Spatio-Temporal analysis of urban growth and its future trends of District Peshawar using CA Markov and SLEUTH Model

By
Arsalan Tahir
(2019-MS-GIS-RS-00000320475)

A Thesis Submitted in partial fulfillment of the requirements for the degree of MASTER of SCIENCE in Remote Sensing and GIS

Institute of Geographical Information Systems
School of Civil and Environmental Engineering
National University of Sciences & Technology
Islamabad, Pakistan
June 2023

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by **Arsalan Tahir (Registration No. MSRSGIS 00000320475), of Session 2019 (Institute of Geographical Information Systems)** has been vetted by undersigned, found complete in all respects as per NUST Statutes/Regulation, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: _____
Name of Supervisor: Dr. Ejaz Hussain
Date: 28.8.2023

Dr. Ejaz Hussain
Associate Dean IGIS, SCEE (NUST)
H-12, ISLAMABAD

Signature (HOD): _____
Date: 28/08/2023

Dr. Javed Iqbal
Professor & HOD IGIS, SCEE (NUST)
H-12, Islamabad

Signature (Associate Dean): _____
Date: 28.8.2023

Dr. Ejaz Hussain
Associate Dean IGIS, SCEE (NUST)
H-12, ISLAMABAD

Signature (Principal & Dean SCEE): _____
Date: 29 AUG 2023

PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

ACADEMIC THESIS: DECLARATION OF AUTHORSHIP

I, **Arsalan Tahir**, declare that this thesis and the work presented in it are my own and have been generated by me as the result of my own original research.

"Spatio-Temporal Analysis of Urban Growth and its trends of district Peshawar using CA Markov and Sleuth Model."

I confirm that.

1. This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text.
2. Wherever any part of this thesis has previously been submitted for a degree or any other qualification at this or any other institution, it has been clearly stated.
3. I have acknowledged all main sources of help.
4. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
5. None of this work has been published before submission.
6. This work is not plagiarized under the HEC plagiarism policy.

Signed: ……

Date: 30-08-2023

DEDICATION

"I dedicate my achievements and success to my family for showing unwavering support, guidance, and love throughout my life's journey.

And

My friends and teachers who motivated me and helped me grow and developed me as a person."

ACKNOWLEDGEMENT

All glory to Allah Almighty, upon Him we ultimately rely for sustenance, wisdom, and guidance.

I would like to express gratitude to my supervisor Dr. Ejaz Hussain, without whom the completion of this thesis could not have been possible. I would like to thank him from the core of my heart for believing in me. He guided me thoroughly throughout my research journey, he was always available and listened to all the problems that I faced in research work and guided me accordingly. He takes time out from his busy schedule and actively participates and helps in student's research in the friendliest way possible.

I would like to thank my committee members Dr. Abdul Waheed (NICE) and Mr. Junaid Aziz Khan (IGIS) for their support, valuable suggestions, thoughtful critique, and guidance during the research phase. I would also like to thank Dr. Salman Atif (IGIS) for valuable suggestions that strengthened my research. I am grateful to Mr. Naqash Taj Abbasi (IGIS) for being supportive during my Masters.

The whole administration of IGIS NUST has been very kind and supportive so I would like to thank them as well.

I would like to thank the people who have inspired and motivated me and helped me achieve much more than I expected.

My parents always believed in me, and they always helped me achieve way more than what I am capable of. So, I take this opportunity to thank them for whatever they have done for me. Thank you for standing by me in the toughest of times.

Lastly, I would like to thank my Siblings, Friends, and Teachers for their prolonged support.

Arsalan Tahir

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Explanation
AOI	Area of Interest
ASCII	American Standard Code for Information Interchange
CA	Cellular Automata
DEM	Digital Elevation Model
ESRI	Environmental Systems Research Institute
ETM	Enhanced Thematic Mapper
GIF	Graphic Interchange Format
GIS	Geographic Information System
LULC	Land Use Land Cover
MC	Markov Chain
MCE	Multi Criteria Evaluation
OLI	Operational Land Imager
OSM	Open Street Map
RS	Remote Sensing
SRTM	Shuttle Radar Topographic Mission
TIRS	Thermal Infrared Sensor
TM	Thematic Mapper
UGM	Urban Growth Models
UN	United Nations
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
WGS-84	World Geodetic System – 1984

Abstract

Land use change and urban expansion are spatially variable and dynamic processes caused by human activities, which causes a significant impact on the urban landscape because of the conversion of natural environment to build environment. In developed countries, strict implementation of zoning regulations results in planned urban growth, however, it is the opposite in the developing countries which expand haphazardly. Pakistan being a developing country, also faces rapid urbanization related issues in many urban areas. District Peshawar is one of these where the urban growth rate is about 14%. It is felt important to track urban growth to address the problems that may arise in future. This study focused on mapping of the Spatio-temporal urban expansion from 1990 to 2021 using Landsat 5 and 8 data. The results show that that the urban area of district Peshawar increased from 50.5 sq km to 260 sq km i.e., about 20.23% of the total district's area. The urban area increases at the cost of vegetative and open land. The further comparison and analysis revealed that out of 260 sq km 147 sq km increase was within the allowable zones while 113sq km was in zones restricted settlements. The research also focused on the future prediction of urban growth trends using CA Markov Chain Model and SLEUTH (Slope, Land use, Urban, Excluded area/land, Transportation, Hill shade) for year 2039. Both models were calibrated and validated using urban extents mapped from 2000 and 2010 remotely sensed data. Simulated urban growth and its comparison with image based urban revealed that CA-MC model over predicted the urban area by 20.50 sq km while the Sleuth model over predicted by about 7.3 sq km area. The performance of Sleuth model was the better in terms of estimating the urban growth. The future simulation using CA Markov and the Sleuth models show that in 2039, the urban area will be about 418 and 476 sq km, respectively. These findings can help the relevant authorities in better monitoring management and formulating policies for the future sustainable urban expansion.

Chapter 1

INTRODUCTION

1.1 General

Land use change and urban growth are spatially varying, and dynamic processes induced by humans. These changes cause rapid urbanization and significantly affects the natural environment because of its conversion to build environment (Clarke, 1995; Pinto, Antunes, & Roca, 2017). The urban growth process increased enormously in the last few decades. This is because the urban areas provide economic opportunities, and this factor drives population towards cities. (Ramachandra, et al., 2014). There are different forms of urban expansion such as infill, edge expansion, and outlying. Infill growth is the predominant form of growth, and it occurs within a built-up area. Expansion growth is primarily focused on infill growth, which is closely linked to existing built-up areas. On the other hand, outlying growth takes place independently from these existing built-up areas. Primarily, peripheral expansion transpires within unoccupied regions and ecologically vulnerable territories situated in the vicinity of the urban center. Outlying growth can be further divided into linear branch, isolated, and clustered branch growths. (Wilson et al., 2003, Stephen. et al., 2005))

Worldwide, urban areas are growing rapidly, the built-up portion of earth had utilized about 400000 square kilometers up to the year 2000, and accounts for 0.3 percent of the total Earth surface (UN, 2005). In 2014, the global urban population was 54 percent which exceeded the rural population for the first time. Presently, over 56 percent of the world's population resides in urban regions, the trend seems to continue, and the human population is predicted to

grow more and migrate to cities. It is expected that about 66 percent of the world's population or 7 out of 10 people will live in urban areas by 2050 (UN,2015; UN, 2019; World Bank, 2020).

Urban Expansion is a very complex dynamic phenomenon, and it is affected by interaction of several parameters in time and space at different scales (Barredo et al., 2003; He et al., 2006). In developed countries, the new settlements occurs according to zoning regulations but in developing countries, it is opposite, and is considered as dispersed urban sprawl (Arsanjani et al., 2011).

1.2 Urbanization in Pakistan

Pakistan is one of the developing countries where haphazard urban expansion occurs. It has the world's 36th-largest landmass (about 881,913 square kilometers) and the sixth-highest populated countryy in the world. Its population was around 207.8 million in 2017 which is 57% more than that in 1998 (132.23 million), (Pakistan Bureau of Statistics, 2017; The World Bank, 2017). Alongside population growth, the country faces an issue of rapid urbanization (Haider & Badami, 2010; Kugelman, 2013; Mustafa & Sawas (2013) and currently 36.4 % of the population resides in cities (Census 2017).

Khyber Pakhtunkhwa (KP) Province has been experiencing rapid increase in its urban population. It raised to 17% in 1998 and to 18.7 % in 2017. In KP, district Peshawar is the 8th largest district of Pakistan in terms of population and is growing rapidly with an average growth rate of 3.99% from 1998 (2 million) to 2017 (4.26 million) with 46.16 percent of the population residing in urban areas. (Census 2017).

1.3 Urban Growth Models

Urban growth at high rates arises many concerns such as loss of agricultural land as well as the environmental impact of increased urbanization. Urban systems are complex and simulating urban dynamics, a challenge for decision-makers and urban planners. To understand the complex urban systems, it is important to understand driving factors, process of urban growth, and then incorporate the current scenario into land use models and enable simulation of the future scenario. (Thapa and Murayama, 2010; Tang et al., 2018). Several static and analytical urban models have been developed to describe urban dynamics and evolving patterns along with prediction of the future urban extent to enable resource managers and urban planners to manage urban expansion and resources efficiently (Tayyebi, 2017; Arentze, 2017)

1.4 Cellular Automata Based Urban Growth Models

Spatially Explicit Models (SEMs) is a location based approach to reproduce the dynamics of geographic phenomena such as epidemic spread, crime pattern land-use change, environmental dynamics, epidemic spread, and urban growth (Verburg and Veldkamp 2001, Brown and Xie 2006, Liu et al. 2017). Cellular Automata (CA) are well known and widely used in past decades for land use and urban growth simulations because they incorporate factors that affect urban development to accurately match the practical on-ground process of urban growth. CA based urban growth models self-organizes to simulate the complex dynamic scenarios in geographic systems (Dragicevic, 2010; Lai, Dragicevic, & Schmidt, 2013; Lu et al. 2013). Five essential components define CA models which includes (a) tessellated cells in space, (b) predetermined cells states, (c) cell neighborhoods, impact of neighborhood cells, number of neighborhood cells, (d) transition rules, determines each entity's transition from one state to another based on the state of neighboring cell, (e) the number of iterations required by

a cell to change its state (White & Engelen, 1993). The CA transition rules among these components determine the cells state defined by the relationship between neighborhood cells.

CA approach is usually integrated with Geographic Information Systems (GIS) to improve the spatial computational capability of the model (Abdullahi & Pradhan, 2018). Geo-CA models are classified as stochastic when they incorporate a random component, resulting in varying outcomes when the model runs with identical inputs (Vermeiren et al., 2016; Mustafa et al., 2018). Conversely, these models are deterministic nature when the random component is omitted, leading to consistent results when model runs with the same inputs. (Omrani et al. 2017, Mustafa et al. 2018). The specification of geographical size, driving variables, neighborhood, and transition probability all have an impact on CA, which results in a variety of simulated outputs with differing accuracy levels and spatial patterns (Feng et al., 2011, Basse et al., 2014, Barreira et al., 2017).

The CA-Markov Chain model is a spatial transmutation and probabilistic, and an efficient simulator which predicts land use land cover changes efficiently compared to urban growth models which are based on linear extrapolation (Aaviksoo, 1993). CA-MC model integrates the stochastic temporal approach of Markov technique with the stochastic spatial cell division automata strategy of cellular automata technique for spatial and temporal pattern simulation of land use changes (Han et al., 2009; Ghosh et al., 2017). CA-MC Model combines and takes advantage of the Markov chain that uses quantity prediction to anticipate future change based on previous trend, while CA identifies the geographic location of changes, and together they create a robust framework. which eliminates the shortcoming of each of the model to predict land use trend efficiently (Eastman, 2009; Huang et al., 2015).

The SLEUTH is an urban model based on whose name is derived from the six input layers namely Slope, Land Use, Excluded layer, Urban, Transportation and Hill shade. It was created by Dr. Keith Clarke and was initially used to model the characteristics and propagation of wildfires. Its first application was simulation of San Francisco Bay area with a 600 meters resolution. SLEUTH is an efficient urban growth simulating model which enhances our knowledge regarding the alterations in the surrounding landscape resulting from urban expansion. (Clarke, 1997). It incorporates different human perceptions to predict future land use (Arthur, 2001).

1.5 Objectives

a. To map and model the spatio-temporal urban growth of Peshawar and its comparison with the master plan.
b. To predict and simulate the future urban growth trends and pattern using CA Markov and SLEUTH models.
c. Compare the performance of CA Markov and SLEUTH models.

1.6 Relevance to National Needs

The increase in urban extent results in the reduction of the agricultural area along with other natural resources and will have negative impact on the environment as well. It also adds to other problems such as increased traffic, traffic congestion as well as increased pollution. Simulating the urban extent will help for better resource management and urban planning.

1.7 Areas of Application

a. Better Urban monitoring, planning and management.
b. Urban Utilities' Management

c. Traffic Management

d. Environmental Impacts

1.8 Literature Review

Hamad, et al.,2018, conducted land use land cover change analysis of Halgurd-Sakran Core Zone (located in the east of Erbil-Iraq) using CA-MC model. They used Landsat 5 images for the years 1993, 1998, 2003 and 2008 and Landsat 8 for the year 2017 for simulating land use land cover map of 2023. The images were classified in four classes with an accuracy of 97%. The satellite images of 1993 and 1998 were used for the optimization, 2003 image for validation, and 2003 and 2008 images to simulate land use of 2017. It was the compared against land use land cover map of 2017 with an accuracy of 80%, and then the model was used to simulate the urban growth of 2023. The CA-MC modelling results were reliable and accurate.

Zhou et al. (2020) attempted to simulate the urban land change of Shanghai using CA-MC. They collected historical land use land cover, constraint factors, driving factors and classified the acquired in six land use classes. CA-MC model was utilizeato simulate urban growth in 2015 which was then compared with the actual land use land cover and found the simulated map to be 94.88% correct and suggested the model to be accurate in simulating land use changes. It was then used to model for future urban simulation which suggested that there will be 33.88% increase in the urban areas by 2030.

Mansouret al., (2020) monitored the land use land cover changes in mountain area of Oman using CA-MC Model. The research used Landsat images of 1998, 2008 and 2018. Digital Elevation Model was employed to create aspect, slope and area elevation. The satellite

images were classified using a supervised classification algorithm. For validation, they projected the land use land cover for the years 2008 and 2018 and compared it with the same year's classified layer with Kappa Coefficient higher than 0.8 which suggests the results were accurate. After validation, the model was used to simulate land use land cover for the years 2028 and 2038. The simulated results showed that the urban land use class increased from 31% to 38.18% in 2028 and rose to 48.25% in 2038 with decline in the bare land and vegetation classes.

Huang et al., (2019) used CA-Markov Chain Model to analyze the future land use land cover change in Beijing, China. The study used land use land cover data for the years 1992, 2001, 2005, 2009 and 2012 acquired from Beijing Municipal Bureau of Land Use and Resource (BMBLUR). The data were converted to a 30*30 meters raster dataset, and categorized in six land use classes. To verify the efficiency of the CA-MC, the year 2012 was simulated based on the land use land cover layer of 2001 and the transition area matric was set to 11 years. A comparative analysis was conducted between the observed area percentage and the projected area percentage and the relative error was found to be less than 5%. The reliability of the model was further tested using scenarios and the simulations for which results were found to be valid. Based on the achieved results, the projection of land use land cover for the years 2020 and 2030 was also termed as reliable. The research suggested that the CA model can be applied to predict land use changes. It quantifies the land use land cover change by category and also predicts the spatial distribution with higher accuracy.

Altuwaijri et al., (2019) predicted urban growth of Arriyadh city, Saudi Arabia using Cellular CA-MC model over a thirty year-time period. The research used Landsat imageries for the year 1987 and 2017. The imageries were classified in four broad categories. The study

considered slope, elevation, distance from the city center and the main road network. These factors were given relevant priorities using the Analytical Hierarchy Process (AHP) and afterwards multi-criteria evaluation (MCE) was applied to produce transition potential maps. A land change modeler was used to forecast the land change map of 2047.. The results suggested that the land occupied by urban class will increase from 38% (2017) to 82.9% in 2047.

Baqa et al., (2021) modeled urban growth trend and pattern of Karachi using CA-Markov Model and Urban Sprawl matrix. Landsat Thematic Mapper (TM.) images from 1990, 2000 and 2010 and Operational Land Imager (OLI) image from 2020 were the primary data sources for assessing the urban patterns. The data was prepared in Google Earth Engine. Land use Land cover classification was carried out for 6 classes, with the overall accuracy classifications between 89% to 91%. In order to distinguish between plain regions and hilly regions, SRTM digital elevation model (DEM) was employed. Topographical maps from Survey of Pakistan were utilized as reference data. The CA-MC model was implemented using four steps. It included land use land cover classification, calculation of transition probability matrix, generation of transition suitability maps and assessment of the model's capacity to forecast future changes. Transition probability matrix for the years 1990-200, 2000-2010 and 2020 were generated and was used to forecast land use land cover of Karachi city for the year 2030. For validation, the year 2020 was simulated which matched the classified layer with Kappa statistics of 0.87. The study concluded that CA-MC is an efficient model in monitoring and evaluating changes in spatial urban patterns.

Wu et al., (2019) examined the spatial scale sensitivity in CA-MC simulation of land use change of Wuhan city, China The research used 30 meters resolution Landsat 5 Thematic

Mapper data for the years 1987, 1996 and 2005. The images were classified in 5 land use land cover classes. A set of transition probabilities matrices with varying cell diameters were calculated for 1987 and 1996 classified layers. A total of 26 transition matrices were calculated which were related to the size of cells and not the size, nature and type of the surrounding neighborhood. The transition probability of construction land was highest (0.63), the second highest was arable land (0.60) followed by water (0.57), forest (0.32) and unused land (0.04). From 1987 to 1996, a substantial portion of the forest and open land was converted to agricultural land. The classified layers from 1987 and 1996 were used to simulate land use land cover for the year 2005 using six transition probability matrices. Different cell sizes ranging from 30 meters to 330 meters were used for the simulation, along with neighborhood filters ranging from 3 to 23 using two distinct types of neighborhoods, Von Neumann and Moore. The accuracy of the simulation decreased continuously during the upscaling of the neighborhood and cell size. The Kappa statistics were computed by comparing land use simulation maps to the 2005 classified land use map which ranged from 0.6 to 0.7 at different cell sizes and neighborhoods. The findings of the study indicated that the CA-MC model is significantly affected by spatial scale change and if the individual component effects are considered, then CA-MC Comparatively to neighborhood type or cell size, exhibits greater sensitivity to neighborhood size.

Mohamed, A., & Worku, H. (2020) simulated the dynamics of urban land use and land cover for Addis Ababa, Ethiopia using CA-MC approach. Landsat images for the years 2005, 2011 and 2015 were classified with an accuracy of 86%, 87% and 87%, respectively. The research used drivers that can affect land use land cover change such as transportation layer, elevation, restriction layer and slope map. Before future land use simulations, the model was pretested by comparing the predicted and actual land use land cover map of 2015. The validate

tool in IDRISI was used which generated 0.85, 0.92 and 0.87 for K-standard, k-location and k-standard respectively. Using the CA-MC model land use land cover for the year 2025 and 2035 were generated under two scenarios.

Singhet al., (2015) used CA-MC model to predict spatial land use land cover changes in district Allahabad, Uttar Pradesh State, India. Landsat images of 1990, 2000 and 2010 were used for mapping purposes. Results of future simulation were compared with present day changes using Kappa coefficient. The study determined that if the efficiency of model reaches more than 80% for initial image, then CA-MC is a strong tool for land use simulation. For validation of the model, the statistical analysis of agreement and disagreement were presented between predicted map and reference map of 2010 which were reasonably similar and the simulated map's overall accuracy was 88.5%.

Al-Shaar, W., (2020) applied CA-MC model to forecast land use pattern in Lebanon. Landsat images for the years 2000, 2009 and 2018 were classified in 4 land use classes using ArcGIS 10.6.1. The accuracy of all the images was more than 85%. The Land use layers for the years 2000 and 2009 were used to simulate land use map of 2018. Kappa statistics were used to validate the accuracy of the model which was 0.80 and were within the specified range. This outcome confirmed the model's accuracy in predicting Lebanon's projected land use change for the year 2036.

Linehan, C. J. (2021) applied SLEUTH model to predict future urban growth of Davidson County, Tennessee. The Landsat TM and ETM images for the years 1986, 1990, 1996 and 2000 were used. The resolution of the images was resampled to 45 meters from 30 meters to decrease array’s size. Land use land cover classification was carried out on these images. If a pixel was classified developed in 1986 and 2000 but was added to other class in

1990 or 1996, it was reclassified as developed in that particular layer. Transportation layers for the years 1986 and 1996 were prepared from primary road network in 1:100,000 scale USGS digital line graphs. Digital elevation model (DEM) was used to prepare slope and hill shade beside this excluded layer was also created which consist of water (100% exclusion) local parks and federal state (80% exclusion). Raster layers of 45-meter resolution were created for all the input layers. The calibration of the model was performed in 3 phases coarse, fine and final. The model was run in test mode utilizing the set of parameters derived from the calibration phase. The model took more than a week to complete the calibration phase. To assess the accuracy of the model, urban growth and extent for the year 2000 was predicted. The model was run with one hundred Monte Carlo iterations. The overall accuracy of the model was 93.1 percent which is quite high but forecasting the exact location of urban pixel was problematic.

Chaudhuri, G., & Clarke, K. C. (2019) modeled urban growth of Kolkata, India using SLEUTH urban growth model. The research used datasets from DIVA-GIS data repository, Open Street Map, 2010 Global Man-made impervious surfaces, Census of India and USGS. Landsat images for 1989, 1999. 2015 and 2010 were classified with an accuracy of 80%-87%. Hillshade along with slope were created from SRTM DEM. Road layers for the years 2004-2005 and 2010 were used from existing shapefiles. Three stages calibration was conducted. The model was employed to forecast urban land use for the years 2011 and 2017, where there was an under prediction of urban pixels by 13% and 6% respectively while in the 2nd experiment it under projected urban pixels by 7%. After validation the model was used to simulate urban growth for the year 2030.

Mallouk et al., (2018) used SLEUTH urban growth model to predict future urban expansion of Casablanca region Morocco. Landsat images of 1984, 2010 and Sentinel images

of 2014 and 2018 were classified. Moreover, the road network and slope excluded area were also obtained from these images. The data were rasterized and exported to SLEUTH 3.0 where the calibration of the model was carried out using three steps which were based on Monte Carlo simulations. For the coarse, fine and final calibration, the data were resampled to 120 m, 60 m and 30 m respectively, after which the simulation for the year 2040 was carried out. The research concluded that SLEUTH model is appropriate for spatially and quantitatively assessing current and the future urban expansion.

Dhanaraj et al., (2022) explored the land use land cover change dynamics and simulated urban growth of Mangaluru city, India through SLEITH model. The research used topo sheets, SRTM DEM, google earth images, transport network and satellite images (Landsat and LISS-III) for the years 2000, 2006, 2011, 2016 and 2020. ArcGIS, QGIS, ERDAS were used for preprocessing, data preparation and land use land cover classification while SLEUTH 3.0 beta was used to run model. The model calibration was checked using the test mode test mode that examines how the model responds to the input data. For suitable coefficients calibration was carried out in three phases. The finest parameters for prediction were used to predict urban growth for the year 2031 with 1000 Monte Carlo iterations. To validate the model, Kappa statistics were used. a simulation was run for the year 2020, and the results were compared to the real land use and cover map for that year. The accuracy of the modeled urban growth was 88.06 percent.

Liu et al., (2019) used SLEUTH model to simulate different urban growth scenarios to Hefei, East China. Four Landsat images (2000, 2005, 2010 and 2015), land use planning map (2006-2020), SRTM DEM, traffic network vector data were used to conduct the research. Water bodies were added to the excluded layer. The model was calibrated using sequential and

automated processes by using forced Monte Carlo iterative method with three phase calibration and 100 iterations. The model produced eight least-square regressions. The 2015 layer was used as seed layer to simulate urban growth for the year 2040. The study regarded future urban simulation important and useful for formulating development plans.

Gonçalves et al., (2019) simulated urban growth of Praia City, Cabo Verde using SLEUTH model. The research used Landsat data for land use land cover classification. ASTER GDEM was used to generate slope and hill shade. The input data needed by the SLEUTH model was created using topographic maps as well as other digital sources. The behavior of SLEUTH model is controlled by five factors. A total of 12 least square regression metrices were produced to calibrate the model. The research was validated by equating urban pixels in the actual land use land cover layer to the number of pixels in the simulated layer.

Chapter 2

Material and Methods

2.1 Study Area

District Peshawar, situated in the Khyber Pakhtunkhwa province of Pakistan, serves as the provincial capital and stands as the most developed district within the region. The location of the site is situated around 160 kilometers to the west of Islamabad, the capital city of Pakistan. The district of Peshawar maintains considerable historical importance due to its role as the focal point for industrial, commercial, historical, and political activities within the province of Khyber Pakhtunkhwa.. District Peshawar is located in close proximity to the border shared by Pakistan and Afghanistan. It shares boundaries with district Khyber and district Mohmand in the West and Northwest of Peshawar. Nowshera and Charsadda are situated to its Northeast and North respectively while Kohat is situated in its South. Its altitude is about 360 meters above mean sea level. It stretches from 33º44´ to 34º15´ North latitudes and 71º22´ to 71º45´ East longitudes. The total area of district Peshawar is 1257 square kilometers, which accounts for roughly 1.69% of the province of Khyber Pakhtunkhwa's overall area. (Adnan & Iqbal 2014). It has only one tehsil and further divided in 4 towns. There are total of 279 villages (Revenue Villages) among which 236 are classified as rural, 15 are categorized as urban, and 28 are designated as partly urban. out of which 236 are rural, 15 are urban and 28 are partly urban. (Figure 2.1).

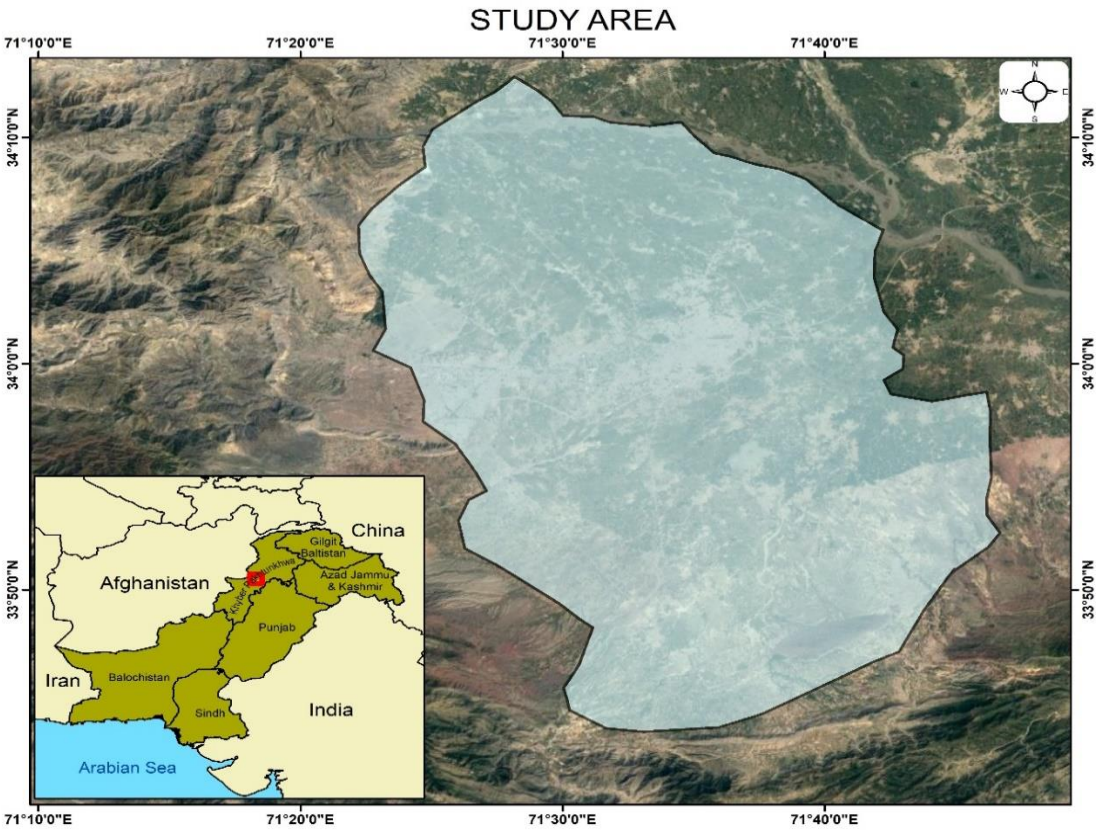

Figure 2. 1. Map of the study area, district Peshawar.

2.2 Demography

District Peshawar is the 8th largest district of Pakistan in terms of population, and it is the fastest growing district as well with an average growth rate of 3.99 from 1998 (2 million) to 2017 (4.26 million) with 46.16 percent of the population residing in urban areas. 99% of the population is Muslim while minority populations include Hindus, Sikhs, and Christians. District Peshawar is an economic hub of Khyber Pakhtunkhwa, therefore people from rural areas migrate to district Peshawar to avail themselves of better employment opportunities.

The population trend observed in district Peshawar throughout several decades is shown in table 2.1. Between the years 1972 and 1981, there was a population growth rate of 3.64%. Subsequently, from 1981 to 1998, this growth rate slightly increased to 3.7%. Furthermore, during the period spanning from 1998 to 2017, the growth rate further escalated to 4%. 46% of the population of district Peshawar lives in urban areas, while 54% lives in rural areas, according to the 2017 Census. (Figure 2.2)

2.3 Climate

Peshawar exhibits semiarid climate, with a scorching summer from May to September and a mild winter from November to March. The highest temperature during the summer temperatures average 25 degrees Celsius and peak at 42 degrees Celsius. The average winter low is 2 degrees Celsius, and the average high is 18 degrees Celsius. The best season starts in March when spring flowers bloom. The monsoon rains and western disturbances increase humidity (UPU, 2014). Peshawar District receives precipitation both in the winter and summer. According to the observed pattern, the mean annual precipitation is estimated to be approximately 420mm. The surface wind speed fluctuates between 2 and 6 knots.

2.4 Geology

River, stream, and flood plain deposits from the Pleistocene age group underlie the district of Peshawar. About 80% of the district's land is made up of these deposits. Riverine, Murree formation, Quaternary Alluvium, and Samana Suk are the main formations. In addition to these, various rock types make up 20% of the land area. The outcrops to the west, south, and southwest of the district frequently expose these rocks. Tertiary rocks, Jurassic rocks, Paleozoic rocks, and Precambrian rocks are a few of them. The district is therefore appropriate for urbanization in terms of geology, assuming that other conditions allow (Figure 2. 3).

2.5 Hydrology

The Kabul River, Bara River, and mountainside streams are the sources of surface water in the district. The river of Kabul, which enters the northern boundary of the Peshawar district and flows south-east, is the principal supply of water for agricultural purposes. It further splits into the Adezai and Naghuman Rivers. The Charsadda district's southern border is formed by the Adezai River. The Naghuman River divides into two separate rivers, which later reunites in the eastern portion of district Peshawar, ultimately converging with the Kabul River in the eastern vicinity of Nowshera region. The district's western and southern regions are home to several ephemeral streams that eventually run into the Kabul River on the district's eastern side. Five canals in the Peshawar District use water from the Kabul River to irrigate the agricultural land. The Kabul River Canal, Warsak Gravity Lift Canal, Hazar Khani Canal, and Joe Sheikh Canal are among them.

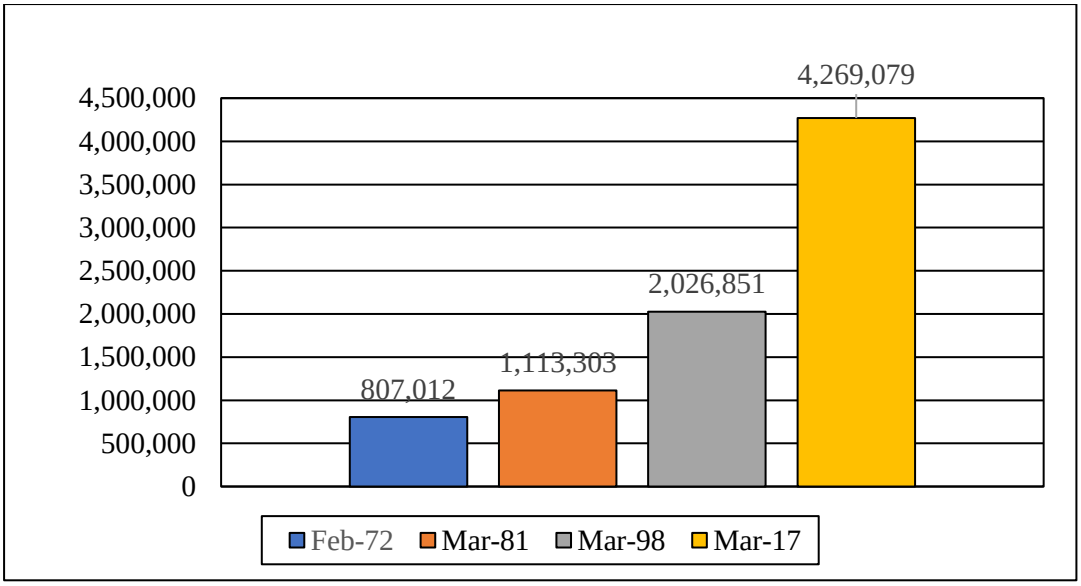

Figure 2. 2. Population of district Peshawar over the years.

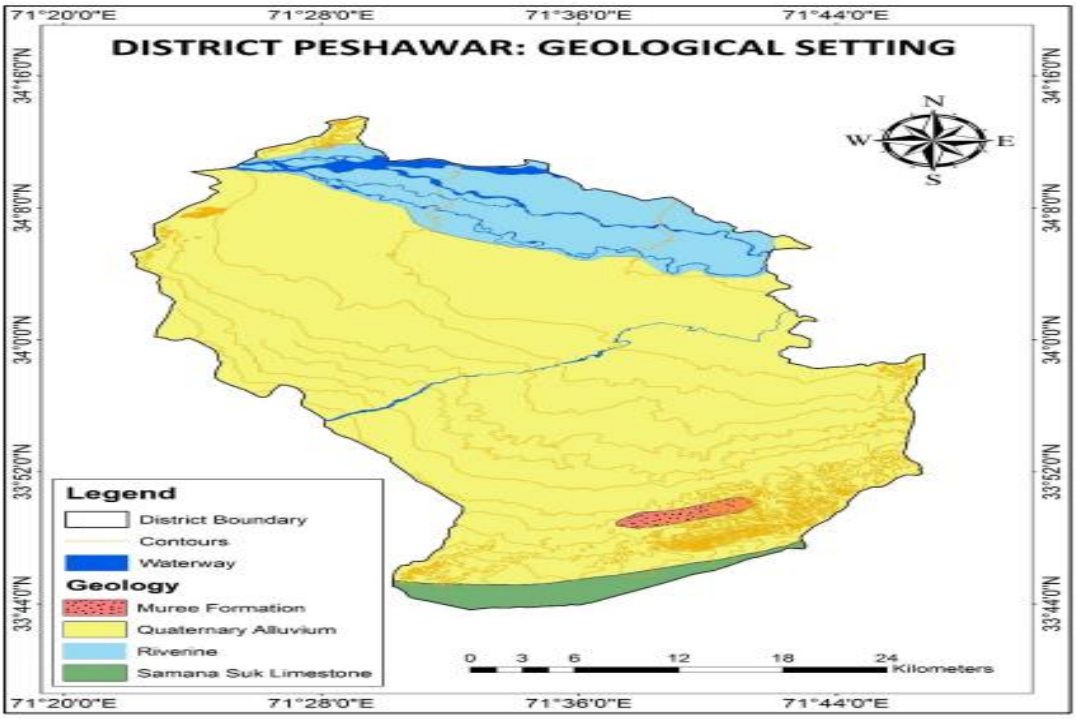

Figure 2. 3. Geological map of district Peshawar. (Map Source: Urban Policy Unit Government of Khyber Pakhtunkhwa).

2.6 Data Sources

Satellites Imageries from Landsat 5 (TM) imageries for the 1990, 2000 and 2010 and Landsat 8 30 meters spatial resolution (OLI) for the year 2020 were taken from United States Geological Survey (USGS). Digital Elevation Model (DEM) was downloaded from Alaska Satellite Facility with spatial resolution of 12.5 meters. Road data was downloaded from Open Street Map (OSM) and Toposheets were acquired from Survey of Pakistan (Table 2.1, Table 2.2)

2.7 Preprocessing

The different bands of Landsat imageries were stacked using Composite Bands tool (Data Management Tools) which combines numerous bands into a single raster dataset. The image contained a larger area than the required, so the study area was extracted from the imagery using the district Peshawar boundary.

2.8.1 Image Classification

Image classification refers to the computational task of categorizing pixels into a limited set of distinct classes or data categories, utilizing the values of their respective pixels as the basis for classification. The classification of remotely sensed data is employed to derive information from a multiband raster and allocate appropriate categories based on shared characteristics, in order to distinguish various objects within the image.

The Supervised Classification method was used to classify imageries of 1990, 2000, 2010 and 2020. The Maximum Likelihood technique assigns each cell in the input to a defined class and this technique was used for land use land cover classification. The imageries were

classified into four classes that are Urban, Vegetation, Water bodies and Open Land. More than two hundred training samples per class were taken for the image classification.

2.8.2 Accuracy Assessment

Any categorization project needs accuracy assessment. The classified image is compared to a credible or ground truth data source. It requires time and money to acquire ground truth through fieldwork. Ground truth data can be generated using high-resolution imagery, categorized imagery, or GIS data layers. The most common method for assessing classified image accuracy is generating random points from ground truth data and comparing them to classified data in a confusion matrix.

The Equalized Stratified Random sampling technique was used to generate a total of 480 sampling points (120 per class). The overall accuracy of the classified layers is 88%, 92%, 91%, and 90% for the years 1990, 2000, 2010 and 2021 respectively. The Kappa coefficient was calculated to be 0.84, 0.90, 0.88 and 0.97 for the years 1990, 2000, 2010 and 2021 respectively. If the value of Kappa Coefficient is more than 0.81 is considered to be near perfect agreement.

2.9 Land Use Land Cover and Urban Layer

Land use land cover layer for the years 1990, 1995, 2000, 2005, 2010 and 2021 were digitized previously (Figure 2. 4). The Urban layer was then extracted using the reclassify tool. The extent of the urban layer 2021 is used as seed layer to initiate SLEUTH model (Figure 2. 4, Figure 2. 5).

Table 2.1. Dataset set.

Data	**Source**
Satellite Imagery	Landsat 5 & 8 (USGS)
Digital Elevation Model (DEM)	Alos Palsar (Alaska Satellite Factory)
Transportation Network	Survey of Pakistan and Open Street Map (OSM)
Excluded Layer	On Screen Digitization
Land Use Plan Map	Urban Policy Unit Khyber Pakhtunkhwa

Table 2.2 Data requirement and preparation.

Layers Preparation in ArcGIS				**IDRISI**
Parameters	Number of Layers Required	Temporal Layers	Conversion of all layers to 8bit gif	Reclassification land use layers
Slope Layer	One	Any		
Hill Shade	One	Any		
Urban Layer	at least four layers	1990, 1995, 2000, 2005, 2010, 2021		
Transportation Layer	at least two layers	1995, 2022		
Excluded Layer	one	Latest		
Land Use	Two Layers	1990, 2021		
Land use plan	One	2017		

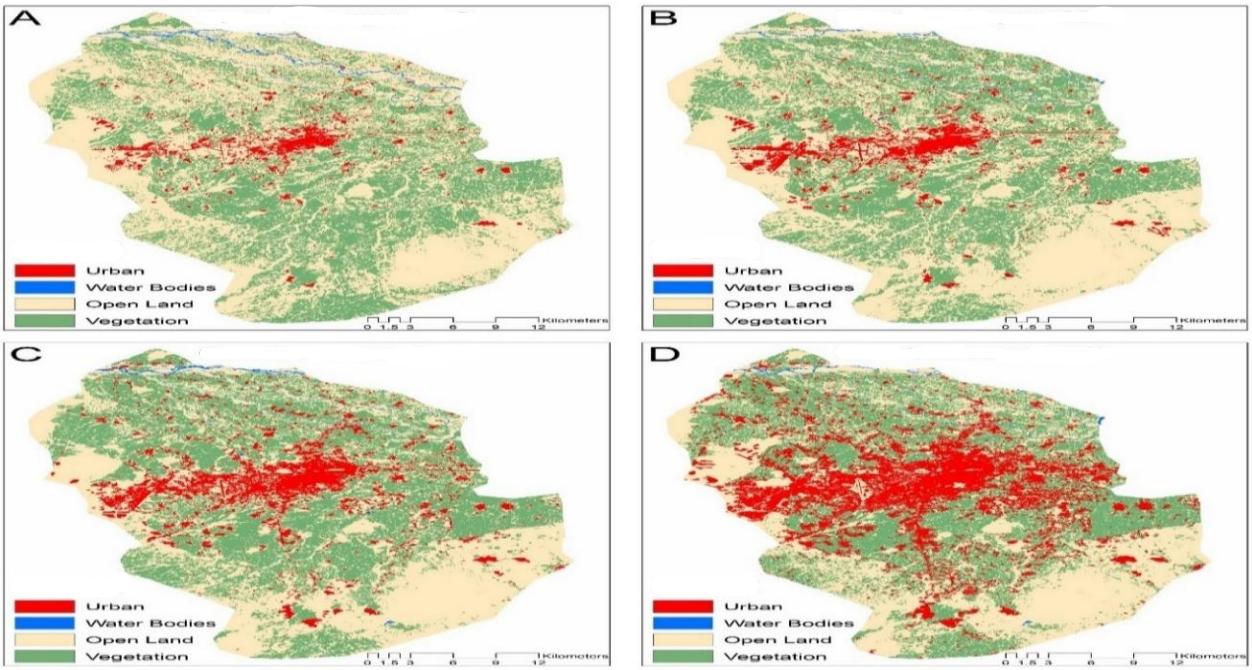

Figure 2. 4. Land use layers of district Peshawar for the year (A) 1990, (B) 2000, (C) 2010 and (D)2021 respectively

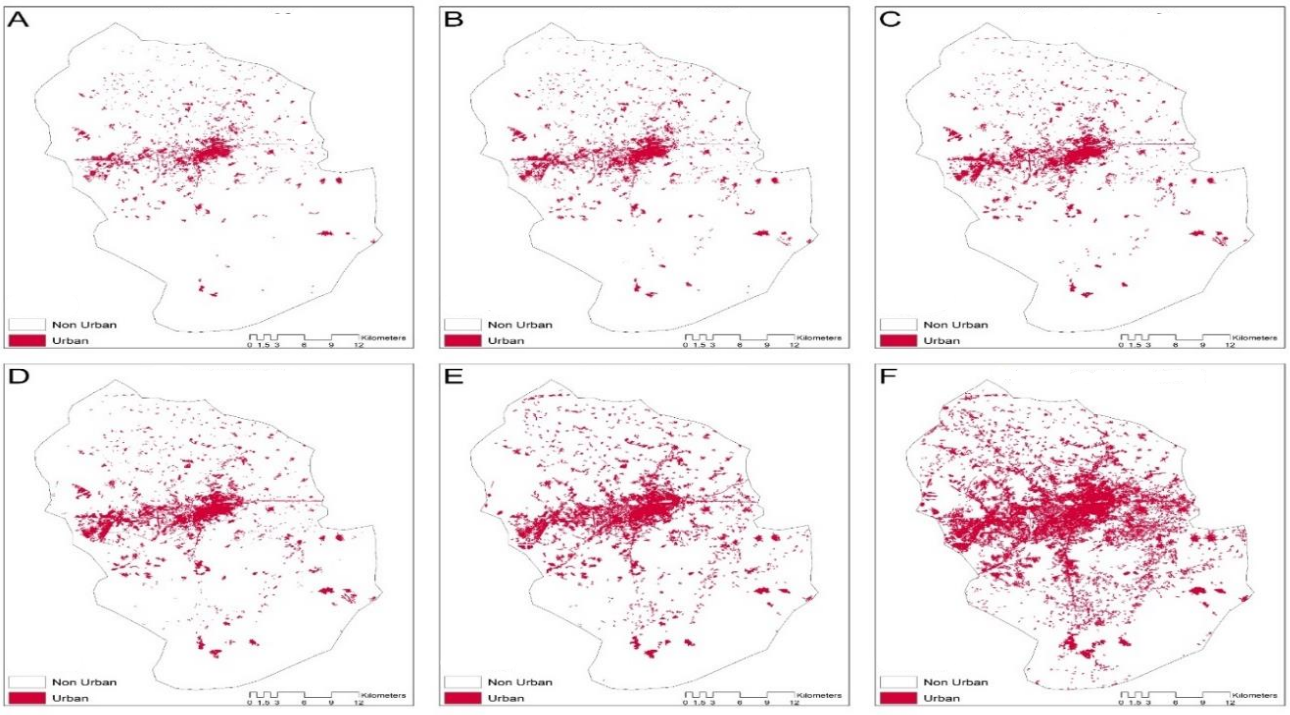

Figure 2. 5 Urban layers of district Peshawar for the year (A)1990, (B)1995, (C)2000, (D) 2005, (E)2010 and (F)2021.

2.10 Transportation Network

Urban sprawl is strongly influenced by the transportation network because cities tend to grow alongside it. However, the degree to which this influence is dependent on the type, status, and relative accessibility of the transportation network—whether it is high, medium, low, or non-existent. One type of road may have a varied impact on two distinct transportation networks. Urban sprawl will be more influenced by a well-developed road than by an un-metaled one. For SLEUTH to function, at least two separate transit tiers are needed.

Topo sheets 38 N/8, 38 N/12, 38 O/5, 38 O/9, 38 O/10, and 38 O/13, which comprised the roads in the research region, were used to generate the transportation network layer for the year 1995. The coordinates mentioned on the topo sheets were used for geo-referencing. After that, digitization of the roadways was carried out.

The latest Transportation Network was obtained from OSM Shape files of the Transportation network were downloaded. After that, the layer was clipped to the research area's boundary (Figure 2. 6).

2.11 Excluded Layer

The excluded layer has an important role to urban growth as it sets constraint and resistant factors on urban growth. In SLEUTH, a certain region can be prohibited from being urbanized by Excluded layer. This is the layer which includes all the possible resistant factors of urbanization. Every city has some areas reserved or prohibited to be urbanized by excluded layer, this layer slows down the urbanization process and can be a factor to alter urban growth trend and pattern. The excluded layer was digitized from Google Earth Pro. Green belts, waterways, parks, playgrounds, graveyards were included in the excluded Layer. (Figure 2.7)

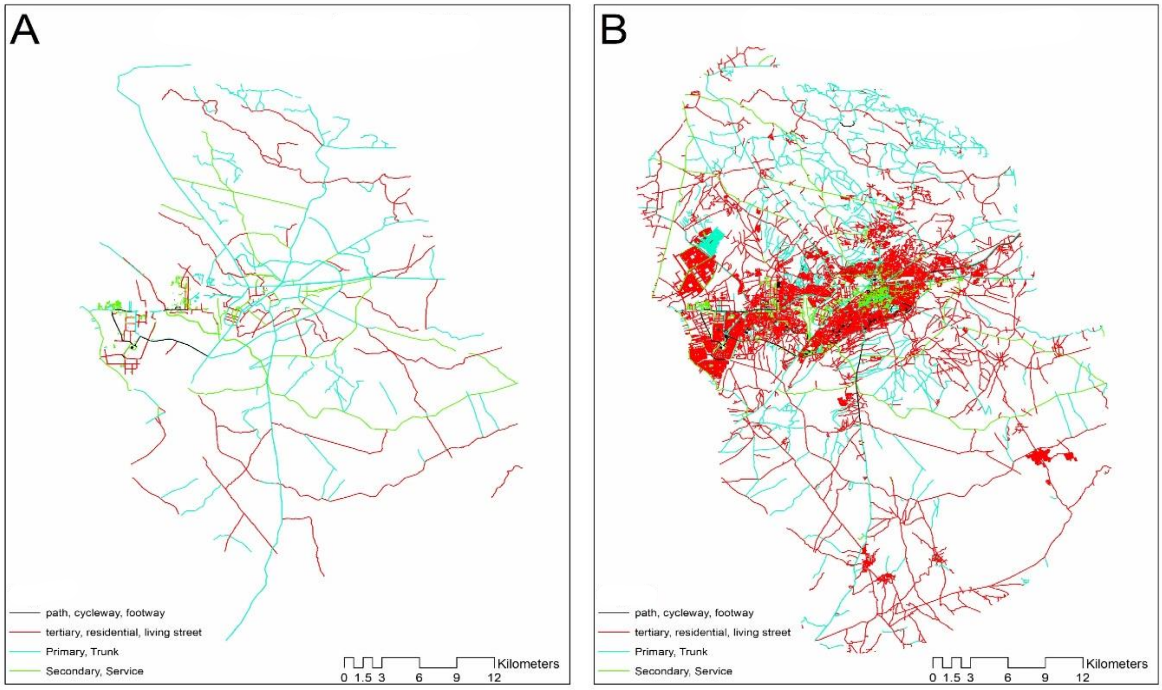

Figure 2. 6. Road layers for the years (A) 1995 and (B)2021 respectively.

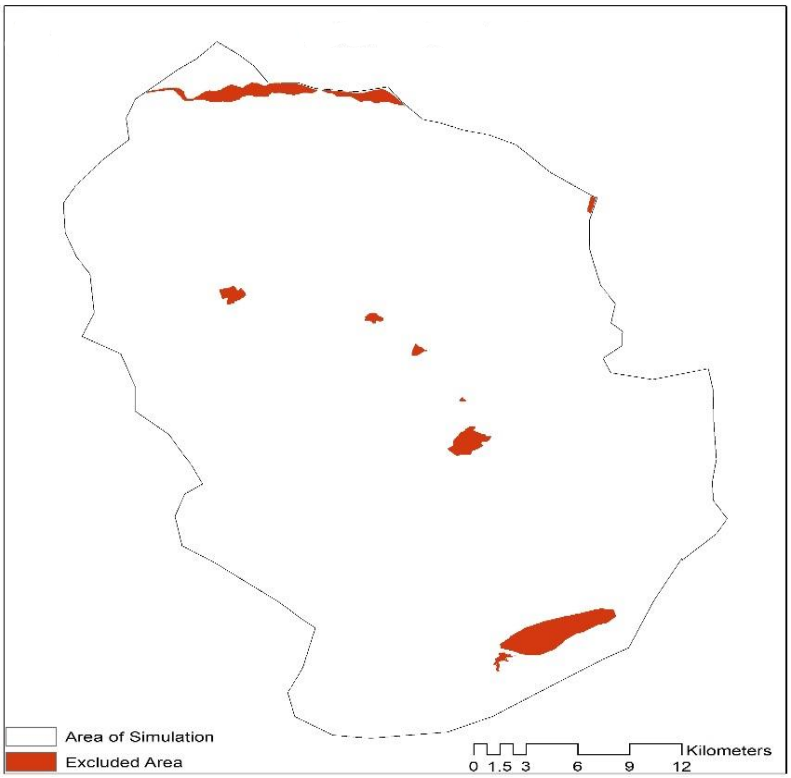

Figure 2. 7. Excluded layer.

2.12 Hill shade Layer

In order to enhance spatial visualization rather than engaging in the simulation, an additional layer known as Hill shade is incorporated into the output image as a backdrop. Hill shade is typically derived from the digital elevation model (DEM). The hill shade layer is employed to add visual impact to the generated layer and animation file. (Figure 2. 8)

2.13 Slope

Topography, which is generally characterized by slope, is one of the fundamental components of urban development. Flat and broad regions are presumably more suitable for urban sprawl (Gilliss et al., 2003). It is defined as 21% for plains and 35% for hilly or mountainous regions. Peshawar is a plain area, so a constraint after 21% slope was applied which prevents urbanization in areas where the slope exceeds 21%. (Figure 2. 9)

2.14 Land Use Plan

The Land use plan of District Peshawar was georeferenced using the coordinates mentioned on the map. The major zones were then digitized which contained 8 major zones named Urban Area (2039), Rural Zones, Mining Zone, Peshawar Model Town, Orchard Zone, Model Environmental Villages, Green Belt and Range Land. The area taken by difference classes is shown in (Figure 2. 10, Table 2. 3.).

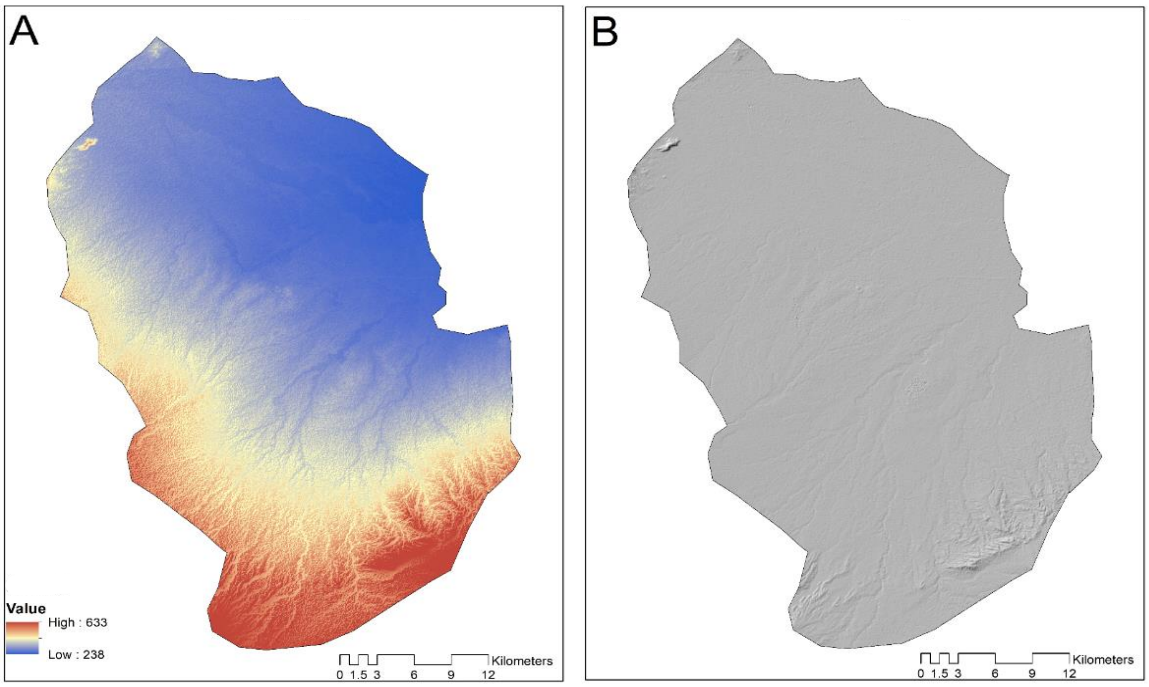

Figure 2. 8. (A) Digital elevation model and (B) hill shade layers respectively.

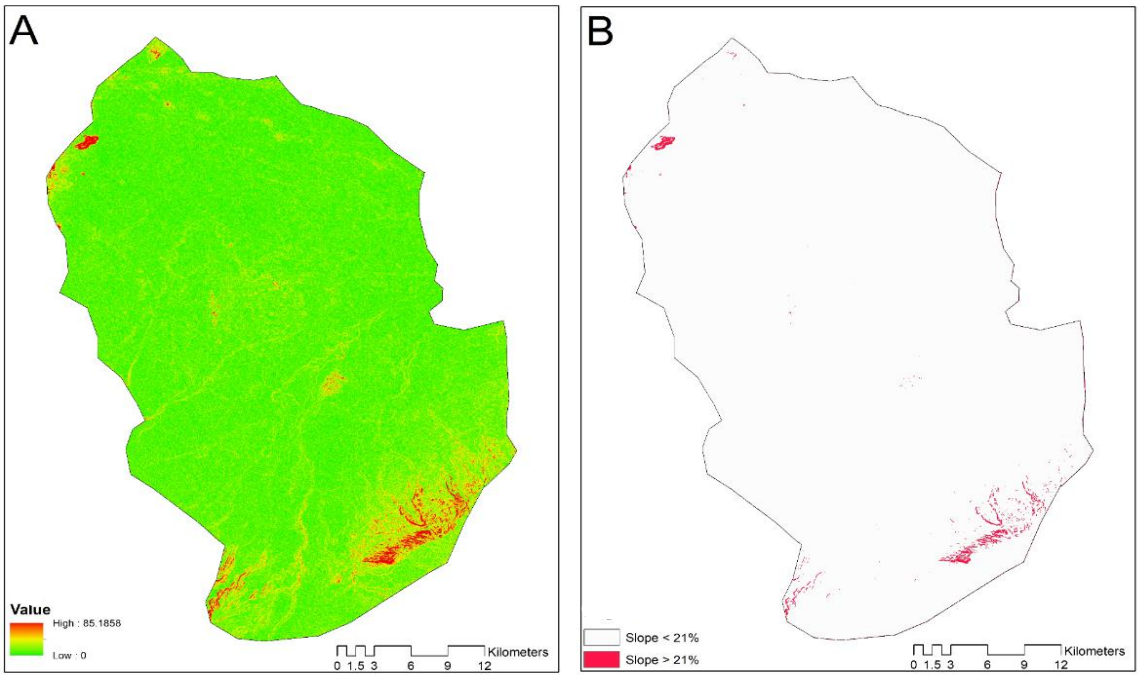

Figure 2. 9. Slope layers.

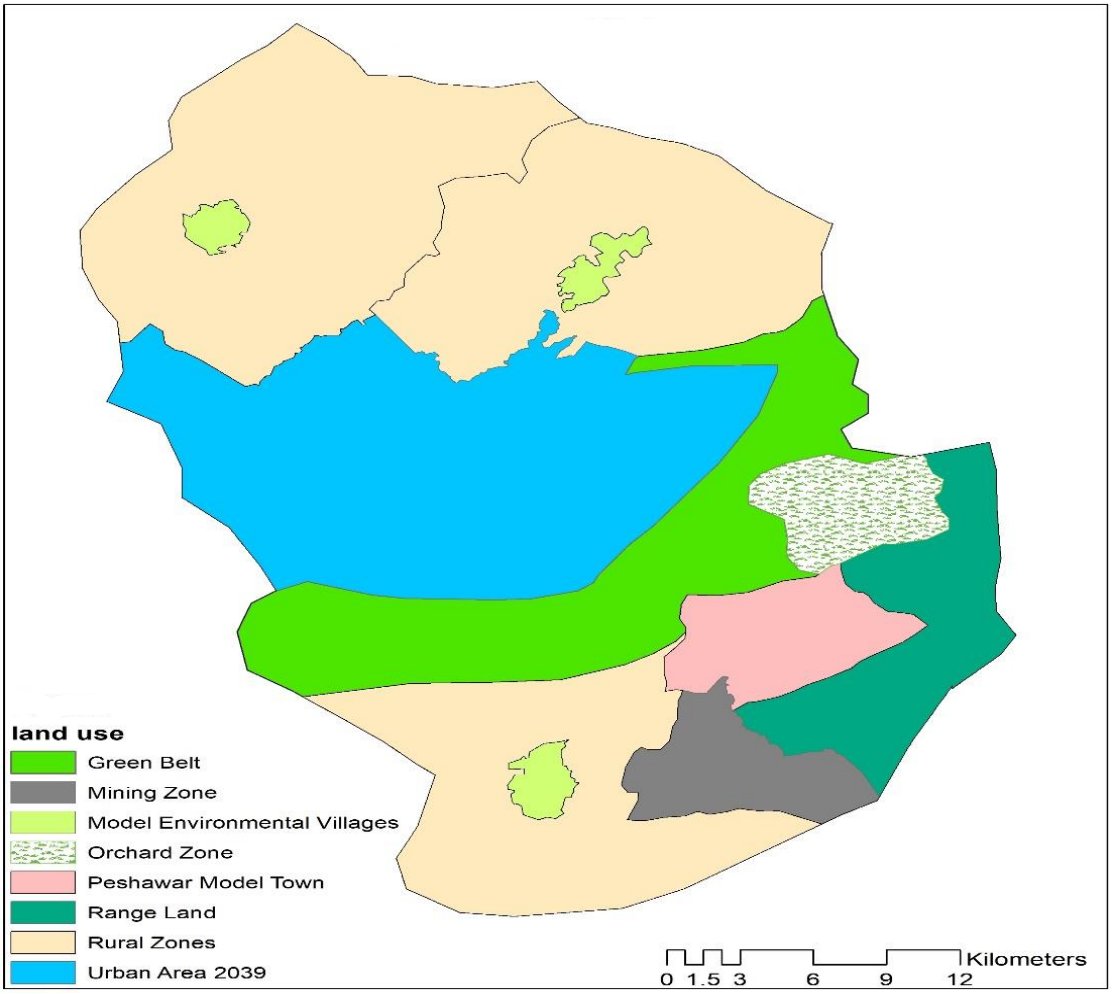

Figure 2. 10. Land use map of district Peshawar.

Table 2. 3. Zones and respective area in land use plan map of district Peshawar.

Sr No	Zone	Area (sq km)
1	Urban Area (2039)	300.0
2	Range Land	88.6
3	Model Environmental Villages	22.6
4	Rural Zones	563.3
5	Orchard Zone	39.4
6	Peshawar Model Town	52.3
7	Mining Zone	47.6
8	Green Belt	171.2
9	**Total Area**	**1284.9**

2.15 Data preparation for SLEUTH Model

The Slope, Land use, Excluded, Urban, Transportation Network and Hill shade layers prepared earlier are used as input in the CA Sleuth model. The SLEUTH model takes its input data in a rectangular preset only, so all the input data were converted accordingly with same dimensions. The projection of all the layers was UTM WGS 84 Zone 42. The prepared layers were converted to raster format as the model requires the data to be pixelated, the data were confined to 8 bits unsigned integer as it must be converted to gif format which is the input data format of SLEUTH model. All the layers were renamed according to the preset defined by the model for input data (Table 2.).

2.16 CA SLEUTH MODEL

CA Sleuth Model requires Linux 9.0 or a higher version and in case of windows-based operating system Cygwin needs to be installed. The model runs in three specified modes that is test, calibration and prediction. A scenario file is used to run the three mentioned modes. Coefficient values modification along with output name and directories can be changed using these scenario files. The model package contains these scenarios files. The Model uses five Coefficients, four growth rules and self- modification during its run to simulate urban growth.

2.16.1 Growth Coefficients

The dispersion, breed, spread, slope, and road gravity are the five model coefficients that influence how the growth rules of this model function. These coefficients are changed during the model's calibration phase to get the best fit. The growth rules' application is determined by their values, which range from 0 to 100. (

Table 2.)

Table 2. 4 The naming format for the Sleuth model.

Data Name
Peshawar.urban.1990.gif
Peshawar.urban.1995.gif
Peshawar.urban.2000.gif
Peshawar.urban.2005.gif
Peshawar.urban.2010.gif
Peshawar.urban.2021.gif
Peshawar.roads.1995.gif
Peshawar.roads.2005.gif
peshawar.roads2022.gif
Peshawar.excluded.gif
Peshawar.slope.gif
Peshawar.hillshade.gif

Table 2. 5. Growth rule along with model coefficients.

Sr No	**Growth Rule**	**Model coefficients**	**Rule Description**
1	Spontaneous growth	Dispersion, Slope	Random conversion of non-urban cell to urban cell
2	New spreading center	Breed, Slope	urban cell from spontaneous growth becomes new spreading centers
3	Edge growth	Spread, Slope	spreading center- edge expansion
4	Road influenced growth	Breed, Road gravity, Dispersion, Slope	The attraction of transportation network to urbanization

a. Dispersion Coefficient: determines how often a pixel will be picked at random for urbanization. It controls how dispersed growth is, for both individual grid cells and the expansion of new settlements through transportation systems.
b. Breed Coefficient: Assesses the likelihood that a newly generated separate spontaneous growth pixel (settlement) will start its own growth cycle and will develop into a new spreading center.
c. Spread Coefficient: This function determines the likelihood that any pixel which segment of spreading center. (a cluster of urban pixels with a count of more than two in a 3x3 neighborhood) will produce a new urban pixel in its vicinity. It controls the degree of outward expansion.
d. Slope Coefficient: Determines whether settlement will extend in higher slopes. When examining a pixel for growing urbanization, the slope of the pixel is taken into account.. Higher slope coefficients reduce the likelihood of urbanization on steeper slopes.
e. Road Gravity: Promotes the growth of new settlements in close proximity to or alongside the transportation network. Based on a fraction of the image dimensions, the maximum search radius for a road from a pixel selected for the trip is computed.

2.16.2 Growth Rules

Spontaneous Growth, New Spreading Center Growth, Edge Growth, Road Influenced Growth, and Self-modification are predetermined growth rules used in the development of the SLEUTH urban model. They are continuously deployed to steer the urban simulation to a particular morphology, and each rule's execution updates the entire cellular space. Growth rules replicate the change in land cover caused by urbanisation are used to estimate future urban extent.

a. Spontaneous Growth Rule: Determines whether a random pixel within the urban matrix will be urbanized. On the matrix, each unurbanized cell "has a definite (low) probability of becoming urban in a time step," The dispersion and slope coefficients are used in this rule.
b. New Spreading Centers: Decides whether a freshly urbanized pixel (as determined by the Spontaneous Growth Rule) will become a new urban center and, if so, whether the adjacent area will also be urbanized. This development rule accounts for the Breed and Slope coefficients.
c. Edge Growth Rule: Creates new growth adjacent to extant urban pixels. The degree of Edge Growth is determined by the Spread and Slope factors.
d. Road Influenced Growth: As indicated by all other rules, this rule takes into account both the current transportation system and urbanization. Selected newly urbanized cells are examined for the presence of roads in their neighborhoods. If a road is found within the maximum defined radius of a selected cell, an urban cell with temporary status is installed at the location on the road nearest to the selected cell. The total number of steps taken in a subsequent walk at random along the road (or roads heading to the initial road) is determined by the dispersion coefficient parameter. Temporary urbanized cell's final location is then taken into consideration as a new urban spreading nucleus. If a nearby cell (on the road) to the temporarily urbanized cell is accessible for urbanization, it will take place (randomly picked among possible candidates). This urbanization will occur (based on a randomly chosen candidate) if two cells that are close to the newly urbanized cell are also eligible. The Dispersion, Breed, Slope, and Road Gravity Coefficients are used to derive this rule.

2.16.3 Self-Modification

Urban growth is not linear; hence the SLEUTH model's coefficients cannot be altered. By linking variables, an extra set of rules allowing the model to dynamically adapt itself is developed to simulate the urbanization accurately. The control coefficients have the ability to self-modify, allowing the CA to adjust to the conditions it created. It is used when growth is either exceptionally high or extremely low in absolute terms. Each coefficient will get multiplied by a number greater than one in order to increase the expansion to prosperity (Boom) during urban explosion, which usually happens during the early development cycles when there are numerous cells available to be urbanized. As soon as urban density stabilizes and sprawl speed declines consistently below the minimum threshold (CRITICAL LOW), all variables must be multiplied by a number less than one. Self-modification is crucial for estimating the usual S-curve growth since without it, linear or exponential growth errors would occur.

2.16.4 Test Mode

As clear from its name, this mode checks the data readiness for the calibration and prediction. It checks that all the layers overlay completely. This phase verifies and validates the data according to the required specifications. For test mode run, the data were set according to the model input requirements. The scenario file was named as *scenario.Peshawar_test*. The command used for test mode run *is../grow test scenario_peshawar_test*. 100 Monte Carlo iterations were predetermined. The model was run with default parameters with no modifications as the purpose of this phase is only to validate the data. The test mode took around 10 minutes to run successfully without returning any errors and exceptions. The successful application of test mode indicates the data and model to be ready for next phase. This phase gives urban extents based on the input data along with average (avg), coefficient

(coeff) and control_stats file which contains values that will be utilized in coarse calibration phase.

2.16.5 Calibration Mode

The calibration procedure generates initializing coefficient values that most closely approximate historical growth in the study area to ensure realism of the model. The model trains on historic data in this phase to match the historic data and prepares the model for future simulation. The goal of the model calibration phase is to compare historical urban extent data with the coefficients of dispersion, breed and spread, slope resistance, and road gravity to attain the best-fit values for the five mentioned growth control parameters. This phase is the model's most crucial stage. The calibration module is carried out in three different phases coarse, fine, and final. Each stage is implemented on a dataset with a different resolution which is decided by the researcher. For this research, coarse calibration was processed at a resolution of 90 meters followed by fine calibration at a 60 meters resolution, and final calibration at 30 meters resolution. The resolution of the data and the size of the area to be simulated have a direct correlation with the computational requirements of the model. The process that takes the longest to complete is calibration.

2.16.6 Coarse Phase Calibration

The process of coarse phase calibration has been executed by utilizing a dataset with coarse resolution. The spatial resolution of the data was resampled to 120 meters. 100 was maintained as the initial value of Monte Carlo iterations. The scenario file for the coarse calibration was *scenario_peshawar90_calibrate*. The names of input data were set according to the model requirements. The start, step and stop values for dispersion, breed, spread, slope,

and road were set to 0,25 and 100 respectively. The Monte Carlo iterations were set to 100. The reason behind using high values is to let the model run in its high efficiency and return the best fit values. The flags of the average file, coefficient file, average file and log file were set to YES as these files will give the most suitable values to be utilized during the fine calibration step. The coefficient values are calculated for each Monte Carlo iteration. The average file's output is measured values of simulated data that have been averaged throughout all runs of Monte Carlo iterations. The standard deviation of averaged data may be found in the standard deviation file's output. The Self Modification Restrictions are next established and used as input. (Table 2.6).

For the model validation, the start date was set to 1990 and the stop date was set to 2005. The start date of calibration was set to 1990 and the stop date was set to the last input urban layer which was 2021 for the prediction of 2030.

2.16.7 Selecting Best Fit Coefficint values from Coarse Calibration

After the successful run, the coarse calibration gives coefficient file which contains the best fit values that can be used in fine calibration mode. The Leesalee metric column present in the control stats file is a shape index used to quantify the level of geographical agreement between the predicted growth patterns and current urban bounds of control years.. It is sorted in descending order. The top three values from the mentioned column are picked . The top values of dispersion, breed, lope, spread and road corresponding to the Leesalee metric column are the most suitable parameters to be used in the fine calibration run. (Table 2.7.)

Table 2.6. Start, step and stop values of coarse calibration.

CALIBRATION_DIFFUSION_START	0
CALIBRATION_DIFFUSION_STEP	25
CALIBRATION_DIFFUSION_STOP	100
CALIBRATION_BREED_START	0
CALIBRATION_BREED_STEP	25
CALIBRATION_BREED_STOP	100
CALIBRATION_SPREAD_START	0
CALIBRATION_SPREAD_STEP	25
CALIBRATION_SPREAD_STOP	100
CALIBRATION_SLOPE_START	0
CALIBRATION_SLOPE_STEP	25
CALIBRATION_SLOPE_STOP	100
CALIBRATION_ROAD_START	0
CALIBRATION_ROAD_STEP	25
CALIBRATION_ROAD_STOP	100

Table 2.7. The best fit values for fine calibration.

Leesalee	**Diffusion**	**Breed**	**Spread**	**Slope**	**Road Gravity**
0.45381	1	1	25	1	100
0.45381	1	1	25	25	100
0.45381	1	1	25	50	100
0.45381	1	1	25	75	100
0.45381	1	1	25	100	100
Fine calibration parameters (start, step, stop)	0, 5, 20	0, 5, 20	20, 5, 40	0, 20, 100	0, 75, 100

2.16.8 Fine Calibration

Fine calibration was performed on 60 meters spatial resolution data. The Monte Carlo iterations were set to 50. The results obtained from the coarse calibration process were utilized for the fine calibration phase. The values were put into the scenario file for the purpose of fine calibration named *peshawar60_calibrate.* A new output directory was made as values from every calibration phase are important and should not be overwritten. For validating the model, the year 1990-2005 were selected for start and stop dates respectively. The start and end dates for future simulation calibration were set to 1990 and 2021 respectively. The coefficient file, log file, average file and standard deviation file are required so the flags of these are set to YES in scenario file named as *scenario.peshawar60_calibrate*. The command used to run calibration mode is *../grow calibrate scenario.peshawar60_calibrate.* The fine calibration phase took 27 hours to complete. The start, step and stop values are given below. (Table 2.8)

2.16.9 Selecting Best Fit Coefficint values from Fine Calibration

After the successful run of fine calibration phase, the best fit and optimum values for the final calibration are derived from the control stats file of the final calibration phase. The Leesalee metric column of the control stats is sorted again as in the previous calibration phase. The top values of breed, dispersion, spread, slope and road gravity corresponding to the Leesalee metric column are extracted and assumed to be the best fit values and used in the final calibration phase (Table 2.9).

Table 2.8. The start, step and stop values for fine calibration.

CALIBRATION_DIFFUSION_START	0
CALIBRATION_DIFFUSION_STEP	5
CALIBRATION_DIFFUSION_STOP	20
CALIBRATION_BREED_START	0
CALIBRATION_BREED_STEP	5
CALIBRATION_BREED_STOP	20
CALIBRATION_SPREAD_START	20
CALIBRATION_SPREAD_STEP	5
CALIBRATION_SPREAD_STOP	40
CALIBRATION_SLOPE_START	0
CALIBRATION_SLOPE_STEP	20
CALIBRATION_SLOPE_STOP	100
CALIBRATION_ROAD_START	0
CALIBRATION_ROAD_STEP	75
CALIBRATION_ROAD_STOP	100

Table 2.9. The best fit values for final calibration.

Leesalee	**Dispersion**	**Breed**	**Spread**	**Slope**	**Road Gravity**
0.48541	1	1	20	1	1
0.48541	1	1	20	20	1
0.48541	1	1	20	40	1
0.48541	1	1	20	60	1
0.48541	1	5	20	80	75
Final calibration parameters (start, step, stop)	1, 1, 5	1, 1, 10	17, 1, 20	1 ,20, 100	1, 3, 75

2.16.10 Final Calibration

For the final calibration, the data were resampled to spatial resolution of 30 meters. All the layers were renamed as stipulated by the model. The scenario file of final calibration was created and named as *scenario.peshawar30_calibrate*. The values taken from the control stats file of fine calibration are used in the scenario file of final calibration. The Monte Carlo iterations were set to 50. The files required for the forecast phase were set to Yes. A new output directory was created for the final calibration results names as peshwawar30_calibrate. The command used to run final prediction phase is *../grow.exe calibrate scenario.peshawar30_calibrate.* The model iterated five thousand times with 50 Monte Carlo iterations.

2.16.11 Selecting Best Fit Coefficient values from Final Calibration

The calibration process ends with the final calibration. The start, step and stop are now narrowed to the best possible values. To select these values for the forcast phase, the control stats of the final calibration is set descending order and the highest five values of the Leesalee metric are selected and the corresponding values of dispersion, breed, slope, spread and road gravity are used as input in the forcast phase.

2.16.12 Forecast

The start, step and stop values from the final calibration are used in the forcast mode. The main objective of the coarse, fine and final calibration is to find a single start, step and stop values for the coefficients. The Monte Carlo interations are set to 100. Scenario file is created

for the forcast phase named as *scenario.peshawar30_forcast*. The coefficient values are modified. For model validation the start and end dates were set to 1990-2010 respectively. The start date of this phase for simulation of 2039 is set to 1990 with and end date of 2039. The flag of average file should be set to yes as the best fit values for prediction comes from average file. The command used to run the forcast phase is *../grow.exe calibrate scenario.peshawar30_forcast.* (Table 2.)

2.16.13 Selecting Best Fit Coefficint values from Forcast phase

When the forcast phase runs successfully it gives avergae file in its output folder which contains the best fit values for the prediction phase. The Leesalee metric column is not used in this phase because for the prediction phase narrowest values should be used and these values are present in the avergae file. The stop date of the avergae file is used as input in the scenario file of predcition phase. The values are in floating format and are rounded up to an integer so that it can be read by the model. 2021 is used as the stop year and the values for the coefficents are taken from the same year. The coefficients used for the prediction mode are mentioned in (Table 2. 1).

Table 2.10. The best fit values for forecast phase.

Leesalee	Dispersion	Breed	Spread	Slope	Road Gravity
0.49312	1	1	20	1	1
0.49312	1	1	20	1	1
0.49312	1	1	20	1	1
0.49312	1	1	20	1	1
0.49312	1	1	20	21	4
Forecast phase parameters (Start, step, stop)	1, 1, 1	1, 1, 1	20, 20, 20	1, 1, 1	1, 1, 1

Table 2. 11. The best fit values for prediction Phase.

Year	Diffusion	Spread	Breed	Slope Resistance	Road Gravity
1995	1.04	20.81	1.04	0.52	1.18
2000	1.09	21.87	1.09	0.46	1.44
2005	1.15	22.99	1.15	1	1.73
2010	1.21	24.16	1.21	1	2.22
2021	1.35	26.96	1.35	1	3.44

2.16.14 Urban Growth Prediction

To run the prediction mode, a scenario file named scenario.peshawar30_predict is created. The suitabe growth factors and forcast coefficients from the forcast phase are used in the prediction phase. The input directory flag is set to the 30 meter resolution data. All the flags in the scenario file are set to yes so that all the statistical are created along with the results. The Monte Carlo iterations are set to 300 for the final mode. 2001 and 2010 were set as the start and end dates for the model validation while the start date is set to 2021 and the end date is set to 2039 for future simulation. The coefficients derived from the forcast calibration are best fit values for prediction but they are in floating format so they are rounded up and converted to interger (as the model cannot read values in floating format) and entered in the prediciton phase. This phase gives the resultant layer of the future urbanization simulation. The command used to execute the prediciton is *../grow.exe predict scenario.peshawar30_predict.* After compppletion of this mode an average, coefficient, memoy, and log files are created which contains all the statictal information of the simulation. Along with these files the urban growth of the year 2010 and 2039 is also created in the prediction output file. (Table 2.1).

2.17 Markov Chain Model

The Markov chain is a succession of random values in which the probability of each random value is determined by the previous value of the corresponding number. A given parcel of land could change from one landuse category to another at any time. All multi-directional landuse changes between mutually incompatible landuse categories are expressed using matrices in Markovian analysis. The landuse distributions at the beginning (Mt) and end (Mt+1) of a discrete time period are used in the MC equation together with

Table 2.12. The best fit values for Sleuth model prediction phase.

Sr No	**Coefficient**	**Values**
1	Dispersion	1
2	Breed	1
3	Spread	27
4	Slope	1
5	Slope	1
6	Road Gravity	3

a transition matrix (MLc) that represents the changes in landuse that took place during that time. These proportional changes are transformed into probabilities of landuse change across the entire sample area, and transition matrices are generated under the assumption that the sample is an accurate representation of the entire area.

Both the MC and the CA are discrete dynamic models in terms of time and state. Markov has a lot of inherent flaws, one of which being a lack of spatial awareness. Although the transition probability may be correct on basis of categories, the geographical distribution of occurrences within each landuse category is unknown. To provide the model spatial character, CA component is added to the model. The Idrisi 15 (Clark Labs, 2006) integrates Markov and CA very well.

2.17.1 Data Preparation for CA Markov Chain Model

Land use land cover layers for the model were prepared in ArcGIS 10.5. The classified images were converted to TXT format using raster to ASCII tool. The TXT format. In Idrisi the images were imported and converted to RST file which is the raster file format of IDRISI. The raster files were reclassified in IDRISI because during the import the classified image transforms to continuous data and in order to successfully run the model the land use classes should be in ditinct classes so discrete number to every class was given and zero to the background and null values so that they are not activley involved in the prediciton.

2.17.2 Defining Spatial Objects

The spatial objects used in the model are already defined from the classification previosuly carried out where each spatial object might manifest in any of the following states. Water Body, Urban, Open land and Vegetaion.

2.17.3 Setting Transition Rules

The CA-MC model runs in two modules, Markov module and the CA-Markov module. Two land use layers of different years serves as input to the Markov tool of the CA-MC model. The difference between the two land use covers is entered and year to be predicted is provided to the tool. This tool gives transition areas and transition probabilty matrix which contains the probabilities of classes to change its current state to a new state over a certain period of time, the transition probabilty matrix is calculated based on the changes occurred in the previous land use layers. The matrix tool only gives these 2 statitical files, and it does not take any spatial aspect into account. The land use layer for the year 1990 and 2000 were used to give probability of change for the year 2010 for validation purpose, Land use layer for the year 1990 and 2021 were used to derive the probability of change for the year 2039. The sum of proabibilty of change of any given class in the Markov module is 1. The model was validated by simulating urban extent for the year 2010 with the probability of urban class to remain urban is 0.8039, The open land can change its state to urban with a probability of 0.3 while probabilty of vegetation class to change to urban land is 0.1 with lowest proabibity given to water class by the module to water bodies which is 0.0036, also indicates the accuracy of the model that water body will remain water body in the prediction (Table 2.1).

After valdating the CA-MC model, the same process was done to simultae the urban growth for the year 2039. Base land use layers for the simulation were 1990 and 2021. The markov module of the CA-MC model was run to give transition areas and transition robabilites matrix which contains information of the classes probabilty of change to other classes during the prediction phase. The probabilty of urban class to remain urban is 0.89. Probability of Open land and vegetation class to change into urban class is 0.25 and 0.1 respectively. The probabilty of water to change its state to urban class is 0.02 which is lowest of them all. (Table 2.).

Table 2.13 Probability of change to other land use classes for 2010 prediction.

Given Probability of Changing to (2010 Prediction)				
Class Name	**Urban**	**Water**	**Open Land**	**Vegetation**
Urban	0.8039	0.0005	0.1777	0.0179
Water	0.0036	0.2203	0.6764	0.0997
Open Land	0.3008	0.0035	0.4608	0.2349
Vegetation	0.1162	0.0006	0.2248	0.6584

Table 2.14 probability of change to other classes for 2039 prediction.

Given Probability of Changing to (2039 Prediction)				
Class Name	**Urban**	**Water**	**Open Land**	**Vegetation**
Urban	0.895	0.0004	0.0877	0.0172
Water	0.028	0.3931	0.5135	0.0655
Open Land	0.259	0.0031	0.5766	0.1615
Vegetation	0.1	0.0007	0.2083	0.691

2.17.4 Determinging CA-Markov Filters

CA filters are able to generate an accurate picture of the space weighting factor, which can be adjusted in accordance with the state of the cells adjacent to the one being analysed. As a matter of course, the definition of the neighbourhood will always be based on the conventional 5 by 5 contiguity filter. Users have the ability to define the neighbourhood in accordance with their own particular interests; nevertheless, in most cases, the standard definition of neighbourhood is utilised. The term "5*5 neighbourhood" refers to the fact that each cellular centre is surrounded by a matrix space that is made of 24 neighbour cells that have a substantial impact on the changes that occur in the cells. For this model, the 5*5 contuity neighbourhood was seleted as it returned least error as compared to 3*3 and 7*7 contuigty filters.

2.17.5 Determining Simulation’s starting point and CA number of iterations.

The base image for the 2010 prediction was classified image of 2000 and for the year 2039 the classified layer of 2021. Different number of iterations were tested to reduce the error between the classified land use layer of 2010 and simulated layer of 2010. 15 iterations with 1100 passes returned the least error where the simulated urban growth of 2010 was nearest to classified layer of 2010.

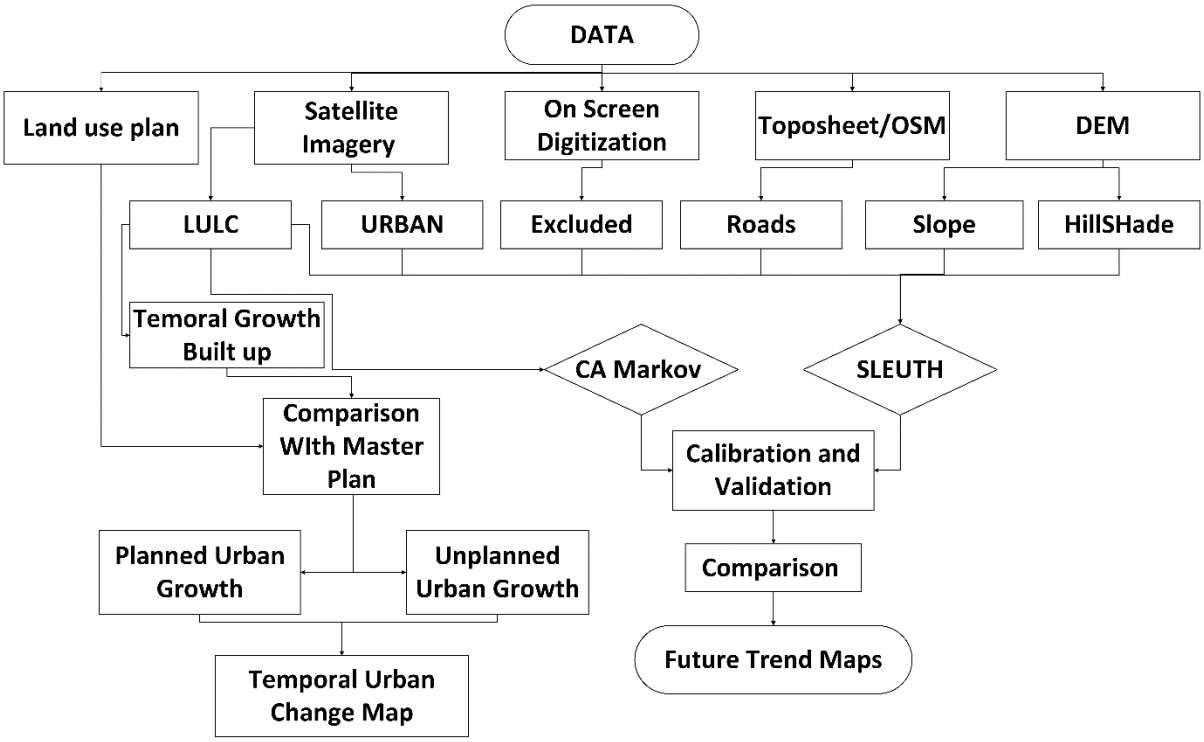

Figure 2. 11 Overall flow chart of the methodology implemented.

Chapter 3

RESULTS AND DISCUSSION

3.1 Temporal Analysis of Urban Growth

The land use land cover analysis was carried out using ESRI's ArcGIS pro 2.8 and ArcGIS 10.5 version. Landsat 5 and Landsat 8 imageries for the years 1990. 2000, 2010 and 2021 were downloaded from USGS Earth Explorer. Supervised classification was carried out. Maximum likelihood classification technique was adopted through which the research area was divided into four classes named as urban, water bodies, vegetation, and open land. The accuracy of the classification was assessed. The overall accuracy of the classified layer for the years 1990, 2000, 2010, and 2021 was 88%, 92%, 91% and 90% with Kappa coefficient of 0.84, 0.90, 0.88 and 0.97 respectively. From 1990 to 2021 the urban class increased from 50.50 sq km (3.93%) to 259.93 sq km (20.23%) where there was abrupt change of 124.91 sq km from 2010 to 2021. The Open Land decreased by 202.79 sq km taking it from 742.04 sq km (57.75%) to 539.25 sq km (41.97%) from 1990 to 2021. The urban area increased by 414% (almost 4 times) as compared to that of 1990. There was a significant decrease in open land from 2010 to 2021 Although the vegetative class remained almost the same from 1990 to 2021 but in actual 67.6 sq km of vegetative class was consumed by urban class and in turn open land was consumed by vegetative class due to which the class area remained almost the same. The urban class increased enormously from 1990 to 2021 while the open land decreased significantly. There were no major changes in the water class over the years. (Figure 3.1).

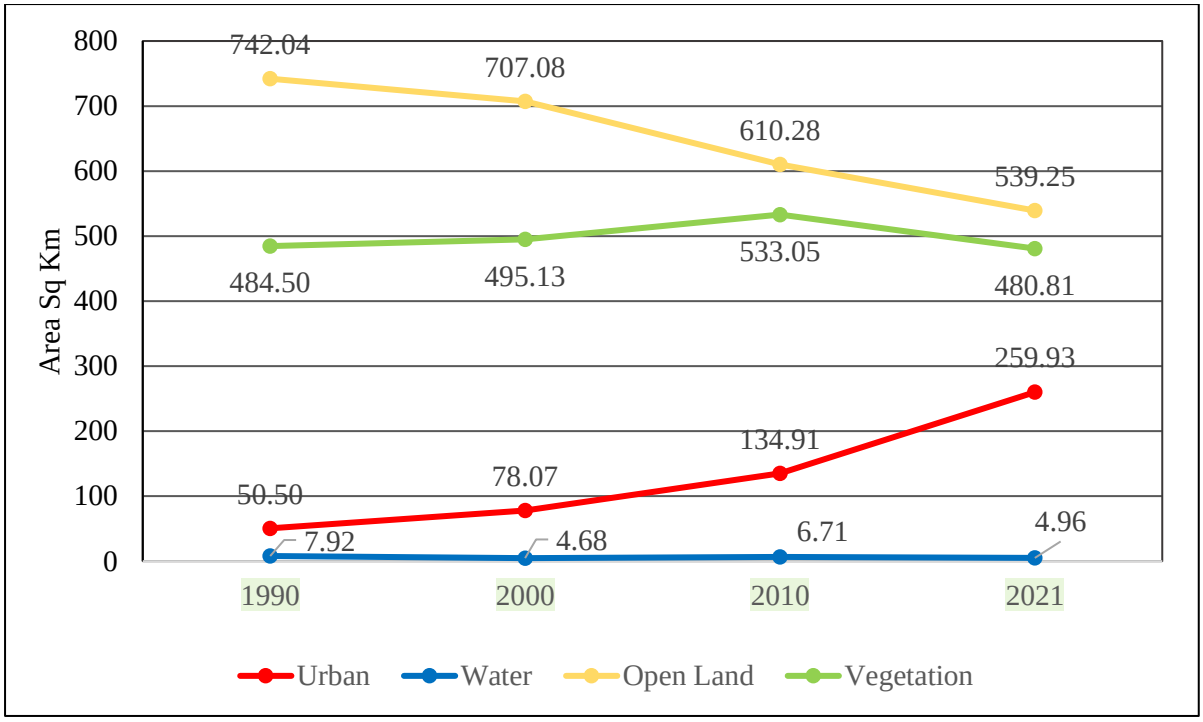

Figure 3. 1. Land use changes from 1990 to 2021.

3.2 Urban Growth Comparison with Land Use Plan of District Peshawar

The urban class of the classified layer for the year 2021 was extracted and settlements were compared with different zones of the land use plan of district Peshawar. The Urban Area (2039), Peshawar Model Town and Model and Model Environmental villages are zones which allow settlements while the zones that restrict settlements are Range Land, Rural, Orchard, Mining and Green Land zones restrict settlements. The Zones that are allowed for settlements have 146.67 sq km settlements out of 374.9 sq km while the zones that are supposed to be for other uses and restricts settlements have 113.67 sq km taken up by settlements out of 910 sq km allotted to these zones. (Table 3. 1, Figure 3. 2).

3.4 CA-MC Model Validation

The CA-MC model was validated by simulating urban growth for the year 2010 and then comparing it to the classified layer of 2010. To minimize the error and reduce the difference between the simulated layer and the classified layer, the model was run under different scenarios. In the first scenario the 3*3 contiguity filter was kept constant while the number of iterations varied from 1 to 17. The model returned the best results when the number of iterations were kept 15 but still there was huge error. In the second scenario the 5*5 contiguity filter was used with varying iterations. The model underpredicted by huge difference till 10 iterations and then started getting near the desired result. The model returned best results with 14 iterations where there was difference of 24.86 sq km only with 13.19% error in the urban and overall error of 7.76%. The overprediction might be because of the anomaly in the urban class for the year 2010, where the urban reduced at some points where there was temporary construction that was classified in 1990 and 2000 classified layers but was not urban in classified layers of 2010 and 2021 (Figure 3.3, Figure 3.4, Table 3.2, Table 3.3).

Table 3. 1 Settlements in different zones of land use plan of district Peshawar.

Sr No	Zone	Area (sq km)	Settlements in Zone (sq km)
1	Urban Area (2039)	300.0	132.18
2	Range Land	88.6	4.00
3	Model Environmental Villages	22.6	7.52
4	Rural Zones	563.3	70.71
5	Orchard Zone	39.4	6.66
6	Peshawar Model Town	52.3	6.55
7	Mining Zone	47.6	1.45
8	Green Belt	171.2	30.85
9	Total Area	1284.9	259.92

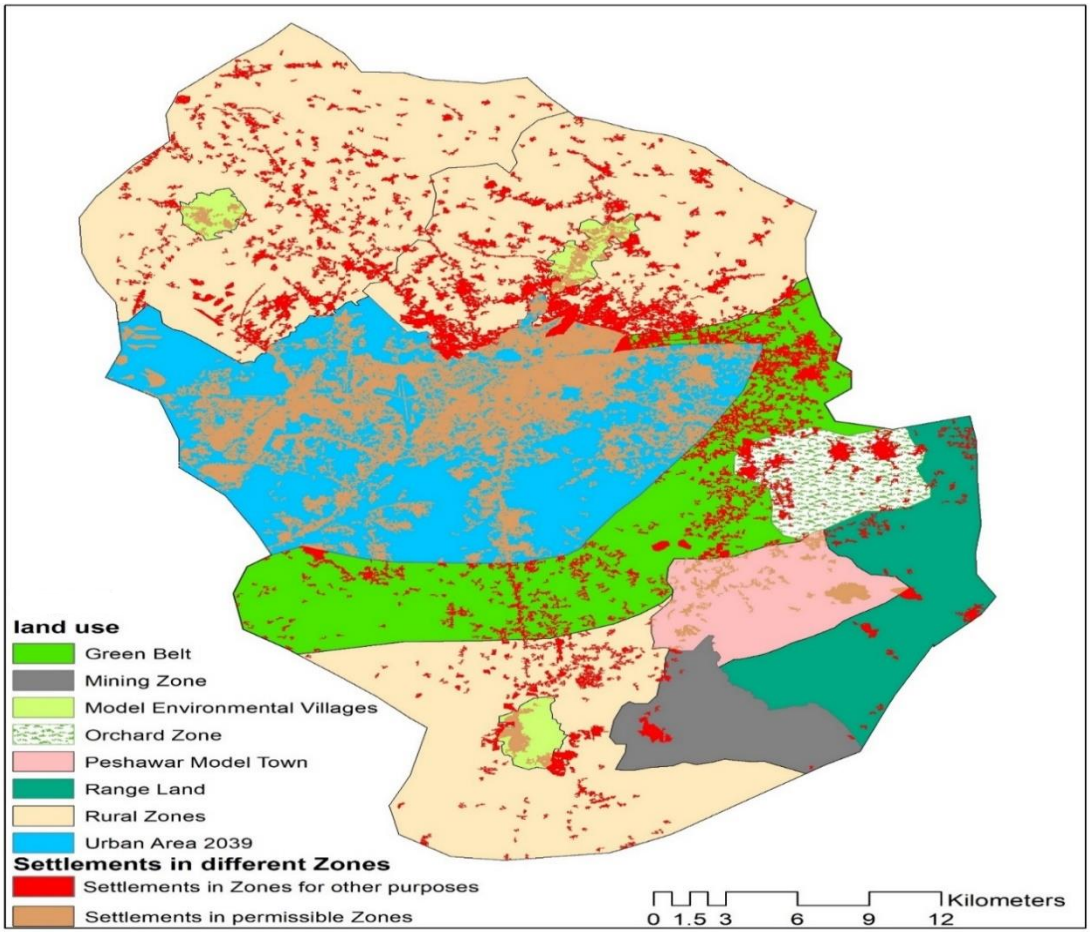

Figure 3. 2 Settlements in different zones of land use plan.

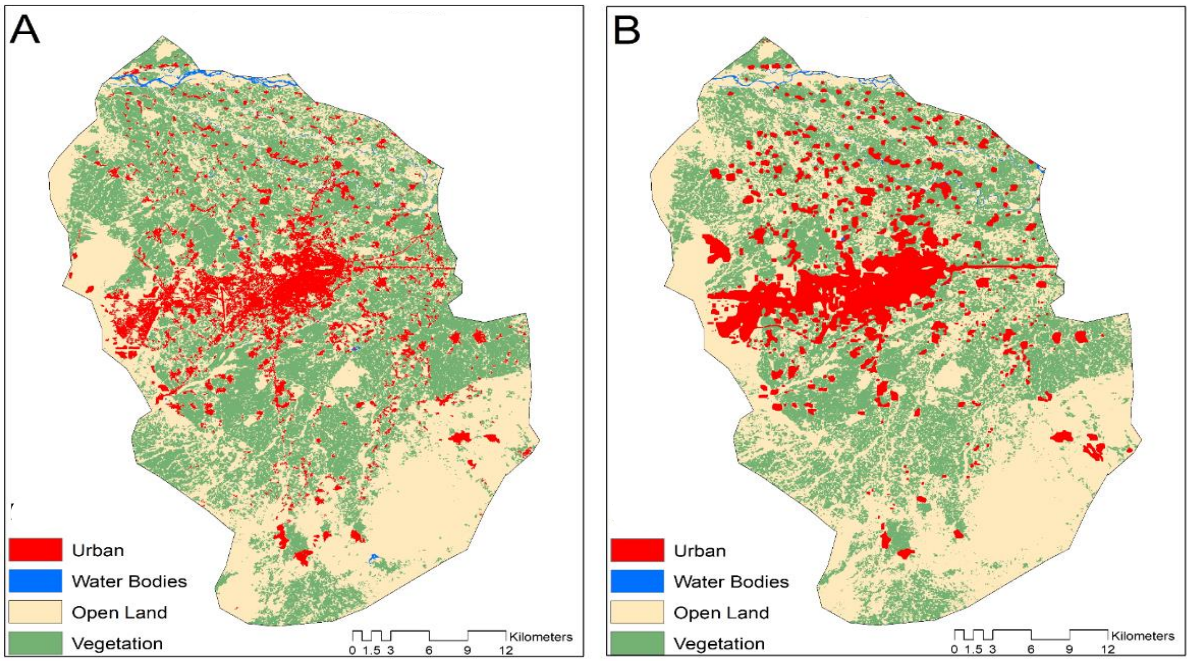

Figure 3. 3 Comparison between classified layer of and simulated layer (CA-MC) 2010.

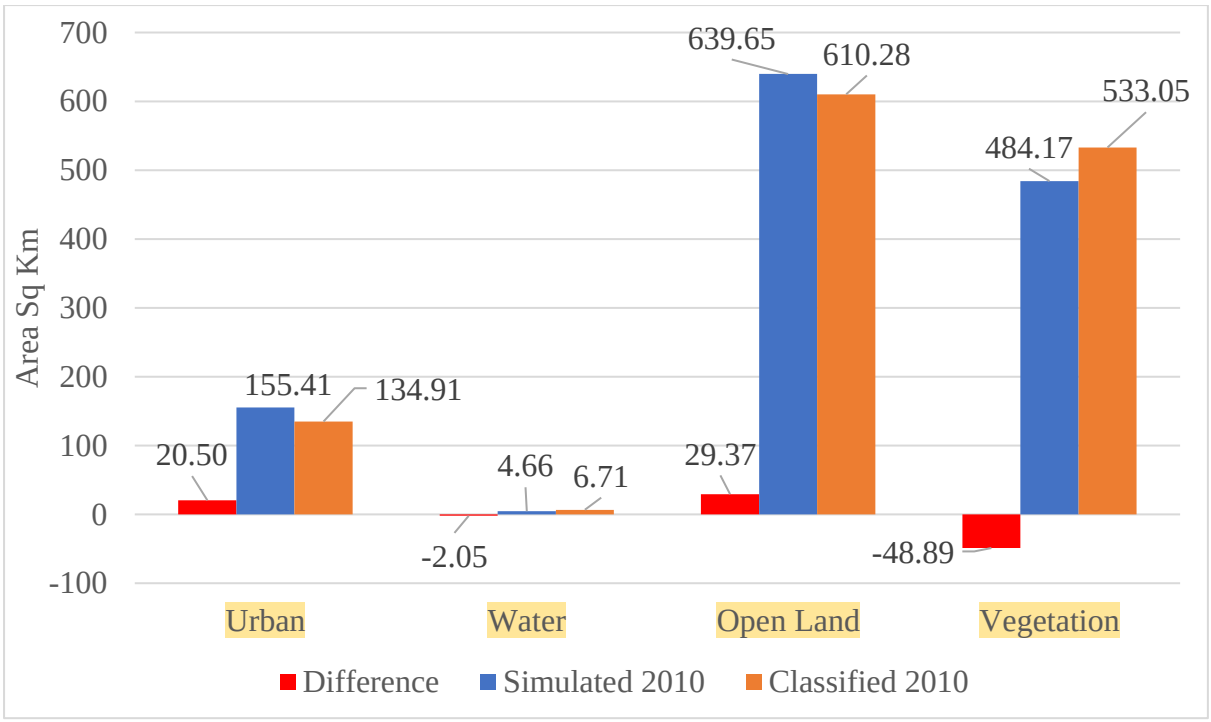

Figure 3. 4. Comparison between areas of different land use classes of classified layer and CA-MC simulated layer for the year 2010.

Table 3. 2 CA-MC calibration using 3*3 filer.

3*3					
Land use	**Urban**	**Water**	**Open Land**	**Vegetation**	**Total**
classified 2010	134.91	6.71	610.28	533.05	1284.96
10 iterations	82.27	4.80	700.97	496.91	1284.96
Difference	-52.64	-1.91	90.69	-36.14	
15 iterations	200.24	4.76	588.48	491.48	1284.96
Difference	65.32	-1.94	-21.81	-41.57	
17 iterations	251.23	4.84	547.30	481.58	1284.96
Difference	116.32	-1.87	-62.98	-51.47	

Table 3. 3 CA-MC calibration using 5*5 filter.

5*5					
Land use	**Urban**	**Water**	**Open Land**	**Vegetation**	**Total**
classified 2010	134.91	6.71	610.28	533.05	1284.96
10 iterations	91.50	3.98	701.79	487.68	1284.96
Difference	43.41	2.73	-91.51	45.37	
13 iterations	98.49	3.57	699.41	483.49	1284.96
Difference	36.42	3.14	-89.13	49.56	
14 iterations	159.77	4.83	631.22	489.13	1284.96
Difference	24.86	-1.88	20.94	-43.92	
15 iterations	155.41	4.66	639.65	485.24	1284.96
Difference	20.50	-2.05	29.37	-47.81	
16 iterations	158.35	6.85	634.72	485.03	1284.96
Difference	23.44	0.14	24.44	-48.02	
17 iterations	193.62	4.72	616.58	470.03	1284.96
Difference	58.71	-1.99	6.30	-63.02	

3.5 Urban simulation of 2039 of District Peshawar using CA-MC Model

The land use and land cover layers from 1990 and 2021 were utilized as input layers in order to model and simulate urban expansion for the year 2039. The environment used to validate the model was used in this simulation. The 5*5 contiguity filter was used. The number of iterations was kept to 15. The markov transition probabilities were calculated using markov module of the CA-MC model. The background cells (null values if any) were given a value of zero so that they do not take part in simulation which might distort the results. The proportional error was set to 0.15 as no classification is 100% accurate. The model iterated 15 times with 180 passes. To analyze the results, the simulated layer was loaded in ArcGIS 10.5. The area of each class was calculated. The simulated layer suggests that by 2039 the urban growth of district Peshawar will increase from 259.93 (20.23%) sq km 418.17 sq km which will cover 32.54% of the total area. The open land and vegetation class will lose 90.64 sq km and 57.56 sq km. and will come down from 549.25 sq and 480.81 sq km to 441.52 and 420.51 sq km respectively (Figure 3.5, Figure 3.6).

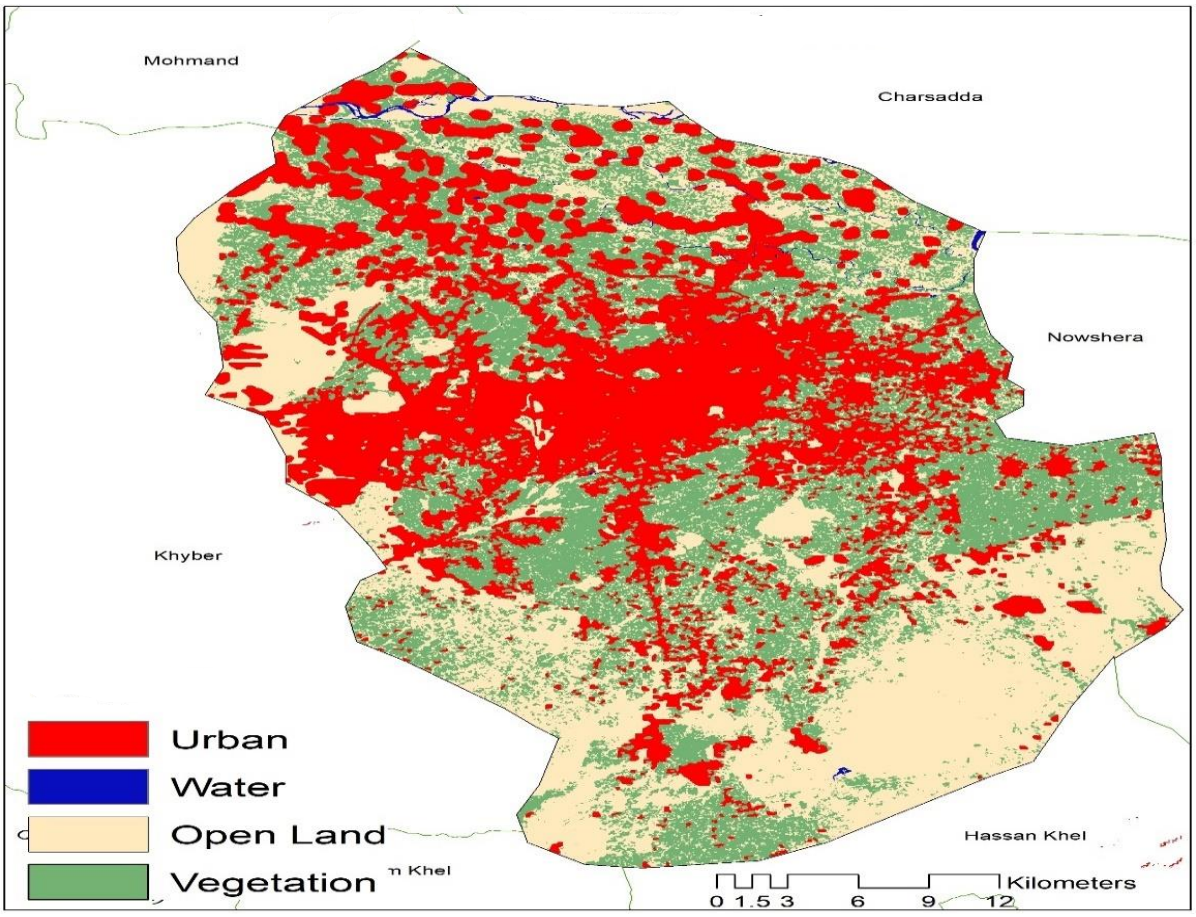

Figure 3. 5 CA-MC simulated layer for the 2039.

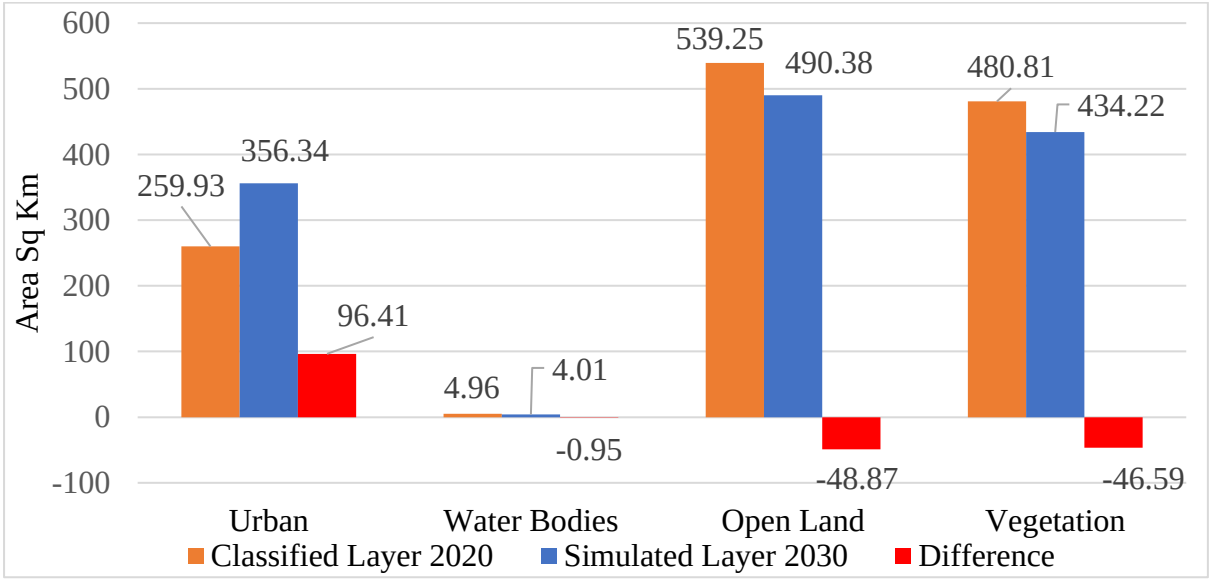

Figure 3. 6. Land use change from 2021 to 2039 using CA-MC model.

3.6 SLEUTH Model Validation

The urban layers for the years 1990, 1995, 2000, 2005 along with the road, excluded, hill shade and slope layers were used to predict the urban growth of the year 2010. The environment in the sleuth model scenario file was continuously changed in order to attain desirable results. The number of Monte Carlo iterations, values of critical high, critical low, boom and bust were changed continuously till the minimum difference between urban area of classified layer of 2010 and simulated layer of 2010 was achieved. The model returned the best results when the number of Monte Carlo iteration were kept 300, the critical high was set to 0.3, critical low to 0.1, boom to 0.2 and bust to 0.09. The urban area in the simulated layer was 142.21 sq km which is 7.3 sq km more than the classified urban in 2010. The model overpredicted urban growth by 5.4% only. (Figure 3.7)

3.7 Urban Simulation of District Peshawar for the year 2039 using SLEUTH Model

Urban layer for the years 1990, 2000, 2010, 2021, road layers for the years 1995, 2022 along with the excluded, slope and hill shade layers used to simulate the urban growth for the year 2039. The best fit parameters from the validation were used as inputs. The values for the coefficients diffusion, breed, spread, slope and road were 1, 1, 27, 1 and 3 respectively. The Monte Carlo iterations were kept 300. The road gravity sensitivity and slope sensitivity were assigned 0.1. After the model successfully completed the future simulation the gif layer was georeferenced in ArcGIS 10.5 for analysis. The prediction suggested that by 2039 the urban area will cover an area of 476 sq km which will take up 37% of the total area. The urban class will increase by 216.35 sq km from 2021 (259.93 sq km). There will be 45.45% increase in urban class from 2021 to 2039. If compared to the urban class of 1990 the model suggests that the urban area will increase 8 times as that of 1990. (Figure 3.8)

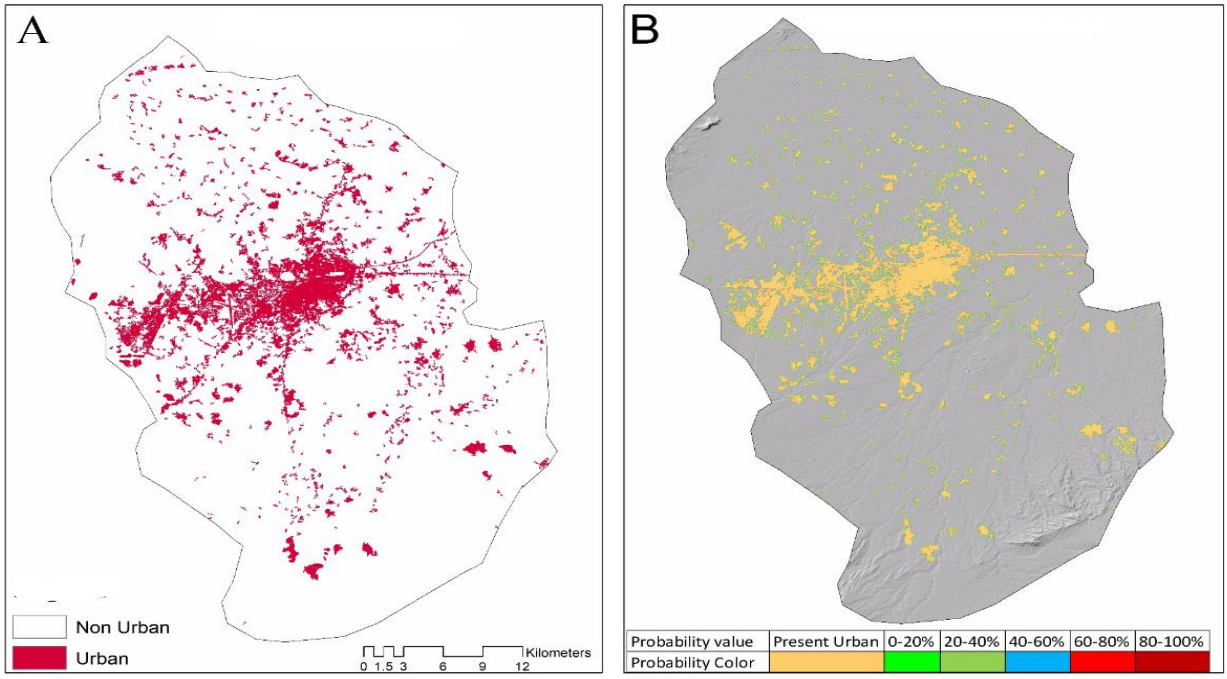

Figure 3. 7 Urban area of 2010 and simulated urban class of 2010.

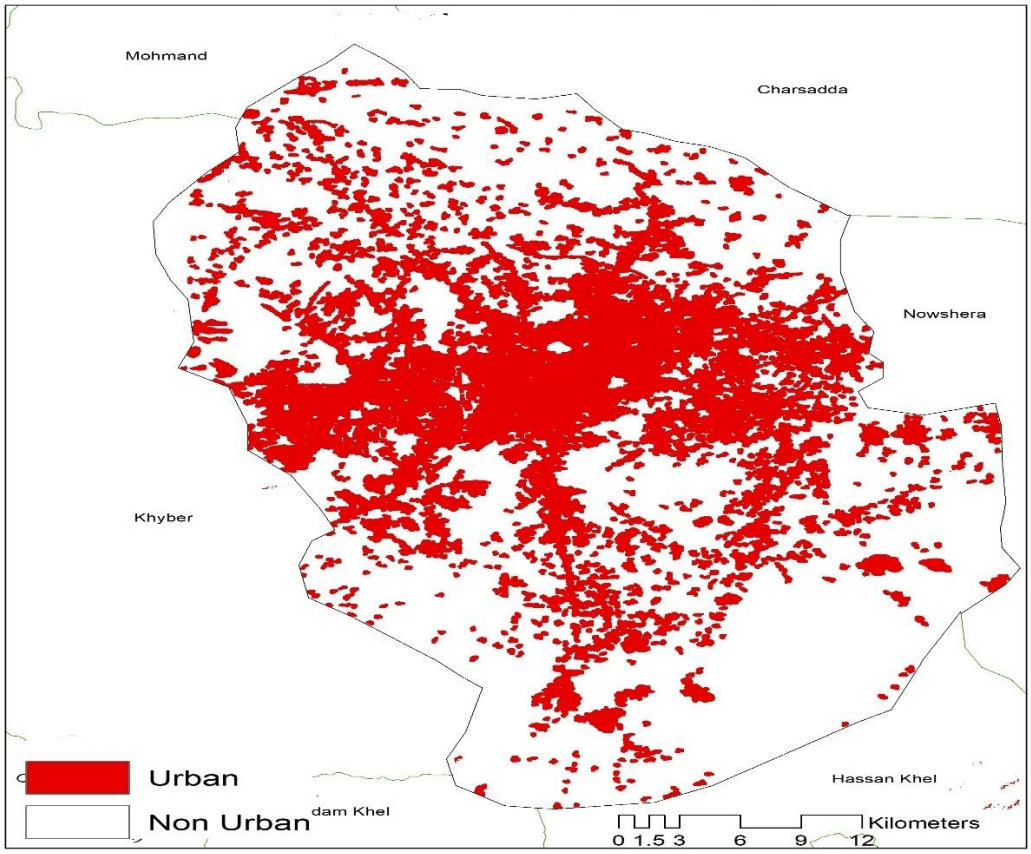

Figure 3. 8 Simulated layer of 2039 using Sleuth model.

3.8 Comparison between the CA Sleuth and CA-MCModel

The CA-MC model uses two land use layers of different years as input data to run the model. The model runs with two modules. The Markov Module which calculates the transtiton probabilities which defines the probability of change from one class to another. It also takes the proportional error into account to deal with the errors induced in the classificaiton phase. The results of Markov module serves as input in the CA Markov module. It asks for a basis image from which the simualtion will take place, number of iterations and the contiguity filters are also defined. As Compared to CA-MC model the CA Sleuth model takes more layers to run. The layers required are 4 urban layers, 2 road layers, excluded layer. Hill shade and slope layer. The Model runs in Linux or Windows based Cygwin. The Model runs in 3 modes, test, calibration and prediction. The test phase checks the data to be valid for the model. The calibration phase calibrates the model and makes it ready for the prediction phase. Calibration mode runs in 3 phases, coarse fine and final. The purpose of these modes is to refine the coefficients that control the urban growth. The calibration gives best fit values for forcast phase. The forcast mode returns single vlaues that serves as input in prediciton phase which gives future simulation as a result. In prediction mode the coefficients best values are used along with the Monte Carlo iterations. The urban growth is also controlled by critical high, critical low, boom, bust, slope sensitivity and road gravity sensitivity values. Both the CA-MC modeland CA Sleuth model were validated by predicting the urban growth for the year 2010 and comparing it with the classified layer of the same year. In CA-MC modelthe number of iterations and contigity filters were changed during every run to so that the model outcome can match the urban class of the classified layer. The best result was achived with 15 iterations using 5*5 contiguity filter. There was an overprediction in the urban class. The urban area in the simulated layer was 155.41 sq km which is 20.50 sq km more than the urban area of 2010

classified layer which is 134.91 sq km. The error in the urban class was 13.18%. If compared to the outcomes of Sleuth model which simulated the ubran area to be 142.21 sq km which is only 7.3 sq km more than the classified urban area. The model predicited the urban pixels by only 5.4%. So by this comparison it is evident the the Sleuth model predicted the urban growth more accurately. The models were then used to simulate the urban growth for thea year 2039. The CA-MCmodel simulated the 2039 urban growth to be 418.17 sq km while the CA Sleuth model predicted the urban class to be 476 sq km by 2039. There is a difference of 57.83 sq km between the two simulations. The Sleuth model predicted the urban growth to be 12.14% more than that of CA-MC simulation for the year 2039. (Figure 3.9, Figure 3.10)

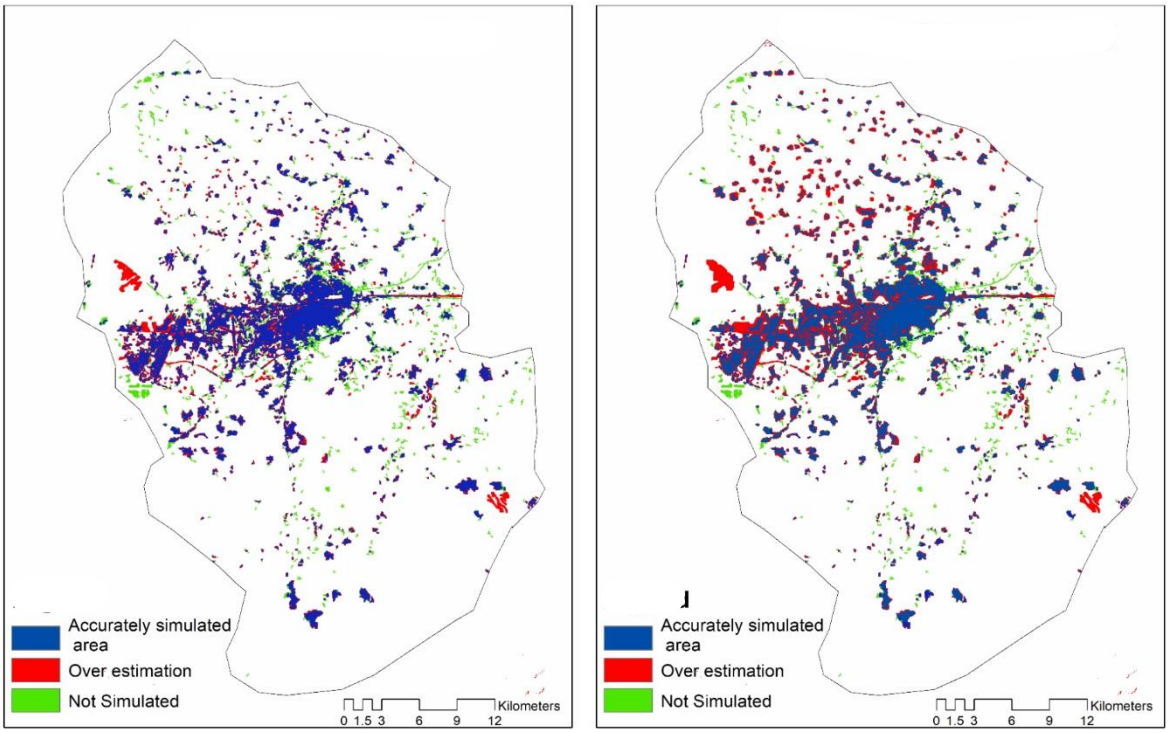

Figure 3. 9 Spatial agreement of Sleuth and CA Markov simulations for the year 2010.

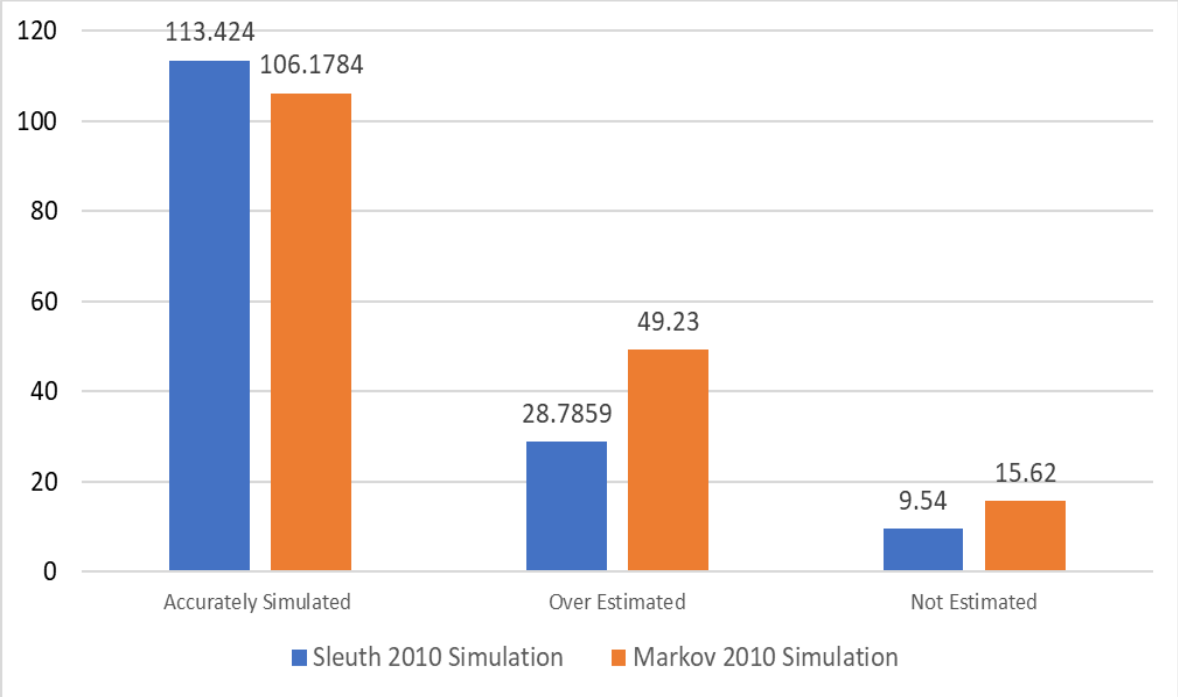

Figure 3. 10. Performance of Sleuth and CA Markov model

Chapter 4

CONCLUSION AND RECOMMENDATIONS

4.1 Conclusion

The urban environment frequently grows in an unsustainable manner, especially in developing countries, which is referred to as haphazard and uncontrolled urban expansion. Urban simulation studies may be quite beneficial in understanding the implications of current planning strategies. Future growth often requires inputs that emphasize existing growth patterns, on which urban development models may simulate future growth.

With nearly 36.4% of its inhabitants living in urban areas, Pakistan is the sixth highest population density country in the world. The province of Khyber Pakhtunkhwa is likewise rapidly urbanizing, 18.7% of the population, as determined by the 2017 census, resides in urban areas. district Peshawar, the capital of the Khyber Pakhtunkhwa province, is expanding rapidly with a growth rate of 14% and is anticipated to reach 18% by 2030. Future issues might result from this, including the loss of agricultural land. Water shortage, traffic congestion, and several other environmental effects.

This study tracked the urban growth of Peshawar for the years 1990, 2000, 2010 and 2021. There was massive increase in the urban area from 1990 to 2021 during these years the urban area grew from 50.50 to 259.93 which is 414.72% more than that of 1990. The growth over the years is as follows. The Urban area grew from 50.50 sq km to 78.07 sq km from 1990 to 2000. From 2000 to 2010 it grew from 78.07 sq km to 134.91 sq. The urban growth from 2010 to 2021 was the highest of them all where the urban area increased from 134.91 sq km to 259.93 sq km taking up 20.23% of the district's total area. The Vegetation class increased from

484.50 to 495.13 from 1990 to 2000 and then increased to 533.05 in 2010 but from 2010 it decreased by 71.03 sq km to stand at 480.81 sq km. The increase from 1990 to 2010 can be justified by the fact that increase in population will increase the agricultural activities as well but the massive urbanization from 2010 to 2021 reduced it as the areas where there used to be fields are taken up by settlements. The class that was most affected by the urbanization is the open land which reduced from 742.04 sq km to 539.25 sq km losing 202.79 sq km from 1990 to 2021. If this continues there will be a point when there will be not enough suitable open land to be taken up by the urban class and urban area will increase on the cost of agricultural land which will have more adverse impacts. The open land is important for the water infiltration as well, the increase of impervious surfaces will restrict the water infiltration and the aquifer will suffer as the water table will not be recharged the way it should be and will result in water table going down further.

The study used two urban growth models CA-MC model and the CA Sleuth Model and gave an in-depth comparative analysis of the models. Both the models were validated by simulating urban growth for the year 2010 and comparing it with the classified layer of 2010. The CA Sleuth model performed better than the CA-MC model as it overpredicted the urban area by only 7.3 sq km as compared to the 20.50 prediction of CA-MC model. Based on the environments used for validation the models were used to simulate urban growth for the year 2039. The CA-MC predicted the urban area to be 418.17 sq km while the CA Sleuth model predicted the urban area to be 476 sq km.

This study effectively illustrated the application of GIS, RS, and urban growth modelling, and the findings were deemed to be encouraging for tracking the changes in land use and land cover and gaining an understanding of what the potential future urban expansion

would imply. It also paves the way for the urban planner to develop strategies and guidelines for sustainable urban expansion and get a thorough understanding of urban sprawl. Also, the models may be used to investigate and simulate the complexity of urban expansion in Pakistan's major cities, which would also be helpful to the relevant authorities.

4.2 Recommendations

The land use plan of district Peshawar should be revised according to the current land use and land cover. The Green belt is right next to the urban plan for the year 2039 which makes the green belt prone to urbanization. There should be zoning regulations to reduce or stop urbanization in the zones that do not permit new settlements. The trend of urbanization in the rural zones is the highest if the current trend continues the prime agricultural land will be consumed by urban class.

The use of high-resolution satellite imagery will help classify the urban area along with major classes accurately. The chance of inaccurate classification due to pixel mixing is considerable in medium resolution imagery since a single pixel represents the average of multiple spectral classes within the region it covers on the ground, each of which is emitted or reflected by a different kind of objects which can lead to urban pixel classified as non-urban and vice versa. The use of high-resolution imagery will reduce this error and help in better image classification.

REFERENCES

1. Aaviksoo, K. (1993). Application of Markov Models in Investigation of Vegetation and Land Use Dynamics in Estonian Mire Landscapes: Theses (Vol. 4). Tartu University.
2. Abdullahi, S., & Pradhan, B. (2018). Land use change modeling and the effect of compact city paradigms: integration of GIS-based cellular automata and weights-of-evidence techniques. Environmental Earth Sciences, 77(6), 251.
3. Abolhasani, S., & Taleai, M. (2020). Assessing the effect of temporal dynamics on urban growth simulation: Towards an asynchronous cellular automata. Transactions in GIS, 24(2), 332-354.
4. Al-Shaar, W., Adjizian Gérard, J., Nehme, N., Lakiss, H., & Buccianti Barakat, L. (2021). Application of modified cellular automata Markov chain model: forecasting land use pattern in Lebanon. Modeling Earth Systems and Environment, 7, 1321-1335.
5. Altuwaijri, H. A., Alotaibi, M. H., Almudlaj, A. M., & Almalki, F. M. (2019). Predicting urban growth of Arriyadh city, capital of the Kingdom of Saudi Arabia, using Markov cellular automata in TerrSet geospatial system. Arabian Journal of Geosciences, 12, 1-15.
6. Arsanjani, J. J., Helbich, M., Kainz, W., & Boloorani, A. D. (2013). Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. International Journal of Applied Earth Observation and Geoinformation, 21, 265-275.
7. Arsanjani, J. J., Kainz, W., & Mousivand, A. J. (2011). Tracking dynamic land-use change using spatially explicit Markov Chain based on cellular automata:

the case of Tehran. International Journal of Image and Data Fusion, 2(4), 329-345.

8. Attaallah, H. (2018, June). Modeling of built-up lands expansion in Gaza Strip, Palestine using Landsat data and CA-Markov model. In IOP Conference Series: Earth and Environmental Science (Vol. 169, No. 1, p. 012035). IOP Publishing.
9. Bajracharya, P., Lippitt, C. D., & Sultana, S. (2020). Modeling urban growth and land cover change in Albuquerque using SLEUTH. The Professional Geographer, 72(2), 181-193.
10. Baqa, M. F., Chen, F., Lu, L., Qureshi, S., Tariq, A., Wang, S., ... & Li, Q. (2021). Monitoring and modeling the patterns and trends of urban growth using urban sprawl matrix and CA-Markov model: A case study of Karachi, Pakistan. Land, 10(7), 700.
11. Barredo, J. I., Kasanko, M., McCormick, N., & Lavalle, C. (2003). Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. Landscape and urban planning, 64(3), 145-160.
12. Barredo, J. I., Kasanko, M., McCormick, N., & Lavalle, C. (2003). Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. Landscape and urban planning, 64(3), 145-160.
13. Barreira-González, P., & Barros, J. (2017). Configuring the neighbourhood effect in irregular cellular automata-based models. International Journal of Geographical Information Science, 31(3), 617-636.
14. Basse, R. M., Omrani, H., Charif, O., Gerber, P., & Bódis, K. (2014). Land use changes modelling using advanced methods: Cellular automata and artificial neural networks. The spatial and explicit representation of land cover dynamics at the cross-border region scale. Applied Geography, 53, 160-171.

15. Bauer, M. E., Yuan, F., & Sawaya, K. E. (2004). Multi-temporal Landsat image classification and change analysis of land cover in the Twin Cities (Minnesota) metropolitan area. In Analysis of Multi-Temporal Remote Sensing Images (pp. 368-375).
16. Berberoğlu, S., Akın, A., & Clarke, K. C. (2016). Cellular automata modeling approaches to forecast urban growth for adana, Turkey: A comparative approach. Landscape and urban planning, 153, 11-27.
17. Chandan, M. C., Nimish, G., & Bharath, H. A. (2020). Analysing spatial patterns and trend of future urban expansion using SLEUTH. Spatial Information Research, 28, 11-23.
18. Chaudhuri, G., & Clarke, K. C. (2019). Modeling an Indian megalopolis-A case study on adapting SLEUTH urban growth model.
19. Clarke, K. C., Gazulis, N., Dietzel, C., & Goldstein, N. C. (2007). A decade of SLEUTHing: Lessons learned from applications of a cellular automaton land use change model. Classics in IJGIS: twenty years of the international journal of geographical information science and systems, 413-427.
20. Dietzel, C., & Clarke, K. (2006). The effect of disaggregating land use categories in cellular automata during model calibration and forecasting. Computers, Environment and Urban Systems, 30(1), 78-101.
21. El Saeid Mustafa, A. M., Nishida, G., Saadi, I., Cools, M., & Teller, J. (2017, March). A Markov Chain Monte Carlo Cellular Automata Model to Simulate Urban Growth. In GEOProcessing 2017. ThinkMind, Nice, France.
22. Firozjaei, M. K., Sedighi, A., Argany, M., Jelokhani-Niaraki, M., & Arsanjani, J. J. (2019). A geographical direction-based approach for capturing the local

variation of urban expansion in the application of CA-Markov model. Cities, 93, 120-135.

23. Fu, X., Wang, X., & Yang, Y. J. (2018). Deriving suitability factors for CA-Markov land use simulation model based on local historical data. Journal of environmental management, 206, 10-19.
24. Gao, C., Feng, Y., Tong, X., Lei, Z., Chen, S., & Zhai, S. (2020). Modeling urban growth using spatially heterogeneous cellular automata models: Comparison of spatial lag, spatial error and GWR. Computers, Environment and Urban Systems, 81, 101459.
25. Ghosh, S., Chatterjee, N. D., & Dinda, S. (2021). Urban ecological security assessment and forecasting using integrated DEMATEL-ANP and CA-Markov models: A case study on Kolkata Metropolitan Area, India. Sustainable Cities and Society, 68, 102773.
26. Gonçalves, T. M., Zhong, X., Ziggah, Y. Y., & Dwamena, B. Y. (2019). Simulating urban growth using cellular automata approach (SLEUTH)-A case study of Praia City, Cabo Verde. IEEE Access, 7, 156430-156442.
27. Haider, M., & Badami, M. G. (2010). Urbanization and local governance challenges in Pakistan. Environment and urbanization ASIA, 1(1), 81-96.
28. Hamad, R., Balzter, H., & Kolo, K. (2018). Predicting land use/land cover changes using a CA-Markov model under two different scenarios. Sustainability, 10(10), 3421.
29. Hassan, M. M., & Nazem, M. N. I. (2016). Examination of land use/land cover changes, urban growth dynamics, and environmental sustainability in Chittagong city, Bangladesh. Environment, development and sustainability, 18, 697-716.

30. Huang, J., Wu, Y., Gao, T., Zhan, Y., & Cui, W. (2015). An integrated approach based on Markov Chain and cellular automata to simulation of urban land use changes. Applied Mathematics & Information Sciences, 9(2), 769.
31. Huang, J., Wu, Y., Gao, T., Zhan, Y., & Cui, W. (2015). An integrated approach based on Markov Chain and cellular automata to simulation of urban land use changes. Applied Mathematics & Information Sciences, 9(2), 769.
32. Huang, Y., Yang, B., Wang, M., Liu, B., & Yang, X. (2020). Analysis of the future land cover change in Beijing using CA–Markov chain model. Environmental Earth Sciences, 79(2), 60.
33. Jain, G. V., & Angadi, D. P. (2022). Land-use change dynamics and urban growth simulations in a medium-sized city of Mangaluru, India through CA-based SLEUTH urban growth model.
34. Jantz, C. A., Goetz, S. J., & Shelley, M. K. (2004). Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area. Environment and Planning B: Planning and Design, 31(2), 251-271.
35. Jantz, C. A., Goetz, S. J., Donato, D., & Claggett, P. (2010). Designing and implementing a regional urban modeling system using the SLEUTH cellular urban model. Computers, Environment and Urban Systems, 34(1), 1-16.
36. Jawarneh, R. N. (2021). Modeling past, present, and future urban growth impacts on primary agricultural land in Greater Irbid Municipality, Jordan using SLEUTH (1972–2050). ISPRS International Journal of Geo-Information, 10(4), 212.

37. Karimi, H., Jafarnezhad, J., Khaledi, J., & Ahmadi, P. (2018). Monitoring and prediction of land use/land cover changes using CA-Markov model: a case study of Ravansar County in Iran. Arabian Journal of Geosciences, 11, 1-9.
38. Kugelman, M. (2013). Pakistan's energy crisis. National Bureau of Asian Research: Washington, DC.
39. Kumar, K. S., Kumari, K. P., & Bhaskar, P. U. (2016, March). Application of Markov Chain & Cellular Automata based model for prediction of Urban transitions. In 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) (pp. 4007-4012). IEEE.
40. Kumar, V., & Agrawal, S. (2022). Urban modelling and forecasting of landuse using SLEUTH model. International Journal of Environmental Science and Technology,1-20.
41. Lai, T., Dragićević, S., & Schmidt, M. (2013). Integration of multicriteria evaluation and cellular automata methods for landslide simulation modelling. Geomatics, Natural Hazards and Risk, 4(4), 355-375.
42. Li, X., & Gong, P. (2016). Urban growth models: progress and perspective. Science bulletin, 61(21), 1637-1650.
43. Liang, X., Liu, X., Li, D., Zhao, H., & Chen, G. (2018). Urban growth simulation by incorporating planning policies into a CA-based future land-use simulation model. International Journal of Geographical Information Science, 32(11), 2294-2316.
44. Liu, Y., Li, L., Chen, L., Cheng, L., Zhou, X., Cui, Y., ... & Liu, W. (2019). Urban growth simulation in different scenarios using the SLEUTH model: A case study of Hefei, East China. PLoS One, 14(11), e0224998.

45. Mallouk, A., Elhadrachi, H., Malaainine, M. E. I., & Rhinane, H. (2019). Using the SLEUTH urban growth model coupled with a GIS to simulate and predict the future urban expansion of Casablanca region, Morocco. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 139-145.
46. Mansour, S., Al-Belushi, M., & Al-Awadhi, T. (2020). Monitoring land use and land cover changes in the mountainous cities of Oman using GIS and CA-Markov modelling techniques. Land Use Policy, 91, 104414.
47. Mohamed, A., & Worku, H. (2020). Simulating urban land use and cover dynamics using cellular automata and Markov chain approach in Addis Ababa and the surrounding. Urban Climate, 31, 100545.
48. Mustafa, A., Bruwier, M., Archambeau, P., Erpicum, S., Pirotton, M., Dewals, B., & Teller, J. (2018). Effects of spatial planning on future flood risks in urban environments. Journal of environmental management, 225, 193-204.
49. Mustafa, D., & Sawas, A. (2013). Urbanisation and Political Change in Pakistan: exploring the known unknowns. Third World Quarterly, 34(7), 1293-1304.
50. Omrani, H., Tayyebi, A., & Pijanowski, B. (2017). Integrating the multi-label land-use concept and cellular automata with the artificial neural network-based Land Transformation Model: an integrated ML-CA-LTM modeling framework. GIScience & Remote Sensing, 54(3), 283-304.
51. Pinto, N., Antunes, A. P., & Roca, J. (2017). Applicability and calibration of an irregular cellular automata model for land use change. Computers, Environment and Urban Systems, 65, 93-102.

52. Rahnama, M. R. (2021). Forecasting land-use changes in Mashhad Metropolitan area using Cellular Automata and Markov chain model for 2016-2030. Sustainable Cities and Society, 64, 102548.
53. Ramachandra, T. V., Aithal, B. H., & Sowmyashree, M. V. (2014). Urban structure in Kolkata: metrics and modelling through geo-informatics. Applied Geomatics, 6(4), 229-244.
54. Sakieh, Y., Amiri, B. J., Danekar, A., Feghhi, J., & Dezhkam, S. (2015). Scenario-based evaluation of urban development sustainability: an integrative modeling approach to compromise between urbanization suitability index and landscape pattern. Environment, Development and Sustainability, 17, 1343-1365.
55. Saxena, A., & Jat, M. K. (2019). Capturing heterogeneous urban growth using SLEUTH model. Remote Sensing Applications: Society and Environment, 13, 426-434.
56. Saxena, A., & Jat, M. K. (2020). Land suitability and urban growth modeling: Development of SLEUTH-Suitability. Computers, Environment and Urban Systems, 81, 101475.
57. Siddiqui, A., Siddiqui, A., Maithani, S., Jha, A. K., Kumar, P., & Srivastav, S. K. (2018). Urban growth dynamics of an Indian metropolitan using CA Markov and Logistic Regression. The Egyptian Jour.
58. Singh, S. K., Mustak, S., Srivastava, P. K., Szabó, S., & Islam, T. (2015). Predicting spatial and decadal LULC changes through cellular automata Markov chain models using earth observation datasets and geo-information. Environmental Processes, 2, 61-78.

59. Tang, R., Zhao, X., Zhou, T., Jiang, B., Wu, D., & Tang, B. (2018). Assessing the impacts of urbanization on albedo in Jing-Jin-Ji Region of China. Remote Sensing, 10(7), 1096.
60. Tayyebi, A., Shafizadeh-Moghadam, H., & Tayyebi, A. H. (2018). Analyzing long-term spatio-temporal patterns of land surface temperature in response to rapid urbanization in the mega-city of Tehran. Land Use Policy, 71, 459-469.
61. Thapa, R. B., & Murayama, Y. (2010). Drivers of urban growth in the Kathmandu valley, Nepal: Examining the efficacy of the analytic hierarchy process. Applied Geography, 30(1), 70-83.
62. Tong, X., & Feng, Y. (2020). A review of assessment methods for cellular automata models of land-use change and urban growth. International Journal of Geographical Information Science, 34(5), 866-898.
63. Verburg, P. H., & Veldkamp, A. (2001). The role of spatially explicit models in land-use change research: a case study for cropping patterns in China. Agriculture, ecosystems & environment, 85(1-3), 177-190.
64. Vermeiren, K., Vanmaercke, M., Beckers, J., & Van Rompaey, A. (2016). ASSURE: a model for the simulation of urban expansion and intra-urban social segregation. International Journal of Geographical Information Science, 30(12), 2377-2400.
65. Wang, H., Hu, Y., & Liang, Y. (2021). Simulation and spatiotemporal evolution analysis of biocapacity in Xilingol based on CA-Markov land simulation. Environmental and Sustainability Indicators, 11, 100136.

66. Watkiss, B. M. (2008). The SLEUTH urban growth model as forecasting and decision-making tool (Doctoral dissertation, Stellenbosch: Stellenbosch University).
67. White, R., & Engelen, G. (1993). Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. Environment and planning A, 25(8), 1175-1199.
68. Wu, H., Li, Z., Clarke, K. C., Shi, W., Fang, L., Lin, A., & Zhou, J. (2019). Examining the sensitivity of spatial scale in cellular automata Markov chain simulation of land use change. International Journal of Geographical Information Science, 33(5), 1040-1061.
69. Yin, C., Yu, D., Zhang, H., You, S., & Chen, G. (2008, November). Simulation of urban growth using a cellular automata-based model in a developing nation's region. In Geoinformatics 2008 and joint conference on GIS and built environment: Geo-simulation and virtual GIS environments (Vol. 7143, pp. 371-378). SPIE.
70. Zhou, L., Dang, X., Sun, Q., & Wang, S. (2020). Multi-scenario simulation of urban land change in Shanghai by random forest and CA-Markov model. Sustainable Cities and Society, 55, 102045.
71. Zhou, Y., Varquez, A. C., & Kanda, M. (2019). High-resolution global urban growth projection based on multiple applications of the SLEUTH urban growth model. Scientific data, 6(1), 34.