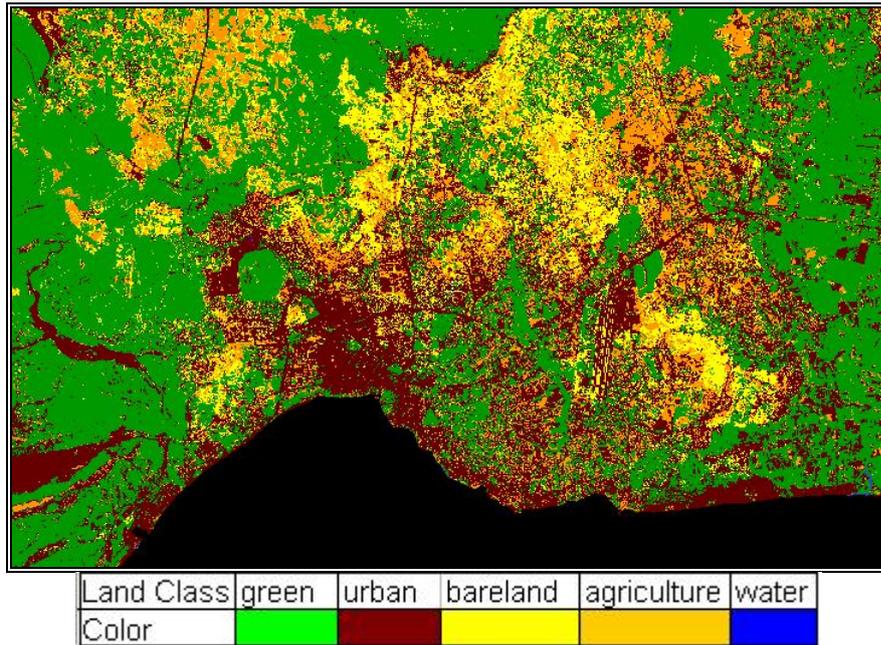
### 3.3.2.1. Post Classification Analyses of the Years 1987 and 1996

After the classification analysis, the post classification steps of aggregation for these two years are performed (Figures 3.21 and 3.22).



**Figure 3.21.** Aggregation of the Maximum Likelihood Classification for the year 1987.

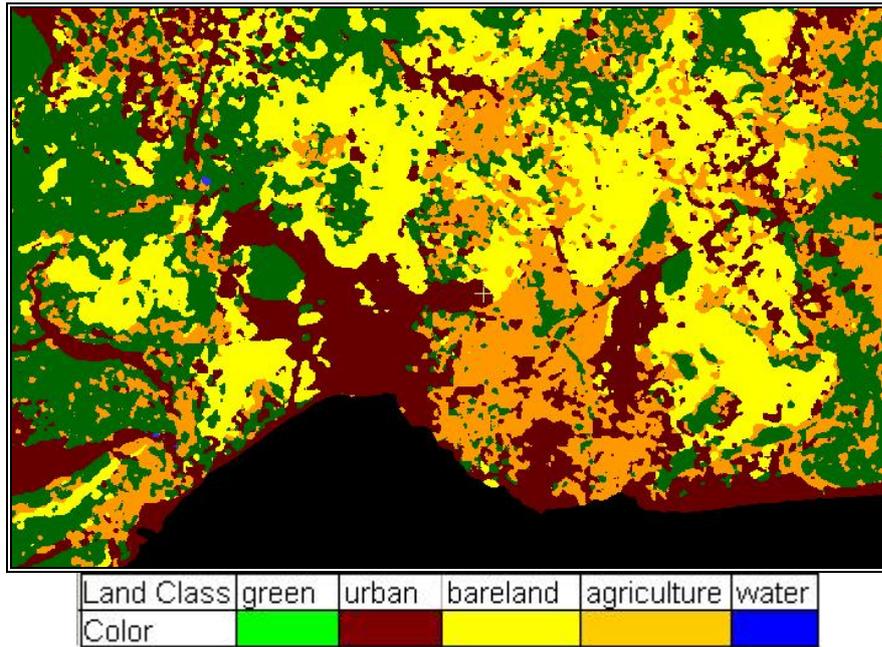
After the aggregation step, number of classes 13 is decreased to 5 as displayed in Figures 3.21 and 3.22. The new classes are green areas with the green color, urban and built up areas with the dark brown color, barelands with the yellow color, agriculture with the light brown color, and the water with the blue color. In addition to these classes, the black color is the bitmap of the Mediterranean Sea. For the integrity of the classification classes of these two years, the next post classification step of filtering is produced (Figure 3.23 and 3.24). The low pass 7 x 7 pixel window filter is applied.



**Figure 3.22.** Aggregation of the Maximum Likelihood Classification for the year 1996.

### 3.3.2.2. Recoding of the Raster Layers of the Years 1987 and 1996

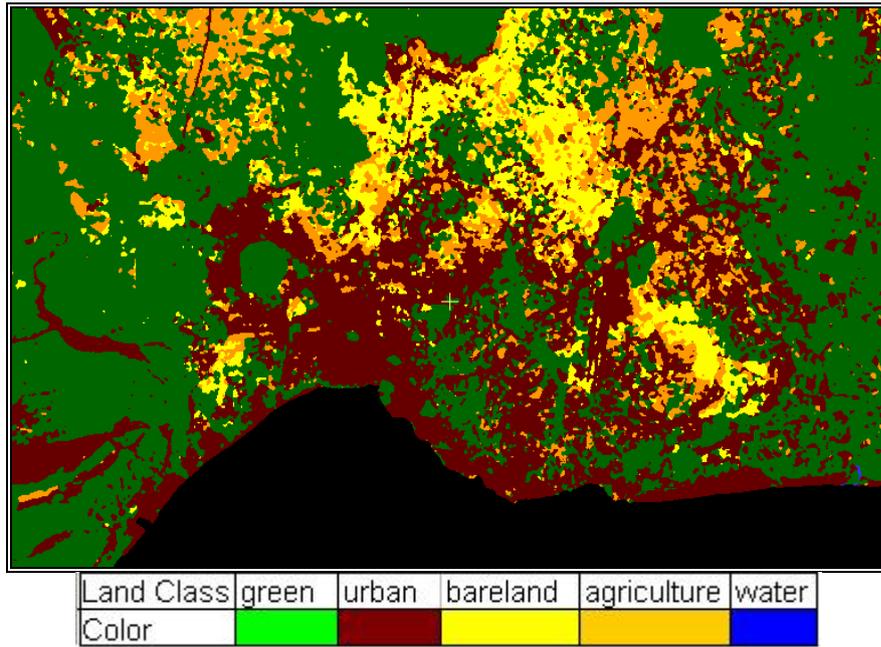
As the last step of the classification methodology, raster conversion is realized in ArcGIS 8.3 and Erdas 8.6 in order to recode the urban extent layers. They are converted to Erdas imagine raster format. Following, these raster layers are recoded in Erdas as binary urban and non-urban pixels (Figure 3.25 and 3.26). The pixel values of all classes except urban are recoded as zero and the urban layer is recoded as one because of the urban extent requirement of the model.



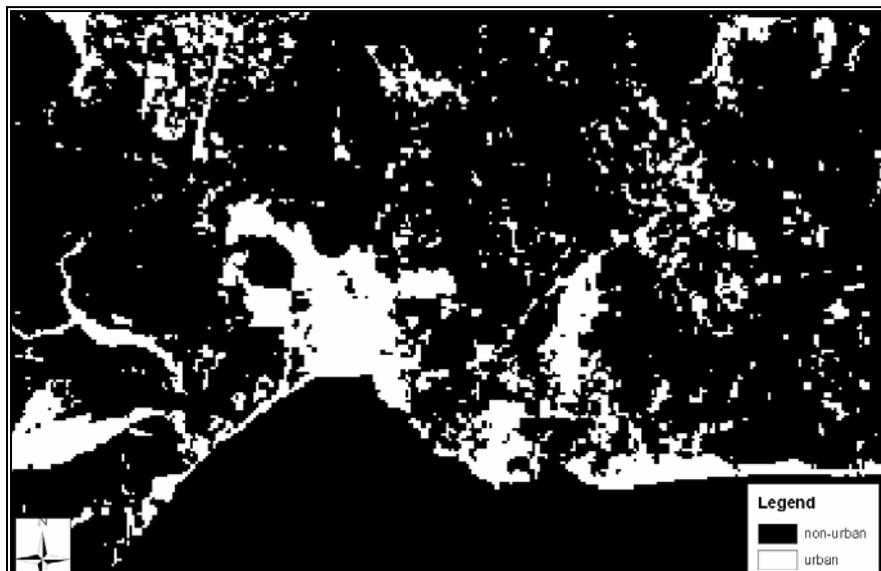
**Figure 3.23.** Mode 7 x 7 Filter Application to the Aggregation Result of the year 1987.

### 3.3.2.3. Generation of the Model Function for the Classification Results of the Years 1987 and 1996

It must be known that the urban extent of the year 2002 is the guide and the correct one. When the Figures 3.25 and 3.26 are examined, it can be seen that some parts of the images 1987 and 1996 are more built up than the ones in 2002 image (Figure 3.17). As an instant, the North West part and East part of the image of 1987 is whiter than the image of the year 2002. In addition, all the east part of the image 1996 is whiter than the image 2002. It can be understood that the classification methodology of the 2002 image with the high accuracy statistics, did not work for the 1987 and 1996 images. It is thought that if the urban pixels of the 1996 image are not present in 2002 image, then these pixels are incorrectly classified and could not be removed by the filtering.



**Figure 3.24.** Mode 7 x 7 Filter Application to the Aggregation Result of the year 1996.



**Figure 3.25.** 1987 Urban Extent Layer after Recoding the Mode Filtered 1987 Image (white is urban, black is non-urban pixels).



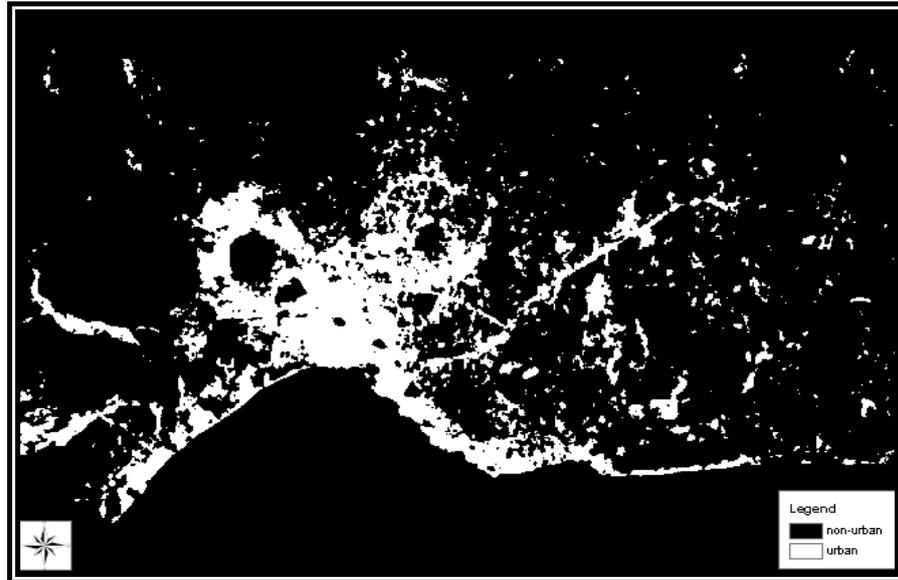
**Figure 3.26.** 1996 Urban Extent Layer after Recoding the Mode Filtered 1996 Image (white is urban, black is non-urban pixels).

For the consistency of the urban extent layers, a function definition is produced;

“EITHER 0 IF (\$n1\_02wgs\_bm== 0 and \$n2\_96wgs\_bm == 1) OR \$n2\_96wgs\_bm OTHERWISE”

The function above stipulates that if the urban pixels found in the 1996 layer, cannot be found in the 2002 layer, then recode them as non-urban pixels, as zero.

According to this model function, the wrongly classified urban pixels in the 1996 image are corrected in respect to the urban pixels in the 2002 image. The urban extent pixels of the year 1996 could not be more than the ones in 2002 because any earthquake or a disaster, which could destroy the built up areas, did not occur during this time. The corrected 1996 urban extent layer is shown in Figure 3.27.



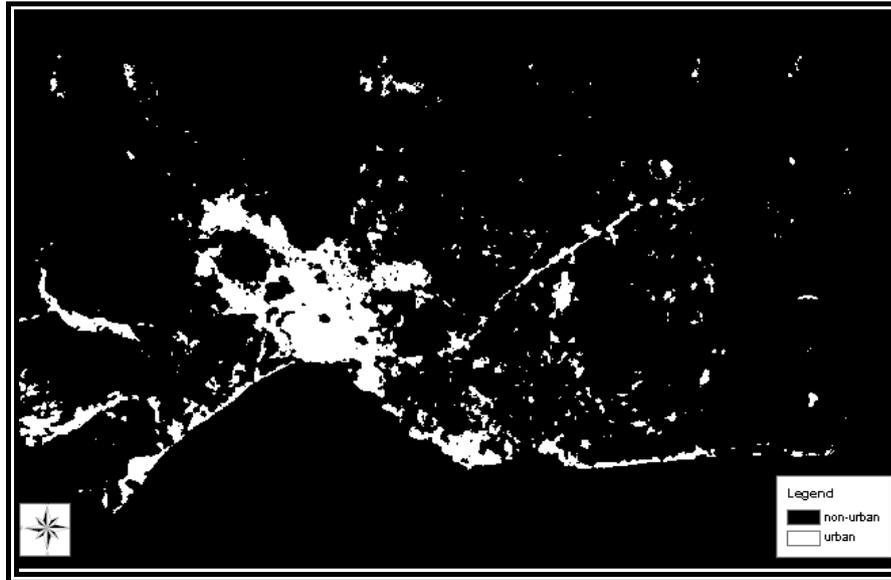
**Figure 3.27.** The Model Applied and Corrected 1996 Urban Extent Layer (white is urban, black is non-urban pixels).

The same model function is applied to the 1987 urban extent layer in respect to corrected new 1996 urban extent layer. Thus, the wrongly classified 1987 urban pixels are removed and converted to non-urban pixels by means of this model function;

“EITHER 0 IF (\$n1\_96new== 0 and \$n2\_87bm == 1) OR \$n2\_87bm OTHERWISE”

The function above stipulates that “if the urban pixels found in 1987 layer, cannot be found in 1996 layer, then recode them as non-urban pixels”.

The new 1987 corrected urban extent layer is displayed in the Figure 3.28. So, the consistency of the urban extent layers is obtained.



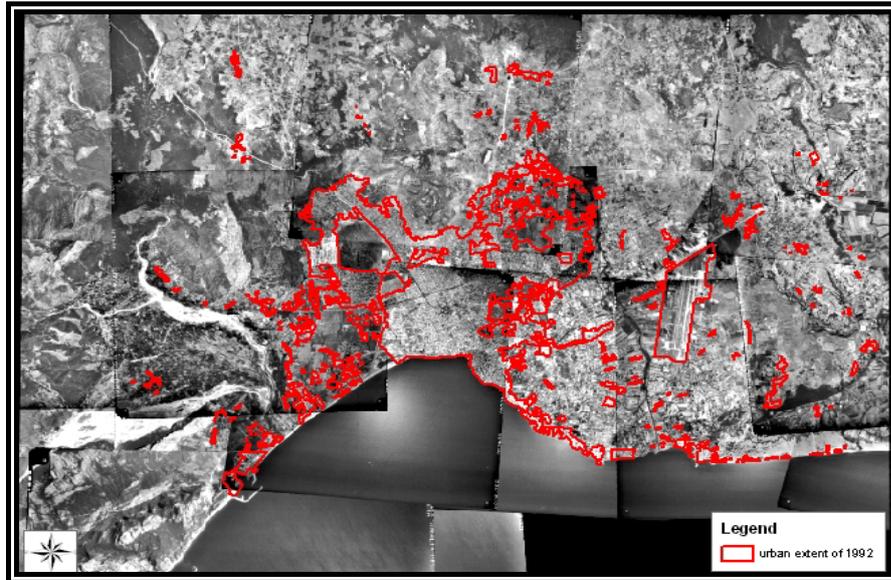
**Figure 3.28.** The Model Applied and Corrected 1987Urban Extent Layer (white is urban, black is non-urban pixels).

### **3.4. Analyses of the 1992 Aerial Photographs**

For urban extent of the year 1992, the satellite images could not be found. Therefore, aerial photographs of the study area were obtained from the General Command of Mapping. First, with the scale of 1/40.000, 18 pieces of aerial photographs were registered one by one. Following, they were clipped and then mosaic application in TNT was done for their integrity of the view (Figure 3.5).

#### **3.4.1. The Generation of the Urban Extent Layer of the Year 1992**

The urban extent of the layer 1992 is precisely digitized from the mosaic image in ArcGIS. The similar process is realized for Istanbul for monitoring the landuse dynamics (Çelikoyan et al., 2003). The urban extent of the year 1992 is displayed in the Figure 3.29.



**Figure 3.29.** The Layer of the Urban Extent of the Year 1992 (Red Polygons are the urban extents)

### 3.4.2. Vector to Raster Conversion and the Recoding of the Urban Extent Layer of the Year 1992

The vector layer of the urban extent of the year 1992 is converted to image format in ArcGIS and Erdas. Following, the image pixels are recoded in Erdas. In order to be read by the model, the urban pixels are recoded as the pixel value one, and the non-urban pixels are recoded as the pixel value zero. The recoded layer is shown in Figure 3.30.

### 3.5. Analyses of the Roads Layer

As it is explained at the beginning of this chapter, for the road theme at least two road layers are required by the model and dates are not expected to match exactly with the urban dates (Yang and Lo, 2003). The road layers are produced for 1995 and 2003. They contain 4 grade road types which are 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> degree of roads in the case study area.



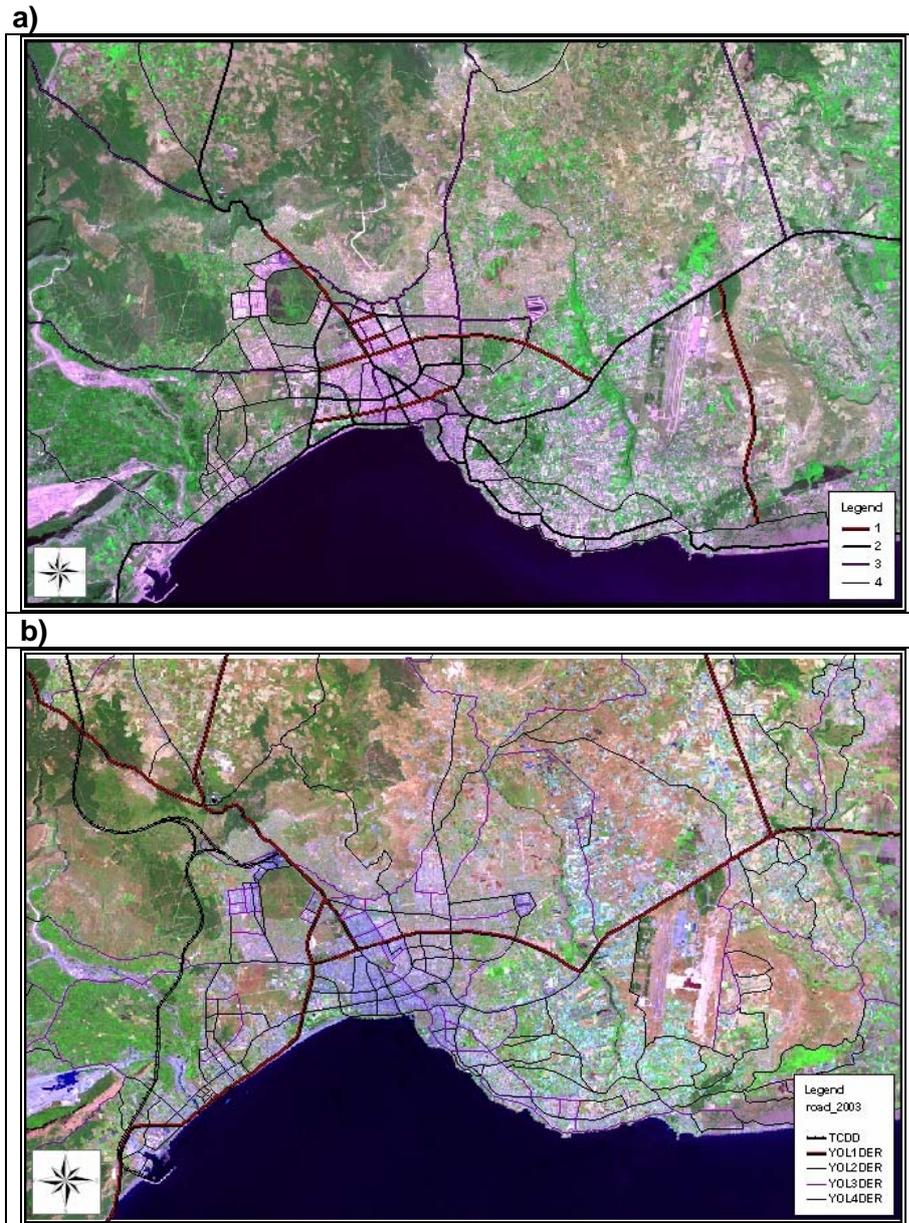
**Figure 3.30.** Recoded Urban Extent Layer of the Year 1992 (white is urban, black is non-urban pixels).

### **3.5.1. The Generation of the Vector Layers of the Roads**

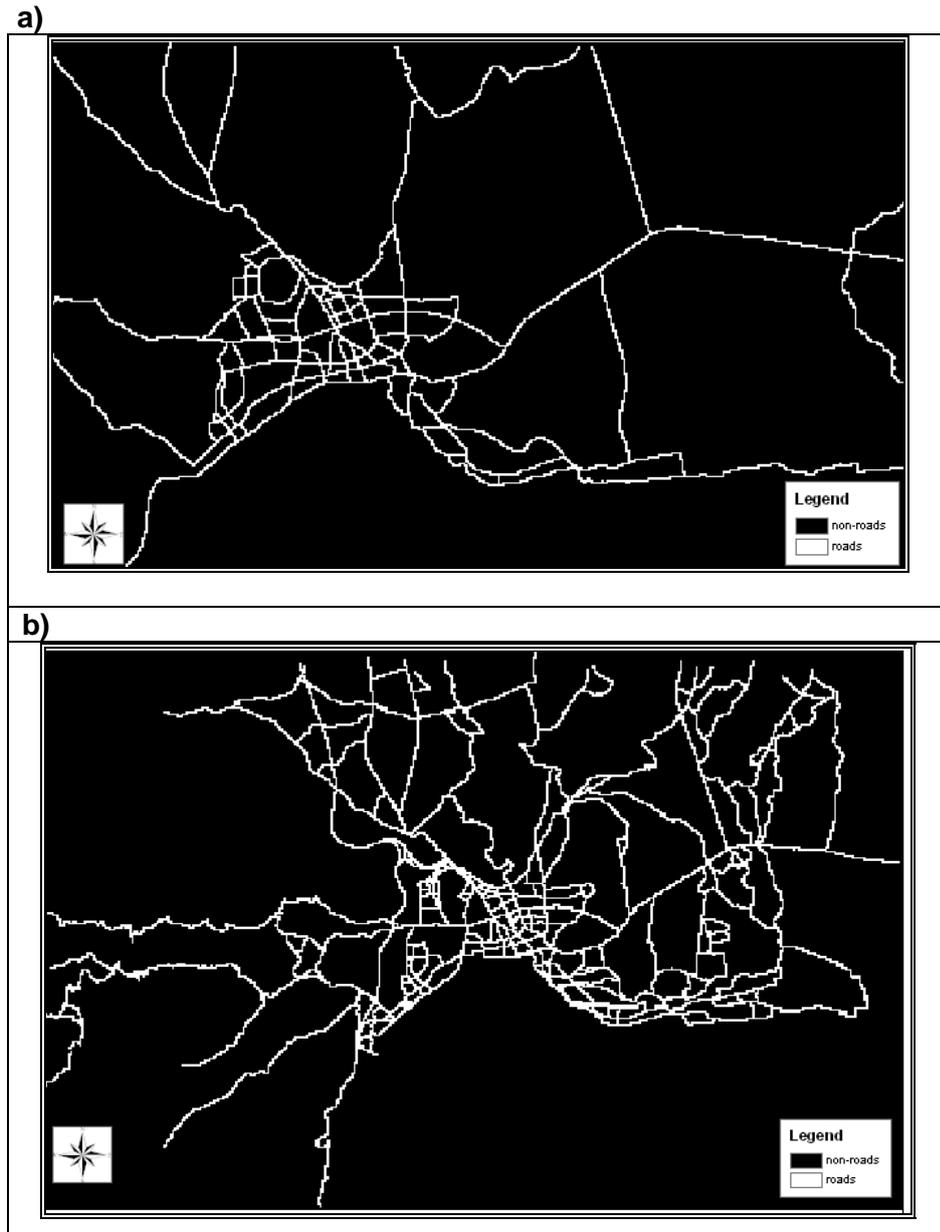
The road layer of the year 1995 is digitized in ArcGIS as 4 degrees from the Antalya Transportation Plan of the year 1995 and the road layer of the year 2003 is digitized in ArcGIS from the transportation plan of Antalya of the year 2003 prepared by DAMPO Planning (Figure 3.31).

### **3.5.2. Vector to Raster Conversion and the Recoding of the Vector Layers of the Roads 1995 and 2003**

The digitized road layers are converted to raster format in ArcGIS and then in Erdas. Then, these images are recoded in Erdas in order to be read by the model. The model requires the road layers as binary images. Therefore, the road pixels are recoded as one and the non-road pixels are recoded as zero (Figure 3.32).



**Figure 3.31.** The Digitized Road Layers of the Years 1995 (a) and 2003 (b)



**Figure 3.32.** The Recoded Road Layers of the Years 1995 (a) and 2003 (b)

### 3.6. Analyses of the Excluded Area Layer

The model requires also an excluded area layer. This layer guides the model not to generate the urban area on these pixels.

### 3.6.1 The Generation of the Excluded Area Layer

The vector layer of the excluded areas is produced from the Antalya 1/5000 Development Plan (DAMPO Planning, 2003) which is approved by the government. In respect of this plan, the unresidential areas as natural conservation areas, forest covered areas, military zones, first and second-degree productive agricultural areas, airport, riverbeds and Mediterranean Sea are digitized (Figure 3.10).

### 3.6.2. Vector to Raster Conversion and the Recoding of the Vector Layer of the Excluded Areas of the Year 2003

The model requires the excluded area layer also as raster format. Therefore, the digitized excluded area layer is converted to raster format. Then, all the excluded areas are recoded as one and the non-excluded areas are recoded as zero. The white pixels represent the excluded areas and the black pixels represent the non-excluded pixels (Figure 3.33).



**Figure 3.33.** The Recoded Excluded Area Layer of the Year 2003 (white is urban, black is non-urban pixels).

### **3.7. Slope and the Hillshade Layers**

The slope and hillshade layers are produced from the SRTM 90 meter DEM data. The model requires the slope layer in percent. Therefore, the slope layer is produced as percent value in ArcGIS. Then the pixel value range is arranged between 0 and 100 as the model required. The hillshade is computed from the 90 meter DEM and arranged as a grayscale image for using as a background image for visualization purposes.

### **3.8. The Resampling and the Subsetting of the Entire Layers**

The model requires the entire input layers in a standardized size. In this thesis for the calibration of model, three different resampling resolutions are created in Erdas 8.6. In the first one, the pixels are resampled with the nearest neighborhood method to 30 meter, in the second one the pixel size is resampled with the same method to 60 meter and in the last, the pixel size is resampled to 120 meter. Then, all the image layers are subsetting to three different sizes for the 3 resampling resolutions; 30, 60 and 120 meters. As a result, three different row and column sizes are created for the three calibration modes; coarse, fine and final, where the calibration of the model is explained in Chapter 4.

### **3.9. The Conversion of the Subsetting Layers to the 8 Bit Unsigned GIF Format**

The model requires all the layers in a standardized format which is 8 bit unsigned GIF. Therefore, all the raster images are converted to 8 bit unsigned format in Erdas and then they are converted to GIF format in TNT.

### 3.10. Methodology of the Study

In this study, a CA model integrated with GIS is used. This model requires all the input data, which are explained in the data preparation part, to be converted into raster format.

To run the model all the input data layers need to be standardized in terms of

- Projection
- Dimension
- Data format
- Resolution

In doing so, all raster input layers are converted to the same projection, UTM WGS 84, Zone 36. Finally, the resultant GIF files are named according to the convention stipulated by the model (Table 3.2).

After the data preparation in respect of the model requirements, the model can be run by these data. The model is operated by Linux version 9.0 on PC microcomputer. As it is mentioned in Chapter 2 the model runs in three modes; test, calibration and then prediction. The general outline of a SLEUTH model run is given in Figure 3.35. By test mode, the data is tested whether it is ready for the calibration then the prediction. The calibration process produces initializing coefficient values that best simulate historical growth for the study area. The purpose of the model calibration phase is to determine the best-fit values for the five growth control parameters including coefficients of diffusion, breed and spread, slope resistance and road gravity with historical urban extent data. It is the most important phase of the model. The five calibration coefficients are all integers and range from 1 to 100. By running SLEUTH in Calibration Mode, the different combinations of coefficients are used to model the historical urban growth. The goal of calibration is to determine which of the  $10^{10}$  (or  $100^5$ ) possible combinations of coefficients gives the best fit of a specific urban region. The term for all the possible coefficient combinations is called the coefficient space. It is unknown, but assumed that the coefficient space is a complex

surface, due to the unique properties of individual urban extents, the spatial scale and resolution chosen for the modeling, but most interestingly, because of the combinatory effects of the calibration coefficients with each other (Silva and Clarke, 2002).

The model is calibrated with the data, by successively narrowing the range of coefficient values. Coarse calibration, takes steps of 25 units through the entire coefficient space, for all coefficients. The second step, fine calibration, takes steps of 5 or 6 units through the coefficient space and the third, final calibration, takes steps of 1 unit through the coefficient space. Self-modification is also important during the calibration phase as it is given in Chapter 2. A growth cycle is the basic unit of SLEUTH execution. It begins by setting each of the coefficients to a unique value. Each of the growth rules is then applied to the raster data. Finally, the resulting growth rate is evaluated. It is applied if the growth rate exceeds or falls short of limit values. Self-modification slightly alters the coefficient values to simulate accelerated or depressed growth that is related with system-wide boom and bust conditions in urban development. A “boom” state occurs if the growth rate exceeds the “critical-high” value, the highest threshold for the urban growth rate, and indicates a period of accelerating growth. Each of the coefficients is increased to encourage the continuation of this trend. A “bust” state occurs when the growth rate is less than the critical-low value, which is the lowest threshold for the urban growth rate. In such an instance, the coefficients will be lowered in order to decrease the rate of growth throughout the system. In addition, a growth rate is:

$$\text{Growth rate} = (\text{number\_growth\_pixels} / \text{total\_number\_urban\_pixels}) * 100$$

(3.1)

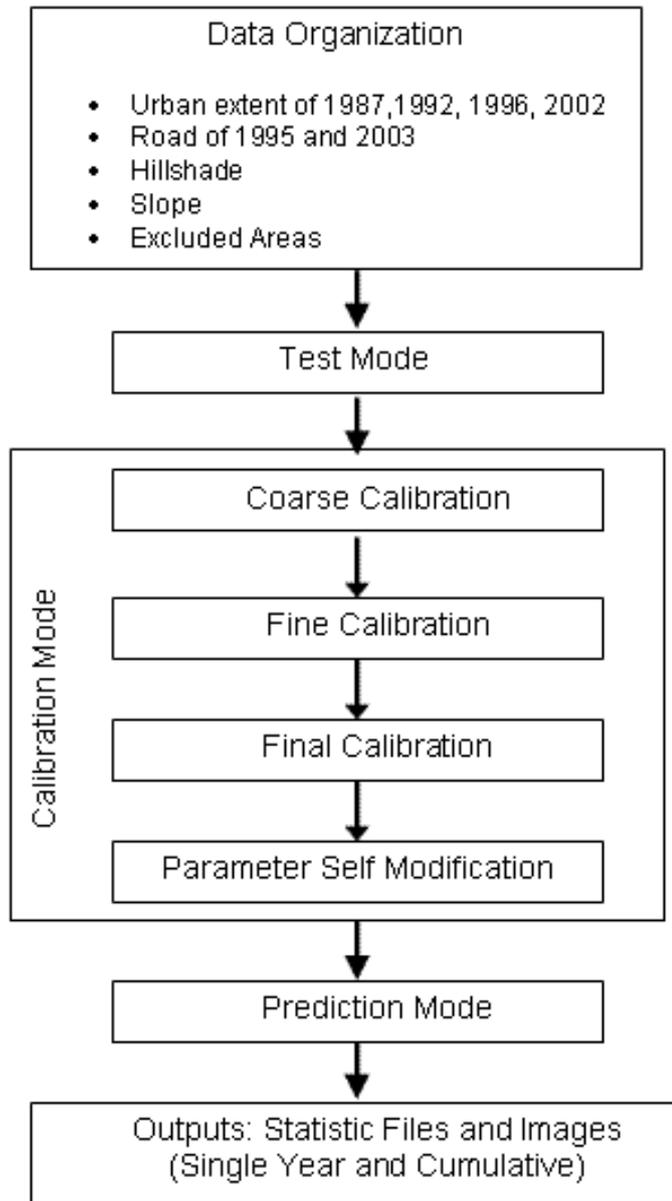
where `number_growth_pixels` is the number of newly urbanized pixels from the current growth cycle, `total_number_urban_pixels` is the amount of urban pixels from the current and previous growth cycle (Candau, 2002).

Results obtained from the calibration phases are sorted, and parameters of the highest scoring model runs are used to begin the next, more refined sequences or permutations over the parameter space. Initial exploration of the parameter

space uses a condensed, resampled and smaller version of the data sets, and as the calibration closes in on the “best” (final) run, the data are increased in spatial resolution. The highest scoring numeric results from each factor that control the behavior of the system from each phase of calibration feed the subsequent phase, with user-determined weights assigned to the different metrics (Goldstein, 2004).

When the calibration mode is completed, the results are used for forecasting studies in the prediction mode. As stated in Chapter 2, SLEUTH forecasts rely on replicating growth trends from the past. Once a coefficient set is found that can best describe how urban change has occurred over time, these values are used to forecast future growth.

The calibration process produces initializing coefficient values that best simulate historical growth for a region. However, due to SLEUTH’s self-modification qualities, coefficient values that initialize the model for a date in the past may be altered by the simulation end date. Therefore, for forecast run initialization, the coefficient values at the simulation end date are used to initialize a new simulation into a future date. Using the best coefficients derived from calibration to run a large number of Monte Carlo simulations will produce a single set of averaged coefficients for the simulation end date. Using the BSS (Best Solution Set) which are the ending coefficient values, parameters derived from the simulation end date, a forecast run may be initialized.



**Figure 3.34.** General outline of a SLEUTH model run examined in this study (Goldstein, 2004).

## CHAPTER 4

### URBAN GROWTH ANALYSES

In this chapter, Urban Growth Analyses for Antalya are examined. First, the structure of the model run and the required coefficients, then the modes of the model (test, calibration, and prediction) adapted to Antalya are explained. In addition, run of these modes and selection of the coefficients from each run, finally, obtaining the outputs of the prediction mode are given in this chapter.

#### 4.1. Model Run

SLEUTH trains the four growth rules to their parameters. The coefficients of the growth rules are Dispersion, Breed, Spread, Slope, and Road Gravity. They can be described together with the five coefficients;

##### 1) Spontaneous Growth Rule

Spontaneous Growth determines if a random pixel in the urban lattice will be urbanized. The Dispersion and the Slope coefficients are used in this rule.

##### 2) New Spreading Centers Rule

The New Spreading Centers Rule determines if a newly urbanized pixel (from the Spontaneous Growth Rule) will be a new urban centre and if so, will urbanize land. The Breed and Slope coefficients factor into this growth rule.

##### 3) Edge Growth Rule

The Edge Growth Rule adds new growth adjacent to existing urban pixels. The Spread and Slope coefficients determine the amount of Edge Growth.

##### 4) Road-Influenced Growth Rule

This final growth rule determines the extent the road (or transportation) network contributes to the urban growth of a city. New urbanized pixels “travel” on the road network and urbanize available pixels. This rule is determined by the Dispersion, Breed, Slope, and Road Gravity coefficients.

#### **4.1.1. Testing the Input Data**

In the test mode application, the entire data are tested if they are ready for the calibration or not. The names stipulated by the model for each layer are set into the test file in Linux and the number of Monte Carlo Iterations are entered as 2. It takes very short time as 5 minutes. If the test mode is finished successively, then the calibration mode is ready for the run. In other words, the data are ready for the calibration and there is not any problem occurred. For the Antalya input data there was not any problem.

#### **4.1.2. Calibration Analyses**

The purpose of the model calibration phase is to determine the best-fit values for the five growth control parameters mentioned in Chapter 3 and above, including coefficients of diffusion, breed and spread, slope resistance and road gravity with historical urban extent data. It is the most important phase of the model. The five calibration coefficients are all integers and range from 1 to 100.

Calibration process is accomplished in three phases as coarse, fine, and final. The first step, coarse calibration, takes steps of 25 units through the entire coefficient space (all the possible coefficient combinations), for all coefficients. The second step, fine calibration, takes steps of 4, 5 or 6 units through the coefficient space and the third, final calibration, takes steps of 1 through the coefficient space. Self-modification is also important during the calibration phase as it is explained in chapter two. A growth cycle is the basic unit of SLEUTH execution. It begins by setting each of the coefficients to a unique value. Each of the growth rules is then applied to the raster data. Finally, the resulting growth rate is evaluated. Self-modification is applied if the growth rate exceeds or falls

short of limit values. It slightly alters the coefficient values to simulate accelerated or depressed growth that is related with system-wide boom and bust conditions in urban development. A “boom” state occurs if the growth rate exceeds the “critical-high” value, the highest threshold for the urban growth rate, and indicates a period of accelerating growth. Each of the coefficients is increased to encourage the continuation of this trend. A “bust” state occurs when the growth rate is less than the critical-low value, which is the lowest threshold for the urban growth rate. In such an instance, the coefficients will be lowered in order to decrease the rate of growth throughout the system. In addition, a growth rate is:

$$\text{Growth rate} = (\text{number\_growth\_pixels} / \text{total\_number\_urban\_pixels}) * 100$$

where number\_growth\_pixels is the number of newly urbanized pixels from the current growth cycle, total\_number\_urban\_pixels is the amount of urban pixels from the current and previous growth cycle (Candau, 2002).

#### **4.1.2.1. Coarse Calibration for the Antalya Input Data**

For the coarse calibration, the pixel size of the data is resampled to 120 meter and renamed as the model stipulated.

For the Antalya study, the first phase of the calibration mode of the model, coarse calibration, is operated by a Linux version 9.0 on PC microcomputer with the properties of 1GB Operator, 512 Mhz Ram. The inputs are entered in the input scenario- calibrate file as stipulated by the model and as shown in Table 4.1.

The flags of the Coefficient file, average file, standard deviation file and the log file are set to YES in order to obtain their output files:

```
WRITE_COEFF_FILE (YES/NO)=YES  
WRITE_AVG_FILE (YES/NO)=YES  
WRITE_STD_DEV_FILE (YES/NO)=YES
```

The output of the coefficient file contains coefficient values for every run, Monte Carlo iteration and year. The output of the average file contains measured values of simulated data averaged over Monte Carlo iterations for every run

**Table 4.1.** The Naming Format of the Data Stipulated by the Model for the Coarse Calibration

antalya.urban.1987.gif
antalya.urban.1992.gif
antalya.urban.1996.gif
antalya.urban.2002.gif
antalya.road.1995.gif
antalya.road.2003.gif
antalya.slope.gif
antalya.excluded.gif
antalya.hillshade.gif

The output of the standard deviation file contains the standard deviation of averaged values. Then the Self Modification Constraints are determined and entered (Table 4.2).

The calibration start date is set to the year of 1987, which is the year of the first satellite datum, and the calibration stop date is set to year of 2003, which is the year of last datum of the area. Scenario file is created and renamed as scenario\_coarse. The inputs are entered in the input scenario\_coarse file as stipulated by the model. Input\_dir flag and output\_dir flag are edited. The flags of the coefficient file, average file, standard deviation file and the log file are set to YES in order to obtain their output files:

**Table 4.2.** Parameters of the Test Run and the Calibartion Runs

<b>Coefficients start-step-stop values</b>	<b>Test Run</b>	<b>Coarse Cal.</b>	<b>Fine Cal.</b>	<b>Final Cal.</b>	<b>Forecast Run</b>
CALIBRATION_DIFFUSION_START	25	0	0	1	1
CALIBRATION_DIFFUSION_STEP	1	25	5	1	1
CALIBRATION_DIFFUSION_STOP	25	100	20	5	1
CALIBRATION_BREED_START	10	0	0	10	20
CALIBRATION_BREED_STEP	1	25	5	2	1
CALIBRATION_BREED_STOP	10	100	20	20	20
CALIBRATION_SPREAD_START	30	0	20	27	32
CALIBRATION_SPREAD_STEP	1	25	5	1	1
CALIBRATION_SPREAD_STOP	30	100	40	32	32
CALIBRATION_SLOPE_START	70	0	75	75	78
CALIBRATION_SLOPE_STEP	1	25	5	3	1
CALIBRATION_SLOPE_STOP	70	100	100	90	78
CALIBRATION_ROAD_START	100	0	25	23	24
CALIBRATION_ROAD_STEP	1	25	5	1	1
CALIBRATION_ROAD_STOP	100	100	50	28	24
<b>Self Modification Constraints</b>					
Critical High	1.500	1.500	1.500	1.500	1.500
Critical Low	0.050	0.050	0.050	0.050	0.050
Boom	1.010	1.010	1.010	1.010	1.010
Bust	0.090	0.090	0.090	0.090	0.090
Critical Slope	21	21	21	21	21
The number of Monte Carlo com	2	100	100	100	100
The Calibration Start Date	1987				
The Calibration Stop Date	2003				

Following, the coefficient settings are arranged as shown in Table 4.2. The coarse calibration execution was run approximately for 1,5 days. During this time the PC was connected with a UPS to take care of electricity interruption. When the computer is stopped accidentally, the execution stops and the run has to start from the beginning. In the coarse phase of calibration, the entire range (0 - 100) of

the five coefficients is explored using large increments (for each coefficient, value = {0, 25, 50, 75, 100}, Table 4.2). The number of Monte Carlo iterations is set to 100 for obtaining a more sensitive result. As the result of the coarse calibration, 7 different files are produced by the model which are avg. file, coeff\_file, control\_stats file, log file, memory file, restart file and the std\_dev file. The averaged coefficient values are derived from the control\_stats (Table 4.3) in order to run the fine and final calibrations and then the prediction mode. SLEUTH generates best-fit statistics for eleven metrics.

These metrics are; product, compare, pop, edges, clusters, cluster size, leesallee, slope, %urban, Xmean, Ymean, and Rad. The description of these metrics are given in Appendix A. SLEUTH generates these metrics for each control year. The simulated data is then compared to the metrics of the historical data and linear regression values are calculated. These best-fit values are written to the control\_stats.log output file. The control\_stats.log file is the main file that is used to derive the coefficient ranges for the next calibration phase.

#### **i. Selecting Coefficient Ranges from Coarse Calibration**

In order to select the coefficient ranges as the start and stop value and the step number, the Leesallee metric column of the control\_stats file is sorted in descending order (Table 4.3). Following, the top three values are picked from the Leesallee metric column. The repeating scores downwards are also picked and evaluated (URL 3). The diffusion, breed, spread, slope and road gravity values corresponds to these Leesallee top values are candidate parameters for the next calibration run (Table. 4.4).

**Table 4.3.** Control\_Stats file from the Coarse (120m) Calibration Run

Compare	Pop	Edges	Clusters	Size	Leesalee	Slope	%Urban	Xmean	Ymean	Rad	Diff	Brd	Sprd	Slp	RG
0.89850	0.99507	0.85311	0.33115	0.68457	0.46756	0.80597	0.99507	0.65836	0.56636	0.99217	1	1	25	100	75
0.89850	0.99507	0.85311	0.33115	0.68457	0.46756	0.80597	0.99507	0.65836	0.56636	0.99217	1	1	25	100	100
0.89902	0.99506	0.85387	0.30691	0.69062	0.46751	0.84216	0.99506	0.66204	0.52187	0.99216	1	1	25	100	50
0.90699	0.99486	0.86592	0.42314	0.67248	0.46740	0.58429	0.99486	0.66783	0.48006	0.99197	1	1	25	75	1
0.90699	0.99486	0.86592	0.42314	0.67248	0.46740	0.58429	0.99486	0.66783	0.48006	0.99197	1	1	25	75	25
0.89762	0.99479	0.84934	0.32239	0.68440	0.46718	0.68654	0.99479	0.65530	0.61839	0.99179	1	1	25	100	1
0.89762	0.99479	0.84934	0.32239	0.68440	0.46718	0.68654	0.99479	0.65530	0.61839	0.99179	1	1	25	100	25
0.90933	0.99475	0.86656	0.33480	0.68685	0.46678	0.10832	0.99475	0.67544	0.55955	0.99184	1	1	25	75	50
0.90933	0.99475	0.86652	0.33668	0.68605	0.46678	0.11325	0.99475	0.67535	0.56685	0.99184	1	1	25	75	75
0.90933	0.99475	0.86652	0.33668	0.68605	0.46678	0.11325	0.99475	0.67535	0.56685	0.99184	1	1	25	75	100
0.91957	0.99471	0.86314	0.34378	0.69121	0.46667	0.23099	0.99471	0.65847	0.37703	0.99187	1	1	25	50	1
0.91957	0.99471	0.86314	0.34378	0.69121	0.46667	0.23099	0.99471	0.65847	0.37703	0.99187	1	1	25	50	25
0.99429	0.99575	0.87695	0.29328	0.70237	0.46657	0.64893	0.99575	0.69073	0.06176	0.99367	1	25	25	100	1
0.99429	0.99575	0.87695	0.29328	0.70237	0.46657	0.64893	0.99575	0.69073	0.06176	0.99367	1	25	25	100	25
0.93086	0.99527	0.86700	0.32275	0.68930	0.46648	0.49025	0.99527	0.66499	0.63159	0.99261	1	1	25	1	1

**Table 4.4.** Resultant Parameters of the Coarse Calibration

LeeSallee	Diffussion	Breed	Spread	Slope	RoadGravity
0,46756	1	1	25	100	100
0,46756	1	1	25	100	75
0,46751	1	1	25	100	50
0,4674	1	1	25	75	25
0,4674	1	1	25	75	1
<b>New Parameters</b>					
<b>For the next calibration</b>	{0-20, 5}	{0-20, 5}	{20-40, 5}	{75-100, 5}	{25-50,5}

Then the high and low values of each of the coefficients are taken. For each coefficient in the scenario file in order to be used for the fine calibration, low values are set to `_START`, high values are set to `_STOP`. The `STEP` value is derived from the difference of the Stop and Start values. The difference is divided by 4, 5 or 6 and then the step value is acquired (Table 4.5). A step value is the increment between the start and stop values 4-6 times. If only one coefficient value is appeared as diffusion (1), breed (1) and spread (25), then the difference between the start and stop values should be 20. For example in the spread column, all the values are 25. Therefore, the start value selected as 20 and the stop value as 40. If all the values in this column were 50, then the start value should be 40 and the stop value 60.

**Table 4.5.** Start, Stop and Step Values Derived from the Coarse Calibration

	Diffussion	Breed	Spread	Slope	RoadGravity
<b>Start Value</b>	0	0	20	75	25
<b>Stop Value</b>	20	20	40	100	50
<b>Step Value</b>	5	5	5	5	5

These values are put in the scenario file in the next (fine) calibration.

#### **4.1.2.2. Fine Calibration for the Antalya Input Data**

For the fine calibration, the pixel size of the data is resampled to 60 meter resolution and renamed as the model stipulated (Table 4.1).

A copy of scenario file used in the previous coarse calibration step is created and renamed as scenario\_fine. The inputs are entered in the input scenario\_fine file as stipulated by the model. Input\_dir flag and output\_dir flag are edited. Then, the gif image files with the file name format as described in Table 4.1 are written. The flags of the coefficient file, average file, standard deviation file and the log file are set to YES in order to obtain their output files.

Following, the coefficient settings are arranged as shown in Table 4.2. These values define a narrowed coefficient range derived from the coarse phase of the calibration. Then the Self Modification Constraints are determined and entered in Linux as done in the previous step. The number of Monte Carlo iterations is set to 100 for obtaining a more sensitive result. The fine calibration start date 1987 and stop date 2003 did not change and entered the same as in the previous step. Because the data set did not change.

The fine calibration execution was run approximately for 3 days. As the result of the fine calibration, 7 different files are produced by the model which are avg. file, coeff\_file, control\_stats file, log file, memory file, restart file and the std\_dev file. The averaged coefficient values are derived from the control\_stats in order to run the final calibration and then the prediction mode. The eleven metrics explained in the previous phase are created again.

#### **i. Selecting Coefficient Ranges from Fine Calibration**

Using the best-fit values found in the control\_stats.log file (Table 4.5) produced in the coarse calibration phase, the range of possible coefficient values are narrowed. In order to select the coefficient ranges as the start and stop value and the step number, the Leesallee metric column of the control\_stats file is sorted in descending order (Table 4.6).

**Table 4.6.** Control\_Stats file from the Fine (60m) Calibration Run

Run	Product	Compare	Cluster Pop	Edges	Clusters	Size	LeesaLee	Slope	%Urban	Xmean	Ymean	Rad	Diff	Brd	Sprd	Slp	RG
624	0.00005	0.84258	0.99395	0.75459	0.92202	0.94626	0.48995	0.57953	0.99395	0.68337	0.00044	0.99015	1	15	30	85	25
444	0.00023	0.83993	0.99403	0.75196	0.75704	0.94731	0.48992	0.54025	0.99403	0.68186	0.00287	0.99023	1	10	30	85	25
798	0.00118	0.84809	0.99401	0.75619	0.86439	0.94245	0.48992	0.57045	0.99401	0.68845	0.01199	0.99029	1	20	30	80	25
630	0.00074	0.84251	0.99388	0.75266	0.90277	0.94664	0.48987	0.58165	0.99388	0.67971	0.00723	0.99006	1	15	30	90	25
792	0.00050	0.84911	0.99414	0.76524	0.64170	0.94314	0.48987	0.63451	0.99414	0.68115	0.00614	0.99046	1	20	30	75	25
824	0.00080	0.84377	0.99401	0.76099	0.88022	0.95325	0.48986	0.52947	0.99401	0.68219	0.00859	0.99024	1	20	30	100	35
816	0.00008	0.84419	0.99396	0.75122	0.80475	0.94665	0.48981	0.54335	0.99396	0.68211	0.00094	0.99017	1	20	30	95	25
817	0.00143	0.84426	0.99392	0.75408	0.89107	0.95145	0.48980	0.55800	0.99392	0.68304	0.01451	0.99013	1	20	30	95	30
445	0.00012	0.84024	0.99386	0.75214	0.81560	0.94097	0.48978	0.54981	0.99386	0.68075	0.00142	0.99002	1	10	30	85	30
612	0.00028	0.84628	0.99398	0.75247	0.82807	0.94704	0.48978	0.58304	0.99398	0.68649	0.00288	0.99024	1	15	30	75	25
804	0.00056	0.84674	0.99404	0.75764	0.72575	0.95376	0.48978	0.56423	0.99404	0.68345	0.00685	0.99032	1	20	30	85	25
810	0.00048	0.84490	0.99398	0.75461	0.77742	0.94702	0.48978	0.57235	0.99398	0.68442	0.00544	0.99021	1	20	30	90	25
438	0.00001	0.84115	0.99380	0.75609	0.75015	0.95102	0.48976	0.56941	0.99380	0.67540	0.00011	0.98995	1	10	30	80	25

Following, the top three values are picked from the Leesallee metric column. if these scores are repeating downwards, they should also be picked and evaluated (URL 3). The diffusion, breed, spread, slope and road gravity values corresponding to these Leesallee top values are candidate parameters for the next calibration run (Table. 4.7).

**Table 4.7.** Resultant Parameters of the Fine Calibration

LeeSallee	Diffussio	Bread	Spread	Slope	RoadGravity
0,48995	1	15	30	85	100
0,48992	1	20	30	80	75
0,48992	1	10	30	85	50
0,48987	1	20	30	75	25
0,48987	1	15	30	90	1
<b>New Parameters</b>					
<b>For the next calibration</b>	{1-5, 1}	{10-20, 2}	{27-32, 1}	{75-90, 3}	{23-28,1}

Then the high and low values of each of the coefficients are taken. For each coefficient in the scenario file to be used for the final calibration low values are set to \_START, high values are set to \_STOP. The STEP value is derived from the difference of the stop and start values. The difference is divided by 4, 5 or 6 and then the step value is acquired (Table 4.8). A step value is the increment between the start and stop values about 4-6 times. If only one coefficient value is appeared as in diffusion (1), and in spread (30), then the average value of the start and stop values should be nearest to the value in the list. For example in the spread column, all the values are 30. Therefore, the start value selected as 27 and the stop value is 32 (27 + 5).  $27 + 32 = 59$ . 5 should be added to 27 because the increment in fine calibration list is 5. The average of these two values is  $29.5 \approx 30$ . Therefore, these values are selected to use them in the next phase.

**Table 4.8.** Start, Stop, and Step Values Derived from the Fine Calibration

	<b>Diffussion</b>	<b>Bread</b>	<b>Spread</b>	<b>Slope</b>	<b>RoadGravity</b>
<b>Start Value</b>	1	10	27	75	23
<b>Stop Value</b>	5	20	32	90	28
<b>Step Value</b>	1	2	1	3	1

These values are put in the scenario file to use them in the next (final) calibration.

#### **4.1.2.3. Final Calibration for the Antalya Input Data**

For the final calibration, the pixel size of the data is resampled to 30 meter resolution and renamed as the model stipulated (Table 4.1).

A copy of scenario file used in the previous fine calibration step is created and renamed as scenario\_final. The inputs are entered in the input scenario\_final file as stipulated by the model Input\_dir flag and output\_dir flag are edited. Then, the gif image files with the file name format are written. The flags of the coefficient file, average file, standard deviation file and the log file are set to YES in order to obtain their output files:

Following, the coefficient settings are arranged (Table 4.2). These values define a narrowed coefficient range derived from the fine phase of the calibration. Then the Self Modification Constraints are determined and entered in Linux as done in the previous step (Table 4.2). The number of monte carlo iterations is set to 100 for obtaining a more sensitive result. The fine calibration start date 1987 and stop date 2003 did not change and entered same as in the previous step. Because the data set did not change.

The final calibration execution was run approximately for 8 days. As result of the final calibration, 7 different files are produced by the model which are avg. file, coeff\_file, control\_stats file, log file, memory file, restart file and the std\_dev file. The averaged coefficient values are derived from the control\_stats in order to run the final calibration and then the prediction mode. The eleven metrics explained in the previous phase are created again.

### i. Selecting Coefficient Ranges from Final Calibration

Using the best-fit values found in the control\_stats.log file (Table 4.8) produced in the fine calibration phase, the range of possible coefficient values are narrowed (Table 4.9).

In order to select the coefficient ranges for the next step as the start and stop value and the step number, the Leesallee metric column of the control\_stats file is sorted in descending order (Table 4.10).

Following, the top three values are picked from the Leesallee metric column. if these scores are repeating downwards, they should also be picked and evaluated (URL 3). This time the diffusion, breed, spread, slope, and road gravity values, which correspond to the top Leesallee value are the parameters for the next run called forecast run (Table. 4.11).

**Table 4.9.** Resultant Parameters of the Final Calibration

LeeSallee	Diffussion	Bread	Spread	Slope	RoadGravity
0,50298	1	20	32	78	24
0,50294	1	20	32	78	26
0,50289	1	18	32	78	27
0,50289	1	12	32	84	24
<b>New Parameters For the next step</b>					
	{1-1, 1}	{20-20, 1}	{32-32, 1}	{78-78, 1}	{24-24,1}

For this step, for each coefficient, the step values are set to 1 and the stop and the start values are set to the same value (Table 4.11).

**Table 4.10.** Control\_Stats file from the Final (30m) Calibration Run

Run	Product	Compare	Pop	Edges	Clusters	ClusterSize	Leesalee	Slope	%Urban	Xmean	Ymean	Rad	Diff	Brd	Spnd	Slp	RG
1267	0.00000	0.68974	0.99428	0.58971	0.55147	0.99980	0.50298	0.42480	0.99428	0.73893	0.00005	0.98904	1	20	32	78	24
1269	0.00000	0.68977	0.99417	0.58278	0.54913	100.000	0.50294	0.39693	0.99417	0.73647	0.00013	0.98888	1	20	32	78	26
415	0.00005	0.68872	0.99445	0.58873	0.56219	0.99998	0.50289	0.40986	0.99445	0.73884	0.00134	0.98923	1	12	32	84	24
1054	0.00001	0.68956	0.99424	0.58995	0.57982	0.99598	0.50289	0.40369	0.99424	0.73438	0.00016	0.98899	1	18	32	78	27
1066	0.00000	0.68847	0.99428	0.58347	0.57704	0.99781	0.50288	0.40212	0.99428	0.73827	0.00006	0.98903	1	18	32	84	27
1275	0.00001	0.68920	0.99423	0.58551	0.52957	0.99938	0.50288	0.42644	0.99423	0.73782	0.00018	0.98897	1	20	32	81	26
1282	0.00003	0.68845	0.99423	0.58912	0.51566	0.99973	0.50288	0.40345	0.99423	0.73916	0.00104	0.98895	1	20	32	84	27
833	0.00000	0.68964	0.99433	0.58472	0.54587	0.99984	0.50287	0.42748	0.99433	0.73855	0.00009	0.98910	1	16	32	75	28
1274	0.00002	0.68942	0.99435	0.58958	0.54979	0.99996	0.50287	0.40814	0.99435	0.73452	0.00052	0.98914	1	20	32	81	25
629	0.00000	0.68779	0.99426	0.59181	0.58068	0.99971	0.50286	0.40996	0.99426	0.73878	0.00000	0.98899	1	14	32	81	28
619	0.00002	0.68830	0.99428	0.58828	0.62866	0.99980	0.50285	0.41723	0.99428	0.73116	0.00047	0.98902	1	14	32	78	24
622	0.00001	0.68845	0.99441	0.58782	0.50861	0.99943	0.50285	0.44357	0.99441	0.73857	0.00039	0.98921	1	14	32	78	27
623	0.00001	0.68850	0.99442	0.59184	0.55304	0.99994	0.50285	0.41625	0.99442	0.73848	0.00022	0.98922	1	14	32	78	28

**Table 4.11.** Start, Stop, and Step Values Derived from the Fine Calibration

	<b>Diffussion</b>	<b>Bread</b>	<b>Spread</b>	<b>Slope</b>	<b>RoadGravity</b>
<b>Start Value</b>	1	20	32	78	24
<b>Stop Value</b>	1	20	32	78	24
<b>Step Value</b>	1	1	1	1	1

These values are put in the scenario file to use them in the forecasting step.

#### **4.1.2.4. Forecast Coefficients Run for the Antalya Input Data**

As it is explained in the previous chapter, the aim of the calibration process is to produce initializing coefficient values that best simulate historical growth for a region. For forecast initialization, the stop date values from the best calibrated coefficients are desired. Using the best coefficients derived from calibration and running SLEUTH for the historical time produce a single set of Stop date coefficients to initialize forecasting. In order to run forecasting phase the scenario file used in the previous step is created. Input gif images with the full resolution are pointed to the Input\_dir flag. Output\_dir flag is pointed as desired and the input format names stipulated by the model are not changed. Monte Carlo Iteration flag is set to 100. The flags of the coefficient file, average file, standard deviation file and the log file are set to YES in order to obtain their output files. Particularly, the Avg\_file flag should set to yes in order to use it for the prediction mode. The file used to store coefficient values is the avg.log file. Following, the coefficient settings are arranged (Table 4.2). These values define a single set of best coefficient values derived from the final phase of the calibration. Then the Self Modification Constraints are determined and entered in Linux as done in the previous step.

##### **i. Selecting Coefficient Ranges from Forecasting Phase**

In order to select the coefficient values for the prediction mode there is not any start, stop and the step value. In addition, the Leesallee metric column is not used. This time the values of the coefficients of the last row of the avg.log file (Table 4.12) which is the stop date year is used. These floating values are rounded to integers in order to be read by the model (Table 4.13).

**Table 4.12.** Average .Log file from the Forecast Run

year	sng	og	rt	pop	area	edges	clusters	xmean	ymean	rad	slope	cl_size	diffus	spread	breed	slp_res	rd_grav	%urban
1992	5.17	2078.22	18.89	67501.97	67501.97	16663.96	1138.68	421.89	382.51	146.58	0.74	58.84	1.04	33.30	20.81	71.40	24.66	17.38
1996	5.72	2388.75	16.26	76683.84	76683.84	18686.76	1327.10	427.96	382.23	156.23	0.73	57.25	1.08	34.65	21.66	64.21	25.38	18.99
2002	5.78	2801.02	13.70	92738.64	92738.64	20731.95	1362.68	437.09	381.57	171.81	0.73	67.55	1.15	36.78	22.99	52.13	26.59	21.80

**Table 4.13.** Coefficient Values Derived from the Forecast Run.

Coefficient Type	Prediction Best Fit Values
dispersion	1
spread	37
breed	23
slope	52
road gravity	27

These values are put in the prediction scenario file to reach the growth prediction.

#### 4.1.3. Urban Growth Prediction Run

In order to run the prediction mode of the model, the scenario file used to derive forecasting coefficients, is modified as scenario\_predict file and contained in the scenario file. The input\_dir flag is set to point the full resolution images. The output\_dir flag is set to point a desired output directory. Then all the output file flags are set to yes in order to create desired statistic files, including average .log file. Monte Carlo Iterations flag is set to 100. It may be greater than 100. Then the coefficients derived from the forecast calibration phase are entered. For the start date, the last urban data layer year of 2002 and for the stop date, the year of 2025 is given to the scenario prediction file.

PREDICTION START DATE = 2002

PREDICTION STOP DATE = 2025

The prediction best fit values derived from the forecast run:

PREDICTION\_DIFFUSION\_BEST FIT = 1

PREDICTION\_SPREAD\_BEST FIT = 37

PREDICTION\_BREED\_BEST FIT = 23

PREDICTION\_SLOPE\_BEST FIT = 52

PREDICTION\_ROAD\_BEST FIT = 27

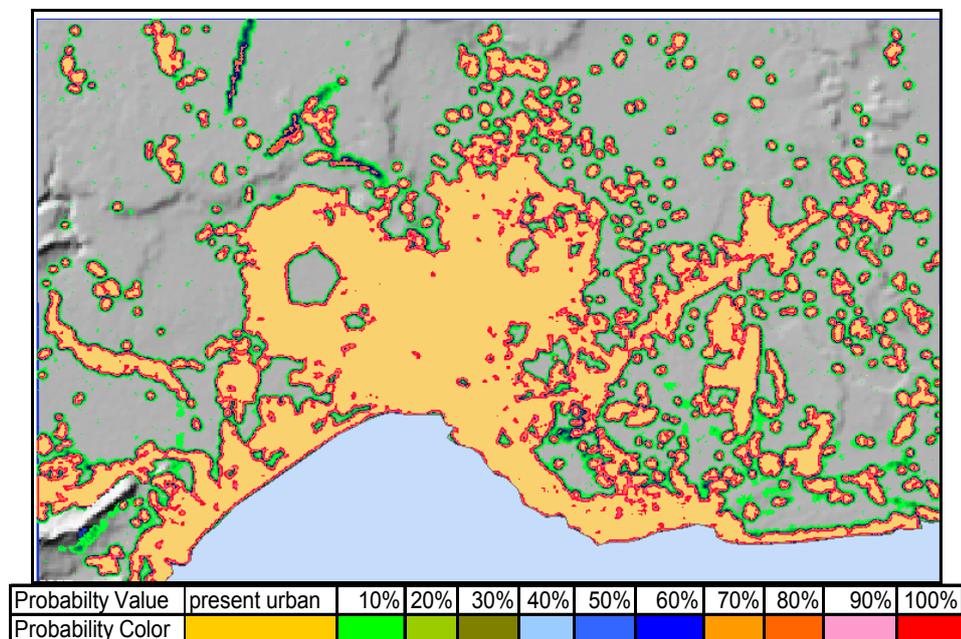
The ECHO flag is also set to yes to print the growth years to the screen as the model executes.

As a result, an average, coefficient, memory, and log files are created as the statistical files. In addition, the urban growths of the years from 2003 to 2025 are created in the prediction output file.

## CHAPTER 5

### RESULTS AND DISCUSSIONS

One planning scenario for future urban development in Antalya is considered in this study. It is predicting the 2025 urban growth of Antalya by protecting all the forest, agricultural and natural conservation areas. In addition, the new urban pixels are not let to grow into the military zones, riverbeds, and water by using the excluded area layer. As the result of this entire urban growth prediction study, the urban area prediction results for each year from 2003 to 2025 are obtained at the end of the prediction run. Each year's growth probability pixels are colored. The higher the probability value, the more likely urbanization is (Figure 5.1). In addition, an average file is produced as an output. Some of the most important statistical measures for this run are given in Table 5.1.



**Figure 5.1.** Urban Growth Prediction of the year 2025 with the Probability Colors

**Tablo 5.1:** Comparison of the Selected Statistical Measures of the Years 2002 and 2025.

Statistical Measures	2002	2025
sng	5,78	5,26
og	2801	3621,26
rt	13,07	28,48
pop	92738,64	201894,3
area	92738,64	201894,3
edges	20731,95	32667,84
clusters	1362,68	1895,96
rad	171,81	253,51
slope	0,73	0,69
diffusion	1,15	1,24
spread	36,78	28,63
breed	22,99	46,05
road_gravity	26,59	37,74
percent_urban	21,8	56,05
growth_rate	3,04	1,81
growth_pixels	2822,23	3659,33

The definitions of the abbreviations in the Table 5.1 are explained below (URL 3)

**sng:** cumulative number of urbanized pixels by spontaneous neighborhood growth (spontaneous growth models the development of urban settlement in undeveloped areas).

**og:** cumulative number of urbanized pixels by organic growth

**rt:** cumulative number of urbanized pixels by road Influenced growth

**pop:** total number of urban pixels

**area:** total number of urban pixels

**edges:** number of urban to nonurban pixel edges

**clusters:** number of urban pixel clusters

**rad:** the radius of the circle which encloses the urban area

**slope:** slope coefficient

**diffusion:** diffusion coefficient

**spread:**spread coefficient

**breed:** breed coefficient

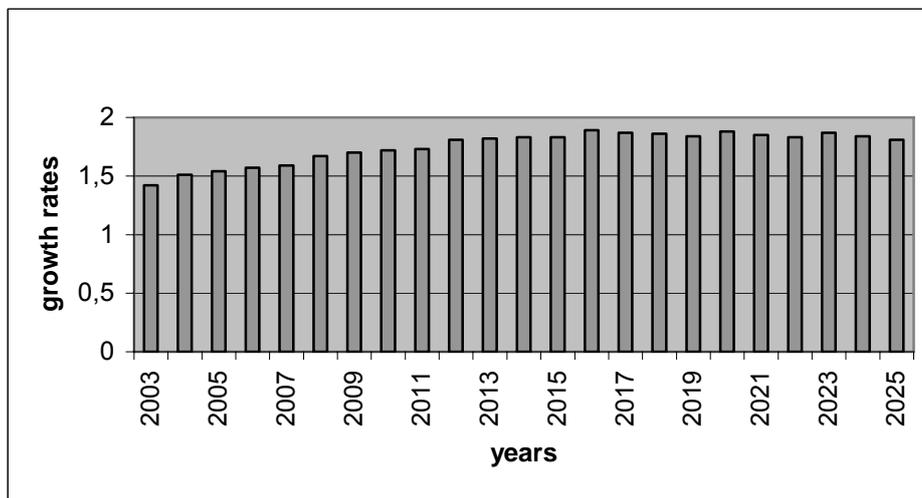
**road gravity:** road gravity value

**percent\_urban:** percent of urbanized pixels divided by the number of pixels available for urbanization

**growth\_rate:**(number\_growth\_pixels/total\_number\_urban\_pixels)\* 100)

**growth\_pixels:** number of growth pixels for each year.

In Table 5.2, the last column indicates the growing urban hectares for each year and calculated by multiplying the growth pixels by one pixel dimension 900m<sup>2</sup>. The decrease of the growing rate through the year 2025 can be seen from the table. The reason of this decrease is the maximum protection of the forest areas, agricultural areas, and natural conservation areas in this study. The growing scenario for Antalya embody an anti-growth strategy, which requires slowing down the growth rate (Figure 5.2) and altering the spatial pattern of growth while maintaining the liveability of the city.



**Figure 5.2.** Urban Growth Rates through the Year 2025

**Table 5.2.** The Output Average File with the Statistical Measures from the Year 2003 through the Year 2025

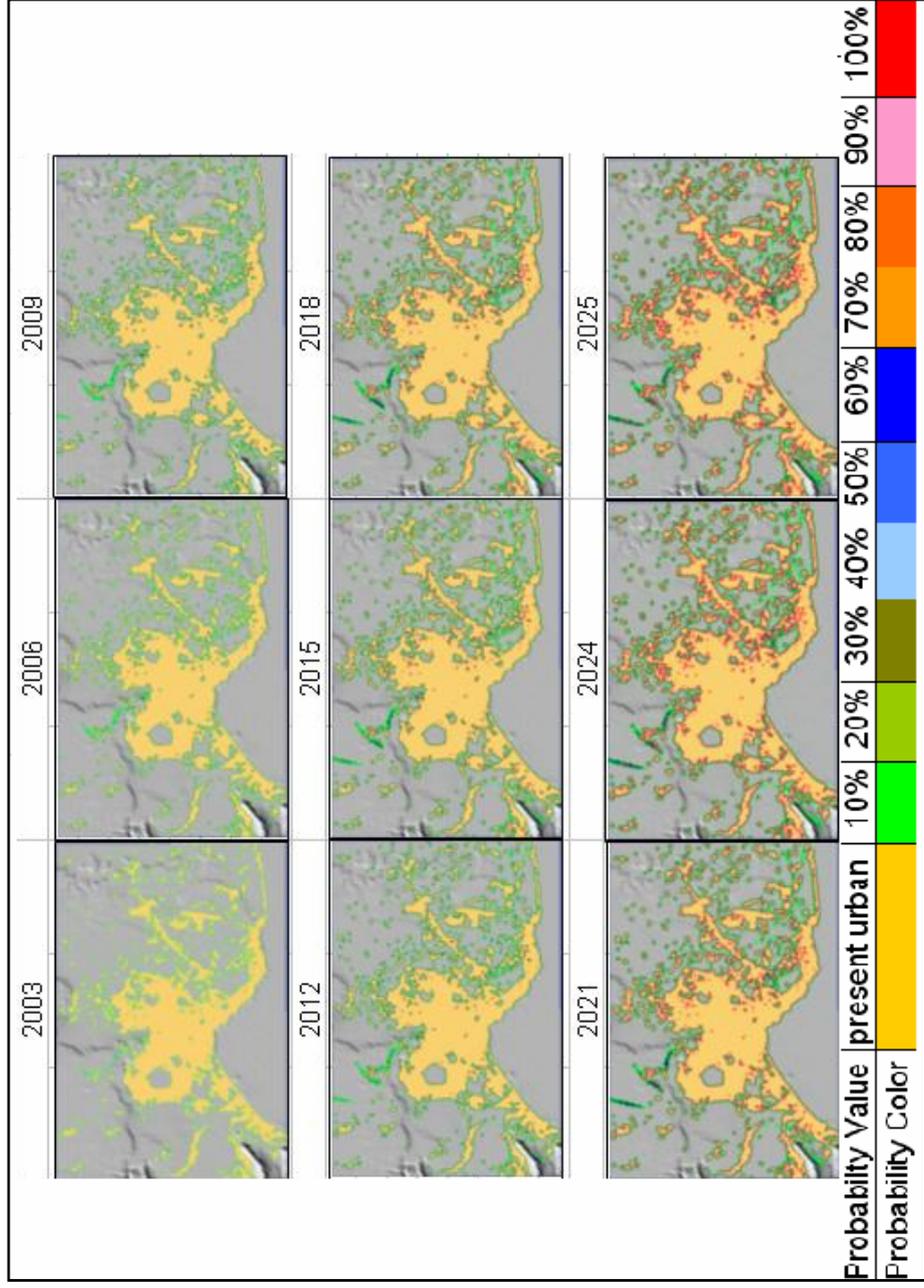
year	sng	og	rt	pop	area	edges	clusters	rad	slope	diffus	spread	breed	road	gr	%urban	grw	pix	grw	ra	grw	ha
2003	4.39	1885.30	49.65	136391.1	136391.1	21470.30	777.28	208.36	0.69	1	23.00	37.00	27.00	27.00	43.36	1942.11	1.42	1.42	175		
2004	4.22	2036.92	49.04	138484.2	138484.2	22258.62	993.93	209.95	0.69	1.01	23.23	37.37	27.43	27.43	43.76	2093.13	1.51	1.51	188		
2005	4.29	2109.46	44.46	140645.6	140645.6	23038.56	1180.47	211.59	0.69	1.02	23.46	37.74	27.87	27.87	44.18	2161.43	1.54	1.54	195		
2006	4.09	2184.97	44.88	142882.6	142882.6	23804.15	1344.43	213.26	0.69	1.03	23.70	38.12	28.31	28.31	44.61	2236.96	1.57	1.57	201		
2007	4.32	2255.87	45.01	145191.2	145191.2	24561.44	1481.93	214.98	0.69	1.04	23.93	38.50	28.76	28.76	45.06	2308.57	1.59	1.59	208		
2008	4.30	2415.91	44.59	147659.3	147659.3	25309.74	1599.09	216.80	0.69	1.05	24.17	38.89	29.21	29.21	45.54	2468.19	1.67	1.67	222		
2009	4.96	2493.89	46.96	150208.9	150208.9	26051.46	1697.27	218.66	0.69	1.06	24.41	39.28	29.67	29.67	46.03	2549.57	1.7	1.7	230		
2010	4.96	2573.85	44.00	152835.0	152835.0	26749.93	1769.37	220.56	0.68	1.07	24.66	39.67	30.13	30.13	46.54	2626.10	1.72	1.72	236		
2011	5.10	2638.53	45.19	155527.6	155527.6	27415.29	1832.18	222.50	0.68	1.08	24.91	40.07	30.59	30.59	47.06	2692.61	1.73	1.73	242		
2012	5.10	2819.07	46.30	158402.1	158402.1	28075.27	1880.39	224.55	0.68	1.09	25.15	40.47	31.06	31.06	47.62	2874.52	1.81	1.81	259		
2013	5.08	2889.59	42.80	161343.8	161343.8	28681.89	1915.20	226.62	0.68	1.10	25.41	40.87	31.54	31.54	48.19	2941.65	1.82	1.82	265		
2014	5.05	2954.31	42.04	164349.1	164349.1	29248.24	1941.92	228.72	0.68	1.12	25.66	41.28	32.02	32.02	48.77	3005.30	1.83	1.83	271		
2015	5.00	3014.18	38.34	167410.4	167410.4	29765.10	1961.04	230.84	0.69	1.13	25.92	41.69	32.51	32.51	49.37	3061.29	1.83	1.83	276		
2016	4.91	3169.43	42.24	170630.8	170630.8	30251.50	1974.18	233.05	0.69	1.14	26.18	42.11	33.00	33.00	49.99	3220.42	1.89	1.89	290		
2017	4.79	3206.60	38.51	173884.5	173884.5	30689.01	1974.48	235.26	0.69	1.15	26.44	42.53	33.50	33.50	50.62	3253.73	1.87	1.87	293		
2018	4.74	3244.94	36.67	177175.2	177175.2	31085.06	1972.85	237.48	0.69	1.16	26.70	42.96	34.01	34.01	51.26	3290.68	1.86	1.86	296		
2019	4.98	3277.14	35.09	180497.0	180497.0	31430.68	1967.41	239.70	0.69	1.17	26.97	43.39	34.52	34.52	51.90	3321.76	1.84	1.84	299		
2020	4.64	3421.73	35.71	183962.6	183962.6	31729.27	1959.84	241.99	0.69	1.18	27.24	43.82	35.04	35.04	52.58	3465.59	1.88	1.88	312		
2021	4.96	3433.56	31.38	187436.5	187436.5	31989.55	1946.02	244.26	0.69	1.20	27.51	44.26	35.56	35.56	53.25	3473.93	1.85	1.85	313		
2022	4.66	3463.90	30.71	190939.6	190939.6	32209.58	1937.11	246.53	0.69	1.21	27.79	44.70	36.10	36.10	53.93	3503.09	1.83	1.83	315		
2023	4.56	3602.17	32.41	194582.8	194582.8	32390.65	1924.98	248.87	0.69	1.22	28.06	45.15	36.64	36.64	54.63	3643.21	1.87	1.87	328		
2024	5.26	3612.19	29.77	198235.0	198235.0	32547.68	1909.71	251.20	0.69	1.23	28.35	45.60	37.18	37.18	55.34	3652.16	1.84	1.84	328		
2025	5.26	3621.26	28.48	201894.3	201894.3	32667.84	1895.96	253.51	0.69	1.24	28.63	46.05	37.74	37.74	56.05	3659.33	1.81	1.81	329		

It can be analyzed that the maximum growth rate is 1.89 by the year 2016. After this year, the decrease in growing rate can be seen clearly because of the environmental protection.

If we examine the urban extent of the years 2003 and 2025 from the Figure 5.3, it can be clearly revealed that the growth pixels are taken place in the east part of Antalya more than the west part. The results show that the east part of Antalya has more potential for development. This may be related to the fact that the eastern part is more suitable for future development according to the model's growth controls.

For another study to encourage the development in the west part, the critical slope can be reduced to 10 from 21 in order to encourage more development in steeper terrain. In this study, the slope coefficient is decreased to 0,69 from 0,73 because the critical slope value is set up to 21 in the scenario file. The spread coefficient measure is also decreasing through the year 2025. This indicates the diffusion occurrence from the existing settlement is decreased. Beside this, the other three coefficient measures (breed, diffusion, and road gravity) are increasing through the year 2025.

As it is seen in Figure 5.4, within the Muratpaşa Municipality the urban pixels are seen in the east part, within Konyaaltı Municipality they occurred in the northeast part, within the Çalkaya Municipality they occurred in north, east and west parts. In the south of Aksu, Varsak, and the Pınarlı Municipalities the urban growth pixels are determined. Doyran, Düzlerçam, and Yurtpınar Municipalities have very small urban growth pixels because their small parts are within the study area.



**Figure 5.3.** Simulation of the Spatial Consequences of Future Urban Growth in Antalya. ( Period of the Years shown in the graphic Is 3. The Final simulation year of 2025 is also shown at the end.

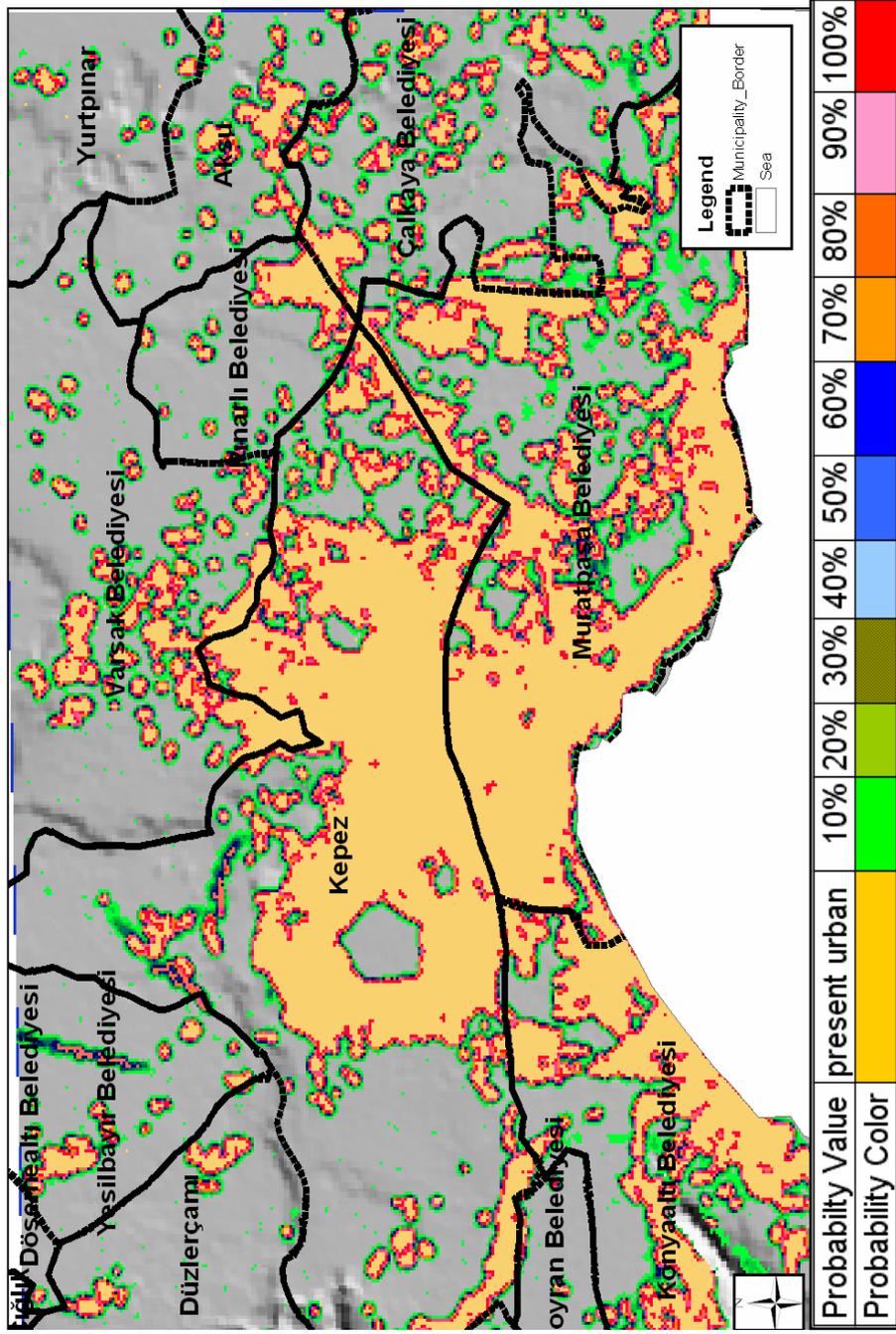
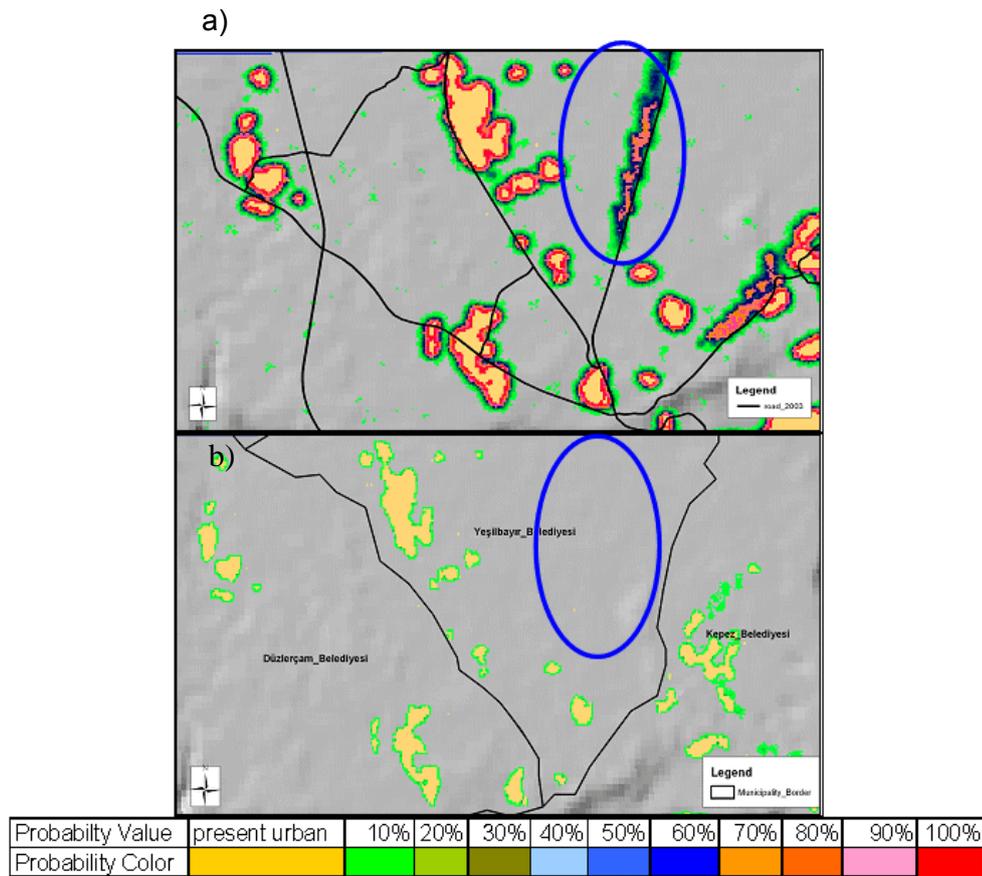


Figure 5.4. The Urban Growth Prediction of the Year 2025 within the Municipalities of Antalya.

In Yeşilbayır Municipality, the road influenced urban growth is obvious in the middle of this municipality (Figure 5.5).

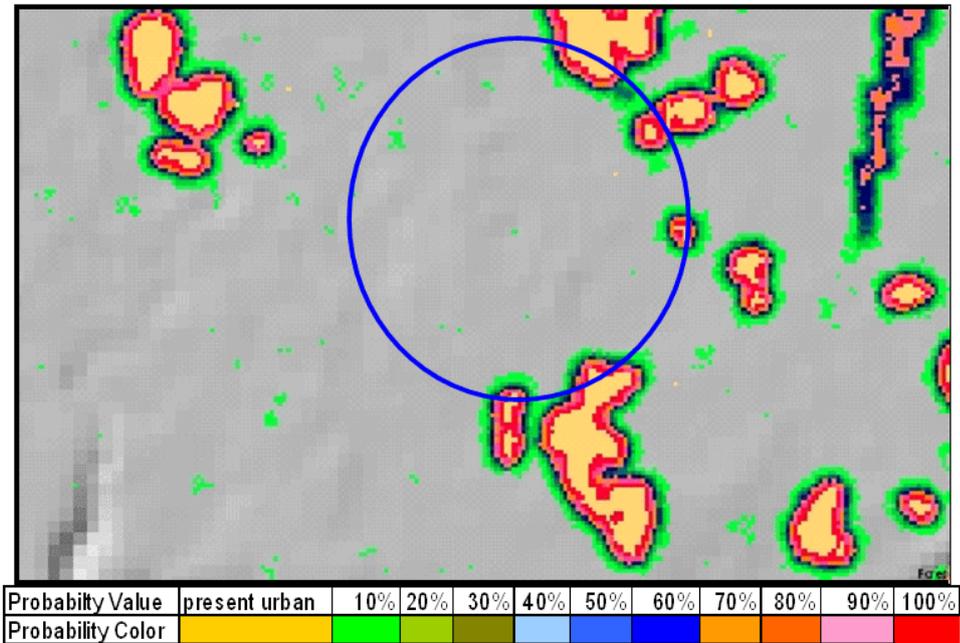
In the figure 5.5.a (2025), the urban growth can be seen obviously, but at the beginning (2003) there is not any urban growth pixels in the corresponding area. This shows us the road influenced growth.



**Figure 5.5.** The Road Influenced Urban Growth in 2025 (a) and 2003 (b).

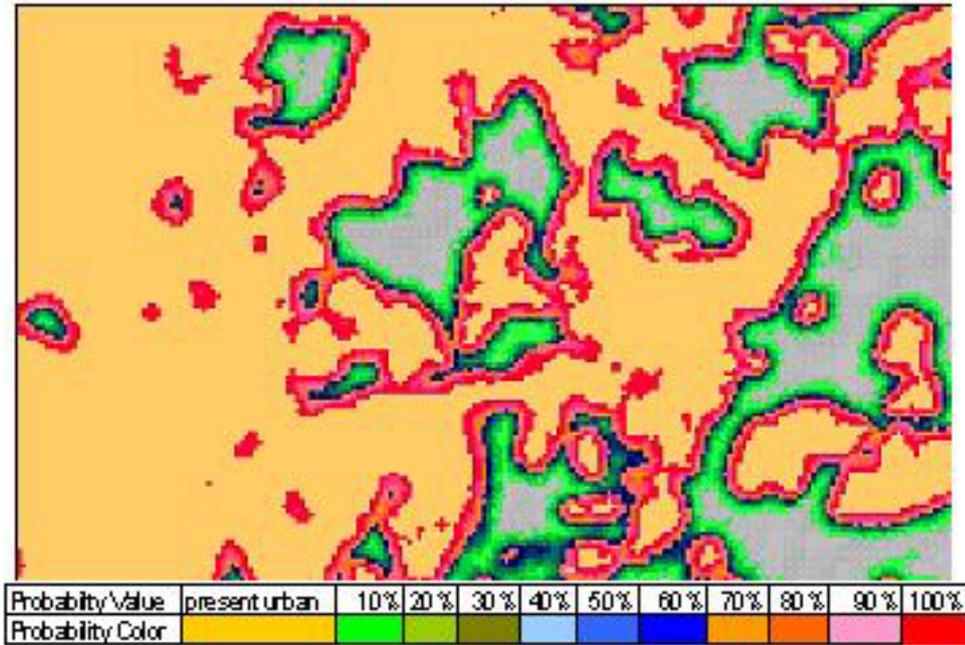
As it is shown in Table 5.1, the sng (spontaneous growth) measure, which is 5,78 in the year 2002, is decreased to 5,26 in 2025. This indicates the decrease of urban development in undeveloped areas.

In Figure 5.6, the area in the circle is the undeveloped urban area. In this area, there is not any growing urban pixel. This can be evidence to the decrease of sng measure. The og (organic growth) measure is increasing from 2801 in the year 2002, to 3621,26 in the year 2025 as shown in Figure 5.7.



**Figure 5.6.** An example of the Decrease of sng (spontaneous growth) Measure in Undeveloped Areas.

The og measure is the organic growth which appears from the existing urban pixels. The increase in og measure explains the expansion of existing urban cells to their surroundings are increased. In Figure 5.7, the yellow pixels are the existing pixels and the massive growth around these pixels can be seen from the figure. The rt (road influenced growth) measure is increased from 13,07 in 2002 to 28,48 in 2025. rt measure is the urban growth pixels influenced by the road. The increase in og and rt measures show that, more residential growth will be encouraged through the 2025 simulation. Figure 5.8 shows the urban growth pixels in 2025 with the current road data. Transportation has always been used as a consistent attractive factor for promoting urban development in the model. In the study area, transportation has been a driving force for housing development.

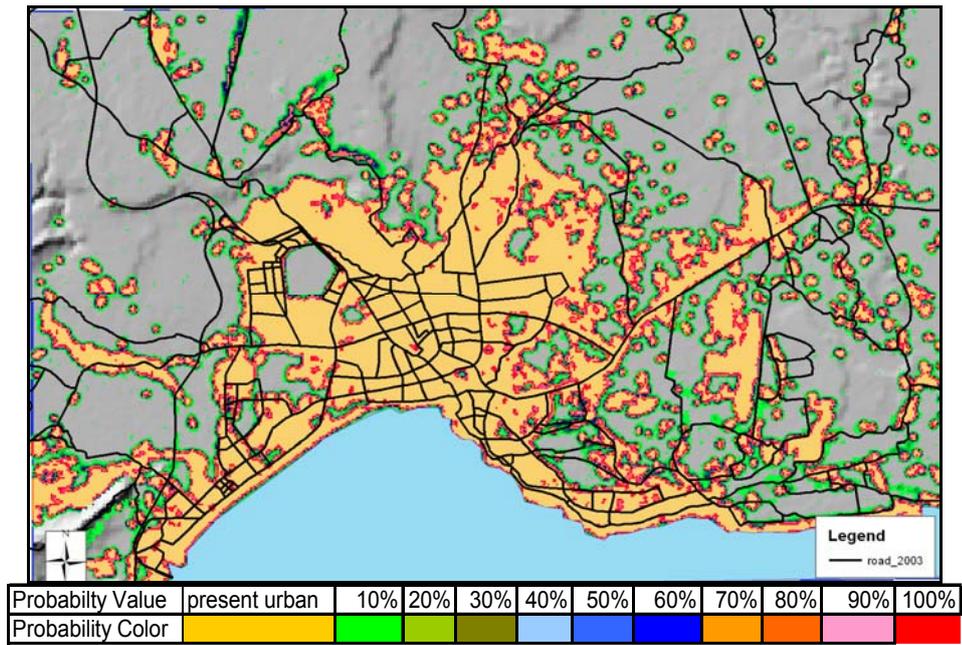


**Figure 5.7.** An Example of the Increase of og (organic growth) Growth Measure.

A low diffusion parameter is obtained for Antalya. This shows that there is not scattering urbanization in this area. Since the spread coefficient is high, it indicates that the transportation invited the urbanization in Antalya.

The pop is the total number of urban growth pixel measure. It is increasing from 92738,64 in 2002 to 201894,3 in 2025. Each pixel dimension is 30m x 30m, 900m<sup>2</sup>. Therefore, in the year 2002, the total urban area is 8347 ha and it is increased to 18171 ha in 2025. The difference of the urban areas between these two years is 9824 ha.

The increase in percent\_urban measure is obvious as 21.80% in 2003 and it is 56.05% in 2025. By 2025, urban settlements would occupy about 56,05% of the entire modeled area.



**Figure 5.8.** The Urban Growth of 2025 with the 2003 Road Data.

## CHAPTER 6

### CONCLUSION AND RECOMMENDATIONS

The escalating urban growth through the world has begun to threaten the environmental and ecological health. To understand the dynamic urban systems together with environment, modeling and simulation with robust methodology and techniques are required. For this reason, dynamic modeling began to gain popularity among geographers and urban planners.

This study aimed to integrate an urban growth model to simulate urban growth in 2025 in the Antalya Metropolitan Area to show the urban growth and urban form of Antalya between the years 2003 and 2025 annually. SLEUTH being a CA model gives the possibility to predict future urban development rather than to explain urban expansion like the other analytical and static urban models.

Being the first application of SLEUTH for a Turkish city, the results are found to be quite encouraging. Main findings of the study can be classified into two dimensions as technological, and application. At the technological theme, the study has demonstrated the usefulness of GIS and the cellular modeling for urban growth prediction. In addition, it is understood that for a long-term urban growth prediction, SLEUTH is suitable. Beside this, the model allows to produce different scenarios. Because the environmental protection is an important concern because of its touristic identity and its being natural heritage, for future urban development planning in Antalya, the scenario used in the study does not let development of the urban growth in the forest, agricultural areas, riverbeds, and natural protective areas, as let in the other ones. However, different scenarios represent different growth strategies that can be adopted by planners. As Yang and Lo (2003) done, in the excluded layer different buffer zones can be applied to the rivers, forests, and

agricultural areas in order to let urban growth around them by using some weights for the buffer distances according to the coast law. By manipulating input data layers, self-modification constraints, and growth control coefficients, different scenarios can be designed. On the other hand, the model is integrated with GIS in order to use the full advantages of this technology. The role of GIS and RS in SLEUTH model is indispensable, especially for input data preparation, model calibration, and developed urban spatial pattern analysis. The model's functionality permits to link the simulation, and the historical and the present growth. This property distinguishes the model from other simulators, which do not have any calibration component. In addition, the model's source code is open to modify the programme system to update the model or add new components for improving the model's performance.

The spatial consequence of urban growth for Antalya is also examined at the application dimension. Rapid change of Antalya in both social and spatial structure during the past 15 years has enabled the city to become an important area for urban dynamic studies. The model's past to present simulation shows an obvious multi-nucleating trend in the evolution of urban spatial form (Figure 5.2). As a result, it is revealed that the growth pixels are taken place in the east part of Antalya more than the west part. The results show that the east part of Antalya has more potential for development. In addition, it is found that the growing rate is decreased through the year 2025. Thus, it is extracted that the protection of the environment affects the growth rate. In addition, the total number of urban growth pixel measures consequently, the total urban area is determined. In the year 2002, the total urban area is 8347 ha and it is increasing to 18171 ha in 2025. Since the model uses the past urban extent information in predicting the future urban development, the development in the 3<sup>rd</sup> dimension is not considered. This factor can be considered in land use predictions of the model. Therefore, the model parameters may change from urban area to another since every urban area has its own properties.

## REFERENCES

Allen, J. and Lu, K., 2003, Modeling and Prediction of Future Urban Growth in the Charleston Region of South Carolina: a GIS-Based Integrated Approach, *Conservation Ecology*, 8(2): 2.

Anderson, J., Hardy, E., Roach, J., and Witmer, R., 1976, A Land Use and Land Cover Classification System for Use with Remote Sensor Data, Geological Survey Professional Paper 964, United States Government Printing Office, Washington: 1976

Arthur, S.T., 2001. A Satellite Based Scheme for Predicting the Effects of Land Cover Change on Local Microclimate and Surface Hydrology. Doctorate Thesis, Penn State, State College.

Avci, M., 2000, A Hierarchical Classification of Landsat TM Imagery for Land Cover Mapping, METU, pp.38.

Bauer, M.E., Yuan, F. and Sawaya, K.E., 2003, Multi-Temporal Landsat Image Classification and Change Analysis of Land Cover in the Twin Cities (Minnesota) Metropolitan Area, Second International Workshop on the Analysis of Multi-temporal Remote Sensing Images, July 16-18, 2003. Ispra, Italy.

Bell, A Model for Forecasting the Location of Fringe Urbanisation, with Gis And 3d Animated Visualisation.

Candau, J., 2002, Temporal Calibration Sensitivity of The Sleuth Urban Growth Model, University of California, Santa Barbara, Department of Geography.

Clarke, K.C., Hoppen, S., Gaydos, L., 1996, Methods and Techniques for Rigorous Calibartion of a Cellular Automaton Model of Urban Growth, University of Santa Barbara, web only ([www.ncgia.ucsb.edu/conf/SANTA\\_FE\\_CDROM/sf\\_papers/clarke\\_keith/clarkeetal.html](http://www.ncgia.ucsb.edu/conf/SANTA_FE_CDROM/sf_papers/clarke_keith/clarkeetal.html), 2006)

Clarke, K.C., Hoppen, S., Gaydos, L., 1997, A Self-Modifying Cellular Automaton Model of Historical Urbanization in the San Francisco Bay Area, *Environment and Planning B*, Vol. 24, pp. 247-261.

Clarke, K.C. and Gaydos, J., 1998, Loose-Coupling a Cellular Automaton Model and GIS: Long-Term Urban Growth Prediction for San Francisco and Washington/Baltimore, *International Journal of Geographical Information Science*, Vol. 12, No. 7, pp.699 – 714.

Çelikoyan, M., Atlan, O., and Kemper, G., 2003, Technical Tools within the Moland-Project For İstanbul ([www.ggs-speyer.de/Moland-Tools.PDF](http://www.ggs-speyer.de/Moland-Tools.PDF), visited in 2005)

Devlet İstatistik Enstitüsü Başkanlığı (DİE), 1990, 1990 Genel Nüfus Sayımı – Nüfusun Sosyal ve Ekonomik Nitelikleri.

Devlet İstatistik Enstitüsü Başkanlığı (DİE), 2000, 2000 Genel Nüfus Sayımı – Nüfusun Sosyal ve Ekonomik Nitelikleri.

DAMPO Planning, 2003, Antalya Development Plan with the scale of 1/5000, ANKARA.

Dietzel, C., and Clarke, K., 2004, Spatial Differences in Multi-Resolution Urban Automata Modeling, *Transactions in GIS*, 2004, 8(4): pp.479 - 492.

Elker, C., 1995, Antalya Ulaşım Planlaması, Bölüm 1: Mevcut Durum, Antalya Büyükşehir Belediyesi.

Esch, T., Roth, A., and Dech, S., 2004, Robust Approach towards an Automated Detection of Built-Up Areas from High Resolution Radar Imagery, ([www.isprs.org/commission8/workshop\\_urban/esch.pdf](http://www.isprs.org/commission8/workshop_urban/esch.pdf)., visited in 2005)

Forster, B.C., 1985, An Examination of Some Problems and Solutions in Monitoring Urban Areas from Satellite Platforms, *International Journal of Remote Sensing*, Vol.6, No.1, pp.139-151.

Goldstein, C, 2004, Brains vs. Brawn - Comparative Strategies for the Calibration of a Cellular Automata - Based Urban Growth Model, Department of Geography, University of California at Santa Barbara ([www.geocomputation.org/2003/Papers/Goldstein\\_Paper.pdf](http://www.geocomputation.org/2003/Papers/Goldstein_Paper.pdf), visited in 2006).

Hung, M., 2002, Urban Land Cover Analysis from Satellite Images ([www.isprs.org/commission1/proceedings02/paper/00099.pdf](http://www.isprs.org/commission1/proceedings02/paper/00099.pdf), visited in 2005).

Karakış, S., Marangoz, A. M., and Büyüksalih, G., 2005, Quickbird Pan Sharpened Görüntüsü Üzerinden Otomatik Detay Çıkarımı ve Coğrafi Bilgi Sistemlerine Uygunluğunun Analizi, TMMOB Harita ve Kadastro Mühendisleri Odası 10. Türkiye Harita Bilimsel ve Teknik Kurultayı, Ankara.

Lecture Notes of the Integration of RS and GIS, 2004.

Li, X and Yeh, A., 1998, Modeling Sustainable Urban Development by the Integration of Constrained Cellular Automata and GIS, *Int. J. of Geographical Information Science*, 2000, Vol. 14, No. 2, pp.131-152.

O'Sullivan, D., and Torrens, P.M., 2000, Cellular Models of Urban Systems, Centre for Advanced Spatial Analysis, University College London, Paper 22, ([www.casa.ucl.ac.uk/cellularmodels.pdf](http://www.casa.ucl.ac.uk/cellularmodels.pdf), visited in 2005)

Pijanowski, C, Long, T., Gage, H., and Cooper, E., 1997, A Land Transformation Model, Conceptual Elements Spatial Object Class Hierarchies, GIS Command Syntax and an Application for Michigan's Saginaw Bay Watershed, Land Use Modeling Workshop, South Dakota.

Silva, E. A., and Clarke, K.C., 2002, Calibration of the SLEUTH Urban Growth Model for Lisbon and Porto, Portugal, *Computers, Environment and Urban Systems*, Pergamon, Vol.26, pp. 525–552.

Torrens, P.M., 2000, How Cellular Models of Urban Systems Work (1.Theory), Centre for Advanced Spatial Analysis, University College London, Paper 28 ([http://www.casa.ucl.ac.uk/how\\_ca\\_work.pdf](http://www.casa.ucl.ac.uk/how_ca_work.pdf), visited in 2005).

Turner, M.G., 1987, Spatial Simulation of Landscape Changes in Georgia: A Comparison of 3 Transition Models, *Landscape Ecology*, Vol. 1, No. 1, pp 29-36.

UTTA Planning, 1995, Antalya Kentsel Bölgesi Gelişme Deseni Açıklama Raporu, Antalya Büyükşehir Belediyesi Kent Planlama Çalışmaları - 1.

Wolfram, S., 1994, Cellular Automata and the Complexity ([www.stephanwolfram.com/publications/books/ca-reprint/](http://www.stephanwolfram.com/publications/books/ca-reprint/), visited in 2006).

Xie, Y., 2003, Integrated Dynamic Urban Evolution Modeling, Institute for Geospatial Research & Education, Eastern Michigan University.

Yang, X. and Lo C.P., 2003, Modeling Urban Growth and Landscape Changes in the Atlanta Metropolitan Area, *Int. J. Geographical Information Science*, Vol. 17, No. 5, pp. 463-488

Zhang, J., and Foody, G.M., 1998, A Fuzzy Classification of Sub-Urban Land Cover from Remotely Sensed Imagery, *Int. J. of Remote Sensing*, Vol. 19, No. 14, pp. 2721- 2738.

## ONLINE REFERENCES

URL1: UN - DESA (United Nations of Department of Economic and Social Affairs) Future World Population Growth To Be Concentrated In Urban Areas Of World, <http://www.un.org/esa/population/publications/wup2001/WUP2001pressrelease.pdf>, (visited on 12.05.2005)

URL2: John Conway, "John Conway's Game of Life", [www.bitstorm.org/gameoflife/](http://www.bitstorm.org/gameoflife/) (visited on 11.15.2005)

URL3: Project Gigalopolis, Urban, and Land Cover Modeling, [www.ncgia.ucsb.edu/projects/gig/](http://www.ncgia.ucsb.edu/projects/gig/) (visited on 01.04.2006)

URL4: Antalya Yerel Gündem 21 Kent Konseyi, [www.antalyakentkonseyi.org.tr/raporlar\\_ant\\_kent\\_envanter.htm](http://www.antalyakentkonseyi.org.tr/raporlar_ant_kent_envanter.htm), 2005 (visited on 12.08.2005)

URL5: Earth Science Data Interface, <http://glcfapp.umiacs.umd.edu:8080/esdi/index.jsp> (visited on 01.13.2005)

URL6: Accuracy Assessment, [www.forestry.ubc.ca/irss/lectures/Accuracy-Assessment.pdf](http://www.forestry.ubc.ca/irss/lectures/Accuracy-Assessment.pdf) (visited on 08.25.2005)

## APPENDIX A

### The Definitions;

**Coarse phase.** See Calibration phase.

**Calibration Mode.** An automated process of searching through the model coefficient space to find a set which best describes historical urban change for a study area. Coefficient sets are generated using the coefficient start, step and stop values defined in the scenario file. Each set initializes a run.

**Calibration phase.** One of three steps in brute force calibration (coarse, fine, and final) through which coefficient ranges are narrowed. See section 3.2 for more information.

**Coefficient start value.** Initial coefficient value for a model run. The low value of a coefficient range.

**Coefficient step value.** In calibration, an increment value which is added to the start value iteratively for all possible permutations of given ranges and increments.

**Coefficient stop value.** Final coefficient value for a model run. The high value of a coefficient range.

**Control year.** A date for which urban data exists in the historical database. An urban layer from the historical database.

**CRITICAL\_HIGH.** The threshold for the urban growth rate above which a boom state exists for the system and self-modification will be applied to the coefficients.

**CRITICAL\_LOW.** The threshold for the urban growth rate below which a bust state exists for the system and self-modification will be applied to the coefficients.

**Derive run.** Initialized with the set of best coefficients selected from the calibration phases, a large number of Monte Carlo simulations are used to simulate growth for the time represented in the historical database. From this run, averaged values of the calibration metrics and coefficient values for the control years may be derived for analysis. Most importantly, the averaged coefficient values from the final control year are used to initialize a calibrated forecast run.

**Final phase.** See Calibration Phase.

**Fine phase.** See Calibration Phase.

**Forecast mode.** Initialized with the most recent image data, will perform a single run, in Monte Carlo fashion, using the calibrated BSS for initialization.

**Growth Cycle.** The basic unit of SLEUTH execution. It begins by setting each of the coefficients to a unique value. Each of the growth rules are then applied. Finally, the resulting growth rate is evaluated. If the growth rate exceeds or falls short of the *CRITICAL\_HIGH* or *CRITICAL\_LOW* values, model self-modification is applied. Selfmodification will slightly alter the coefficient values to simulate accelerated or depressed growth that is related with boom and bust conditions in urban development.

**Prediction Mode.** See prediction mode.

**ROAD\_GRAV\_SENSITIVITY.** A change value that is applied to the road gravity coefficient during self-modification.

**Run.** An execution of SLEUTH that begins with a single set of coefficient values, and performs a designated number of Monte Carlo iterations. May be followed

by another run (as in calibration) or finish at the end of the Monte Carlos (as in forecasting).

**Scenario file.** The SLEUTH execution file.

**Self-modification.** A process of slightly altering SLEUTH coefficient values to simulate accelerated or depressed growth that is related with system-wide boom and bust conditions in urban development.

**SLOPE\_SENSITIVITY.** A constant value that is applied to the slope resistance coefficient during self-modification.

**Start date.** The first year represented by SLEUTH simulation. In Calibration this date corresponds with the date of the earliest (most historical) urban layer. In Forecasting, it will correspond to the START\_DATE value defined in the run's scenario file, which must also be the date of the most recent urban layer.

**Stop date.** The final year represented by SLEUTH simulation. In Calibration, this date corresponds with the date of the most recent urban layer. In Forecasting it will correspond to the STOP\_DATE value defined in the run's scenario file.

**Simulation.** A simulation is a series of growth cycles that begins at a start date and completes at a stop date (Candau, 2002).

**Products:** They are the values obtained from the calibration modes. These are the values of the five calibration metrics best-fit scores, which are used to sort and identify the best performing coefficient sets.

**Compare:** It compares the amount of modeled urban area to known urban area for the stop date year where  $P_{modeled}$  is the modeled urban area and actual area for final year is  $P_{actual}$ : if  $(P_{modeled} < P_{actual})$  {compare =  $(P_{modeled} / P_{actual})$ } else {compare =  $1 - (P_{modeled} / P_{actual})$ }

**Pop:** It is the least squares regression score for the modeled urban area compared to actual urban area for the control years.

**Edges:** It is the least squares regression score for the modeled amount of urban perimeter or edge, compared to actual urban perimeter for the control years.

**Clusters:** It is the least squares regression score for the modeled number of urban clusters compared to known number of urban clusters for the control years.

**Cluster\_size:** It is the least squares regression score for the modeled average urban cluster size compared to known average urban cluster size for the control years.

**Leesallee:** It is a shape index, a measurement of spatial fit between the model's growth and the known urban extent for the control years 1 being a perfect match and 0 representing a spatial disconnect:

$$S = (A \cap B) / (A \cup B)$$

where  $A$  is modeled and  $B$  is actual urban area.

**Slope:** It is the least squares regression of average slope for modeled urbanized cells compared to average slope of known urban cells for the control years

**%Urban:** It is the least squares regression of percent of available pixels urbanized compared to the urbanized pixels for the control years.

**Xmean:** It is the least squares regression of average longitude (calculated using column values) for modeled urbanized locations compared to average longitude of known urban locations for the control years.

**Ymean:** It is the least squares regression of average latitude (calculated using row values) compared to average latitude of known urban locations for the control years.

**Rad:**  $\sqrt{(\text{std}_x^2 + \text{std}_y^2)}$  is a measure of urban dispersal.

**Avg.log.** A SLEUTH output file. It contains measured values of simulated data averaged over Monte Carlo iterations for every run and control year.

**Base statistics.** The measurements taken from the control year data.

**Best solution set.** (BSS) The goal of model calibration. The average of each of the five coefficient values from the final simulated year of the derive run. These are used to initialize a forecast run.

**BOOM.** Self-modification parameter applied to the dispersion, breed, and spread coefficients when the system is in a boom state.

**Boom state.** A state of accelerating urban growth entered into the when the urban growth rate exceeds the *CRITICAL\_HIGH*.

**BUST.** Self-modification parameter applied to the dispersion, breed, and spread coefficients when the system is in a bust state.

**Bust state.** A state of decelerated urban growth entered into the when the urban growth rate goes below the *CRITICAL\_LOW*.