Masters Program in Geospatial Technologies

LAND COVER MAPPING ANALYSIS AND URBAN GROWTH MODELING USING REMOTE SENSING TECHNIQUES

Case Study: Greater Cairo Region - Egypt

Yasmine Megahed

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

NOVA IMS Information Management School
Universitat Jaume I
Westfälische Wilhelms-Universität Münster

LAND COVER MAPPING ANALYSIS AND URBAN GROWTH MODELING USING REMOTE SENSING TECHNIQUES

Case Study: Greater Cairo Region - Egypt

Dissertation supervised by

Mário Caetano, PhD

Pedro Cabral, PhD

Flililberto Pla, PhD

Joel Silva, MSc

February 2015

ACKNOWLEDGMENTS

I would like first to thank God, The most Merciful, The Most Generous and The Most Gracious, for helping me accomplish this work. This thesis would not have existed without God's guidance and grace.

Then, I would like to express my deep gratitude to all my supervisors for their cooperation and aspiring guidance that allowed me to finalize this thesis and to make it as perfect as we can. I am grateful for the efforts you all made, your support, immense knowledge and your insightful comments that helped me to improve the work and finish the thesis in the best way possible. I am really honored to have you as my supervisors.

Special thanks to Joel Silva for his help, patience, continuous encouragement through this journey to finish the thesis, and his advice to solve the problems I faced during this work, as well as his valuable suggestions regarding the way results are presented. I am lucky to have you among my co-supervisors.

I would like to express my warm thanks to Information Management School – IMS's staff members who facilitated all services and provided me with the aids I asked for in order to carry out this work. My sincere thanks goes to Professor Marco Painho, the coordinator of this master's program, who made valuable comment suggestions on my work during the thesis follow-up sessions which he managed. His advice and feedback always gave me an inspiration to improve my work. This type of professionalism and commitment to the highest level of student's satisfaction has to be acknowledged.

Last but not the least, I would like to thank my mother Mrs. Shadia Khougaly. Every success in my life is a direct result of your kindness, efforts, prayers and your full dedication to me. No words can fully describe your role and effect in my life. You deserve to be the one I finalize this acknowledgement with.

LAND COVER MAPPING ANALYSIS AND URBAN GROWTH MODELING USING REMOTE SENSING TECHNIQUES

Case Study: Greater Cairo Region – Egypt

ABSTRACT

The rapid growth of big cities has been noticed since 1950s when the majority of world population turned to live in urban areas rather than villages, seeking better job opportunities and higher quality of services and lifestyle circumstances. This demographic transition from rural to urban is expected to have a continuous increase. Governments, especially in less developed countries, are going to face more challenges in different sectors, raising the essence of understanding the spatial pattern of the growth for an effective urban planning.

The study aimed to detect, analyse and model the urban growth in Greater Cairo Region (GCR) as one of the fast growing mega cities in the world using remote sensing data. Knowing the current and estimated urbanization situation in GCR will help decision makers in Egypt to adjust their plans and develop new ones. These plans should focus on resources reallocation to overcome the problems arising in the future and to achieve a sustainable development of urban areas, especially after the high percentage of illegal settlements which took place in the last decades.

The study focused on a period of 30 years; from 1984 to 2014, and the major transitions to urban were modelled to predict the future scenarios in 2025. Three satellite images of different time stamps (1984, 2003 and 2014) were classified using Support Vector Machines (SVM) classifier, then the land cover changes were detected by applying a high level mapping technique. Later the results were analyzed for higher accurate estimations of the urban growth in the future in 2025 using Land Change Modeler (LCM) embedded in IDRISI software. Moreover, the spatial and temporal urban growth patterns were analyzed using statistical metrics developed in FRAGSTATS software.

The study resulted in an overall classification accuracy of 96%, 97.3% and 96.3% for 1984, 2003 and 2014's map, respectively. Between 1984 and 2003, 19 179 hectares of vegetation and 21 417 hectares of desert changed to urban, while from 2003 to 2014, the transitions to urban from both land cover classes were found to be 16 486 and 31 045 hectares, respectively. The model results indicated that 14% of the vegetation and 4% of the desert in 2014 will turn into urban in 2025, representing 16 512 and 24 687 hectares, respectively.

KEYWORDS

Urban change Detection

Binary Maps

Land Change Modeler

Support Vector Machines

Urban Growth Modeling

Urban Spatial Patterns

ACRONYMS

AUC – Area Under Curve

CAPMAS – Central Agency for Public Mobilization and Statistics

CCDM – Comprehensive Change Detection Method

DEM – Digital Elevation Model

FOM – Figure Of Merit

GCR – Greater Cairo Region

KI – Kappa Index

LCLU – Land Cover Land Use

LCM – Land Change Modeler

LUCC – Land-Use & Cover Change

MC-CA – Markov Chain – Cellular Automata

MCE – Multi-Criteria Evaluation

ML – Maximum Likelihood

MLP – Multi-Layer Perceptron

MMU – Minimum Mapping Unit

NDBI – Normalized Difference Built-up Index

NDVI – The Normalized Difference Vegetation Index

NIR – Near Infrared

OLI – Operational Land Imager

RGB – Red Green Blue

ROC – Relative Operating Characteristic

SIS – State Information Service

SVM – Support Vector Machines

TM – Thematic Mapper

TRIS – Thermal Infrared Sensor

UN – United Nations

UNFPA – United Nations Fund for Population Activities

WHO – World Health Organization

INDEX OF THE TEXT

INDEX OF TABLES

INDEX OF FIGURES

1. INTRODUCTION

Urbanization is the demographic transition from rural to urban which is associated with shifts from an agriculture-based economy to mass industry, technology, and service (WHO, 2014). Recent studies indicate the fact that our world is undergoing the largest wave of urban growth in history (UNFPA, 2014). The world urban residents began to increase significantly since 1950s with a population expansion of more than 3% per year, nearly 60 million every year nowadays (WHO, 2014). In the future, the growth rate of the urban population is expected to grow approximately 1.84% per year between 2015 and 2020, 1.63% per year between 2020 and 2025, and 1.44% per year between 2025 and 2030, and by 2050 the urban population is expected to almost double, increasing from approximately 3.4 billion1 in 2009 to 6.4 billion (WHO, 2014). Figure 1 shows that in 1980 around 40% of the total world population had lived in cities, and this percentage increased to 50% (half of the world population) in 2010 (UN, 2014). In other words, one hundred years ago, 2 out of every 10 had lived in an urban area, growing to be 6 in 2030, and 7 out of 10 is the estimation for the urban residents in 2050 (WHO, 2014).

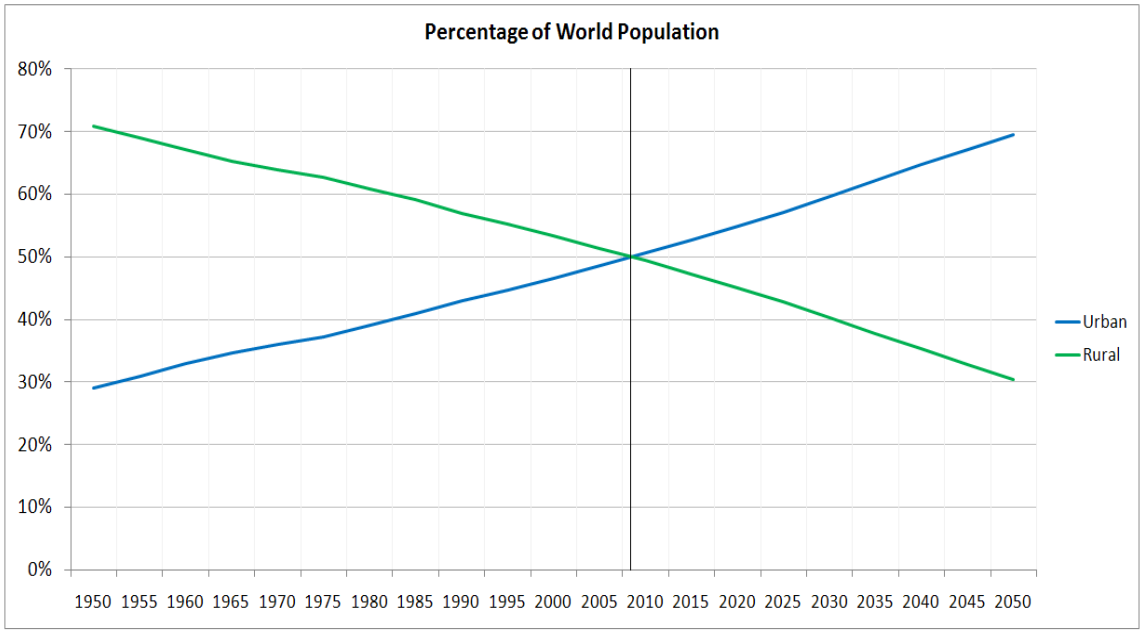


Figure 1. Historical shift of the urban/rural population ratio

The proportion of people living in urban areas is larger in developed countries than in less developed ones. Both are expected to increase in 2050, but at a more moderate rate in developed countries than in developing ones (Figure 2), because population densities of cities in developing countries are generally three times higher than in industrial countries (Thorpe, 2014). The urban population in less developed

[1]American billion = 10 ^ 9

countries has an average growth rate of 165 000 person per day and is estimated to keep increasing from 2.5 billion in 2009 to more than the double in 2050 (5.2 billion) (WHO, 2014).

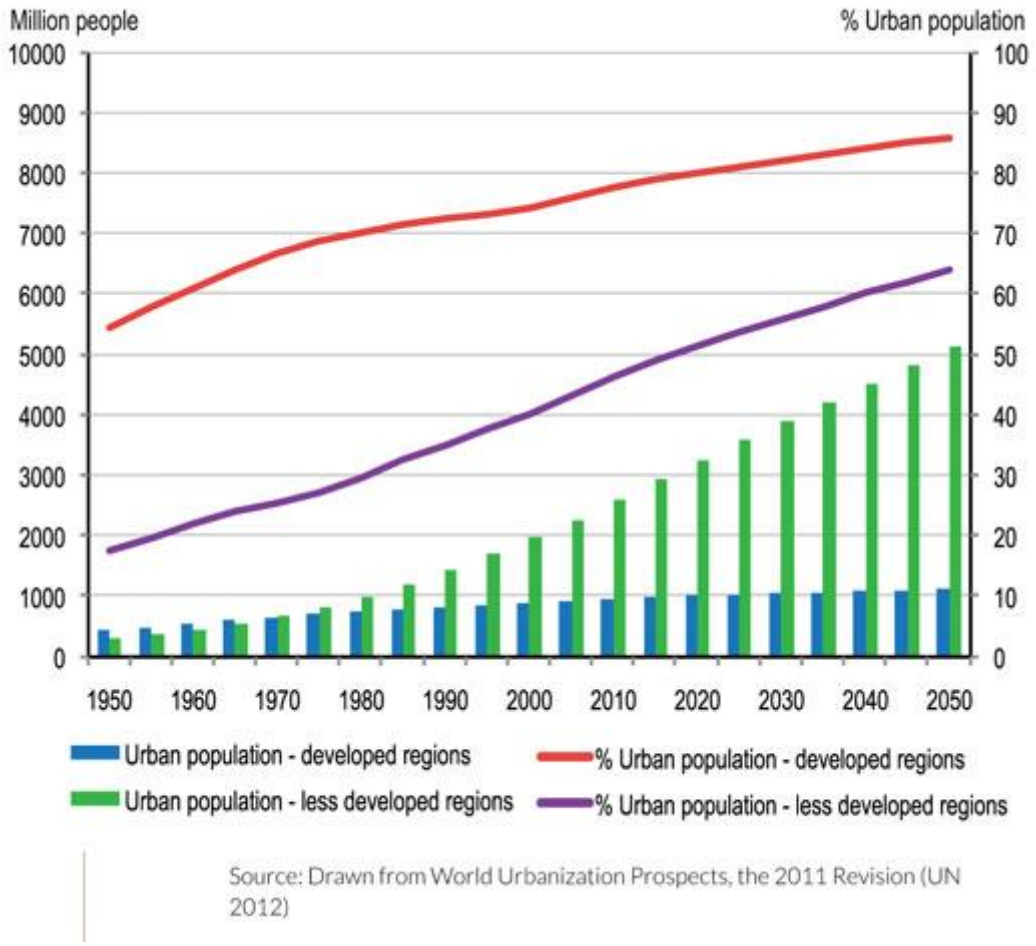


Figure 2. Contribution of urban and rural population growth to total population growth 1950–2030

Consequently, this massive increase in urban population in developing countries makes governments, policy makers and civil society organizations face many challenges in different fields; accommodation, poverty, employment, and other administrative issues, and despite they have already reacted to some of them, still no longer enough to come up with this significant increase (UNFPA, 2007).

As one of the most populous regions among the developing countries, GCR (Greater Cairo Region) has witnessed a massive urban growth during the last decades. As a result, critical housing pressures began to appear, as more than 75% of population is living within 30 km from the city center. In addition to this, the region's urban development has been characterized by the rapid expansion of densely populated informal settlements that reached 40% in 2009 (Figure 3) (Kipper and Fischer, 2009).

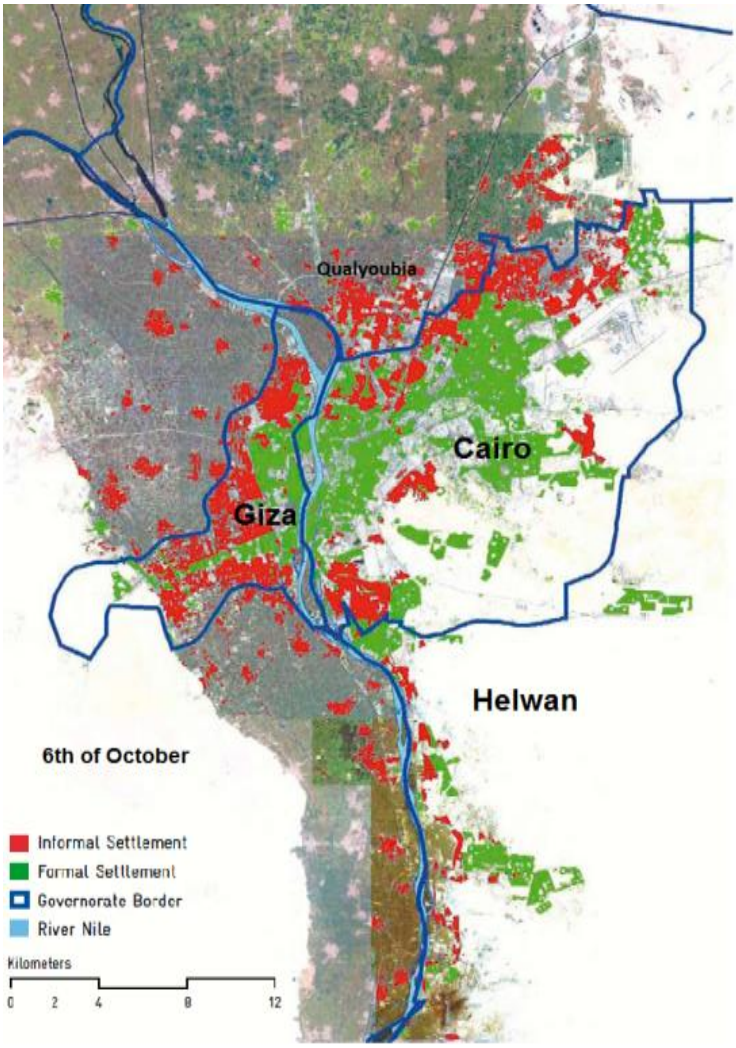


Figure 3. Greater Cairo's legal and illegal urbanization[2]

Consequently, an effective shortage in public services such as water, wastewater, electricity, energy, and other services have been noticed. These problems are interrelated with the uncontrolled urbanization invading GCR nowadays, which requires definitive administrative plans to detect, analyze and estimate its magnitude and extent.

Major and critical LUCC (Land-Use & Cover Change) in areas that contain big cities of high urbanization trends can be described as other type of land-use converting into urban land. Unfortunately, the conventional survey and mapping techniques are expensive and time consuming for urban expansion estimations. Such information is not available for most of the urban centers, especially in developing countries (Huang *et al.*, 2008). Thus, governmental and private research centers have turned to use GIS (Geographic Information Science) and remote sensing tools in monitoring, detecting, and analyzing urban growth. They were found to be cost effective and technologically efficient (Epstein *et al.*, 2002), and in some cases, they can be the only reliable source for a sufficient monitoring. Satellite images provide a synoptic overview for large regions recorded always with a

[2]Helwan and 6th October Cities were officially independent governorates in 2008. In 2011they went back again to be only cities that belong to Cairo and Giza, respectively.

standardized and calibrated monitoring system (Möller, 2005). A historic view back to the past can be viewed mostly for free using image archives that store a large number of satellite remote sensing imagery starting in the early 1970 and increase the data in daily means (Möller, 2005). This data can be analyzed and used for urban growth prediction and modeling purposes.

2. THEORETICAL FRAMEWORK

A simulated LCLU (Land Cover Land Use) map is a result of a sequence of different procedures that precede the prediction phase (Figure 4). Multiple satellite scenes for the same study area, obtained in different time stamps, are classified in order to produce LCLU maps. This classification is validated through an accuracy assessment process which is performed with the aid of validation data (*e.g.* reference maps) to ensure that the classification matches the ground truth classes. The validation is followed by the change detection step in which the amount of each class in time t1 that turned to another class in time t2 is determined. These transitions are recorded in a change matrix which represents the input to the subsequent step; calibrating and modeling the transitions of interest. The previously classified LCLU maps contribute in validating the predictive capacity of the model, which once validated, yields a LCLU map of a future date.

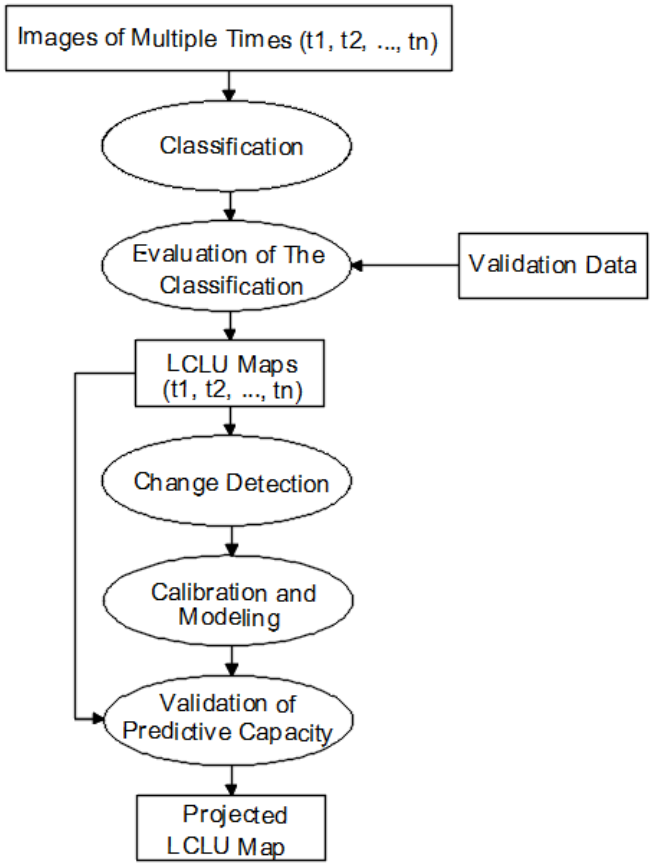


Figure 4. General work approach

2.1 Image Classification

Because classification results are the basis for many environmental and socioeconomic applications (Lu and Weng, 2007), classifying remotely sensed data into a thematic map is an essential step towards further analysis and applications such as LUCC detection and simulation prediction models. Figure 5 represents major steps involved in the image classification procedure (Lu *et al*., 2011).

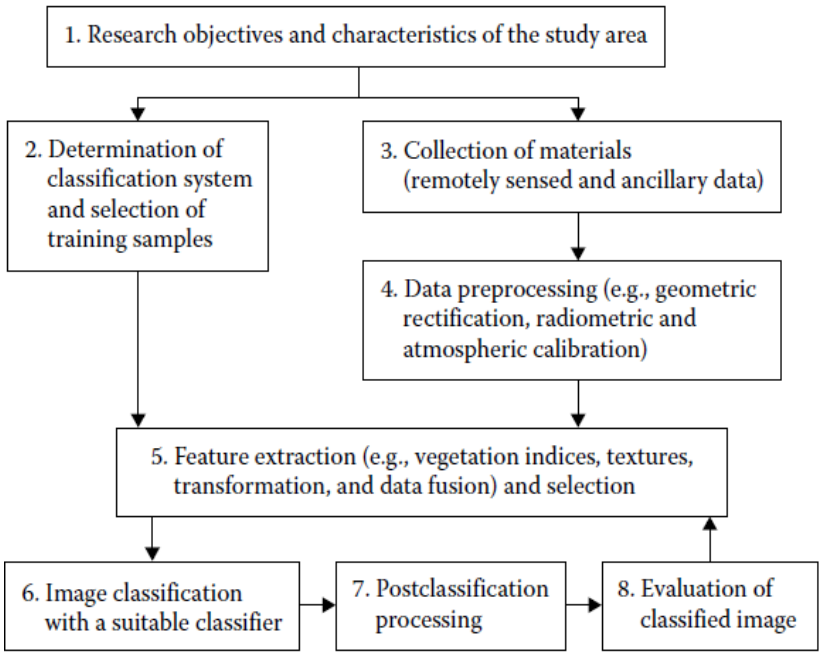


Figure 5. Image classification procedure

The deep awareness of the study area and the study objectives contribute in the determination of most interesting classes, minimum allowed accuracy for each class and for the whole image, MMU (Minimum Mapping Unit), available and required data, and time cost and labor constraints (Lu *et al*., 2011). Classification system should be informative, exhaustive, and separable, based on the user's need and imagery spatial resolution. Training samples should be provided in sufficient number and degree of representativeness, to ensure consistent accuracy assessment process after classification. They are usually collected from fieldwork, or fine spatial resolution aerial photographs and satellite images (Lu and Weng, 2007).

In image prepressing, conversions among different sources or formats are done to avoid geometric, radiometric and atmospheric errors before classification (Lu *et al*., 2011). Feature extraction is done to reduce the data redundancy inherent in remotely sensed data or to extract specific LCLU information depending on some potential variables like spectral signatures, vegetation indices, and ancillary data. Too many variables in a classification procedure may decrease classification accuracy, thus it is recommended to select only variables that are most useful in separation (Lu *et al*., 2011).

Image classification approaches can be grouped into different categories: supervised versus unsupervised based on type of learning, parametric (*e.g.* ML (Maximum Likelihood)) versus nonparametric (*e.g.* decision tree and neural network) based on assumption on data distribution, hard versus soft (fuzzy) based on the number of outputs for each spatial unit, in addition to per-pixel, sub-pixel, and per-field classifications (Lu and Weng, 2007). The selection of the classifier is affected by the image spatial resolution, data sources, classification system, and classification software (Lu *et al*., 2011).

SVM (Support Vector Machines) represent a noticeable development in machine learning research (Pal and Mather, 2005) particularly appealing in the remote sensing field due to their ability to well generalize, even with limited training samples, which is a common limitation for remote sensing applications (Mountrakis *et al*., 2011). SVM are supervised non-parametric statistical learning approach in which a hyperplane is built to separate examples of different classes, maximizing the distance (margin) of the examples lying nearby it (support vectors) (Sáez *et al*., 2013). Figure 6 illustrates a simple scenario of a two-class separable classification problem in a two-dimensional input space (Mountrakis *et al*., 2011).

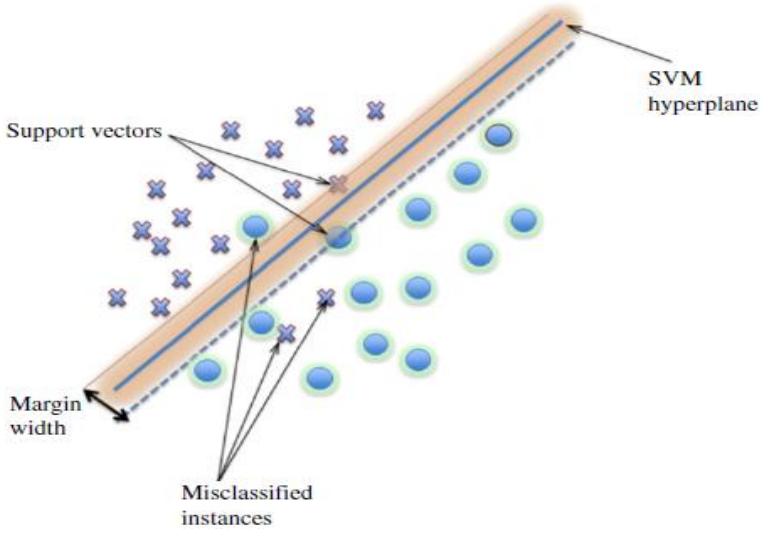


Figure 6. Linear support vector machine example

A better generalization is achieved when the distances from the examples of both classes to the hyperplane are larger (Sáez *et al*., 2013). Pal and Mather (2005) compared SVM with ML and ANN (Artificial Neural Network) algorithms in terms of classification accuracy. The study indicated that SVM can achieve high classification accuracy with high dimensional data, even if the size of the training dataset is small. Benarchid and Raissouni (2013) used SVM to automatically extract buildings in suburban areas in Tetuan city in Morocco, using very high resolution satellite images. 83.76% of existing buildings have been extracted by only using color features. The study stated that the result can be improved by adding other features (*e.g.*, spectral, texture, morphology and context). Singh *et al*.

(2014) applied SVM classifier to estimate the LCLU of Pichavaram forest, in India, using a multi-temporal Landsat images captured in 1991, 2000, and 2009. The classified images recorded high accuracy of 89.4, 94.1 and 94.5%, respectively.

Post-classification process is applied to enhance the classification process previously done. It includes the recoding of LCLU classes, removal of "salt-and-pepper" effects, and the modification of the classified image using ancillary data or expert knowledge (Lu *et al*., 2011). Finally, the classified images are evaluated based on expert knowledge (*e.g*. qualitative evaluation) or sampling strategies (*e.g*. quantitative accuracy assessment) to generate error matrix and other derived accuracy assessment elements, such as overall accuracy, omission error, commission error, and KI (Kappa Index) (Lu *et al*., 2011).

2.2 LUCC Detection

Land-cover refers to the physical characteristics of earth's surface, captured in the distribution of vegetation, water, soil and other physical features of the land, including those created solely by human activities. While land-use refers to the way in which land has been used by humans and their habitat, usually with accent on the functional role of land for economic activities. It is the intended employment of management strategy placed on the land-cover type by human agents, and/or managers (Ramachandra and Kumar, 2004).

Change detection is the process of identifying differences in the state of an object or phenomenon by observing it at different times (Singh, 1989). Generally, it involves the application of multi-temporal data sets to quantitatively analyze the temporal effects of the phenomena of interest. It acts as a main base towards a better interpretation of the relationships and interactions between human and natural phenomena to ensure better resource management and usage (Lu *et al*., 2011). Figure 7 illustrates the main two groups of change detection techniques: binary change/no-change information in which the output has only two possibilities; weather the class has changed on not changed within two specific time stamps. The second approach yields a detailed "from-to" change trajectory which results into a complete matrix with the exact amount of pixels turned from one class into another. This change could easily be presented in any units of area (*e.g.* hectare, squared kilometers) knowing the spatial resolution of the image (Lu *et al*., 2011).

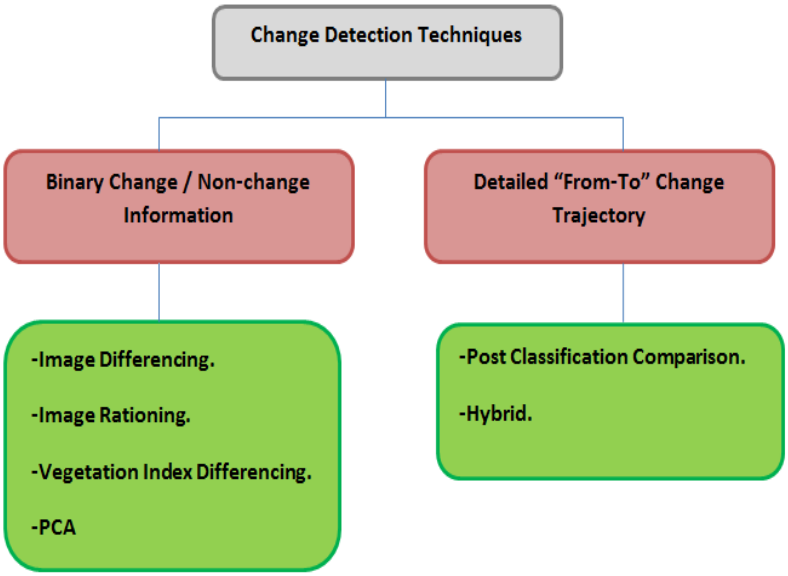


Figure 7. LUCC detection techniques

Singh (1989) described briefly each technique in both groups. In image differencing, spatially registered images of two time stamps are subtracted mathematically, pixel by pixel, to produce a further image which represents the change between both of them. Pixels with high change are found at the end of the histogram that represents pixel distribution in each band, while pixels with no change are grouped around the mean. In regression technique, pixels from time t1 are assumed to have linear relation with those from time t2. It accounts for differences in the mean and variance between pixel values for both dates, whereas in rationing, two images from different dates are rationed, band by band and are compared on a pixel by pixel basis. In areas of change the ratio value would be significantly greater or less than 1, depending on the nature of the changes between the two dates.

The idea of vegetation index differencing is the strong vegetation absorbance in the RED and strong reflectance in the NIR (Near Infrared) part of the spectrum, thus vegetation is more likely to appear darker in the RED part and brighter in the NIR part. There are several vegetation indices commonly used in vegetation studies when using Landsat MSS (Multispectral Scanner) data, the ratio vegetation index is one of the popular, which represents the ratio between NIR band 4 and red band 2. So, if the difference between band4/band2 for two images obtained in different time stamps is significant, that is an indication of vegetation change occurrence. PCA (Principal Components Analysis) is similar to image differencing and image regression techniques, except that in PCA two four-band Landsat scenes of the same area but different dates, are treated as a single eight band data set, aiming at reducing the number of spectral components to fewer principal components accounting for the most variance in the original multispectral images.

The problem while applying change or non-change detection techniques is the difficulty to precisely identify the thresholds. Usually methods used to select thresholds lead to external influences on the differences caused by atmospheric conditions, sun angles, soil moistures, and phonological differences

in addition to the threshold itself is highly subjective and scene dependent, depending on the analyst's familiarity with the study area (Lu *et al*., 2011).

On the other hand, post classification comparison technique consists of an independent classification of each image, followed by a thematic overlay of the classifications resulting into a complete "from-to" change matrix of the conversion between each class on the two dates (Tewolde and Cabral, 2011). The problem in this technique is the high effect of errors resulting from image classification on the change map. For example, if two images classified with 80% accuracy might have only a 0.80 x 0.80 x 100= 64% correct joint classification rate (Singh, 1989). Combination image enhancement/post-classification technique was mentioned in (Mas, 1999), where the change image is recoded into a binary mask consisting of areas that have changed between the two dates. The change mask is then overlaid onto the second time stamp image and only those pixels that were detected as having changed are classified in the t2 imagery. A traditional post- classification comparison can then be applied to yield complete "from-to" change information. This method may reduce change detection errors.

A lot of work has been conducted extensively regarding urban growth. Tewolde and Cabral (2011) studied the spatiotemporal LUCC in the Greater Asmara Area – Eritrea. 1989, 2000, and 2009 satellite images were classified by NN (Nearest Neighbor) algorithm using eCognition Developer 8. Overall accuracy and KI for the three classified images were above the minimum acceptable level of accuracy (85%). LUCC detection was performed using post-classification comparison technique. Jin *et al*. (2013) presented a new CCDM (Comprehensive Change Detection Method) for updating the NLCD (National Land Cover Database). It integrates MIICA (Multi-Index Integrated Change Analysis) model and a novel change model called Zone, which extracts change information from two Landsat image pairs. MIICA uses four spectral indices to obtain the changes that occurred between two images of different time stamps. CCDM contains a knowledge-based system, which uses critical information on historical and current land cover trends combined with the likelihood of the land cover to change, in order to gather the changes from MIICA and Zone. CCDM was recommended because of its simplicity and high capability of capturing disturbances associated with land cover changes. On the other hand, CCDM suffers some limitations, as it detects change for only one class, in addition, it can produce certain amount of commission errors. Haas and Ban (2014) investigated land cover changes in China's three largest urban agglomerations: JJJ (Jing-Jin-Ji), YRD (Yangtze River Delta) and PRD (Pearl River Delta). Six images (two images per region, in 1990 and 2010) were classified using random forest decision tree ensemble classifier. The average overall accuracy for JJJ, PRD, and YRD for 1990 and 2010 images were 85% and 87%, while KI was 0.83 and 0.86, respectively.

To evaluate LUCC detection results, error matrix-based accuracy assessment method is very commonly used, however, Lu *et al*. (2011) reviewed other methods from previous studies, such as: accuracy assessment curves, area-based accuracy assessment for the change maps analysis, and accuracy assessment of remotely sensed-derived change detection monograph.

2.3 LUCC Modelling

Models are simplifications of reality, they are theoretical abstractions that represent systems in such a way that essential features crucial to the theory and its application are identified and highlighted (Batty, 2009). LUCC models are tools to support the analysis of the causes and consequences of LUCC for better understanding of the system functionality, and to support land-use planning and policy. Models are useful for simplifying the complex suite of socioeconomic and biophysical forces that influence the rate and spatial pattern of LUCC and for estimating the impacts of changes (Verburg *et al*., 2004).

Verburg *et al*. (2004) listed six important concepts worth to be taken into consideration while modelling LUCC. They are: level of analysis, cross-scale dynamics, driving forces, spatial interaction and neighborhood effects, temporal dynamics, and level of integration.

Level of analysis is directly related to the used perspective; micro-level perspective, which relies on simulation of individuals behavior, or macro-level perspective, that relies on macro-economic theory or apply the systems approach. Both perspectives refer to the issue of scale, which is known by extent and resolution. Extent is the magnitude of a dimension used in measuring (*e.g.* area covered by a map), while resolution is the precision used in this measurement. Several LUCC models are structured hierarchically, so multiple levels are taken into consideration. Pure cellular automata models determine the number of cells that change in each step of the simulation based on cellular dynamics. Figure 8 represents the different levels of top-down allocation procedure (Verburg *et al*., 2004).

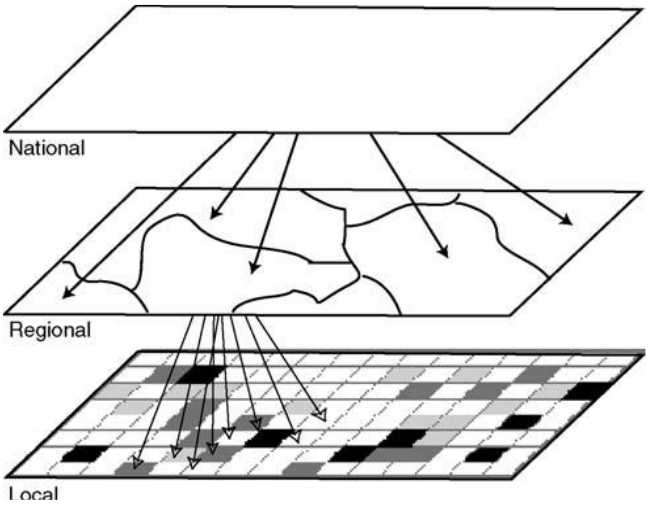


Figure 8. Top–down allocation procedure

Driving forces are the socioeconomic and biophysical variables that affect LUCC, directly or indirectly and influence allocation decisions. Distance to road networks, land slope, distance to water bodies, land value, and transportation conditions are just few examples of these factors. Different driving forces are used in case of modeling in different scales of analysis (local and regional levels). Generally, relations

between land-use and driving forces can be determined using empirical methods or expert knowledge (Verburg *et al*., 2004).

Spatial autocorrelation exist as a result of spatial interactions between land-use types themselves (*e.g.* urban expansion is often situated in areas next to the already existing ones). Spatial autocorrelation in land use patterns is scale dependent, for example in small scale, residential areas have a positive spatial autocorrelation, while in a larger scale, a parcel's probability of development decreases as the amount of existing neighboring development increases (negative spatial autocorrelation), due to the negative spatial effects generated from the development, such as crowd (Verburg *et al*., 2004). Spatial connectivity also occurs as a result of spatial interactions acting over larger distances (*e.g.* land-use in the upstream part of a river affects the land-use in the downstream of the same river). Cellular automata is a common method that considers spatial interactions, in which the state of a pixel is calculated based on its initial state, the structure of the surrounding pixels (Figure 9), and a set of transition rules (Verburg *et al*., 2004).

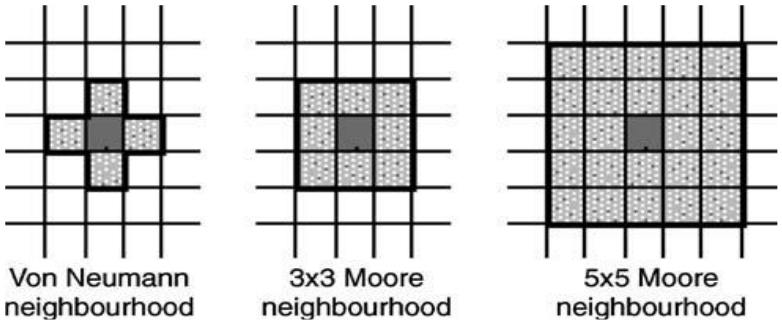


Figure 9. Alternative neighbourhoods used in cellular automata models

The temporal dimension of the model is related to model validation, which is most often based on the comparison of model results for a historic period with the actual changes in land-use as they have occurred. Such a validation makes it necessary to have land-use data for another year than the data used in model parameterization. The difference between the two timestamps used in model validation should be the same of the difference for which future scenario simulations are made. Finally, the determination of integration level is important for LUCC modelling, which usually controlled by model scale and purpose (Verburg *et al*., 2004).

Brown *et al*. (2013) identified five types of modeling approaches: machine learning, cellular sector-based economic, spatially disaggregated economic and agent-based approaches. Machine learning approaches rely on algorithms that encode relationships between LUCC and the characteristics of locations where they are most likely to occur. Cellular approaches model and simulate temporal changes by combining maps of likelihood based on relationships of observed patterns. Sector-based economic models focus on inputs and outputs, and trade among regions and sectors, to identify the demand for land of different types. Spatially disaggregated economic models are designed according to microeconomic theory to help investigators to understand LUCC as an outcome of individual decisions.

Agent-based models are defined at a micro-level and can consist of one or more types of agents (individuals or institutions), as well as an environment in which the agents are embedded. Therefore, systems can be studied at many scales and parts (Brown *et al*., 2004).

MC-CA (Markov Chain – Cellular Automata) integrated model is one of the most common tools for modelling urban land-use expansion. CA is a spatially explicit model which relies on the iteration of a given dimensional cell based on supporting socio-economic and geographical data, to change into urban or nonurban form within a given time frame. On the other hand, MC model determines the actual amount of change between land use categories non-spatially, in other words, MC is a stochastic process model that describes the probability that one state (*e.g.*, cropland) changes to another state (*e.g.*, built-up areas) within a given time period, while CA determines where this conversion to urban will take place (Vaz *et al*., 2012). In CA, there are cellular entities that independently vary their states, as well as their immediate neighbours, according to predefined transition rules. Various ways of defining transition rules make CA models function differently and, consequently, produce dissimilar outputs (Arsanjani *et al*., 2013).

Selecting the driving forces that affect LUCC is essential for modeling process, on which transitions rules rely to determine the cell state, varying along with the variation of its neighborhood states. MCE (Multi-Criteria Evaluation) is one of the popular methods for selecting the most suitable driving forces that are believed to highly effect urban growth in the future. AHP (Analytic Hierarchy Process) is one approach that allows weighting of land-use transition potential on the basis of a set of potential maps. It incorporates growth constraints and determines the weights of the (fuzzy) potential maps by means of pairwise assessments. Values are standardized from 0 to 1 indicating least and most suitable sites, respectively. The weighting parameters are usually determined by expert knowledge or qualitative interviews, and consistency ratio are used to verify meaningless of the selected weights, consequently, the suitability of the weighting schema. Sigmoid, J-shaped, and Linear are examples of membership functions to determine degree of suitability between the control points (Moghadam and Helbich, 2013).

Gong *et al*. (2015) used MC-CA model in IDRISI, for urban growth prediction in 2007 using classified imageries of years 1989, 2001 and 2007 for Harbin, China. They used MCE module to determine driving forces, some of them were soil properties, elevation, slope, rainfall, water, population density, accumulative temperature, distance from roads and settlements. In that way suitability maps for each land-use classes were produced. The VALIDATE module results in KI with values above 0.8, when compared to the actual maps of 2007.

Another well-known modeling tool is LCM (Land Change Modeler), which is embedded as well in IDRISI software (Tewolde and Cabral, 2011). Only thematic raster images with the same land cover categories listed in the same sequential order can be inputted in LCM for analysis, and background areas must be identified on maps coded with 0. It evaluates land cover changes between two different times, calculates the changes, and displays the results with various graphs and maps. Then, it predicts future LCLU maps on the basis of relative transition potential maps (Roy *et al*., 2014).

Tewolde and Cabral (2011) applied LCM for the urban growth prediction in Asmara, Eritrea, in 2020 using 1989, 2000, and 2009 satellite images, previously classified in the same study. MLP (Multi-Layer

Perceptron) neural network was applied to generate transition potential maps, which were used with MC modeler and transition probability grid to predict year time-t (2009). The driving forces for the study were distance to existing urban areas and distance to road network. The accuracy of 2009 simulated map was examined against the previously classified one of the same year, with an average KI of 83%.

Roy *et al*. (2014) used LCM to predict LUCC in 2011 in a Mediterranean catchment in South Eastern France. The predicted images were compared to the real 2011 map. Different times were used for predictions: short (2003-2008), intermediate (1982-2003), and long (1950-1982). Driving forces were selected by Cramer's coefficient. However, altitude, slope, and distance from roads had the greatest impact among other tested variables. The results indicated that shorter time scales produce better prediction accuracy. Stable land covers are easier to be predicted than cases of rapid change, and quantity is easier to be predicted than location for longer time periods.

Vega *et al*. (2012) compared biodiversity loss modelling results in western Mexico using a combined unsupervised / supervised approach using two spatially explicit models: DINAMICA model that uses the Weights of Evidence method, a supervised approach in which the weights can be selected and edited by a user, and LCM which relies upon neural networks. DINAMICA had better results at the per transition level, but the overall change potential map generated using LCM is more accurate, because neural networks outputs are able to express the change of various land cover types more adequately than individual probabilities obtained through the weights of evidence method.

LUCC models are tested through three different processes: calibration, validation, and verification, in order to prove and confirm that the theory matches with facts. Calibration is the process of dimensioning a model in terms of finding a set of parameter values that enable the model to reproduce characteristics of the data in the most appropriate way. This is different from validation, which seeks to optimize a model's goodness of fit to data, such as how close the predictions are to the observed data. It could be measured by KI and / or sensitivity analysis. On the other hand, model verification accords to testing the model for internal consistency and is often separate from testing how good the model's predictions are (Batty, 2009).

3. STUDY AREA

Egypt is located in the northeastern corner of the African continent. It is bordered by Libya to the west, Sudan to the south, the Red Sea to the east, and the Mediterranean Sea to the north, with an approximate area of one million square kilometers (Figure 10).

Figure 10. Egyptian borders

Egypt has the largest and most densely settled population among the Arab countries. It has a total inside population of 86 million with a density of 1100 capita / km^2 in populated areas, and 8 million of outside population (CAPMAS, 2014). The selected area for the study is the metropolitan area of Cairo; the political capital and Qalyubiyah city, beside parts of Giza and Sharqiyah, cities that belong to GCR (Figure 11). The study area is located at 30° 02' N and 31° 21' E, in the middle of Delta Region, and covers an area of 8 942 km^2

Figure 11. Study area: Greater Cairo – Egypt

GCR is considered one of the fastest growing mega cities worldwide, with the highest population and population density among other Egyptian governorates (SIS, 2014) (Figure 12). Cairo city was the most populous among other Egyptian cities in 2013 (SIS, 2014), with almost 9 million capita, representing 10.7% of total population recorded in the same year.

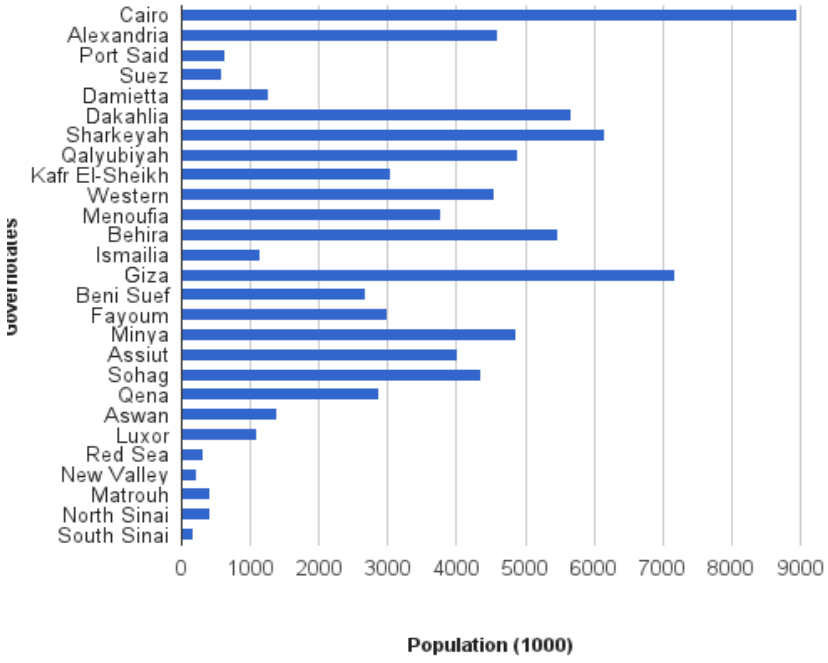


Figure 12. Egypt governorates population in 2013

The internal migration from Upper Egypt and the Delta began just after the Second World War, especially in the middle of 1950s due to the massive industrialization policy launched by President Nasser, which encouraged people to migrate, seeking better job opportunities and higher quality of

education (Kipper and Fischer, 2009). In 1947, GCR hosted around 3 million representing 12.5% of total Egyptian population at that time. This number kept growing to 13 million in 1996, representing 17.3% of total Egyptian population. (Figure 13) (Kipper and Fischer, 2009). In 2006, the total population in GCR reached 16.1 million capita (CAPMAS, 2014).

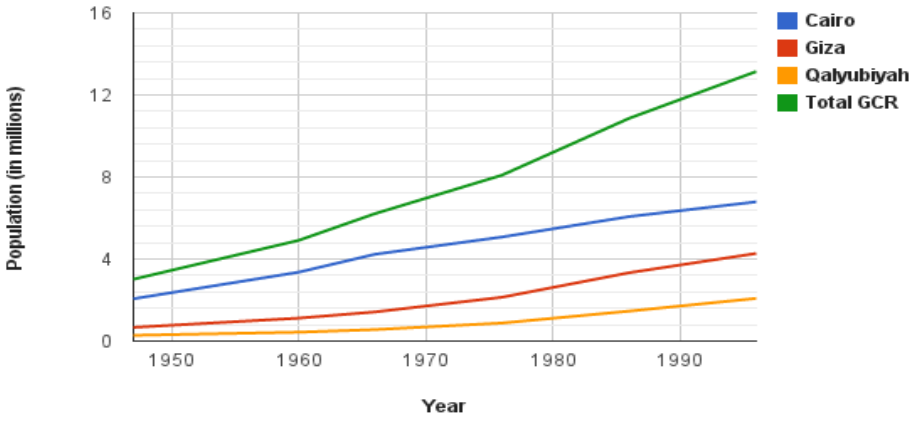


Figure 13. GCR's population growth

Population data about GCR is always a problematic issue, due to the change of administrative boundaries of the governorates, thus its population vary considerably, depending on the geographical definition used to identify the region. Tables 1, 2 and 3 obtained from (CAPMAS, 2014) list the districts in Giza, Qalyubiyah, and Sharqiyah cities that belong to GCR, respectively, based on the latest official geographic boundaries declarations.

District	Qism / Markaz
Dokki	Qism
Ahram	Qism
Aguza	Qism
Badrashin	Markaz
Hawamdeyah	Qism
Giza	Qism
Giza	Markaz
Omraneyah	Qism
Warrak	Qism
Sheikh-Zayed	Qism
6th October First	Qism
6th October Second	Qism

Table 1. Districts belong to GCR in Giza city

District	Qism / Markaz*
Qanater AlKhayreya	Markaz
Obour	Qism
Ossim	Markaz
Bulak AlDakrour	Qism
Imbaba	Qism
Kirdasa	Markaz

Table 2. Districts belong to GCR in Qalyubiyah city

District	Qism / Markaz
Qalyoub	Qism
Qalyoub	Markaz
Shubra AlKhima 1st	Qism
Shubra AlKhima 2nd	Qism

Table 3. Districts belong to GCR in Sharqiyah city

Different studies have been carried out previously for LUCC detection and modelling in GCR. Yin *et al*. (2005) used ISODATA clustering procedure for image classification and image differencing technique for the LUCC detection between 1986 and 1999 with an overall accuracy of 87% for both images. The study indicated that urban areas increased from 344.4 km^2 in 1986 to 460.4 km^2 in 1999. At the same time, the population density increased from 7 158 person / km^2 to 9 074 person / km^2. The spatial pattern distribution of urban and population was compared in order to reveal the relation between urban areas and population. It was found that population per unit of urban land decreased from 27.188 person / km2 in 1986 to 25.799 person / km2 in 1999.

Mohamed (2012) used ML classifier using ERDAS for the LUCC detection in 1973 and 2006 with an overall accuracies of 92% and 86%, and KI of 0.87% and 0.78, respectively. The study applied post-classification comparison technique for LUCC detection. Results showed that urban areas expanded from 223.8 km^2 in 1973 to 557.9 km^2 in 2006, with total agricultural cut-offs and urbanized desert of 136.7 km^2 and 187.3 km^2, respectively.

Vaz *et al*. (2011) used Landsat images in years 1984, 2000 and 2008 for image classification using nearest distance classifier, and pixel quantification technique for LUCC detection. They used MC-CA integrated model with MCE procedure to select and evaluate the driving forces: distance from roads, slope, proximity to built-up areas, and elevation. For model validation, 2008 map was produced and compared to actual one from Google maps. Model accuracy was 88% indicating the validity of model parameters, and the model to predict urban growth in 2038.

4. METHODS

Figure 14 illustrates the flowchart of the methodology that was applied during the study. Three satellite images of different time stamps: 1984, 2003 and 2014 were classified using supervised-nonparametric SVM classifier. For accuracy assessment, random points over the study area were generated and visually compared to Google Earth images to validate the maps. On the other hand, for LUCC detection, a hybrid technique was applied, in which binary maps (areas of change / no change) for both periods 1984-2003 and 2003-2014 were produced using NDVI (The Normalized Difference Vegetation Index), NIR, and RED difference method. Then only areas of change were thematically overlaid to produce the change matrix in both periods; 1984 – 2003 and 2003 - 2014. The transitions between 1984 and 2003 were modelled using LCM, after the driving forces had been set, to produce a predicted map of 2014 that was compared to the LCLU map of the same year, in order to validate the model. Once validated, the model was used with the same driving forces and the same prediction time difference (11 years) using 1984 and 2014's LCLU maps to predict the LCLU map of 2025.

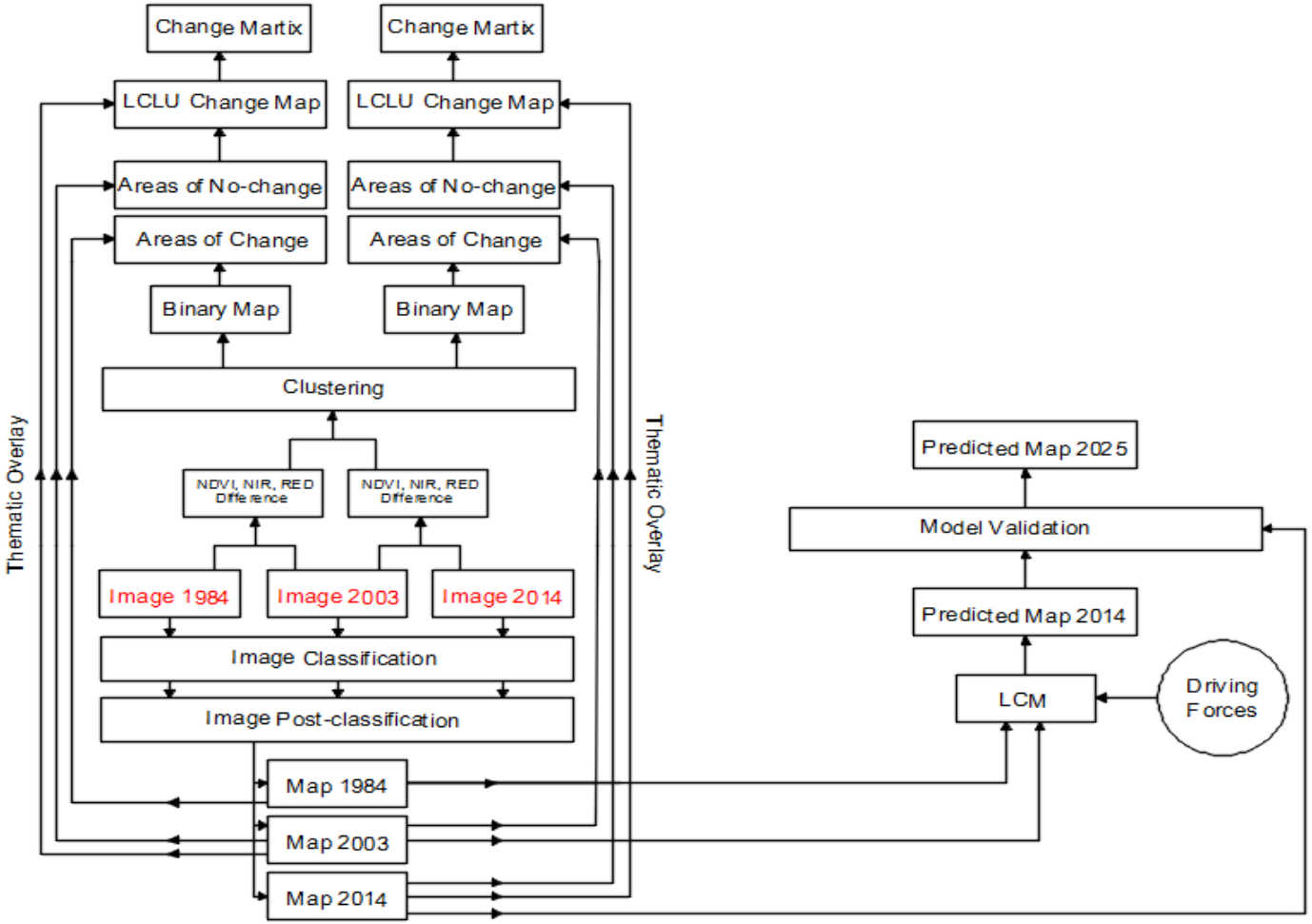


Figure 14. Methodology

4.1 Data

Three cloud free satellite imageries for years 1984, 2003, and 2014 that cover the entire study area were obtained from Landsat scenes available for free download on EarthExplorer, accessible on http://earthexplorer.usgs.gov/ . Further information about the imageries is illustrated in Table 4, which appears below.

Acquisition date	Sensor	Spatial resolution	Path/Row	Landsat	Number of bands[3]	Radiometric resolution
15/03/2014	OLI - TIRS	30 m	176/39	Landsat 8	11	16 bits
07/07/2003	TM	30 m	176/39	Landsat 5	7	8 bits
02/07/1984	TM	30 m	176/39	Landsat 5	7	8 bits

Table 4. Imageries attributes

4.3 Binary Maps Production

To produce maps of change / no change, NDVI difference technique was first applied. The difference in NDVI between same satellite images obtained in different dates is a common method to produce binary maps of change / no change information. Radiometric normalization between same images recorded with the same sensor, but in different times, is essential for change detection, because theoretically, they are assumed to store similar digital levels, but practically, they do not, due to the effect of different atmospheric conditions and sun geometry from different recorded dates, so that pixels from the same terrain can show different radiance values, and, therefore, different values in their digital levels (Broncano *et al*., 2010). Linear normalization is the technique that was applied in NDVI normalization. Afterwards, two NDVI differences were calculated to produce the binary maps of the change / no change occurred between 1984 – 2003 and 2003 – 2014 as well. For thresholds determination, firstly, around 500 points of change and no change per each difference where collected by the visual comparison of each pair of images; 1984 with 2003, and 2003 with 2014, in order to adjust the accuracy of the thresholds values. These thresholds determine the range of NDVI differences mandatory to decide whether the pixel had been subjected to change or not (Figure 15). After this, thresholds were tuned to match most of the change / no change points, previously collected.

[3]Resolution of each band is available from: http://landsat.usgs.gov/band_designations_landsat_satellites.php

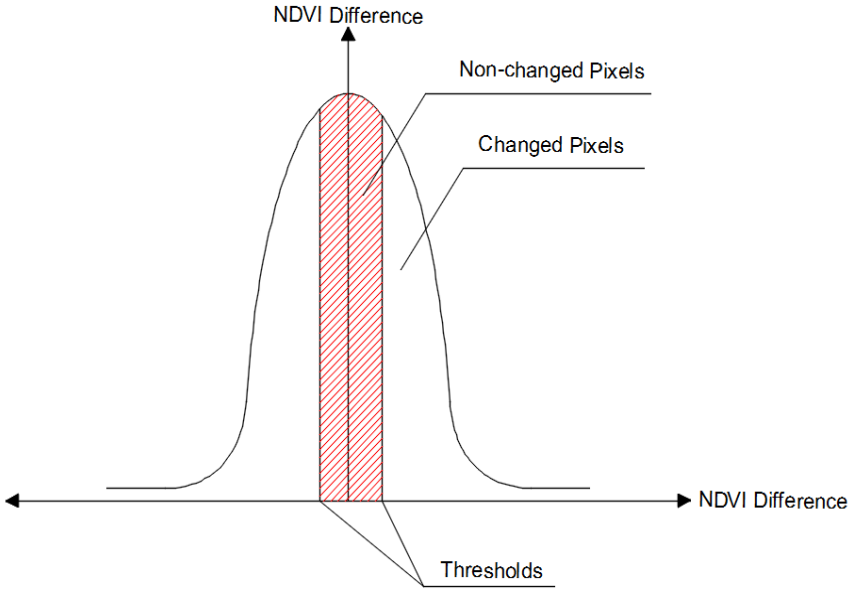


Figure 15. Thresholds determination

4.4 Image Classification

To determine LCLU classes, urban was considered as an essential class because it represents the main interest and focus of the study. For simplicity, three non-urban classes were considered: Vegetation, Desert and Water. Agriculture is one of the main activities in the study area which is concentrated around the Nile River banks and other secondary water streams. Figure 16 shows the steps that had been followed to determine the areas in which the training samples were collected.

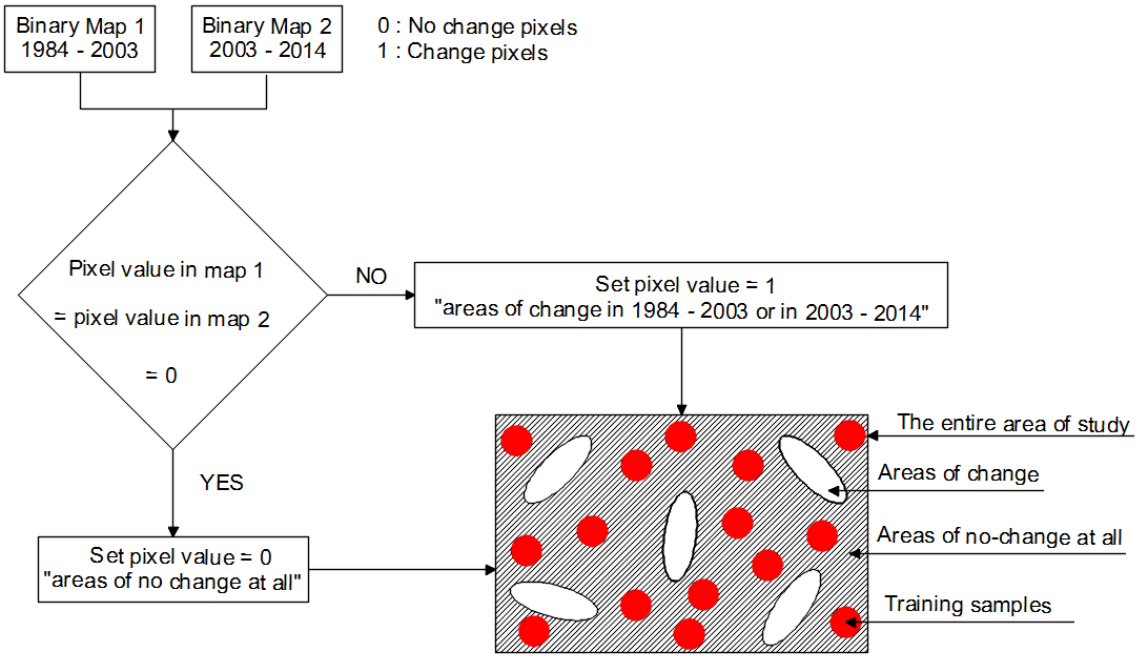


Figure 16. The applied procedure in collecting training samples

First, training samples were collected for all classes, based on 2014 image, in areas of no change over the study period, 1984 – 2014. In that way only one training dataset was used in the classification of the three images, to save effort and time.

Afterwards images were classified using SVM classifier, with the LIBSVM[4] (SVM Library) enabled in Matlab 8.3. SVM is a supervised non-parametric statistical learning approach, meaning that it has multiple parameters, hence, the classification was carried out in several trials per image, each with different values of the same parameter, aiming at producing maps that best represent the reality. Table 5 gives brief descriptions of the different parameters used to obtain the best representative maps, while Table 6 shows the optimum values of each parameter.

Parameter	Description
s	SVM type: 0 for C-SVC, 1 for nu-SVC, 2 for one-class SVM, 3 for epsilon-SVR (regression), 4 for nu-SVR (regression) type.
t	Kernel function: 0 for linear, 1 for polynomial, 2 for exponential, 3 for sigmoid, and 4 for pre-computed kernel. All except linear, are functions of gamma, coefficient, and degree.
d	Degree of the kernel function.
g	Gamma of the kernel function.
r	Coefficient of the kernel function.
c	Cost parameter for C-SVC, epsilon-SVR, and nu-SVR types.

Table 5. Parameters description[5]

[4]Available from: http://www.csie.ntu.edu.tw/~cjlin/libsvm/
[5] Data Source: SVM library documentation.

Image	s	t	d	g	r	c
2014	0	1	3	0.25	1	1
2003	0	1	3	0.25	1	1
1984	0	0	-	-	-	1

Table 6. Optimum parameter values

Areas of no change were only classified once in order to guarantee time consistency. For each image, areas of change were classified separately by using the training samples defined on the no-changed areas.

4.2 Image Pre-processing

This phase generally includes geometric, atmospheric, and radiometric corrections, however both, geometric and atmospheric corrections were skipped. Firstly, satellite scenes from Landsat are usually offered with geometric correction done, secondly, all three images have clear sky and there was no need for an absolute atmospheric correction.

4.5 Image Post-classification

Image post-classification focus on enhancing the quality of the produced maps to be more representative to the landscape. Each raster map was converted to map of polygons, and then was generalized to a MMU of one hectare, by selecting all polygons of an area less than one hectare and eliminating them. In the elimination process, each selected polygon is merged to the largest adjacent polygon. Finally, the maps were converted to raster again.

4.6 Accuracy Assessment

To validate the maps, 100 random points per class were generated for each map and visually classified via Google Earth maps as ancillary data. The “Historical View” tool in Google Earth engine was used to validate the old maps; 1984 and 2003.

Figure 17 shows an example of how the base error matrix, as a method of evaluating the classification, is done. The values in the diagonal represent the number of points in which its classification matches the reference dataset, whereas the other values are the number of the misclassified points. The accuracy of any class is the probability that the classifier has labeled an image pixel into the reference dataset, or in other words, it is the probability of a reference pixel being correctly classified. There are two types of error: commission and omission, the first represents pixels that belong to another class but are labeled as belonging to the class, while the second represents pixels that belong to the truth class but failed to be classified into the proper class. The overall accuracy is the total classification accuracy that is obtained by dividing the summation of the diagonal (correctly classified points) by the total number of testing points.

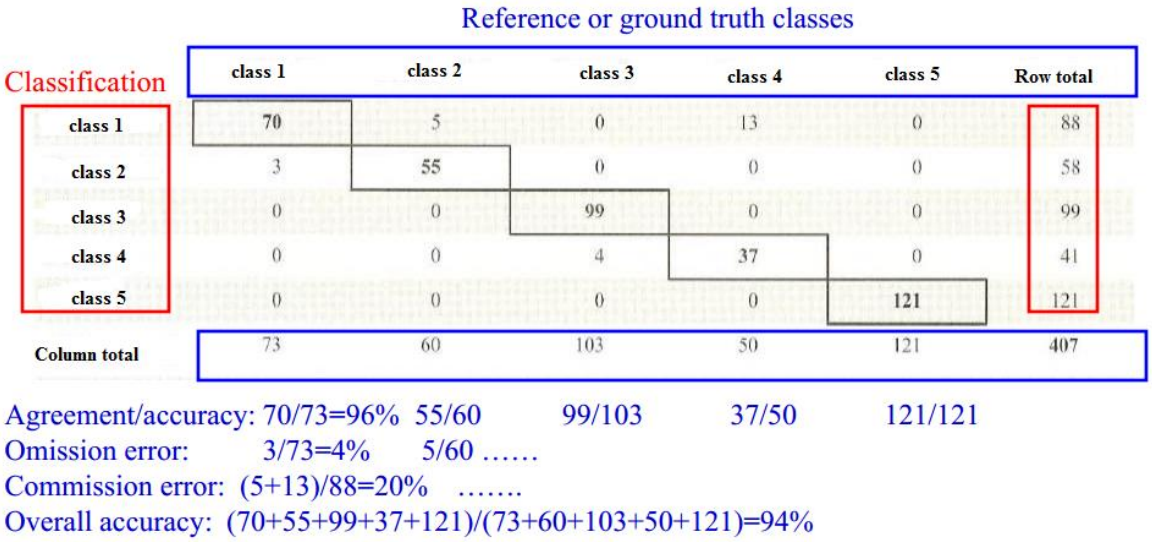


Figure 17. Accuracy assessment example

4.7 LUCC Detection

For LUCC determination, only areas of change were thematically overlaid with each present map; areas of change between 1984-2003 were overlaid with 2003 map, likewise, areas change between 2003-2014 were overlaid with 2014 map, thus change matrices were produced, giving information about the amount of land use that turned from a class to another.

4.8 Analysis of Spatial Urban Growth Pattern

For better urban sprawl interpretation, FRAGSTATS software, version 4.2, was used to calculate some statistical spatial metrics for the urban class over 30 years, from 1984 to 2014, based on the LCLU maps of 1984, 2003 and 2014. The selected subset of matrices which was applied in the study is given in Table 7. They are the most commonly used and explored matrices in similar studies (*e.g.* Araya *et al*., 2010; Herold *et al*., 2003).

Metrics	Description	Units	Range
CA—Class Area	The sum of the areas of all urban patches, that is, total urban area in the landscape.	Hectares	CA > 0, no limit
NP—Number of Patches	The number of urban patches in the landscape.	None	NP ≥ 1, no limit
ED—Edge Density	The sum of the lengths of all edge segments involving the urban patch type, divided by the total landscape area.	Meters/ m^2	ED ≥ 0, no limit

LPI—Largest Patch Index	The area of the largest patch of the corresponding patch type divided by total area covered by urban.	%	$0 < LPI \leq 100$
ENN_MN—Euclidian Mean Nearest Neighbor Distance	The distance mean value over all urban patches to the nearest neighboring urban patch, based on shortest edge-to-edge distance from cell center to cell center.	Meters	EMN_MN > 0, no limit
AWMPFD—Area Weighted Mean Patch Fractal Dimension	Area weighted mean value of the fractal dimension values of all urban patches, the fractal dimension of a patch equals two times the logarithm of patch perimeter divided by the logarithm of patch area; the perimeter is adjusted to correct for the raster bias in perimeter.	None	$1 \leq AWMPFD \leq 2$
CONTAG—Contagion	Measures the overall probability that a cell of a patch type is adjacent to cells of the same type.	%	$0 < CONTAG \leq 100$

Table 7. Spatial metrics

4.9 LUCC Modelling

LCM is an integrated software environment within IDRISI for the analysis of LUCC. It is a powerful tool to assess historical land cover data and to use that assessment to predict future scenarios. LCM embedded in IDRISI 17.0 was used in this study to predict the LCLU map in 2025 through the procedure shown in Figure 18.

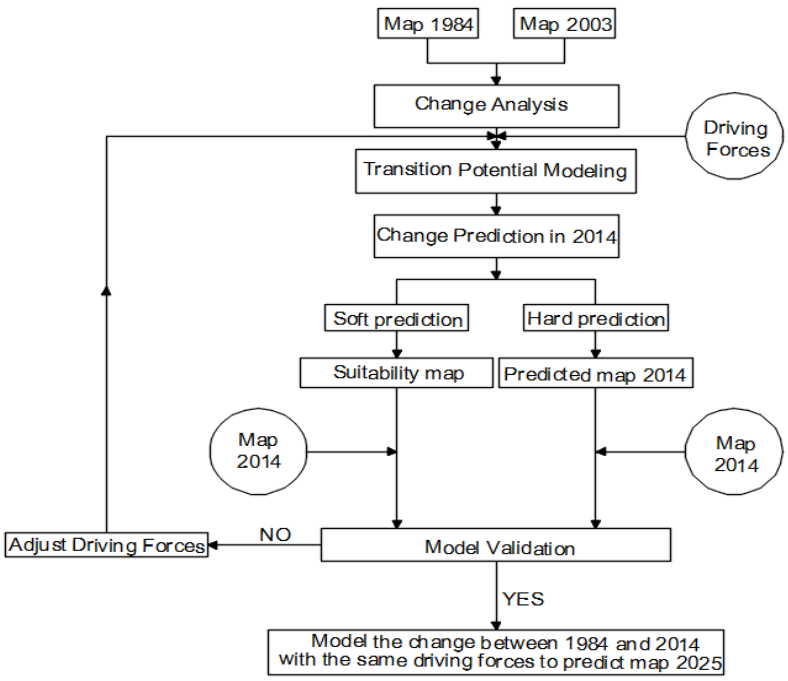


Figure 18. Modelling Process

Both LCLU maps of 1984 and 2003 were the input to the LCM to predict a map of 2014 and then compare it with 2014's LCLU map to validate the model. In the first step, the change analysis process was carried out, in which the changes were assessed between 1984 and 2003. These changes represent the transitions from one class to another, which are important to identify the dominant transitions to urban and target them for modeling.

The second step was the transition potential modeling using MLP neural network. This step is responsible for determining the location of the change. It results into a number of transition potential maps equal to the significant transitions to urban, considered in the first step (Eastman, 2012). These transition potential maps represent the suitability of a pixel to turn to urban in each transition based on a group of factors, named "Driving Forces", that are used to model the historical change process. This step is usually managed through a transition sub-model that contains a group of all LCLU transitions, in case they are thought to have the same underlying driving forces. Otherwise, multiple transition sub-models will be used, containing a single LCLU transition, each with different group of driving forces. The driving forces can be either static or dynamic. The static variables are those which remain unchanged over time, while the dynamic ones are time-dependent drivers, consequently they are recalculated over the prediction period (Eastman, 2012). LCM provides an optional quick test of the potential explanatory power of each driving force represented by Cramer's V. This value varies from 0 to 1, indicating a discarded variable and an excellent potential one, respectively. Although this test is not precise, it acts as a guide to determine whether the driving force is worth to be considered or not (Eastman, 2012). Once the variables are selected, each transition sub-model can be modeled.

In this way, the model was calibrated, as the parameter values (driving forces and transition potential maps) which enable the model to reproduce characteristics of the data appropriately, were determined. The model was ready for the third step; future scenarios prediction. In this step, LCM uses the change rates calculated from the first step as well as the transition potential maps produced from the second step, to predict a future scenario for 2014. This step is responsible for determining the quantity of change to urban in each transition in 2014 using MC analysis, in which the class of a pixel is determined by knowing its previous state in 2003 and the probability of transitioning from it to urban. It figures out exactly how much land would be expected to transition from 2003 to 2014 based on a projection of the transition potentials into the future.

There are two basic types of prediction: the hard and soft predictions (Eastman, 2012). The hard prediction yields a projected map of 2014, where each pixel is assigned one land cover class, the class that is most likely to change to. The soft prediction is however different, it produces a vulnerability map in which each pixel is assigned a value from 0 to 1 indicating the probability of the pixel to turn to urban in 2014. The hard prediction yields only a single realization, while the soft prediction is a comprehensive assessment of change potential.

The validation process aims to determine the quality of 2014's predicted map in relation to 2014's LCLU map (the map of reality). There are two endorsed approaches to validate a model: the visual and statistical approaches. The visual examination is the quickest way to reveal spatial patterns, which the

statistical method may fail to detect. However, it is subjective and can be misleading, therefore the statistical approach is essential to be carried out (Pontius *et al*., 2006).

In the visual validation, a 3-way cross tabulation between 2003's LCLU map, 2014's predicted map, and the map of reality was run to illustrate the accuracy of the model results. The output is a map of four categories:

1- Hits: Model predicted change and it occurred in reality.
2- False alarms: Model predicted change to urban while it persisted in reality.
3- Misses: Model predicted persistence and it changed to urban in reality.
4- Null success: Model did not predict change and it did not occur in reality.

False alarms and misses represent the errors that resulted from the model as a disagreement between the simulated map (2014's predicted map) and the reference map (map of reality), while hits and null success represent the model correctness.

On the other hand, there are an infinite number of different ways to compare maps statistically. They varies in terms of ease of interpretation, offering a valuable output which can contribute in the enhancement of the model because the purpose of the assessment is to find ways to improve the method of making the predicted map agrees more closely with the map of reality, and showing the similarities between the predicted and the reference maps (Pontius *et al*., 2006). A good statistical validation model should show the degree of agreement between both maps, in each LCLU class, in terms of quantity of cells and location of cells. IDRISI enables a hard statistical validation using the VALIDATE module, and a soft statistical validation using the ROC (Relative Operating Characteristic) module.

The VALIDATE module examines the agreement between two pair of maps that show any categorical variable, which can have any number of categories (Pontius *et al*., 2006). The map of reality acted as the reference map, while 2014's predicted map was the comparison map. On the other hand, the ROC module is more concerned about examining only the urban concentration in areas of relatively high suitability. In other words, it focuses on how well the suitability map predicts the locations of new urban settlements in 2014, based on the changes in urban between 1984 and 2003. In measures the maps agreement in terms of cells locations in a certain LCLU class rather than their quantity in each class.

To check the agreement between the new urban areas in 2014's predicted map and the map of reality in the ROC approach, the reference map was a Boolean map of new urban development in which all cells of new urban settlements (since 2003) in 2014's LCLU map were assigned a value of 1, whereas 0 elsewhere. The comparison map was the suitability map resulted previously from the LCM with a suitability of 1 for the urban cells in 2003's LCLU map, to ensure the highest likelihood for them to remain urban in 2014. The ROC module considers many possible thresholds, for each one, it reclassifies the suitability map so that the cells of a value higher than the threshold are reclassified as predicted urban, while the remaining cells are reclassified as predicted persistence (Pontius *et al*., 2006). Then, for each threshold, the ROC examines the agreement between the reference map and the

reclassified suitability map and plots the results into the ROC curve, where the AUC (Area Under Curve) represents the overall agreement between the two maps. An AUC value of 0.5 indicates complete randomness between the two maps, whereas a value of 1 indicates perfect spatial agreement (Olmedo *et al*., 2013).

Although this is the most common type of ROC analysis, its restricted focus on predicting a certain type of change (urban development) by masking parts of the study area may be misleading, as it can hide weaknesses or strengths of the model. Hence a ROC analysis of a null model was performed to predict persistence only. In this run, the reference map remained exactly the same as the first run, while the comparison map changed to be a Boolean map of the urban areas in 2003's LCLU map that were assigned 1, and 0 elsewhere.

Once the model was validated, it was used to predict 2025's LCLU map, with the same driving forces, modelling the changes between 1984 and 2014's LCLU maps.

5. RESULTS AND DISCUSSION

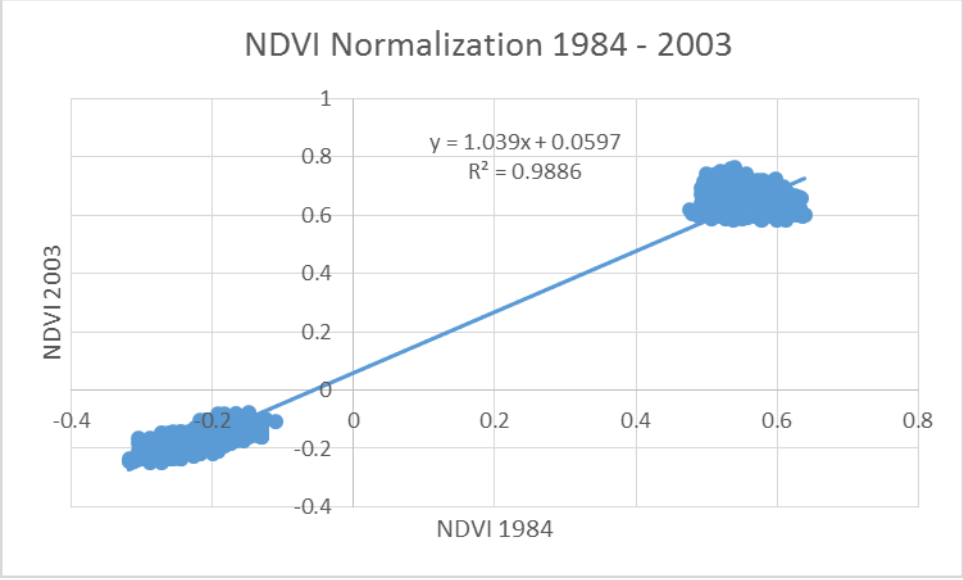
5.1 Change / No-change Maps

In this study, NDVI (ranges from -1 to +1) was calculated by Equation 1, for each image., then a linear normalization was performed, in which the relationships between NDVI 1984 – NDVI 2003 and NDVI 2003 – NDVI 2014 were both assumed to be linear (Broncano *et al.*, 2010). NDVI 2014 acted as a master against NDVI 2003 (O'Connell *et al.*, 2013), likewise, NDVI 2003 against NDVI 1984. Meaning that values of NDVI 2003 were corrected to follow the linear equation with its master and NDVI 1984 followed the same procedure as well. To do so, an average of 1260 points of very high NDVI values (*e.g.* dense vegetation areas) and very low ones (*e.g.* desert and water areas) were collected for each NDVI, in order to draw the linear relation (Figure 19). After normalization, difference in NDVI (ranges from -2 to +2) could be calculated to produce binary maps required for LUCC detection.

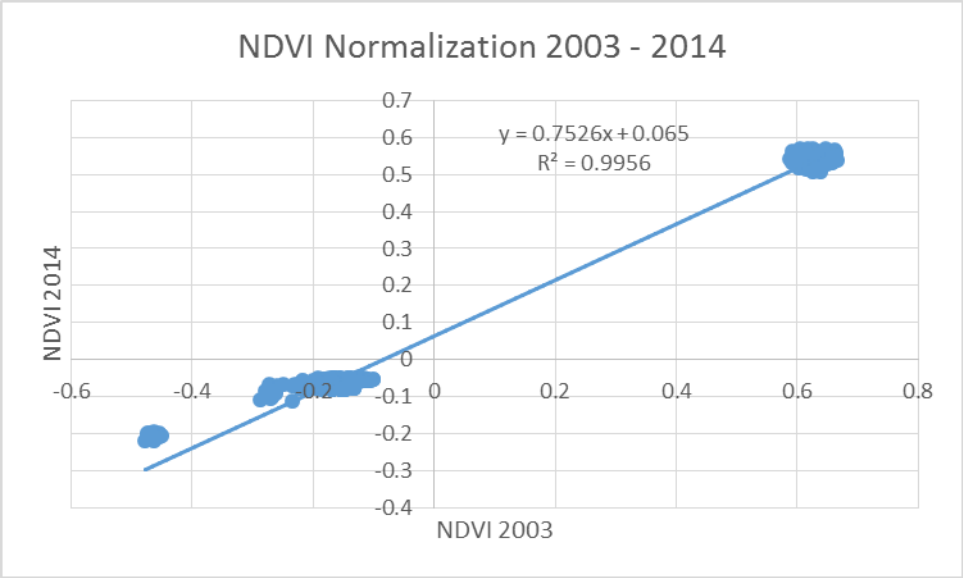
$$NDVI = (NIR - RED)/(NIR + RED) \quad (1)$$

Where:

NIR is the near infrared range, represented by band 4 in Landsat 5,
RED is the red range, represented by band 3 in Landsat 5.



a. NDVI normalization 1984 – 2003

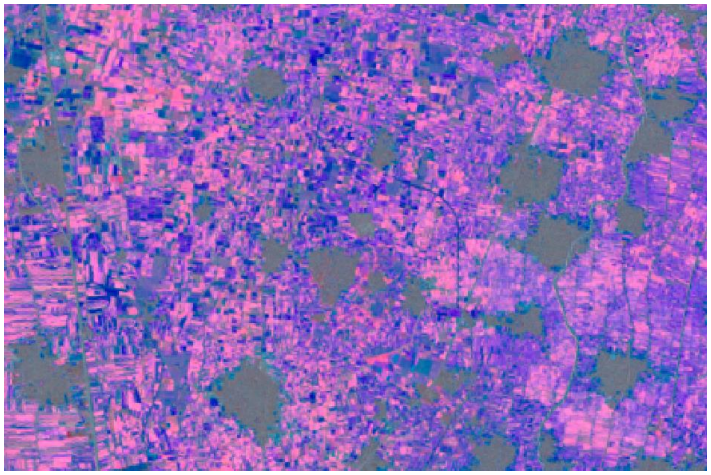


b. NDVI normalization 2003 – 2014

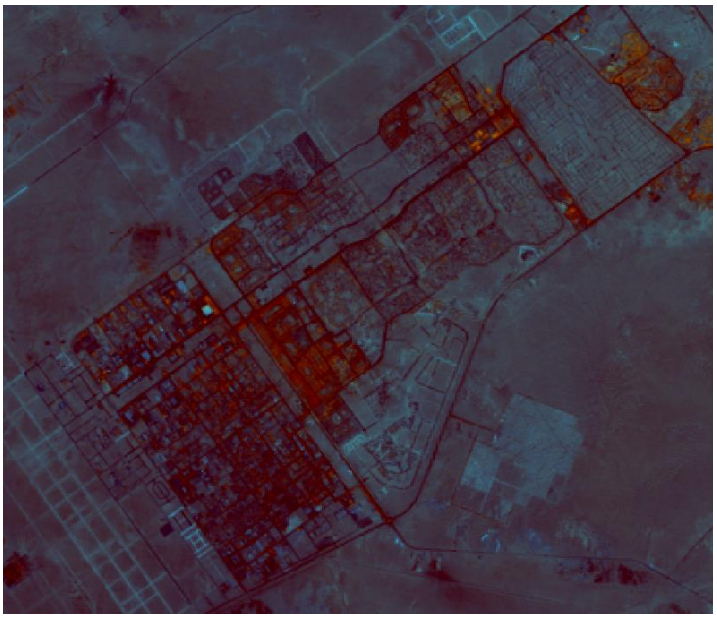
Figure 19. NDVI normalization

The tuning process resulted in thresholds values of ±0.06 for the NDVI difference between 2003 and 2014. These values matched 80% of the previously collected set of check points (see section 4.3). When applying NDVI difference method, and checking 2003–2014 binary map, unfortunately it was not able to capture all the pixels that turned from desert into urban. Worth to mention that decreasing the no-change range (±0.06) is not recommended at all, as when tried, wide parts of the desert were considered as areas of change, while in fact they are not, meaning that the method itself was incompatible with the aim of the study, the urban growth detection, leading to inaccurate results in LUCC, if it had been applied.

Veettil (2012) compared multiple urban change detection techniques, and found that the direct use of NDVI difference by applying threshold values failed to detect many settlements, while when combined with NIR difference and Red difference, it gave a better result in urban change. In this study, this approach was applied, in which NDVI difference, NIR difference and Red difference were calculated for each pair of change, and then displayed in RGB (Red Green Blue) color guns repetitively. In this band composition, new urbanized areas appear in bluish color, the increase in vegetation in yellow or fuchsia, the decrease in vegetation in dark blue and the new urbanization in desert is viewed in very dark red and very dark blue. Figure 20 shows examples of this composition in RGB, where urban areas of no change appear in gray, surrounded by bluish rings, representing new urbanized areas.



a. New urbanization in agriculture



b. New urbanization in desert

Figure 20. NDVI, NIR and Red difference combination approach

The different radiometric resolutions between 2003 and 2014's imageries required performing a mathematical normalization in NIR and RED values before computing the differences. An 8-bit digital number (*e.g.* 2003's imagery) ranges from 0 to 255 (2 ^ 8 - 1), whereas a 16-bit digital number (*e.g.* 2014's imagery) ranges from 0 to 65535 (2 ^ 16 - 1). Hence, a simple normalization was carried out according to Equation 2 so that both; NIR and RED range from 0 to 65535, consequently NIR difference and RED difference could be computed.

$$NV = (V - \min RV)/(\max RV + MRV) \quad (2)$$

Where:

NV is the normalized value,
V is the value to be normalized,

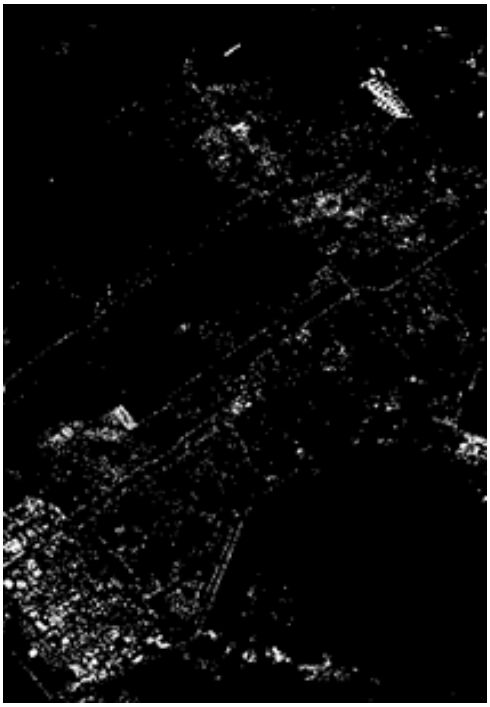
minRV is the minimum range value (255),
maxRV is the maximum range value (65535).

Unsupervised classification was run for multiple times for both images. In each trial, every cluster was visually examined and compared to the satellite scenes in order to determine whether it was a cluster of change or no change. The points of change and no change previously collected for the thresholds tuning in NDVI difference technique (see section 4.3) were used to more accurately decide about the cluster; change or no change. The unsupervised classification of both image of differences matched around 93% of these points. In that way, binary maps were produced; pixels that belong to clusters of change had a value of 1, and those belong to clusters of no change had a value of 0.

Figure 21 shows a new city (6th of October city, west of GCR) which was built in 2003, and expanded in 2014. The figure illustrates how the expansion was hardly captured using NDVI difference method, whereas it was precisely determined using the combination of differences of NDVI, NIR and Red.

a. 6th October City in 2003 – RGB colour composition

b. 6th October City expansion in 2014 – RGB colour composition



c. Changes between 2003 and 2014 in 6th of October City in binary map using NDVI difference approach

d. Changes between 2003 and 2014 in 6th of October City in binary map using NDVI, NIR and Red differences approach

Figure 21. Two approaches for binary maps production - changed pixels are in white

5.2 Image Classification

Figure 22 shows the training samples in green colour, which were collected in areas of no change at all over the study period (1984 - 2014). The same samples were used in the classification of the three images.

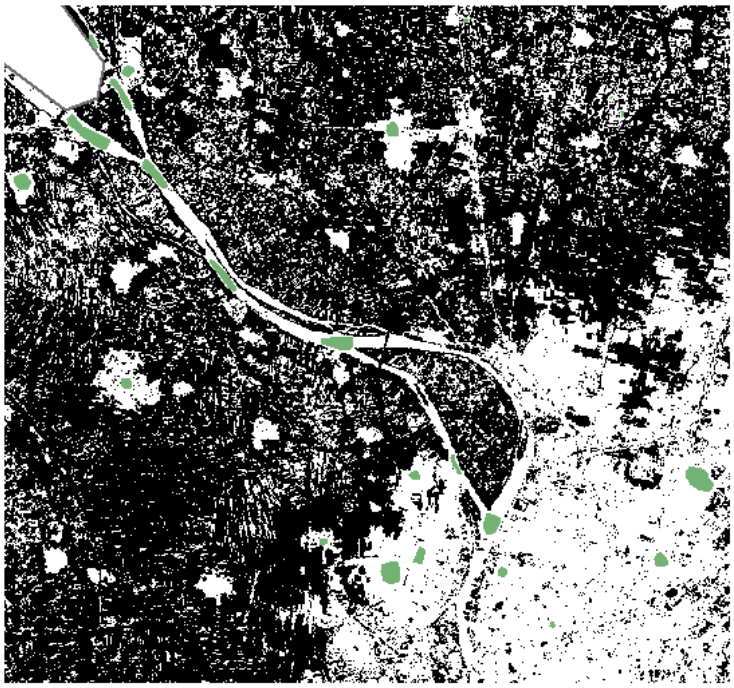
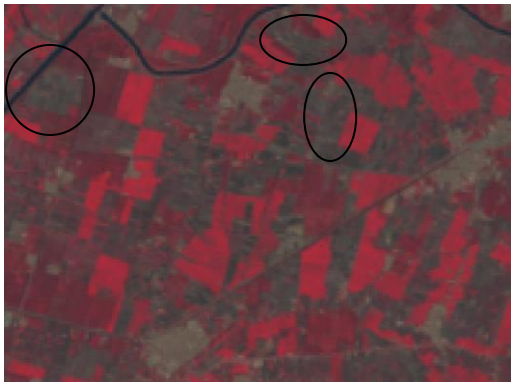
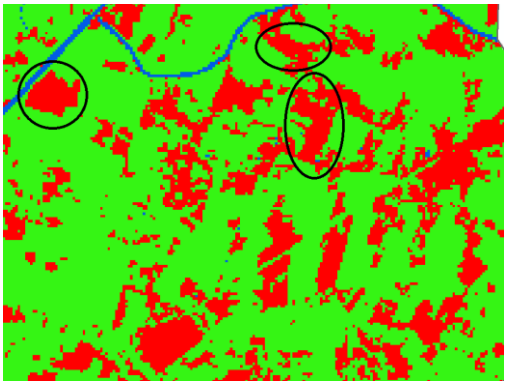


Figure 22. Training samples in areas of no change – no change areas are in white

The visual inspection of the produced maps indicated the efficiency of the classifier in separating between pixels that belong to different land uses, however, it could not separate between urban areas and unplanted agricultural fields, that existed in 1984 and 2003 images, not because of the classifier, but because training samples were taken based on 2014's image which had almost all agricultural fields planted. These unplanted agricultural fields shared a similar gray appearance with urban areas, but a little bit darker, thus were classified as urban instead of vegetation (Figure 23).



a. Examples of unplanted agricultural fields in 1984 image, viewed in RGB colour composition



b. Misclassification of unplanted agricultural fields in 1984 map

Figure 23. Misclassification of unplanted agricultural fields

In order to avoid accuracy decrease in the classification, LUCC and modeling that might have resulted from such misclassification, it was essential to extract these unplanted agricultural fields from the urban class and reclassify them as vegetation. One suggested approach was to identify a new class for them, with additional training samples in images where they existed (1984 and 2003), but because they had relatively small areas and were surrounded by real urban pixels, it was difficult to take precise and adequate training samples for them. Another procedure that was successfully applied was to identify the urban areas using the NDBI (Normalized Difference Built-up Index) (Equation 3), in which urban areas appear very bright, and these unplanted agricultural fields remained included.

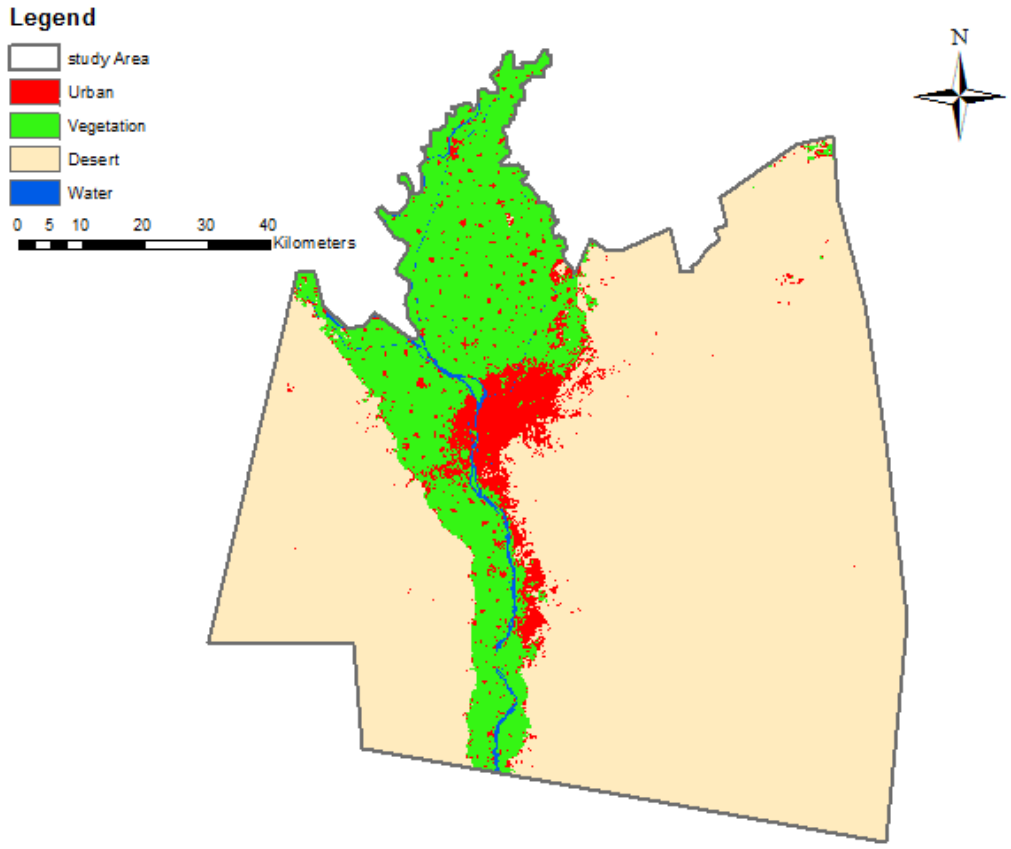
$$NDBI = (MIR - NIR)/(MIR + NIR) \quad (3)$$

Where:

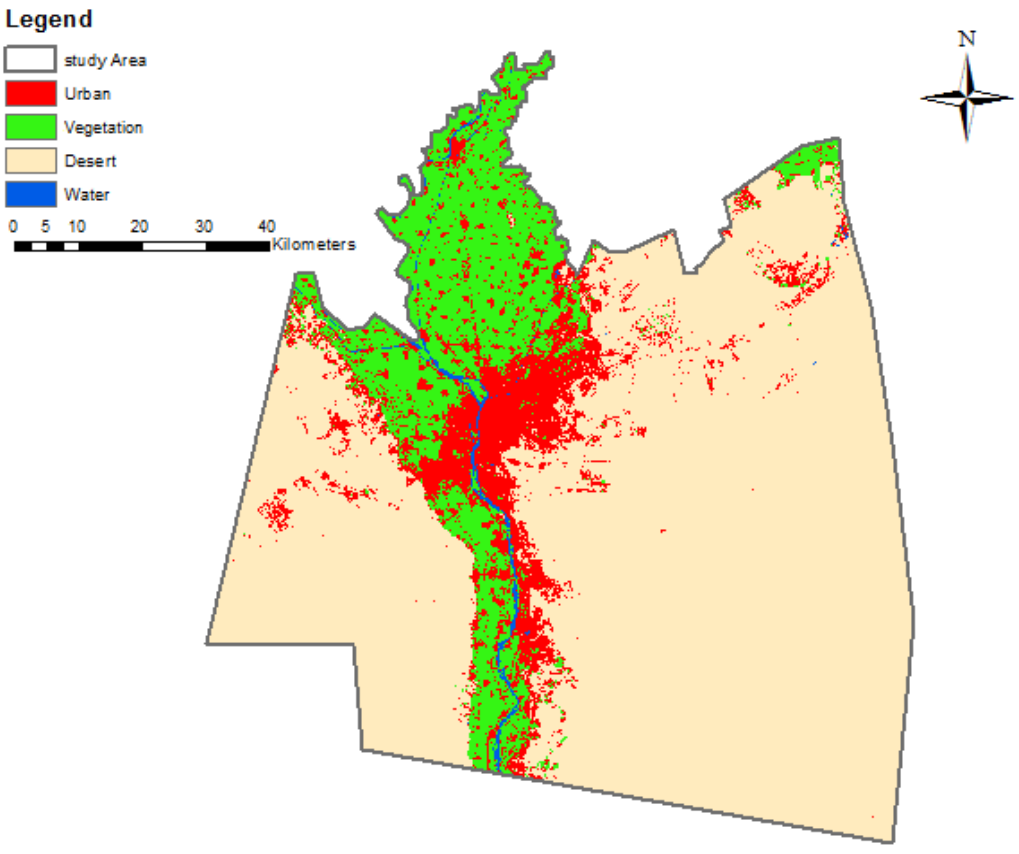
MIR is the mid infrared range, represented by band 5 in Landsat 5,
NIR is the near infrared range, represented by band 4 in Landsat 5.

Several unsupervised classifications were tried until it was possible to visually identify in which cluster(s) these unplanted agricultural fields were mostly included. Eventually a query could be performed to target all the pixels that were misclassified as urban and in the same time existed in the identified cluster(s), to the vegetation class. This query was applied in certain areas of the images, only where the problem of misclassification existed, in order to avoid turning some real urban pixels into vegetation.

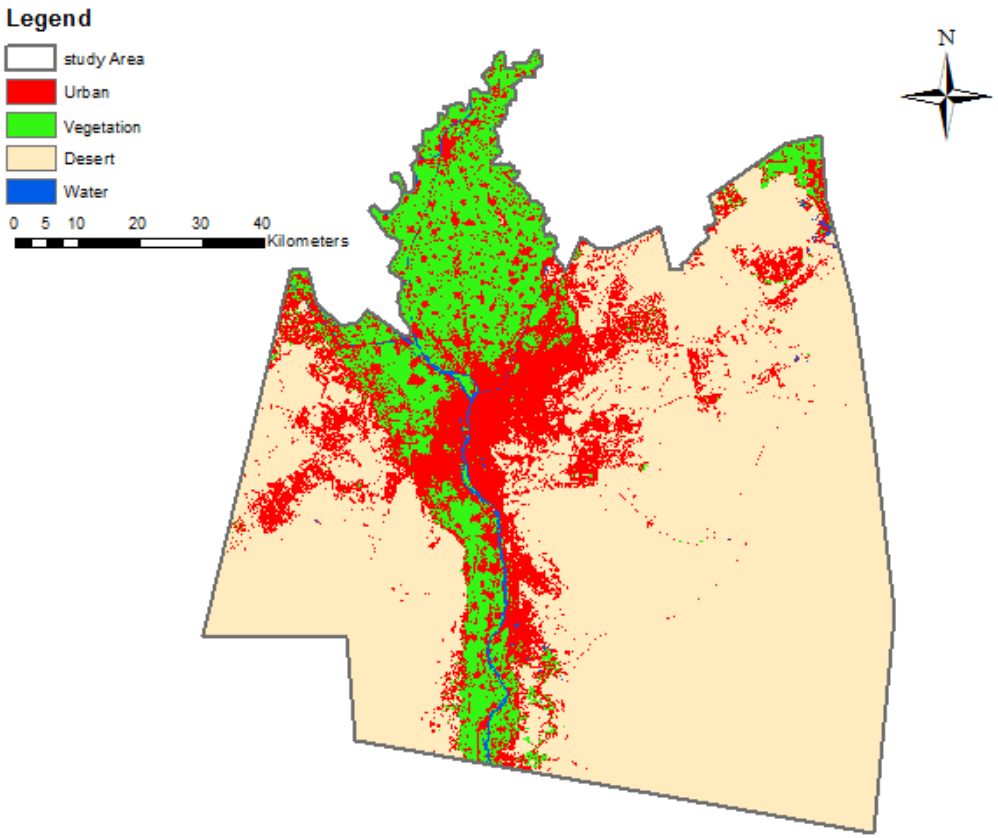
Figure 24 shows the LCLU maps. Generally, vegetation is concentrated mainly in Qalyubiyah city, in the North of the study area, and along the western bank of Nile River in Giza city, with less density along the eastern bank of the river in Cairo city. On the other hand, urban areas are concentrated in the heart of Cairo and extend along the eastern bank of Nile River, with some scattered urban communities inside the agricultural areas, and new recent built cities in the desert as well.



a. LCLU map - 1984



b. LCLU map - 2003



c. LCLU map - 2014

Figure 24. LCLU maps

Figure 25 gives an overview of the urban expansion that occurred over the 30 years, from 1984 to 2014. New urban areas in 2003 were built mainly around the urban communities which already existed in 1984, with new built-up cities in the South of Cairo (*e.g.*15th May city) and to the South of Giza along the Nile River. Small urban settlements in the desert parts of Cairo and Giza cities were built in 2003 and are located to the East (*e.g.* Nasr city) and to the West of the study area (*e.g.* Sheikh-Zayed and 6th October cities). Whereas the new urbanization in 2014 took place significantly in the desert where a dramatic expansion in urban areas occurred. Beside this, new settlements were built in the agricultural parts towards the North of the study area.

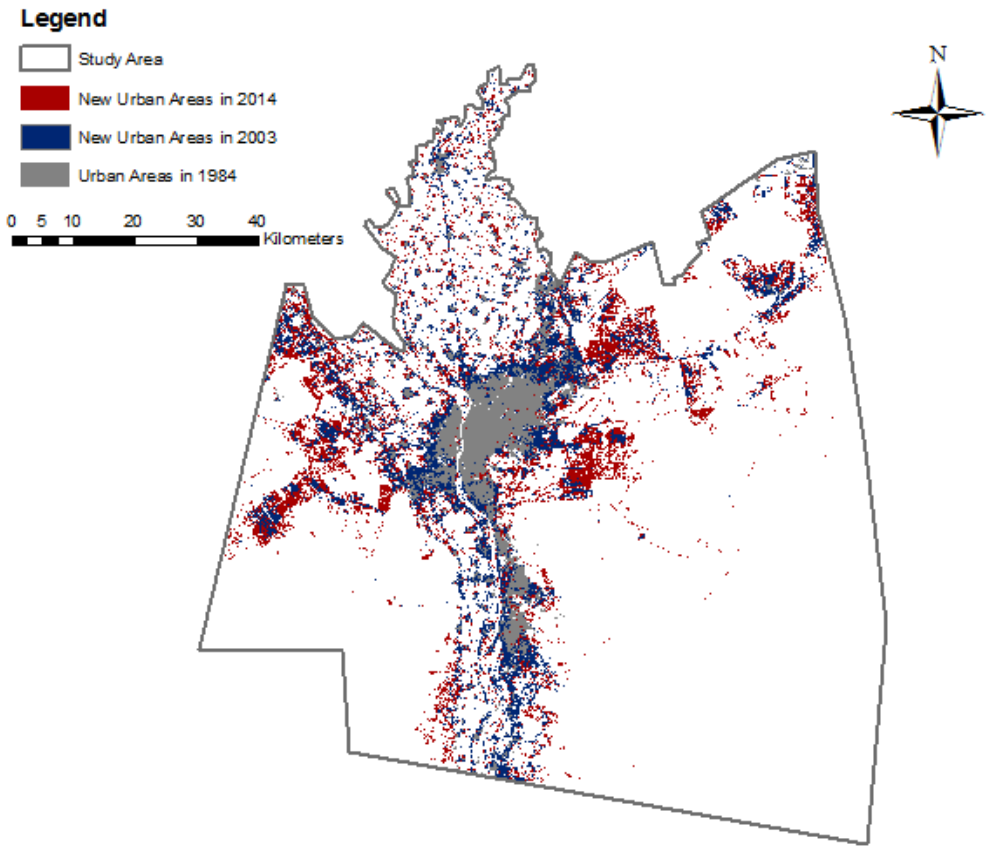
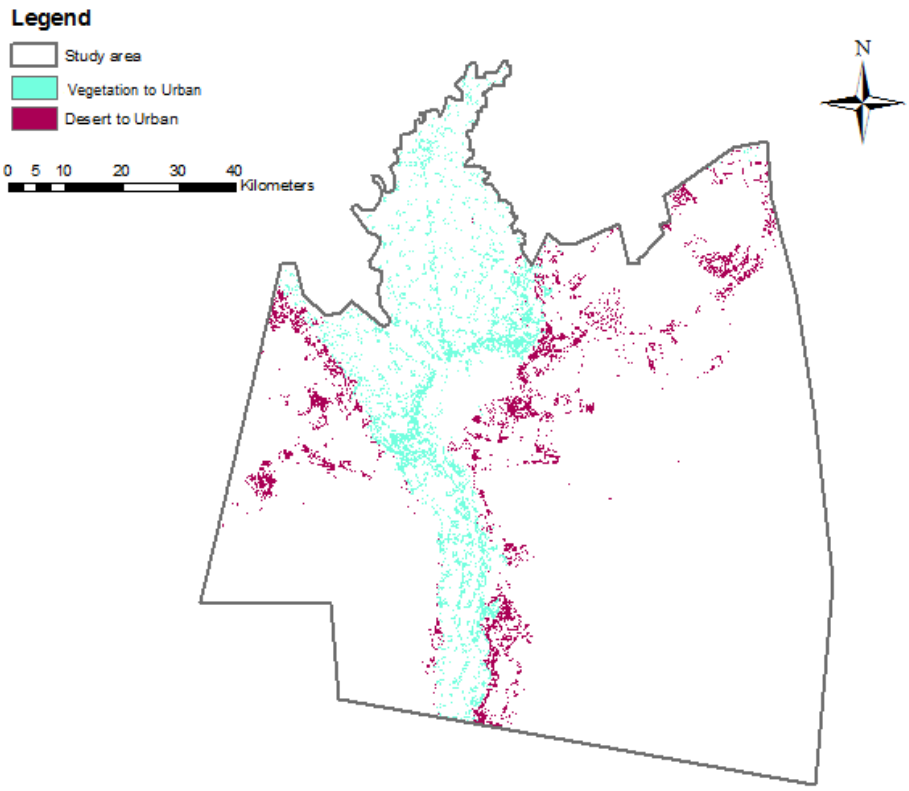
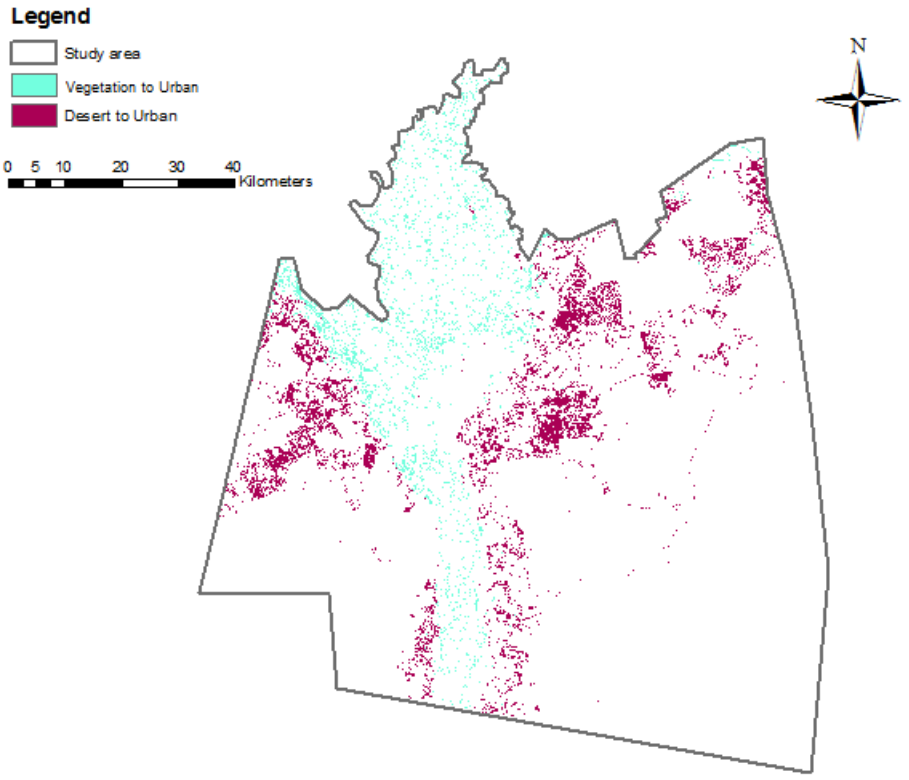


Figure 25. Urban expansion between 1984 and 2014

The urban areas were gained mostly from vegetation and desert. However, the trend of urbanization between 1984 and 2003 is quite different from what occurred between 2003 and 2014 (Figure 26). The vegetation that turned into urban areas was dense and concentrated along the water bodies (mainly the Nile River) during the first period, while it was uniformly distributed over the agricultural areas in the second period. On the other hand, the urbanized desert had much more expansion and density during the second period especially in the north eastern and north western parts of GCR.



a. Spatial trend of change to urban 1984 – 2003



b. Spatial trend of change to urban 2003 – 2014

Figure 26. Urbanization spatial trends

5.3 Accuracy Assessment

Table 8 illustrates the accuracy assessment results of the LCLU maps for 1984, 2003 and 2014. The tolerance was calculated from Equation 4.

$$Tolerance = 1.96 \times \sqrt{A \times (1 - A) \div T} \quad (4)$$

Where:

A is the accuracy of the class,
T is the total number of points per class.

The overall accuracy showed very good results, as all of them exceeded 95% which is the average of the overall accuracies that had been achieved in previous researches, indicating the suitability and the precision capabilities of SVM classifier. Moreover, the accuracies per each class from both; user and producer perspectives were very high as well. Despite the high user accuracy of the vegetation class in 1984's LCLU map (93%), it is the lowest amongst other user accuracies, and this is a side effect of using NDBI. Although NDBI helped in extracting unplanted agricultural fields and reclassifying them as vegetation, instead of urban, it affected some real urban pixels, and considered them as vegetation. This resulted in lowering the producer accuracy of the urban class in the same map, 1984 (90.7%) relative to other producer accuracies. Hence, NDBI is not considered as a very precise method to differentiate unplanted agricultural fields from urban areas, because it may affect the classification accuracy of the vegetation class itself, but with no significant effect towards the overall accuracy. Generally, using NDBI for extracting urban pixels requires paying attention to the accuracy of the other class which is mixed with.

	2014		2003		1984	
	User Accuracy	Producer Accuracy	User Accuracy	Producer Accuracy	User Accuracy	Producer Accuracy
Urban	95% ± 4.3%	94.1% ± 4.6%	97% ± 3.3%	92.4% ± 5.1%	97% ± 3.3%	90.7% ± 5.5%
Vegetation	97% ± 3.3%	94.2% ± 4.5%	97% ± 3.3%	97% ± 3.3%	93% ± 5%	95.9% ± 4%
Desert	97% ± 3.3%	97% ± 3.3%	97% ± 3.3%	100% ± 0%	98% ± 2.7%	99% ± 2%
Water	95% ± 4.3%	99% ± 2%	98% ± 2.7%	100% ± 0%	97% ± 3.3%	100% ± 0%
Overall Accuracy	96% ± 1%		97.3% ± 0.8%		96.3% ± 0.9%	

Table 8. Accuracy assessment results of the LCLU maps

5.4 LUCC Detection

Tables 9 and 10 are the change matrices that represent the transition from one class to another in percentage and in hectares as well, which occurred between 1984 and 2003, and between 2003 and 2014.

		2003 (%)			
		Urban	Vegetation	Desert	Water
1984 (%)	Urban	97	2	1	0
	Vegetation	13	86	0	0
	Desert	3	1	96	0
	Water	3	7	0	90

a- LUCC in percentage

		2003 (hectare)			
		Urban	Vegetation	Desert	Water
1984 (hectare)	Urban	54 990	1 093	792	96
	Vegetation	19 179	125 790	344	959
	Desert	21 417	5 761	656870	192
	Water	179	467	3	5 990

b- LUCC in hectares

Table 9. LUCC 1984 – 2003

		2014 (%)			
		Urban	Vegetation	Desert	Water
2003 (%)	Urban	100	0	0	0
	Vegetation	12	87	0	0
	Desert	5	0	95	0
	Water	6	9	0	85

a- LUCC in percentage

		2014 (hectare)			
		Urban	Vegetation	Desert	Water
2003 (hectare)	Urban	10 6938	61	343	110
	Vegetation	16 486	115 497	122	222
	Desert	31 045	3099	612 533	424
	Water	393	1 322	13	55 66

a- LUCC in hectares

Table 10. LUCC 2003 - 2014

The most significant changes in both periods are the transitions from vegetation and desert to urban. Over 19 years, from 1984 to 2003, the vegetation lost 13% to urban, representing 19 179 hectares, and almost the same percentage (12%) within only 11 years, from 2003 to 2014, representing an amount of 16 486 hectares (Figure 27). This indicates the massive leveling of agricultural lands in GCR for urbanization purposes especially during the last decade, because of the absence and / or the inactivation of farmlands protection laws. While 3% of desert areas turned to urban between 1984 and 2003 which is equivalent to 21 417 hectares, larger than the amount of transition from vegetation to urban during the same period despite the lower percentage, due to the larger total desert area compared to the total agricultural ones. This percentage increased to 5% between 2003 and 2014, representing 31 045 hectares, resulting from the application of desert reconstruction strategies to build new communities outside the Nile Valley. The gains and losses in LCLU classes are given in Figure 28, where the significant gains in urban and the dramatic losses in vegetation and desert are noticeable.

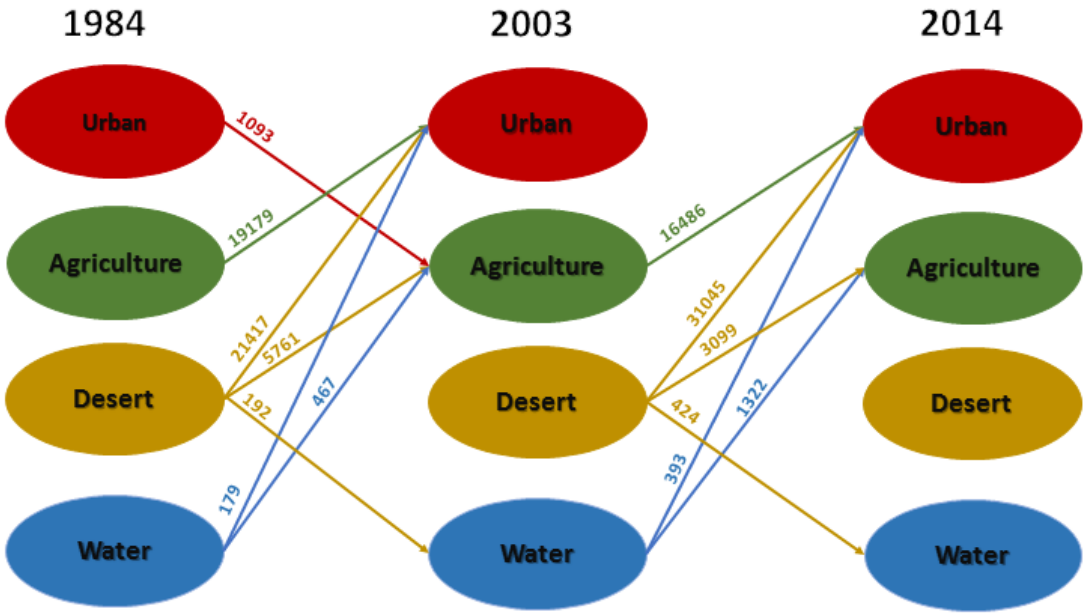


Figure 27. LCLU transitions in hectares

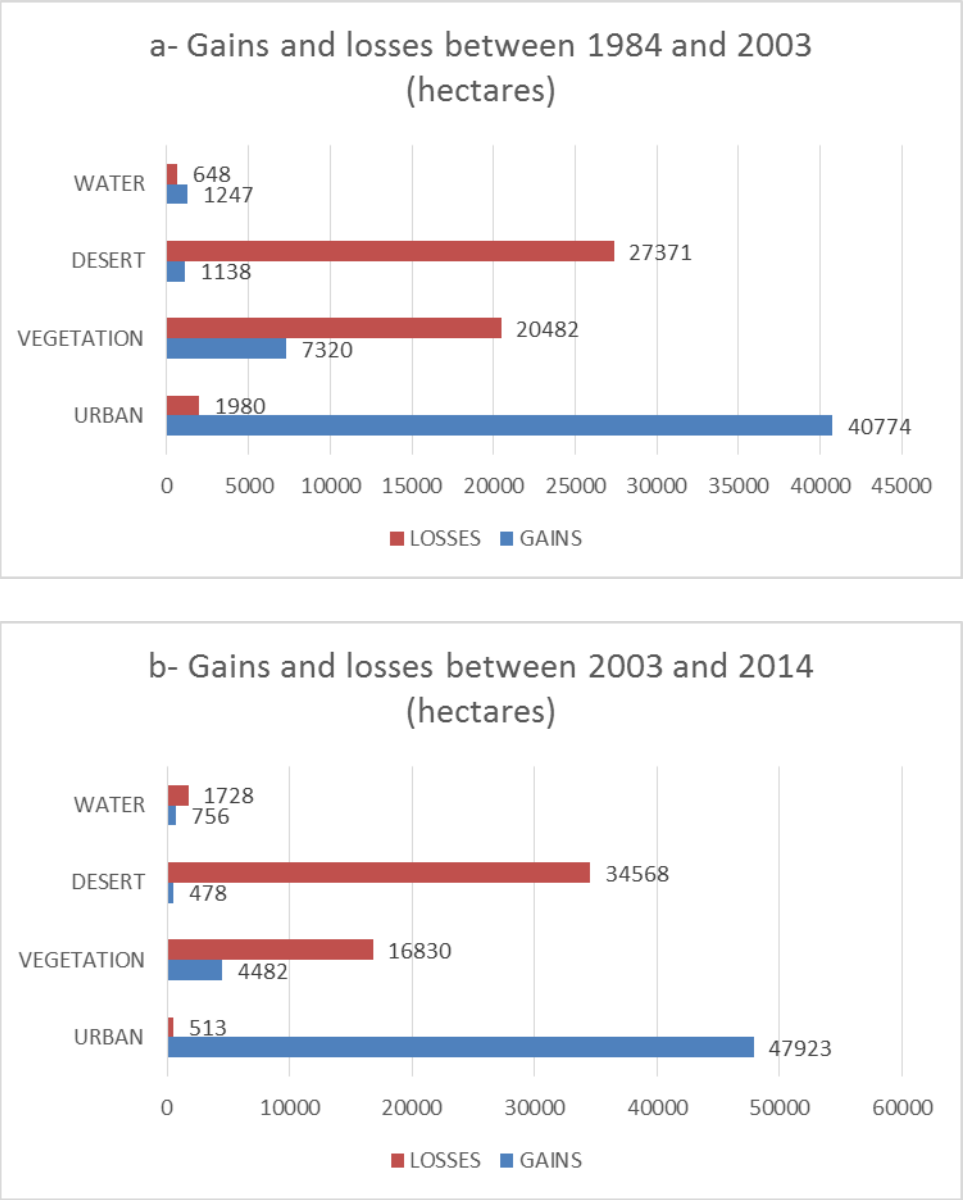


Figure 28. LCLU gains and losses in hectare

On the other hand, there are some minor changes such as; 1% and less than 1% turned from desert to vegetation, which represent 5 761 and 3 099 hectares during 1984 – 2003 and 2003 – 2014 respectively, due to the governmental reclamation plans in the north eastern and north western parts of GCR. Likewise, less than 1% of desert turned to water which is equivalent to 192 and 424 hectares in both periods respectively, resulting from the construction of new unconventional waste water treatment ponds in the north eastern area in GCR (Figure 29).

1984 **2003** **2014**

Figure 29. Waste water treatment ponds expansion – all at a fixed scale

Changes from water to urban were 3% and 6% or in other words, 179 and 393 hectares in both periods, respectively. Some of these changes are due to fill works in some canals, and others are resulted from capturing the satellite image while some wastewater treatment ponds were empty.

The change from water to vegetation is a false change, because of the growth of some unseasonal plants over the Nile River in one image while they were uprooted in another image. Figure 30 shows some vegetation pixels in yellow color which grew over the Nile River in one image and were uprooted in other image.

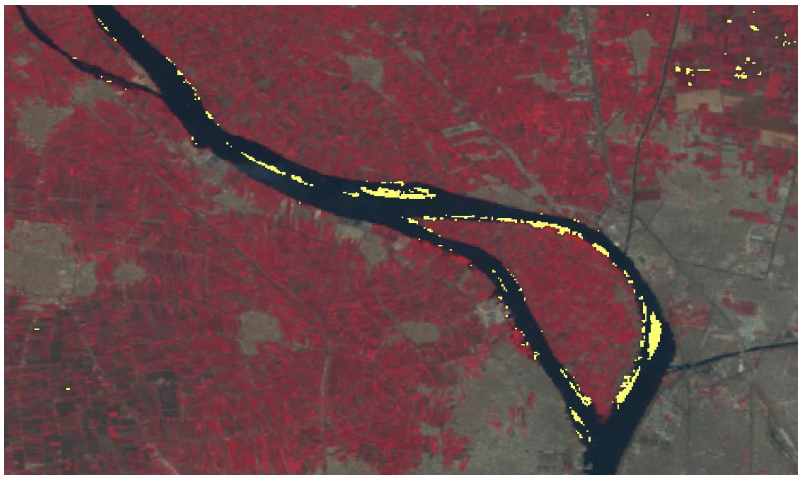
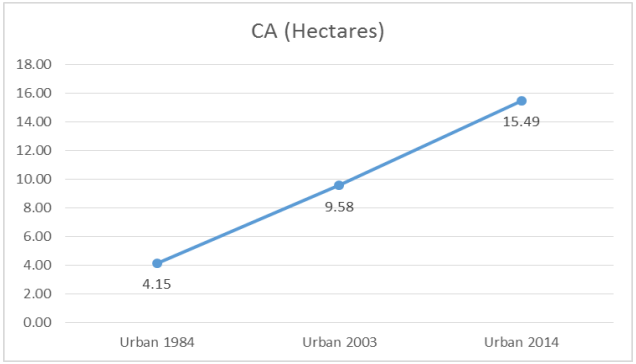


Figure 30. False change from water to vegetation

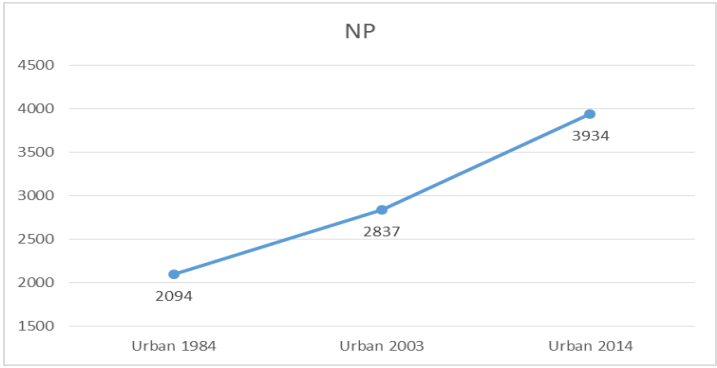
5.5 Spatial Urban Growth Pattern

The temporal urban growth signatures of spatial metrics are illustrated in Figure 31. As a result of the continuous urban expansion over the study period (1984-2014), CA and NP have boomed between 1984 and 2003 with a dramatic increase in 2014, indicating a higher urbanization rate between 2003 and 2014. ED in 2014 was almost four times of what had been in 1984, thus indicating an increase in

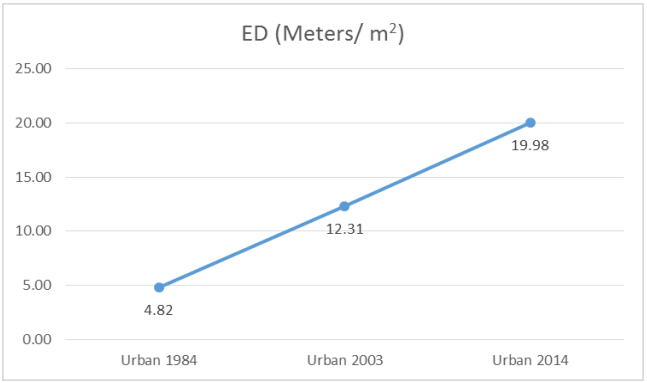
the total length of the edge of the urban patches, as due to land use fragmentation. Moreover, the increase of LPI emphasises the proportion growth of the total landscape area comprised by the largest urban patch. By contrast, ENN_MN has dipped from 1984 to 2014, meaning that the space between urban neighbours is shrinking by time, as a result of higher urbanization density. On the other hand, more complex metrics such as AWMPFD, were computed. Generally speaking, the fractal dimension describes the complexity and the fragmentation of a patch by a perimeter–area proportion. The values range between 1 and 2. Low values are derived when a patch has a compact rectangular form with a relatively small perimeter relative to the area. If the patches are more complex and fragmented, the perimeter increases and yields a higher fractal dimension (Herold *et al*., 2003). AWMPFD averages the fractal dimensions of all patches by weighting larger land cover patches (Herold *et al*., 2003). It climbed from 1.31 to 1.45 between 1984 and 2003, while later it has increased steadily to reach 1.46 in 2014. This means that the level of complexity and fragmentation is increasing for the landscape patches, The CONTAG indicates the heterogeneity of the landscape throughout a given probability that determines patches adjacency. The lower CONTAG values are, the more heterogeneous the landscape becomes, so that the drop in CONTAG values between 1984 and 2014 can have resulted from higher fragmentation due to more individual urban units. Finally, Shannon's Entropy is a spatial concentration indicator in which the values vary between 0 and 1. Lower values imply higher distribution concentration in one region. These values have gradually increased over the study period because of the more dispersed distribution taking place in GCR.



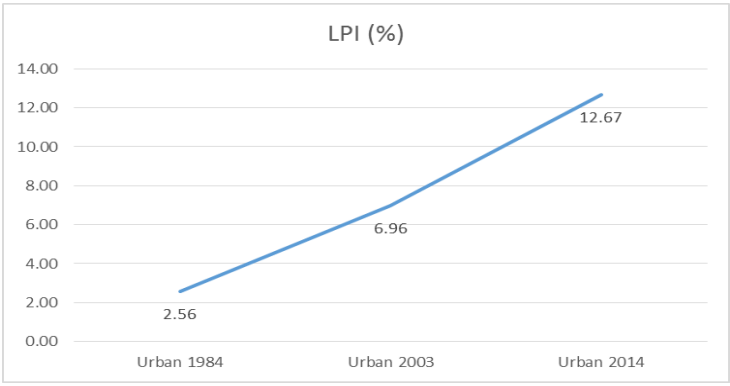
a- Class Area



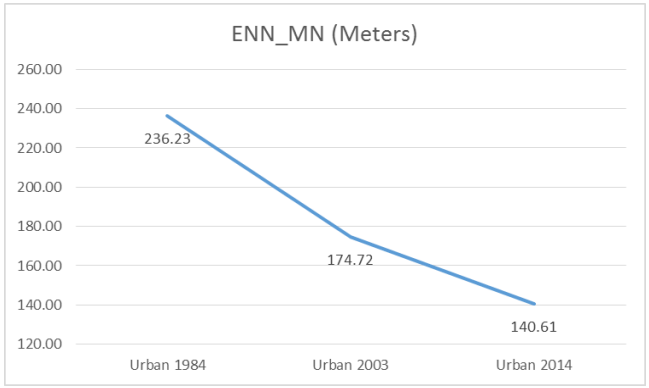
b- Number of Patches



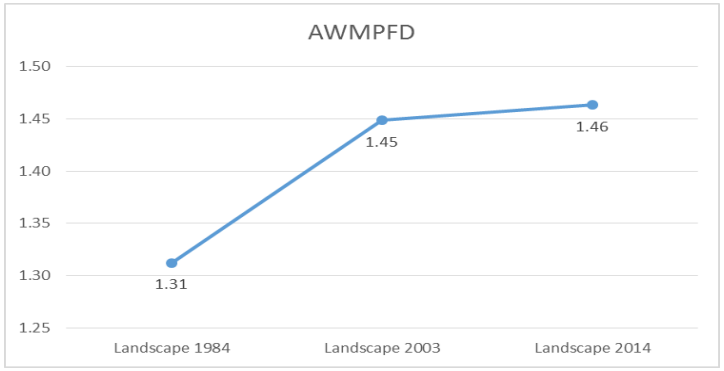
c- Edge Density



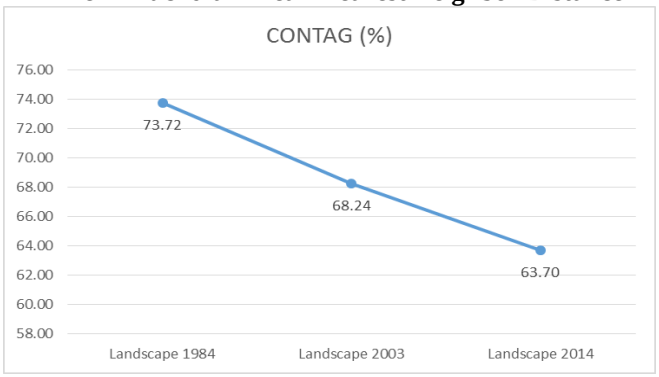
d- Largest Patch Index



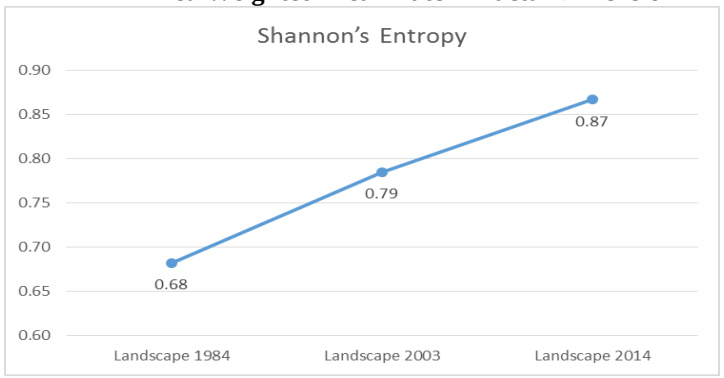
e- Euclidian Mean Nearest Neighbor Distance



f- Area Weighted Mean Patch Fractal Dimension



g- Contagion



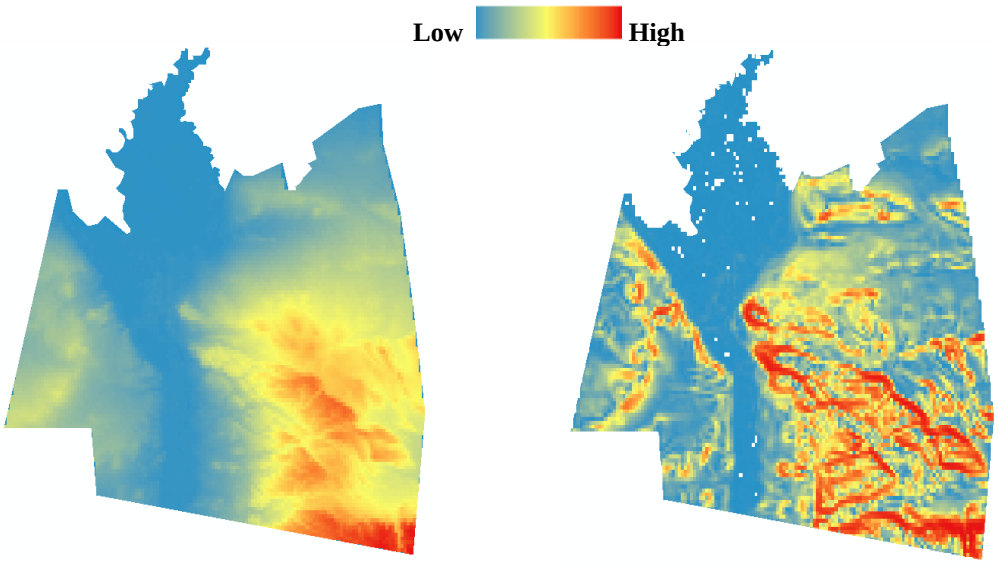
h- Shannon's Entropy

Figure 31. Temporal urban growth signatures of spatial metrics

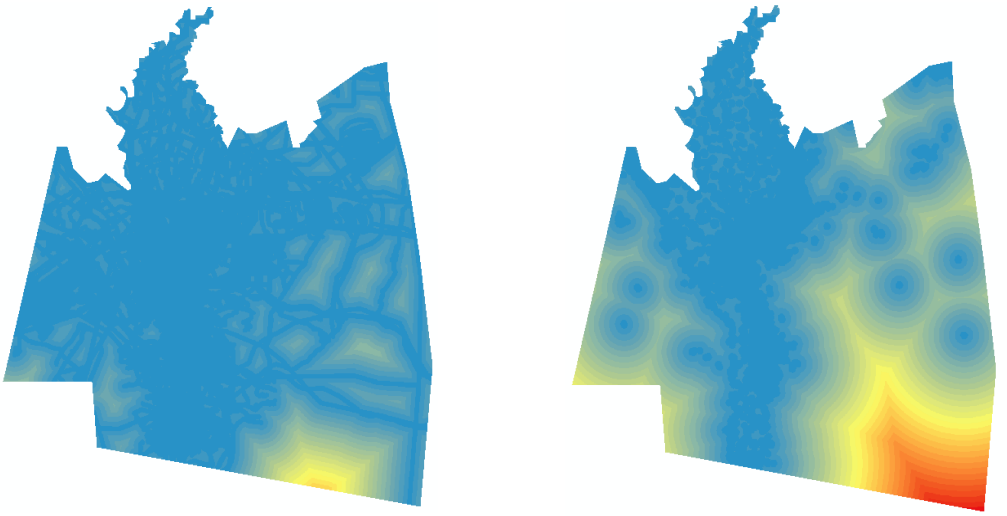
5.6 LUCC Modelling

The LUCC results indicated two significant changes to urban: from vegetation and from desert, consequently both were the model's major transitions. Four driving forces were considered in this study: elevation, slope, distance to roads in 2014 (since it was not possible to obtain them in 2003), and distance to existing urban areas in 1984. These predictor variables were chosen based on recent similar studies (*e.g.* Vaz *et al*., 2011) in which they were found to highly affect urban sprawl and based on the LUCC between 1984 and 2003 where the new urban settlements were noticed to occur nearby the built-up areas in 1984, and the road network in 2014. The distances from the urban areas and the roads were calculated using the DISTANCE tool, enabled in IDRISI. The closer the pixel is to both of them, the higher likelihood to turn to urban. The maps of the selected driving forces are shown in Figure 32.

The elevation varies over the study area, regions of low elevation are concentrated in Qalyubiyah city, in the North, and around the perimeter of the Nile River, while elevation values increase significantly in the East of GCR and reach an extreme in the South-East. The western parts are areas of high elevation as well, but lower than the eastern ones. The minimum and maximum elevations are 8 and 793m above sea level, respectively. Because slope is a function of elevation, it more or less follows the same trend of DEM (Digital Elevation Model). The vast majority of GCR has a close connection up to 3 kilometres to the road network, however, some parts to the East and to the South are far from the road network where the closest access is 15-20 kilometres. In terms of distance to urban settlements, Figure 32 - d shows that in 1984 majority of GCR was of a close distance from nearby built-up areas, up to 8 kilometres, increasing gradually especially in the South-East direction to reach a maximum of 45 kilometres.



a- DEM **b- Slope**



c- Distance to roads in 2014 **d- Distance to urban in 1984**

Figure 32. The model's driving forces

All driving forces were set as static variables as they were assumed not to change over time, except distance to urban in 1984, it was defined as a dynamic factor and was recalculated over the prediction period (11 years, from 2003 to 2014), as the urban expands, consequently, distance to urban in 1984 will vary over time. Likewise, distance to roads, it should have been considered as a dynamic factor, because new roads are assumed to be built between 2003 and 2014. This could be applied in case the roads were obtained in 2003, but as long as the roads were already in 2014, there was no need for this assumption. Both transitions to urban had the same driving forces depending on the visual examination of the urban spatial trend which indicated that the selected predictor variables affect both of them.

Table 11 illustrates the potential explanatory power of each driving force represented by Cramer's V. Eastman (2012) stated that the variables that have a Cramer's V of about 0.15 or higher are useful while those with values of 0.4 or higher are good. Thus, the selected factors were found to be relevant and worth consideration. It is important to emphasis the fact that this test is just an indicator of factors refusal or consideration, only the validation results are the differential indicators, whether the factors are acceptable or need modifications.

Driving force	Cramer's V
DEM	0.55
Slope	0.49
Distance to roads in 2014	0.25
Distance to urban in 1984	0.52

Table 11. Cramer's V values of the selected driving forces of change

After the selection of the predictor variables, both transitions were modeled in one transition sub-model, as they had the same driving forces, with the aim of producing the transition maps. There are

many methods to model the transitions, however a MLP neural network allows for modeling more than one transition at a time, hence, it was applied in this study. Figure 33 shows the resulting transition potential maps from vegetation and desert to urban.

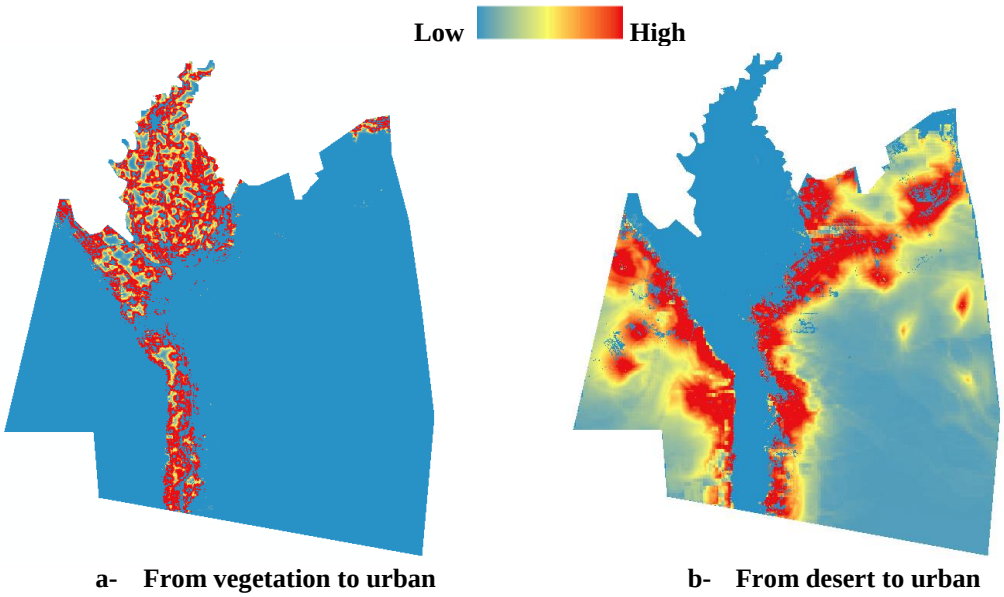
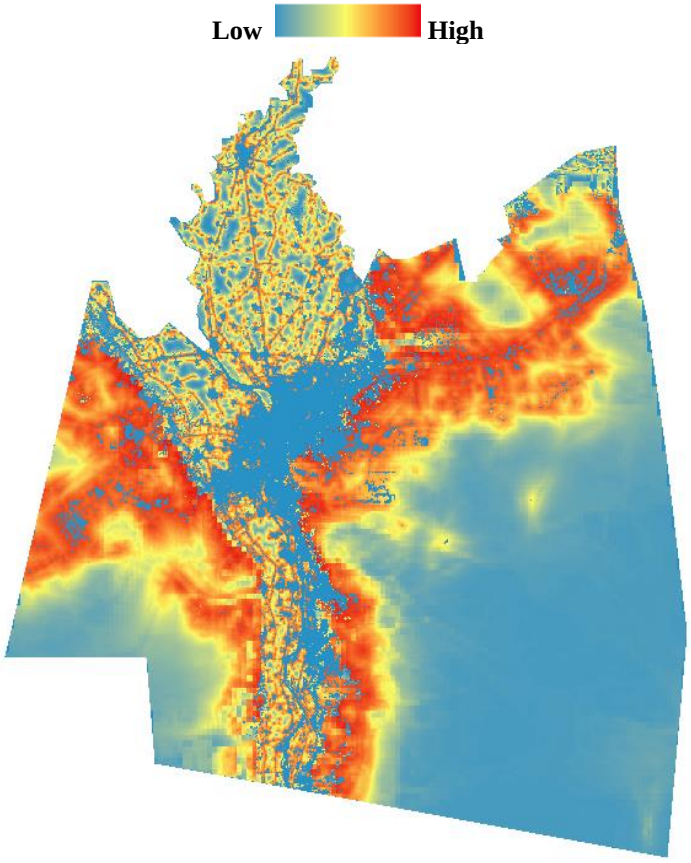
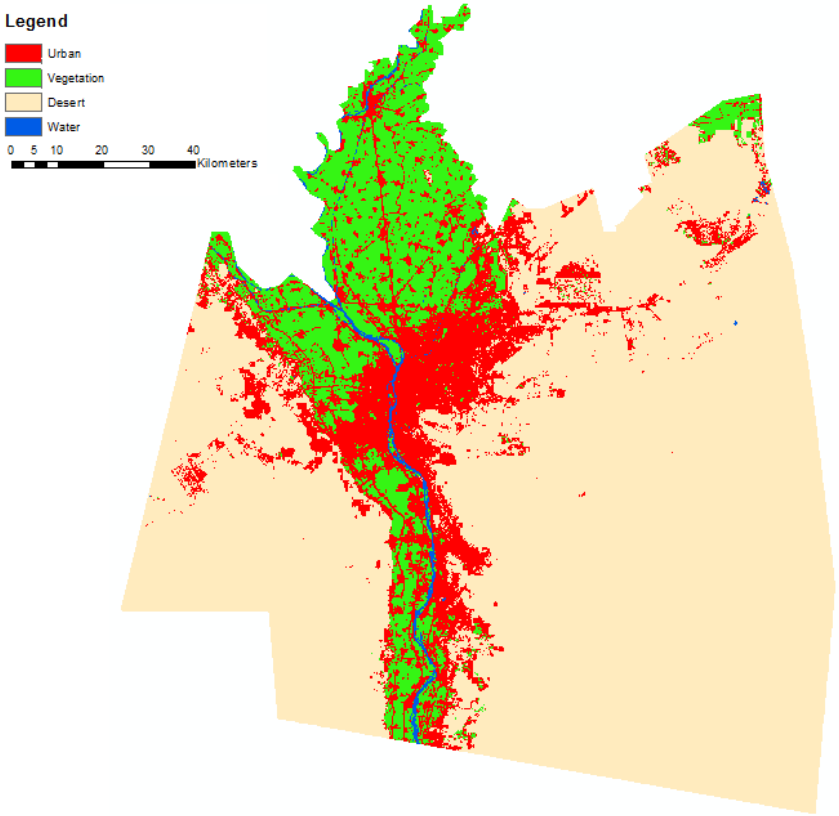


Figure 33. Transition potential maps

Up to this stage, the model was ready to predict the urbanization scenarios for 2014 based on the changes occurred between 1984 and 2003, and the location of the possible future changes determined by the transition potential maps. Both soft and hard prediction results are illustrated in Figure 34, below.



a- Soft prediction map



b- Hard prediction map

Figure 34. 2014's prediction results

The soft prediction map shows that the western, north - eastern, and the southern parts along the Nile River banks had the highest likelihood to turn into urban in 2014, while the North showed less vulnerability towards urban development. Beside the urban areas already existed in 2003, and assumed to keep its urban state in the future, the south-eastern part of GCR did not show sensitivity to change to urban, thus it was expected to keep its desert LCLU class in 2014.

The visual validation of the simulated change in 2014 resulted in a map of correctness and error, given by Figure 35. The map consists of 1.31% hits, 89.73% null successes, 2.30% false alarms, and 6.66% misses. The simulated change is 3.61% of the landscape, less than the observed change which is 7.97% of the landscape, and this is why the error occurred. The FOM (Figure Of Merit) is a ratio between hits and the summation of hits, misses and false alarms. It ranges from 0%, meaning no overlap between observed and predicted change, to 100%, meaning perfect overlap between observed and predicted change (Martins *et al.*, 2012). In this case, the FOM is 12.76% which is a low performance, however still higher than some recent studies such as (Rodríguez *et al.*, 2013) and (Olmedo *et al.*, 2013) where FOM was 2.9 and 10.4%, respectively.

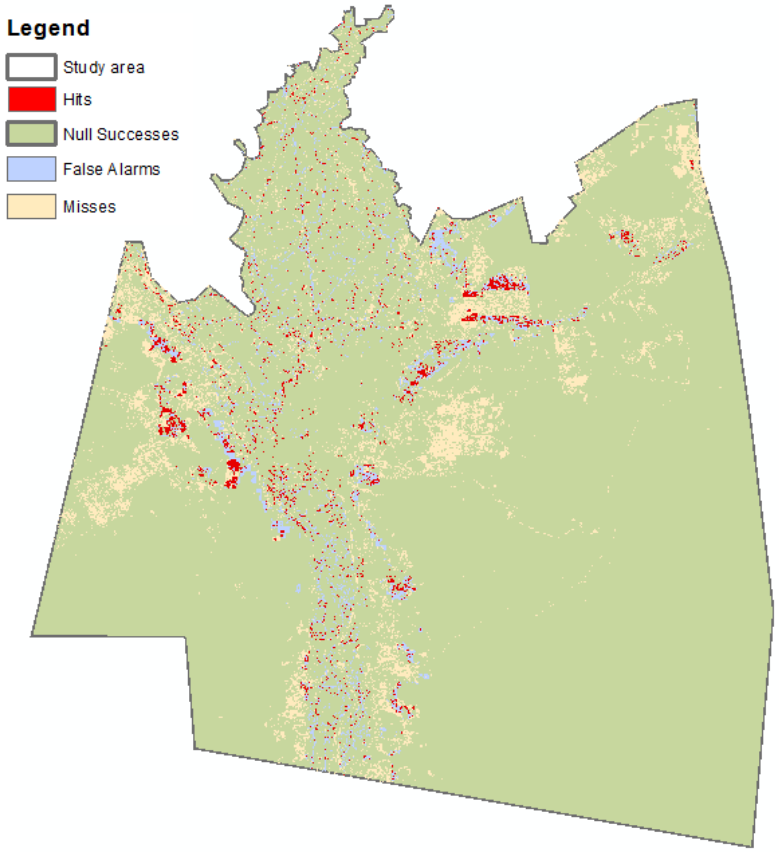


Figure 35. Visual validation - Map of correctness and error based on 2003 (reference), 2014 (reference) and 2014 (simulated) LCLU maps

On the other hand, the statistical hard validation using the VALIDATE module, which compared the projected LULC in 2014 with actual LULC in 2014, resulted in Kstandard, Kno and Klocalion of 91.26%, 92.88% and 94.24%, respectively, where Kno is the overall accuracy of the simulation run, Klocation indicates the level of agreement of location (Nadoushan *et al*., 2012), and Kstandard is the proportion assigned correctly versus the proportion that is correct by chance (Kim *et al*., 2011). Since all Kappa statistics (K) that were obtained in previous studies (*e.g.* Regmi *et al*., 2014; Subedi *et al*., 2013) were well above 80%, the coefficients obtained from this module in this study indicate a perfect model, as all of them are even above 90%.

The soft statistical validation was done by ROC analysis and resulted in AUC value of 0.71 in the first run, to check the agreement of the urban development in 2014. While in the second run the agreement of the non-urban areas in 2014 was found to be 0.62 indicating that the accuracy of the model to predict urban is higher than the accuracy of the null model. AUC value of 0.71 indicates a fair level of discrimination according to the approach recommended by (Swets, 1988). It is better than the mean AUC value of 0.64 that was obtained by (Thies *et al*., 2014), however, lower than some studies in which it exceeded 0.80 (*e.g.* Sangermano *et al*., 2012).

The lower accuracy of the soft validation (71%) versus the hard validation (> 90%) highlights a limited ability in the projection of urban transitions in terms of location. The detected location of change does not perfectly match with the selected driving forces, thus additional driving forces may need to be considered, or maybe a MCE procedure needs to be carried out to produce more accurate suitability map. Nevertheless, the model shows a reasonable performance even according to the ROC results.

The same driving forces were later used to model the change between 1984 and 2014 using the real LCLU maps, in order to predict the LCLU map of 2025 (Figure 36), with the same prediction period (11 years). 14% of the vegetation and 4% of the desert in 2014 are expected to transition to urban in 2025, which is equivalent to 16 512 and 24 687 hectares, respectively.

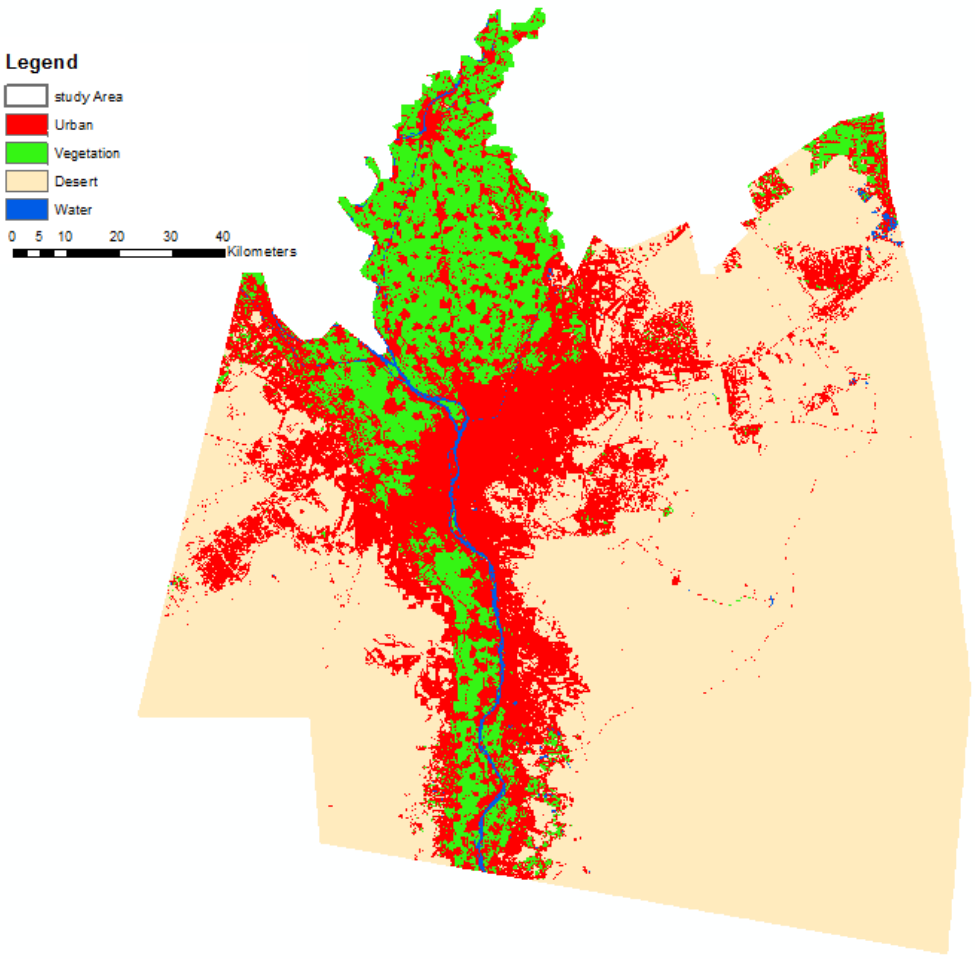
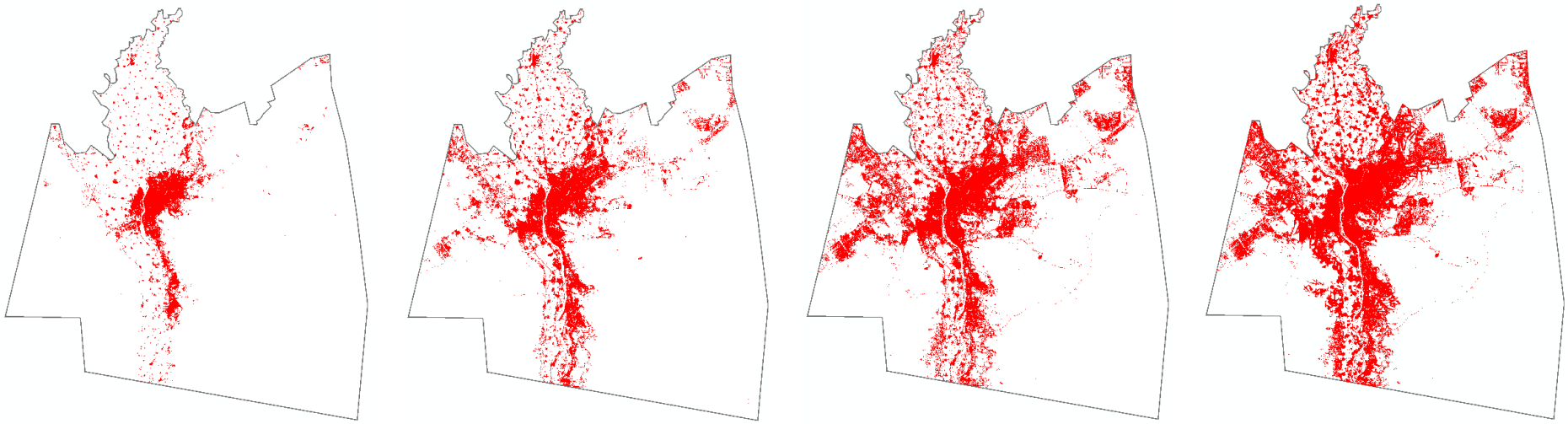


Figure 36. 2025's LCLU estimated map

Figure 37 shows the urban growth in GCR in 1984, 2003, 2014 and the estimated urban settlements in 2025. The urban expansion has boomed over 30 years, from 1984 to 2014, and the modeling results confirms that it will be increasing in 2025. The urban areas were 41 488, 95 793 and 154 861 hectares in 1984, 2003 and 2014, representing 4.64%, 10.71% and 17.32% of the total area of GCR, respectively. In 2025, according to the model estimations, the urban in GCR will expand to 196 047 hectares which is about 21.93% of the region. This unplanned vast growth is a serious threat towards the ecological system, as there is an obvious tendency of a continuous agricultural loss with increasing rates: 13%, 12% and 14% within 19, 11 and 11 years between 1984 – 2003, 2003 – 2014 and 2014 – 2025, respectively. This implies a current and upcoming shortage in food production, especially with the increasing population in the Egyptian capital.



a- Urban settlements in 1984 **b- Urban settlements in 2003** **c- Urban settlements in 2014** **d- Predicted urban settlements in 2025**

Figure 37. Urban growth in GCR

On the other hand, and if this urban sprawl trend continues, the cultural heritage in the study area will be very soon fully surrounded by glassy, steel and reinforced concrete buildings. Actually this behaviour unfortunately has already started around some monuments in GCR. The summation of the total urban patches was calculated over time in a 5-km buffer zone around some of the vital cultural heritage places in GCR as illustrated in Figure 38. These places are: first, the Great Giza Pyramids and the Sphinx area representing the Pharaonic history. Second, Al-Hussien Mosque, Khan Al-Khalily area, Al-Moez Street and the Citadel, all in one area representing an essential part of the ancient Islamic Cairo. Finally, the third one is Al-Baron Palace which reflects the modern Belgian architecture.

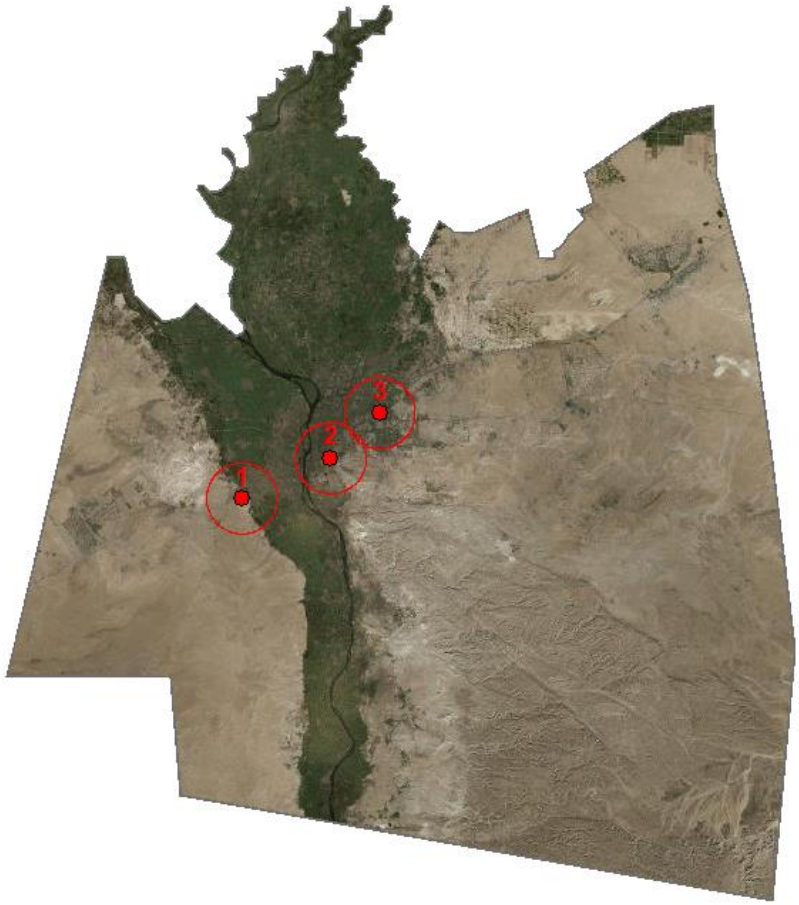


Figure 38. Three different cultural heritage areas representing 1- Pharaonic, 2- Islamic and 3- Modern history

Figure 39 shows the total area of urban settlements within a 5-km buffer around the major monuments previously illustrated in Figure 38. Islamic Cairo and Al-Baron Palace are surrounded by much more dense urban than the Pyramids area, which had almost one-fifth and one-sixth of the urban areas around both of them in 1984 (9 271 hectares). This value had a dramatic increase over time to the extent that it is expected to have an equal surrounding urban to the Islamic Cairo in 2025 (70 512 hectares). This calls for a preserving plans from the Egyptian policy makers, not only for agriculture, but also for monuments and cultural heritage.

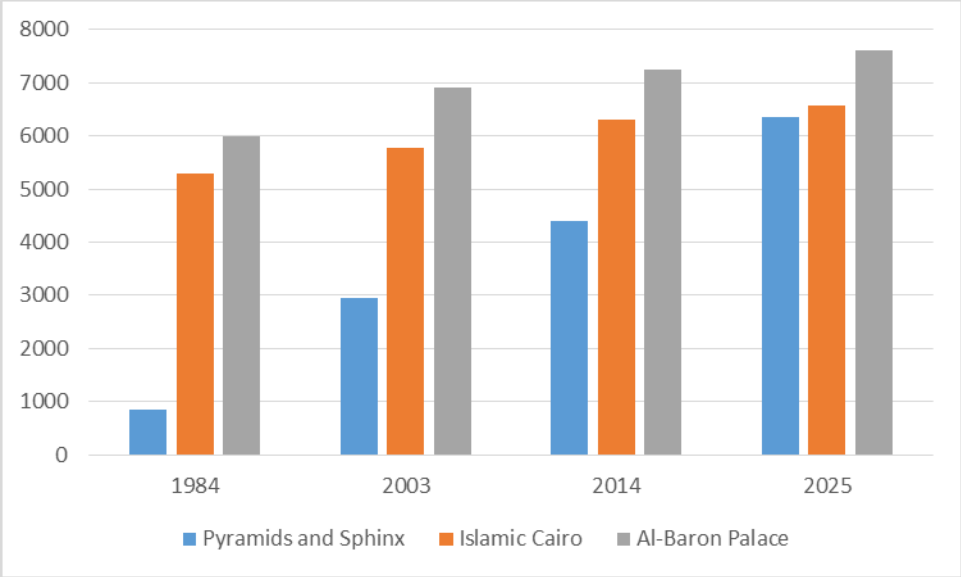


Figure 39. Total area of urban patches within 5-km buffer (hectares)

6. CONCLUSION

This study was performed to detect, analyse and model the urban growth in GCR using remote sensing data over a period of 30 years, from 1984 to 2014. Three Landsat scenes obtained in 1984, 2003 and 2014 were classified using SVM with an overall accuracy of 96%, 97.3% and 96.3%, respectively. These high values highlight the valuable capabilities of the SVM classifier, superior to other available techniques. NDBI was used to extract the unplanted agricultural fields, which were misclassified as vegetation in 1984's and 2003's LCLU maps, and it successfully targeted them for urban reclassification. However, it can affect the accuracy of both classes; urban and vegetation, relative to the remaining classes, but with no significant influence on the overall accuracy. For best practice, this technique should be applied only in areas where the problem exists. Dividing the image may help as well.

Eight metrics were computed to analyze the spatial and temporal urban growth from 1984 to 2014: CA (Class Area), NP (Number of Patches), ED (Edge Density), LPI (Largest Patch Index), ENN_MN (Euclidian Mean Nearest Neighbor Distance), AWMPFD (Area Weighted Mean Patch Fractal Dimension), CONTAG (Contagion) and Shannon's Entropy. All emphasis the more dense urbanization taking place in GCR as well as the dispersed and fragmented landscape as a result of individual urban establishment.

LUCC detection was determined using a high level land cover mapping technique in which combines binary maps of change / no change information with the post-comparison approach. In binary maps production, NDVI difference technique performed well in detecting pixels that changed from vegetation, but failed to detect new urban settlements from other classes, especially desert. Whereas a combination of NDVI, NIR and RED differences, when tried instead, it proved a much better performance in capturing urban change. The most significant changes are the transitions from vegetation and desert to urban. From 1984 to 2003, 13% of the vegetation has been lost to urban, and 12% has been lost between 2003 and 2014, representing 19 179 and 16 486 hectares, respectively. While 3% of desert areas turned to urban between 1984 and 2003, raised to 5% between 2003 and 2014, which are equivalent to 21 417 and 31 045 hectares, respectively. The absence of low activation against agricultural land levelling and the application of desert reconstruction strategies to build new communities outside the Nile Valley, are the main reasons for such massive urbanization.

Both were considered the main transitions in modeling the change between 1984 and 2003, with the same driving forces: DEM, slope, distance to road network, and distance to existing urban areas. The validation results of the simulated 2014 map against 2014's real LCLU map showed perfect (KI > 090%), fair (AUC = 71%) and low (FOM=12.76%) accuracies using VALIDATE module, ROC module and the visual validation approach, respectively. These results indicate that the correct localization of the projected LULCC is more pronounced than the quantity of the projected LULCC, so additional predictive factors (*e.g.* distance to markets) might help. Also applying MCE (*e.g.* AHP) might result in more accurate suitability map, as it weights each factors based on expert knowledge. Nevertheless, the model is considered reasonable for prediction even with the ROC validation results. The projected 2025 LCLU map estimates an urban transition of 14% from vegetation and 4% from desert, between 2014 and 2025, which are equivalent to 16 512 and 24 687 hectares,

respectively. The study suggests that vegetation and desert areas are being dramatically urbanized. On the other hand, the area of urban patches were calculated over time (from 1984 to 2025) in a 5-km buffer around three main historical places in GCR: the Great Pyramids, Islamic Cairo and Al-Baron Palace. The results showed similar gradual patterns of increase for Islamic Cairo and Al-Baron Palace, as they both have been surrounded by dense urban settlements. This trend is estimated to keep increasing in 2025 Al-Baron Palace, whereas no significant urban development is expected to occur any more within a 5-km ring around the Islamic Cairo area, as it is already full of urban, showing no promising vacancies of further settlements. The Pyramids area showed a booming urbanization since the urban density in 2014 was five times what was in 1984, moreover, this trend is estimated to even increase in 2025 to be almost equal to the Islamic Cairo records. These results imply that the Egyptian cultural heritage will be fully surrounded with glassy, steel and reinforced concrete buildings in the near future, especially around the Pyramids area. This indicates that these monuments are gradually losing their value and unique appearance, merging with new urban. If this trend continues, protection policies have to be undertaken to not only preserve agricultural fields, but also cultural heritage in addition to other resources (*e.g.* aquifers and ground water activities) that may be affected as well.

BIBLIOGRAPHIC REFERENCES

ARAYA, Y. H. and CABRAL, P., *Analysis and modeling of urban land cover change in Setúbal and Sesimbra, Portugal*. Remote Sensing 2.6 (2010): 1549-1563.

ARSANJANI, J., HELBICH, M., KAINZ, W., and BOLOORANI, A. D., *Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion*. International Journal of Applied Earth Observation and Geoinformation 21 (2013): 265-275.

BATTY, M., *Urban modeling*. International Encyclopedia of Human Geography, Elsevier, Oxford (2009).

BENARCHID, O. and RAISSOUNI, N., *Support Vector Machines for Object Based Building Extraction in Suburban Area using Very High Resolution Satellite Images, a Case Study: Tetuan, Morocco*. IAES International Journal of Artificial Intelligence (IJ-AI) 2, no. 1 (2013): 43-50.

BROWN, D. G., DANIEL, G., VERBURG, P. H., JR, R. G. P., and LANGE, M. D., *Opportunities to improve impact, integration, and evaluation of land change models*. Current Opinion in Environmental Sustainability 5, no. 5 (2013): 452-457.

BROWN, D. G., WALKER, R., MANSON, S., and SETO, K., *Modeling land use and land cover change*. In Land Change Science, pp. 395-409. Springer Netherlands, 2004.

Central Agency for Public Mobilization and Statistics. [Online]. Available from: http://www.capmas.gov.eg/Default.aspx?lang=2 [Accessed 20th September 2014].

EASTMAN, J. R., 2012. *Idrisi Selva Manual: Version 17.01*. Idrisi Production Source Code 1987–2012, Clark University, 1–322.

EPSTEIN, J., PAYNE, K., KRAMER, E., 2002. *Techniques for mapping suburban sprawl*. Photographer Engineering Remote Sensing. 63 (9).

GONG, W., LI, Y., FAN, W., and STOTT, P., *Analysis and simulation of land use spatial pattern in Harbin prefecture based on trajectories and cellular automata—Markov modelling*. International Journal of Applied Earth Observation and Geoinformation 34 (2015): 207-216.

HAAS, J. and BAN, Y., *Urban growth and environmental impacts in Jing-Jin-Ji, the Yangtze, River Delta and the Pearl River Delta*. International Journal of Applied Earth Observation and Geoinformation 30 (2014): 42-55.

HEROLD, M., GOLDSTEIN, N. C., and CLARKE, K. C., *The spatiotemporal form of urban growth: measurement, analysis and modeling*. Remote sensing of Environment 86.3 (2003): 286-302.

HUANG, W., LIU, H., LUAN, Q., JIANG, Q., LIU, J., LIU, H., *Detection and prediction of land use change in Beijing based on remote sensing and GIS*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37 (2008): 75-82.

J. Pontius, R. Gilmore, and H. Chen. "Land change modeling with GEOMOD." *Clark University* (2006).

JIN, S. L. Y., DANIELSON, P., HOMER, C., FRY, J., and XIAN, G., *A comprehensive change detection method for updating the national land cover database to circa 2011*. Remote Sensing of Environment132 (2013): 159-175.

KIM, I. G. Y. J., PARK, S., and TENHUNEN, J., *Predicted Land Use Change in the Soyang River Basin, South Korea*. In 2011 TERRECO Science Conference, Karlsruhe Institute of Technology, Garmisch-Partenkirchen, Germany, pp. 2-7. 2011.

KIPPER, R., FISCHER, M., *Cairo's informal areas: Between urban challenges and hidden potentials – Facts, voices and visions*. GTZ Egypt, Participatory Development Program in Urban Areas, Cairo (2009).

LU, D. and WENG, Q., *A survey of image classification methods and techniques for improving classification performance*. International Journal of Remote Sensing, 28 (5) (2007).

LU, D., WENG, Q., MORAN, E., LI, G., and HETRICK, S., *Remote sensing image classification*. Advances in environmental remote sensing: sensors, algorithms, and applications. Boca Raton, FL: CRC Press/Taylor & Francis Group Publishers Part II (2011): 219-240.

MARTINS, V. N., P., and CABRAL, SOUSA E SILVA, D., *Urban modelling for seismic prone areas: the case study of Vila Franca do Campo (Azores Archipelago, Portugal)*. Natural Hazards and Earth System Science 12.9 (2012): 2731-2741.

MAS, J. F., *Monitoring land-cover changes: A comparison of change detection techniques*. International journal of remote sensing 20, no. 1 (1999): 139-152.

MOGHADAM, H. S. and HELBICH, M., *Spatiotemporal urbanization processes in the megacity of Mumbai, India: A Markov chains-cellular automata urban growth model*. Applied Geography 40 (2013): 140-149.

MOHAMED, E., *Analysis of urban growth at Cairo, Egypt using remote sensing and* GIS. Natural Science 2012 (2012).

MÖLLER, M., *Remote Sensing for the Monitoring of Urban Growth Patterns*. In The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences Edited by: Moeller, M and Wentz. 2005.

MOUNTRAKIS, G., JUNGHO, I., and OGOLE, C., *Support vector machines in remote sensing: A review*. ISPRS Journal of Photogrammetry and Remote Sensing 66, no. 3 (2011): 247-259.

NADOUSHAN, M. A., SOFFIANIAN, A., and ALEBRAHIM, A., *Predicting Urban Expansion in Arak Metropolitan Area Using Two Land Change Models*. World Applied Sciences Journal 18.8 (2012): 1124-1132.

O'CONNELL, J., CONNOLLY, J., VERMOTE, E. F., and HOLDEN, N. M., *Radiometric normalization for change detection in peatlands: a modified temporal invariant cluster approach.* International Journal of Remote Sensing34, no. 8 (2013): 2905-2924.

OLMEDO, M. T. C., PAEGELOW, M., and MAS, J. F., *Interest in intermediate soft-classified maps in land change model validation: suitability versus transition potential*. International Journal of Geographical Information Science 27.12 (2013): 2343-2361.

PAL, M., and MATHER, P. M., *Support vector machines for classification in remote sensing.* International Journal of Remote Sensing 26, no. 5 (2005): 1007-1011.

PINILLA, C. J., C., CRESPO, R. G., and CASTILLO, A., *Relative radiometric normalization of multitemporal images*. International Journal of Interactive Multimedia and Artificial Intelligence 1, no. 3 (2010).

RAMACHANDRA, T. V. and KUMAR, U., *Geographic resources decision support system for land use, land cover dynamics analysis*. In Proceedings of the FOSS/GRASS Users conference, 2004.

REGMI, R. R., SAHA, S. K., and BALLA, M. K., *Geospatial Analysis of Land Use Land Cover Change Modeling at Phewa Lake Watershed of Nepal by Using Cellular Automata Markov Model.* (2014).

RODRÍGUEZ, N. E., ARMENTERAS-PASCUAL, D., and ALUMBREROS, J. R.. *Land use and land cover change in the Colombian Andes: dynamics and future scenarios*. Journal of Land Use Science 8.2 (2013): 154-174.

ROY, H. G., DENNIS, M. F., and EMSELLEM, K., *Predicting Land Cover Change in a Mediterranean Catchment at Different Time Scales.* InComputational Science and Its Applications–ICCSA 2014, pp. 315-330. Springer International Publishing, 2014.

SÁEZ, J.A., GALAR, M., LUENGO, J., and HERRERA, F., *Tackling the problem of classification with noisy data using Multiple Classifier Systems: Analysis of the performance and robustness.* Information Sciences 247 (2013): 1-20.

SANGERMANO, F., TOLEDANO, J., and RONALD, J., Eastman. *Land cover change in the Bolivian Amazon and its implications for REDD+ and endemic biodiversity*. Landscape ecology 27.4 (2012): 571-584.

SINGH, A., *Review article digital change detection techniques using remotely-sensed data*. International journal of remote sensing 10, no. 6 (1989): 989-1003.

SINGH, S. K., SRIVASTAVA, P. K., GUPTA, M., THAKUR, J. K., and MUKHERJEE, S., *Appraisal of land use/land cover of mangrove forest ecosystem using support vector machine.* Environmental Earth Sciences 71, no. 5 (2014): 2245-2255.

State Information Service. Egypt Gateway [Online]. Available from: http://www.sis.gov.eg/En/Default.aspx [Accessed 26th September 2014].

SUBEDI, P., SUBEDI, K., and THAPA, B., *Application of a Hybrid Cellular Automaton–Markov (CA-Markov) Model in Land-Use Change Prediction: A Case Study of Saddle Creek Drainage Basin, Florida*. Applied Ecology and Environmental Sciences 1.6 (2013): 126-132.

SWETS, J., 1988, *Measuring the accuracy of diagnostic systems*. Science, 240, 1285–1293.

TEWOLDE, M. G. and CABRAL, P., *Urban sprawl analysis and modeling in Asmara, Eritrea*. Remote Sensing 3, no. 10 (2011): 2148-2165.

THIES, B., MEYER, H., NAUSS, T., and BENDIX, J., *Projecting land-use and land-cover changes in a tropical mountain forest of Southern Ecuador.* Journal of Land Use Science 9, no. 1 (2014): 1-33.

THORPE, D., *Should We Curb City Growth to Feed the Planet?*. [Online]. Available from: http://sustainablecitiescollective.com/david-thorpe/222431/should-we-curb-city-growth-feed-planet [Accessed 28th December 2014].

United Nations Fund for Population Activities. *State of World Population 2007: Introduction*.[Online].Availablefrom:https://www.unfpa.org/swp/2007/english/print/introduction.html [Accessed 23th July 2014].

United Nations Fund for Population Activities. Urbanization: A Majority in Cities. [Online]. Available from: http://www.unfpa.org/pds/urbanization.htm [Accessed 23th July 2014].

United Nations, Department of Economic and Social Affairs. [Online]. Available from: http://upload.wikimedia.org/wikipedia/commons/c/c8/Percentage_of_World_Population_Urban_Rural.PNG [Accessed 23th July 2014].

VAZ, E. D. N., CAETANO, M., and NIJKAMP, P., *Trapped between antiquity and urbanism–a multi-criteria assessment model of the greater Cairo Metropolitan area*. Journal of Land Use Science 6, no. 4 (2011): 283-299.

VAZ, E. D. N., NIJKAMP, P., PAINHO, M., and CAETANO, M., *A Multi-Scenario Forecast of Urban Change: A Study on Urban Growth in the Algarve*. Landscape and Urban Planning, Vol. 104, No. 2, 2012, pp. 201-211.

VEETTIL, B. K., *A Comparative Study of Urban Change Detection Techniques Using High Spatial Resolution Images.* Proceedings of the 4th GEOBIA, May 7-9, 2012 - Rio de Janeiro - Brazil. p.029.

VEGA, P. A., MAS, J. F., and ZIELINSKA, A. L., *Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest.* Environmental Modelling & Software 29, no. 1 (2012): 11-23.

VERBURG, P. H., SCHOT, P. P., DIJST, M. J., and VELDKAMP, A., *Land use change modelling: current practice and research priorities*. GeoJournal 61, no. 4 (2004): 309-324.

World Health Organization. Urbanization: Urban Population Growth. [Online]. Available from: http://www.who.int/gho/urban_health/situation_trends/urban_population_growth_text/en/ [Accessed 23th July 2014].

YIN, Z. Y., D. STEWART, J., BULLARD, S., and MACLACHLAN, J. T., *Changes in urban built-up surface and population distribution patterns during 1986–1999: A case study of Cairo, Egypt.* Computers, Environment and Urban Systems 29.5 (2005): 595-616.