Masters Program in Geospatial Technologies

EXPLORING LAND USE LAND COVER CHANGE TO UNDERSTAND URBAN WARMING EFFECT IN HANOI INNER CITY, VIETNAM

Tran Xuan Duy

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

EXPLORING LAND USE LAND COVER CHANGE TO UNDERSTAND URBAN WARMING EFFECT IN HANOI INNER CITY, VIETNAM

Dissertation supervised by

Filiberto Pla Bañón, PhD

Professor, Dept. Lenguajes y Sistemas Informaticos

Universitat Jaume I, Castellón, Spain

Dissertation co-supervised by

Soe W.Myint, PhD

Professor, School of Geographical Sciences and Urban Planning,

Arizona State University

Mario Caetano, PhD

Professor, Instituto Superior de Estatística e Gestão da Informação

Universidade Nova de Lisboa, Lisbon, Portugal

March 2016

ACKNOWLEDGMENTS

I would like to express my deepest and sincere gratitude to my supervisors, Prof. Dr. Filiberto Pla Bañón, Prof. Dr Soe W.Myint, and Prof. Dr Mario Caetano, for their kindness in providing an opportunity to be their advices. I am grateful for their valuable supervision, suggestions, encouragement, helpful comments and constructive criticism throughout the period of this research work.

Significantly, I would like to express my sincere gratitude to Dr Pedro Latorre Carmona, Department of Computer Languages and Systems, Jaume I University for his support, and guidance to my research.

My thanks also go to all my professors, lectures and administrators of the Master program in Geospatial Technologies in Jaume I University (UJI), Institute for Geoinformatics (ifgi) - Muenster University, and NOVA Information Management School (NOVA IMS) for their teaching, helpful suggestions, and kindly support during my master.

My heartfelt gratitude and sincere thanks go to all my professors, lectures and members of the Faculty of Geography, Hanoi National University of Education.

Finally yet importantly, I would like to convey my heartfelt love, respects and gratitude to my beloved family members and my dear friends for their never-ending encouragements, inspiration and sacrifices given to me.

EXPLORING LAND USE LAND COVER CHANGE TO UNDERSTAND URBAN WARMING EFFECT IN HANOI INNER CITY, VIETNAM

ABSTRACT

Recently, urbanization is occurring rapidly in Hanoi, the second largest city in Vietnam. The process is profoundly reflected in Hanoi inner city where the socio-economic development is faster than in other areas. This has led to the acquisition of agricultural land that in turn, has resulted in land use changes, and subsequently increasing the residential, commercial and industrial land. The transformation between different land use types especially the urban expansion will crucially influence the land surface temperature pattern (LST). This will severely affect to the community in relation to people's health and energy consumption. Exploring land use land cover (LULC) change to understand urban warming effect is a necessary work for community and local government. The research can be used as a scientific basis for urban planners in urban planning and management as well as to increase the community awareness in urban warming effect. The purpose of this research is to determine and analyze the relationship between LULC change and LST pattern. To achieve the research goal, we need to accomplish a series of specific objectives. First, we perform supervised maximum likelihood classification method and change detection to determine the patterns and rate of change, and land cover and land use transformation within and around Hanoi inner city. Then we explore the relationship between land surface temperature and a) vegetation, b) man-made features, and c) crops land using normalized vegetation, and built-up indices within each LULC type. After that, we employ a Markov chains model to simulate future LULC change using different environmental and planning scenarios. Finally, we apply linear and non-linear regression to predict future urban climate patterns in Hanoi inner city using the predicted land cover and land use change.

KEYWORDS

Land use land cover

Urban heat island

Regression analysis

Landsat

Remote Sensing

Geographical Information Systems

Spatial analysis

ACRONYMS

LULC – Land use land cover

UHI – Urban heat island

GIS – Geographic Information System

NDVI – Normailzed Differences Vegetation Index

NDBI – Normailzed Differences Built-up Index

KRR – Kernel Ridge Regression

NDWI – Normailzed Differences Water Index

LST – Land Surface Temperature

CONTENT

INDEX OF TABLES

INDEX OF FIGURES

1. INTRODUCTION

1.1. Background and Motivation

A great number of studies have shown that the increase in the heat storage capacity of urban surfaces has created the so-called urban heat island effect (UHI), in which built up areas are hotter than nearby rural areas (Oke, 1982; Taha, 1997; Rizwan, Dennis & Chunho, 2008). For instance, the surface temperature inside a city is normally higher than the surrounding areas by 2 - 3 Celsius degrees due to urban heat island phenomenon. This difference creates a negative impact on people and environment because it decreases the air quality, increases the energy consumption, loses the biological control, and affects people's health (Kikegawa, Genchi, Yoshikado & Kondo, 2003; Grimmond, 2007; Meineke, Dunn & Frank, 2014; Plocoste, Jacoby-Koaly, Molinié & Petit, 2014).

Advances in thermal remote sensing, Geographical Information Systems (GIS), and statistical methods have enabled the research community to characterize and examine the relationship between UHI and landscape. A great number of studies related to the UHI issue have been carried out, giving a significant contribution both in theory and practice for policy makers and researchers (Quattrochi & Luvall, 1999; Yuan & Bauer, 2007; Rizwan et al., 2008; Junxiang, Conghe, Lu, Feige, Xianlei & Jianguo, 2011; Kumar, Bhaskar & Padmakumari, 2012; Radhi, Fikry & Sharples, 2013; Myint, Wentz, Brazel & Quattrochi, 2013; Zhou, Qian, Li, Li & Han, 2014; Adams & Smith, 2014; Coseo & Larsen, 2014; Juer, Shihong, Xin & Luo, 2014; Chun & Guldmann, 2014; Rotem-Mindali, Michael, Helman & Lensky, 2015; Kourtidis, Georgoulias, Rapsomanikis, Amiridis, Keramitsoglou, Hooyberghs & Melas, 2015). These studies cover a wide range of topics, carried out in a variety places, and diverse in terms of scale. Recent UHI studies can be categorized in different kinds of topics, including urban ventilation and surface material alteration, outdoor thermal and wind comfort during the exacerbated UHI by the local heat waves, UHI spatial–temporal variation, model evaluation and enhancement, future temperature forecast, and building energy saving (Mirzaei, 2015).

Besides air temperature, land surface temperature (LST) derived from remote sensing data has been widely used as an indicator for UHI research. With the benefit of thermal remote sensing, LST has been provided from numerous sources (such as

Landsat, MODIS, and ASTER) which cover a wide range of geographical surface, and alter from global to local scale. These abundant data significantly contribute to the revolution of UHI research. In its relationship with landscape, surface temperature has a direct interaction with land use land cover (LULC) characteristics (Quattrochi & Luvall, 1999). Therefore, studying of the relationship between LULC and LST is a key to understand UHI. Exploring the relationship between LULC and LST does not only help researchers explain the cause, effect, and distribution of UHI, but also provides a certain way to model the UHI trend from a spatio-temporal perspective.

For the LULC indicators, two main groups have been generally used in LULC - UHI relationship analysis. The first group is called land cover indices extracted by using band ratios obtained from different single bands in satellite images, such as Normalized Difference Vegetation Index (NDVI), Normalized Difference Built-up Index (NDBI), or Bareness Index (NDBnI). These indices are an effective alternative for quickly and objectively land mapping (Zha, Gao & Ni, 2003). Each land cover index takes advantage of the unique spectral response to specific LULC types such as built-up or vegetation (Guo, Wu, Xiao, Chen, Liu & Zhang, 2015). These indices also can be used to examine the seasonal variability in the relationship between land cover and UHI (Zhou et al., 2014).

However, land cover indices vary day by day and do not have a good prediction performance. In addition, in tropical areas, using high temporal resolution of land indices faces the challenges from weather issues such as clouds and water vapor. Moreover, these indices have some restrictions if we want to have better hints about how landscape affects UHI. For instance, by using NDBI, we cannot distinguish very detailed urban features. Therefore, LULC types (for example urban, impervious surface, vegetation, crops land), have been applied as another indicator in UHI research (Walawender, Szymanowski, Hajto & Bokwa, 2014; Coseo & Larsen, 2014). This group is more stable and provides a better alternative for future value prediction. In addition, from an urban planning and land use management perspective, the result related to LULC types in the UHI research is more practical. To achieve the best result, a general trend of many studies is that they used both land cover indices and LULC types in the landscape-UHI relationship analysis. Regression has been commonly used as the best solution to gain insight into such correlation (Yuan &

Bauer, 2007; Adams & Smith, 2014; Zhou et al., 2014; Coseo & Larsen, 2014; Kim & Guldmann, 2014; Juer et al., 2014; Rotem-Mindali et al 2015;).

Land use/land cover change is strongly related to urbanization and industrialization, which are the processes of transformation from rural to urban land, as well as from agriculture to industrial areas. These processes are happening fast in Hanoi, the second largest city of Vietnam, leading to the expansion of urban space and growth of the urban population. Urban area occupied nearly 30 percent in 2010 and the increasing rate has reached 3.8 percent per year (world bank, 2011). The process is profoundly reflected in Hanoi inner city where it covers only 7 percent in total area but counts more than 40 percent of population (Hanoi portal, 2015). Urbanization and industrialization have resulted in the loss of agricultural land and subsequently in the increase of residential, commercial, and industrial lands (Thuy, Duy, Duc & Luong, 2014). The transformation between different land use types (especially the urban expansion) has crucially lead to the change in land surface temperature pattern which is the main cause of the urban heat island effect. Recently, urban warming effect is considered as one of the major problems in Hanoi city due to large concentration of population and economic activities. According to the climate change effect (Niem, Wen, Renee, Darlene & Duy, 2013), the influence of urban heat island will be much more severe in Hanoi city in the future.

Research activity related to the relationship between land use land cover and urban heat island in Hanoi city has been limited so far. However, they have succeeded in showing the existence of urban heat island in the city as well as the effect of land use change in urban temperature change (Andhang, Han, Phuong, Tetsu & Takahiro, 2014; Nam, Kubota & Trihamdani, 2015). Nevertheless, the temperature data used in these studies was collected from a series of weather stations. The limited number of these stations has created the main issues related to the accuracy and scale of the analysis. At first, this raised the question about how accurate conclusions can be in the areas where no weather stations exist. Consequently, these researches cannot reveal a comprehensive and detailed picture of large scale urban heat island pattern. In terms of urban planning and management practices, high resolution of urban heat island map and analysis will be much more efficient. To improve the drawbacks of these works, our research will use temperature derived from satellite remote sensing thermal band. Compared to the temperature collected from weather stations, thermal imagery has

complemented in situ measurements by providing spatial richness across multiple temporal and spatial scales and in vertical direction (Myint et al., 2013). In this research, we used both land indices and LULC types to have a better gaining insight into the relationship between LULC and its change and UHI in Hanoi inner city. In terms of the spatio-temporal context, we put our effort to draw a broad and comprehensive picture about the change in LULC and UHI patterns in the past, present, and future. We also emphasized on the urbanization and urban fragmentation to examine how this process affected UHI.

This research is essential for the local government and the people living in the city. It will provide quantitative evidence about the major issues in LULC change and UHI in the city. The results also highlight the urban growth and its negative effects to the environment. Effectively, the research provides a prediction of urban land use change and climate patterns in different scenarios. One key aim is that the research will make an important impact to the policy makers and community awareness. These research outcomes will provide a scientific basis for a sustainable urban planning and management. Due to the similar magnitude of its impact on LST as LULC change (*Luyssaert et al., 2014),* the improvement in land management can significantly contribute to UHI mitigation. It also provides a strong argument to change the culture practices toward contributing more in urban environment improvement and protection.

1.2. Research Objectives

- To apply supervised maximum likelihood classification and change detection to determine the patterns and rate of change, and LULC transformation within and around Hanoi inner city in 2003-2015.
- To gain insight into the relationship between land surface temperature and a) vegetation, b) man-made features, and c) crops land using normalized vegetation and built-up indices within each LULC type analyzed afterwards.
- To employ a Markov chains model to simulate future LULC change using different environmental and planning scenarios.
- To apply non-parametric regression to predict future urban climate patterns in Hanoi inner city using the predicted land cover and land use change.

1.3. Research structure

The research is structured as follows: section 2 briefly describes about the data and study area. Section 3 focuses on the methods used during this research. Section 4 presents the main results and discussion about them. Section 5 contains the conclusion about the achievements and limitation of the research, as well as proposed future work.

2. DATA AND STUDY AREA

2.1. Data used

Surface Reflectance High Level Data Products imagery including Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imager (OLI) acquired in May 05, 2003; May 24, 2007; and July 01, 2015 are the main data used for this research. These free accessed data are processed by NASA that generates radiometric calibration and atmospheric correction algorithms to Level-1 products (http://earthexplorer.usgs.gov/). Satellite images were used to classify the LULC, retrieve LST, and calculate NDBI, NDVI, and NDWI indices. The limitation of these data is that all of them were acquired around 10 AM. At that time the LST was quite low (less than 3-6 Celsius degrees compared to the maximum LST at 13 PM). Hence, the absolute LST value retrieved in the morning data did not strongly reflect the severity of UHI in the study area. However, this issue did not significantly affect on the result because the research focused on the UHI pattern rather than as the absolute value. Besides satellite images, the research also used the reports, statistical data, and geographical information from institutions and organizations from the Vietnamese government, including the General Statistics Office (GSO), the Ministry of Natural Resources and Environment (MONRE), and the National Centre for Hydro-Meteorological Forecasting (NCHMF).

2.2. The study area

Our research focused on ten inner districts in the Hanoi city (and the area surrounding them - Figure 1), which is the capital of Vietnam, located in the center of the Red river delta, being the second largest delta in the country. Hanoi inner city covers only 7% of the total area but contains more than 40% of population, and 20% of the city urban land (Hanoi portal, 2015). Urbanization has occurred faster in and around inner city than in other areas due to the advantages of location and human resources. This has led to the acquisition of agricultural land that in turn, has resulted in land use changes, and subsequently increasing the built-up area such as residential, commercial/industrial, and impervious surface. Moreover, the effect is more severe when urbanization takes place without sustainable planning and control. The consequence is the distribution of industrial parks inside the residential area, lack of open and green spaces, and growing up of very high urban density.

The study area is a small and flat plain with an average elevation that is less than 10 meters above the sea level (Yonezawa, 2009). Due to this characteristic, the temperature within the study area is independent of the elevation or latitude change. It is an important condition to reduce the elevation and latitude impacts when assessing the relationship between surface temperature and land use land cover, as well as modeling the future temperature pattern. The city has a great number of water bodies. This dense system of rivers and lakes gives a cooling effect to the urban center in the summer. However, these open water areas are decreasing due to urban expansion and environment pollution.

Figure 1: Map of Vietnam, Hanoi, and Hanoi inner city

Located within the warm humid subtropical climate zone, the city has the typical climate of northern Vietnam with hot and humid summers and cold and dry winters. The summer season starts from May to August with an average temperature

of 29^0C (NCHMF). As a low aptitude center, with the impact of the Foeln (a type of dry, warm, down-slope wind occurred in the downwind side of a mountain range), the city often experiences several hot period in summer time. The consequences of high built-up density - hot and humidity weather combination creates a tremendous effect in the city. Extreme hot temperature events have happened during the last few years. In June 16, 2010, the mean temperature in the city reached 34.6^0C (night temperature was 30.4^0C, and day temperature was 39.6^0C), which was the highest record since 1961 (NCHMF). In May 28, 2015 the temperature exceeded 40^0C, being the highest temperature since the beginning of records in history (NCHMF). Under the effect of rapid urbanization, industrialization and global warming effect, the summer will be much more severe in the city in future.

3. RESEARCH METHODS

The methods applied in this research include supervised maximum likelihood classification and change detection analysis, land surface temperature retrieval method, land use land cover indices, High/Low Clustering Getis ord statistics, urban landscape analysis, Multi-Layer Perceptron neural network and Markov chain model, and regression analysis. These methods aim to answer the research questions related to LULC and LST characteristics and LULC-LST relationship in the study area.

3.1. Supervised Maximum likelihood classification

Supervised Classification method using Maximum Likelihood algorithm was used to classify land use land cover of the study area. This method assigns a pixel to the class of highest probability based on a statistical distance which is determined by the mean values and covariance matrix of the clusters (Klaus Tempfli, Norman Kerle, Gerrit Huurneman & Lucas Janssen, 2009). Composite imagery used False Color band combination of bands 7,6,4 in Landsat 8 and 7,5,3 in Landsat 5 TM and Landsat 7 ETM+ to have better visualization of urban environments. Training samples for each land use/land cover type were determined by comparing the False Color composition image, Google Earth, and in-situ information. The classification result included eleven classes of land use/land cover within six categories: vegetation, agriculture (high-density, low-density, wetland), urban (high-residential, low-residential, industrial and commercial, impervious surface), vacant land, water, and sandbars. Impervious surface classification in this research means pavements as roads, sidewalks, driveways and parking lots that are covered by impenetrable materials such as asphalt, concrete, brick, and stone. Vacant land or bare land includes bare soil and other clear surface such as pre-construction area and open area in industrial parks where does not covered by vegetation or impervious material.

Accuracy assessment for each classification was made by uploading 350 points taken from each classified image to GoogEarth Pro to compare their similarity. Based on this, we calculated the Overall accuracy, Producer's Accuracy, User's Accuracy and Kappa index to evaluate the classification accuracy. We then compiled these classified images by using the overlay tool in ArcGIS to assess the land use change from 2003 to 2015.

3.2. Land surface temperature retrieval

Our research used the thermal infrared bands of different types of Landsat images (band 6 of Landsat 5 TM, Landsat 7 ETM+ and band 10 of Landsat 8) to estimate the land surface temperature of Hanoi inner city. The main steps are taken as below based on the guidelines in Landsat data users handbook published by United States Geological Survey (Landsat 7, 2011; Landsat 8, 2015):

(1) Conversion of pixel values from Digital number units to radiation values

Landsat 5TM and Landsat 7 ETM+

$$L\lambda = Grescale * QCAL + Brescale, \text{ is also expressed as}$$

$$L\lambda = ((LMAX\lambda - LMIN\lambda)/(QCALMAX-QCALMIN)) * (QCAL-QCALMIN) + LMIN\lambda$$

Landsat8

$$L\lambda = ML* QCal + AL$$

(2) Removing the atmospheric effects

Removing the atmospheric effects in the thermal region is the process to convert the space-reaching radiance to a surface-leaving radiance. This study applied the method proposed by Barsi, Barker, & Schott (Barsi et al., 2003):

$$LT = \tau\varepsilon L\lambda + Lu + (1-\varepsilon) Ld$$

The atmospheric transmission (τ), upwelling (Lu), and downwelling (Ld) parameters were calculated by using the "Atmospheric Correction Parameter Calculator" online tool (http://atmcorr.gsfc.nasa.gov). This tool applies the National Centers for Environmental Prediction (NCEP) modeled atmospheric global profiles for a particular date, time and location as the input data (Barsi et al., 2003). The emissivity of the surface was calculated by using the method proposed by Sobrino, Jiménez-Muñoz, & Paolini (Sobrino et al., 2004).

(3) Converting the surface-leaving radiance to apparent surface temperature by using the Landsat specific estimate of the Planck curve:

$$T = K2/\ln(K1/LT + 1)$$

(4) Transfer land surface temperature values from Kelvin units to Celsius units (^{0}C):

$$T\ (^0C) = T\ (Kelvin) – 273.15.$$

Table 1 presents all the parameters introduced at the beginning of section 3.2.

Table 1: The parameters in LST retrieval

Parameters	Definition
Lλ	the spectral radiance at the sensor's aperture
Grescale	the rescaled gain (the data product "gain" contained in the Level 1 product header or ancillary data record)
Brescale	the rescaled bias (the data product "offset" contained in the Level 1 product header or ancillary data record)
QCAL	the quantized calibrated pixel value
LMINλ	the spectral radiance that is scaled to QCALMIN
LMAXλ	the spectral radiance that is scaled to QCALMAX
LMAXλ	the spectral radiance that is scaled to QCALMAX
QCALMIN	the minimum quantized calibrated pixel value (corresponding to LMINλ)
QCALMAX	the maximum quantized calibrated pixel value (corresponding to LMAXλ)
ML	the radiance multiplicative scaling factor for the band (*RADIANCE_MULT_BAND_n* from the metadata)
AL	the radiance additive scaling factor for the band (*RADIANCE_ADD_BAND_n* from the metadata)
τ	the atmospheric transmission
ε	the emissivity of the surface
LT	the surface leaving radiance of a blackbody target with kinetic temperature of T
Lu	the upwelling or atmospheric path radiance
T	the apparent surface temperature in Kelvin
K1, K2	The calibration constants
LTOA	the space-reaching or TOA radiance measured by the instrument

3.3. Land use land cover indices

The most common land use/land cover indices used during this research are NDVI (Normalized Difference Vegetation Index), NDBI (Normalized Difference Built- up Index), and NDWI (Normalized Difference Water Index). These indices can be extracted from Landsat images according to the formulae:

NDVI = (NIR - R)/ (NIR + R) (Rouse, Haas, Schell, & Deering, 1974).

NDBI = (MIR – NIR)/ (MIR + NIR) (Zha et al., 2003)

NDWI = (G - MIR)/(G + MIR) (Xu, 2006)

Where: NIR, R, G, MIR mean Near infrared, Red, Green, and Mid-infrared band respectively.

We applied NDVI and NDBI as LULC indicator to investigate the relationship between LULC and LST in general and within each LULC types. NDWI was used to get a higher accuracy for water bodies. We then integrated the extracted water to the LULC classification to achieve a better result.

3.4. Land use land cover change prediction

To simulate the LULC change in the future, we apply land change prediction using Multi-Layer Perceptron (MLP) neural network (Palit & Popovic, 2006) and Markov chain model (Baker, 1989). The process of future LULC prediction included two main steps: (1) modeling the potential for change and (2) change prediction.

First, we used five potential variables including distance to main road, distance to existing urban center, population, distance to water bodies, and elevation to model the LULC transition in the periods of 2003-2007-2015. The transition map was inferred by applying MLP neural network - the most widespread network structure for efficient time series forecasting (Palit & Popovic, 2006). MLP neural network training is based on the Backpropagation (BP) training strategy. From a set of LULC output values, the network learns from potential inputs to create a potential map for each transition. For instance, the areas closer to main roads, urban center, water bodies, and with a higher elevation may be more suitable for urban growth than others.

In the second process, we applied Markov Chain analysis to model the future LULC by specifying the prediction date. Markov Chain will determine the amount of change using two land cover maps from two different times. The procedure determines transition probability given by a matrix that records the probability each land cover category has to change to every other category. In the future LULC change simulation, we considered three different scenarios: slow, medium, and fast urban growth development. We set the constrain areas to avoid some unreasonable transformations, for example from water or sandbars to urban. We did not use low land/wetland and vegetation area to get better prediction result. In addition, we associate the prediction input with infrastructure planning of the city such as foreseen transportation projects or new residential area planning.

3.5. Regression analysis

Single and Multiple linear regression analysis was applied to determine the correlation between land surface temperature and land use land cover. We extracted NDVI, NDBI, and LST values from each pixel in the study area to point data type. All of these points were used in the linear regression model. This model gives a general idea about the correlation/relationship between LST and LULC. We can easily get an overall picture of the study area. However, in the case of LST prediction, we considered many LULC variables that compose a very complicate structure. This may create non-linear correlations in some areas. For instance, the general trend in the correlation between LST and urban area is mainly linear, and therefore the urban rate growth would lead to an LST increase. Nevertheless, at some small urban rate threshold, we may observe an opposite trend. This is because of the effect from other LULC types such as water or vegetation. We need a better and more flexible regression method, which can perform the local variation of LULC-LST correlation. Hence, we applied a non-linear regression method called Kernel Ridge Regression algorithm (Saunders, Gammerman & Vovk, 1998) to predict the future surface temperature.

Ridge regression is a generalisation of least squares regression. In particular, let us assume the aim is to fit the following linear function to our data: $y = \omega \cdot x$ Linear regression aims at assessing $\omega = \omega_0$ such that the following function is minimised:

$$L_T(\omega) = \sum_{i=1}^{T}(y_i - \omega \cdot x_i)^2 \quad (1)$$

where $\{(x_1, y_1), \dots, (x_T, y_T)\}$ is the training set group of points. Ridge regression slightly modifies this equation to:

$$L_T(\omega) = a\|\omega\|^2 + \sum_{i=1}^{T}(y_i - \omega \cdot x_i)^2 \quad (2)$$

being a is a fixed positive constant.

There are different methods to obtain the $\{(a, \omega)\}$ parameters. One of them consists of applying constrained minimisation methods on the so-called dual version of Eq. (2) (read the paper from Saunders, Gammerman & Vovk (Saunders et al.,, 1998) for further details). Under this framework, it can be shown that the estimation of the $\{(a, \omega)\}$ parameters depends on the dot products of the x elements, i. e., $x_i \cdot x_j$.

Kernel Ridge Regression (KRR) is a modification of Eq. (2) in such a way that non

linear functions can be implicitly fitted. It can be shown (Saunders et al.,, 1998) that in this case the aim is related to the estimation of a function (mapping) ϕ that "transforms" the training points to higher dimensional spaces $\left(x_i \rightarrow \phi(x_i)\right)$ where the problem may be tackled as a linearlisation of the non-linear lower dimensional space where the x_i points lie. It can also be shown that the dot products of the x elements, i. e., $x_i \cdot x_j$ are transformed into $\phi(x_i) \cdot \phi\left(x_j\right)$ (and this group of products forms the so called transformation kernel).

The LST prediction process consisted in the following steps. First, we calculated percentage cover of five classified land use types (urban, vegetation, crops land, water, and vacant land) using three different window sizes (5x5, 10x10, and 20x20 which means 150, 300, and 600 meters resolution, respectively). Then we extracted the mean surface temperature value in the same resolution as we did for LULC. After that, we performed the regression analysis between LST and all five LULC variables at different spatial resolutions for a group of data points called the training set. The second part of the process uses the training set to predict the LST. As the data we had, we used LST in 2003 and 2007 data, and LULC in 2015 in 5x5, 10x10, and 20x20 window sizes to predict LST in 2015. At the third step, we compared the result to choose which data is best LST input. "The best" in this case was evaluated based on comparison of similarity between predicted value and test value. After deciding the best-input time and resolution, we perform the future LST prediction, based on the simulation of future LULC scenarios.

3.6. High/Low Clustering Getis ord statistics

We used the Hot Spot Analysis tool in ArcGIS which calculates the Getis-Ord Gi* statistic in ArcGIS to determine how concentrated the high or low values of LST are in the study area. The Getis-Ord Gi* local statistic is given as (ESRI, 2016):

$$\boldsymbol{G_i^*} = \frac{\sum_{j=1}^{n} \omega_{i,j} x_j - \bar{\boldsymbol{X}} \sum_{j=1}^{n} \omega_{i,j}}{\boldsymbol{S}\sqrt{\frac{n\sum_{j=1}^{n} \omega_{i,j}^2 - \left(\sum_{j=1}^{n} \omega_{i,j}\right)^2}{n-1}}} \quad (3)$$

where x_j is the attribute value for feature j, $w_{i,j}$ is the spatial weight between feature i and j, n is the equal to the total number of features and:

$$\overline{\boldsymbol{X}} = \frac{\sum_{j=1}^{n} x_j}{n} \quad (4)$$

$$\boldsymbol{S} = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \overline{\boldsymbol{X}}^2} \quad (5)$$

From the Gi score, we divided the LST pattern into seven categories: very hot spot, hot spot, warm spot, no significant, cool spot, cold spot, and very cold spot. In the final statistics of the hot spot and cold spot, we took the values that are significant at a 95% confidency or higher.

3.7. Urban landscape analysis

In this research, we used an urban fragmentation index proposed by Burchfield, Overman, Puga, & Turner (Burchfield et al., 2006) and Angel, Parent, & Civco (Angel et al., 2012). In terms of urban heat island analysis, understanding urban landscapes structure helps us to have a deeper insight into how urban fragmentation affects heat island in the city. To understand urban landscape fragmentation, we applied the following fragmentation metrics (Angel et al., 2012):

Fringe open space: It consists of all open space pixels within 100 meters of urban or suburban pixels;

Captured open space: It consists of all open space clusters that are fully surrounded by built-up and fringe open space pixels and are less than 200 hectares in area.

Exterior open space: It consists of all fringe open space pixels which are less than 100 meters from the open countryside.

Urbanized Open Space: It consists of all fringe open spaces (including exterior open spaces) and all captured open spaces.

Rural Open Space: It consists of all open spaces that area not urbanized open spaces.

Infill is defined as all new development that occurred between two time periods within all the open spaces in the city footprint of the earlier period excluding exterior open space.

Extension is all new development that occurred between two time periods in contiguous clusters that contained exterior open space in the earlier period and that were not infill

Leapfrog development is all new construction made between two prescribed time periods and was built entirely outside the exterior open space of the earlier period.

The intermediary metrics used to assess these fragmentation metrics (shown in italics above) are defined as follows:

- Urban built-up pixels are pixels that have a majority of built-up pixels within their walking distance circle;

- Suburban built-up pixels are pixels that have 10–50 per cent of built-up pixels within their walking distance circle;

- Rural built-up pixels are pixels that have less than 10 per cent of built-up pixels within their walking distance circle;

- Fringe open space consists of all open space pixels within 100 meters of urban or suburban pixels;

- Captured open space consists of all open space clusters that are fully surrounded by built-up and fringe open space pixels and are less than 200 hectares in area; exterior open space consists of all fringe open space pixels that are less than 100 meters from the open countryside;

- Urbanized open space consists of all fringe open space, captured open space and exterior open space pixels in the city; and

- The urban landscape area consists of all the built-up area of the city and all its urbanized open space.

4. RESULTS AND DISCUSSION

This section describes the LULC characteristics, its change, and urbanization in the study area from 2003 to 2015. The LST pattern, NDVI, and NDBI distribution is analyzed as well. Based on LULC and LST, we estimated the relationship between them and analyzed the UHI effect. Finally, we explored the future UHI pattern based on the predicted LULC to have a comprehensive understanding about the impact of LULC on UHI from the past to the future.

4.1. Land use land cover change and urbanization in Hanoi inner city

Table 2 provides the producer's accuracy, user's accuracy, overall classification accuracy, and Kappa coefficient that were generated for the LULC classification accuracy assessment in 2003, 2007, and 2015.

Table 2: Accuracy assessment of LULC classification in 2003, 2007, and 2015

	2003		2007		2015	
LULC types	Pro Acc	Usr Acc	Pro Acc	Usr Acc	Pro Acc	Usr Acc
Vegetation	87.5	87.5	76.2	88.9	87.0	90.9
High-density crops land	94.2	98.0	96.1	96.1	90.2	94.9
Low-density crops land	97.9	92.0	97.7	91.3	100.0	91.7
High residential density	95.8	92.0	96.0	96.0	85.1	90.9
Low residential density	97.1	89.5	100.0	96.0	96.7	90.8
Industrial/Commercial	80.0	100.0	81.3	86.7	85.7	92.3
Impervious surface	93.8	83.3	96.6	93.3	90.0	93.1
Water	96.2	100.0	97.6	100.0	96.7	96.7
Sand bars	100.0	100.0	87.5	100.0	100.0	100.0
Vacant land	85.7	90.0	84.0	92.6	93.3	90.0
Wet land agriculture	94.7	90.0	100.0	92.6	94.7	90.0
	Overall Acc = 93.8		*Overall Acc = 94.4*		*Overall Acc = 92.3*	
	Kappa = 0.93		*Kappa = 0.94*		*Kappa = 0.91*	

The overall classification accuracies for 2003, 2007, and 2015 were 93.8%, 94.4%, and 92.3% respectively. Water and sandbars LULC types got the highest accuracy (higher than 97%) in all classified images. In 2003, impervious surface and vegetation gave the lowest user's accuracy (83.3% and 87.5%). In 2007, industrial/commercial land gave the lowest user's accuracy (87.7%). In 2015, all the LULC types presented high accuracy ranging from 90 to 100%. The Kappa coefficient from these classified images was 0.93, 0.94, and 0.91 respectively. Figures 2-4 show the 2003, 2007, and

2015 LULC maps.

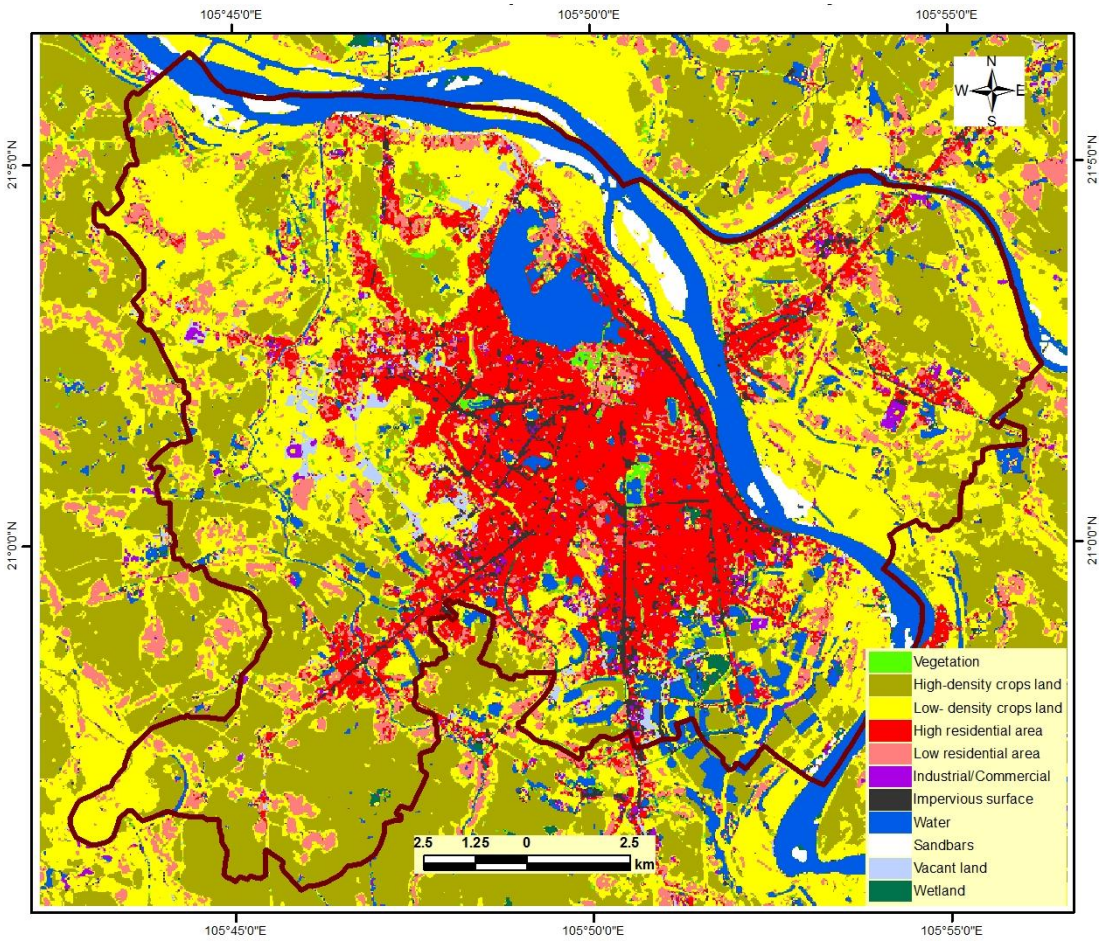


Figure 2: Land use land cover map of Hanoi inner city in 2003

The LULC map in 2003 clearly demonstrated the Hanoi inner city LULC pattern. High residential area was highly agglomerated in the center. Quite independent from that huge cluster was the low residential area and some industrial parks. Agriculture was located around the urban area and along the river. Water areas were highlighted by the Red River and the West Lake. Besides, there was a dense system of lakes and pools located in the southeast part of the city center. In addition, within the urban center some water area also existed that was part of the city parks. The Hanoi inner city land use structure in 2003 showed agriculture land covering 15762.68 hectares and accounting for 54.36 % of the total land area. The second largest land use type (urban) covered 8831.53 hectares and accounted for 30.46 % of the total area. This class included a large proportion of residential land (25.2%) while industrial/commercial land and impervious surface occupied only a small part (1.62 % and 3.63 %). Water occupied 10.37 % and other land cover types including

vegetation, vacant land, and sandbars occupied a small proportion of the area accounting for less than two percent each.

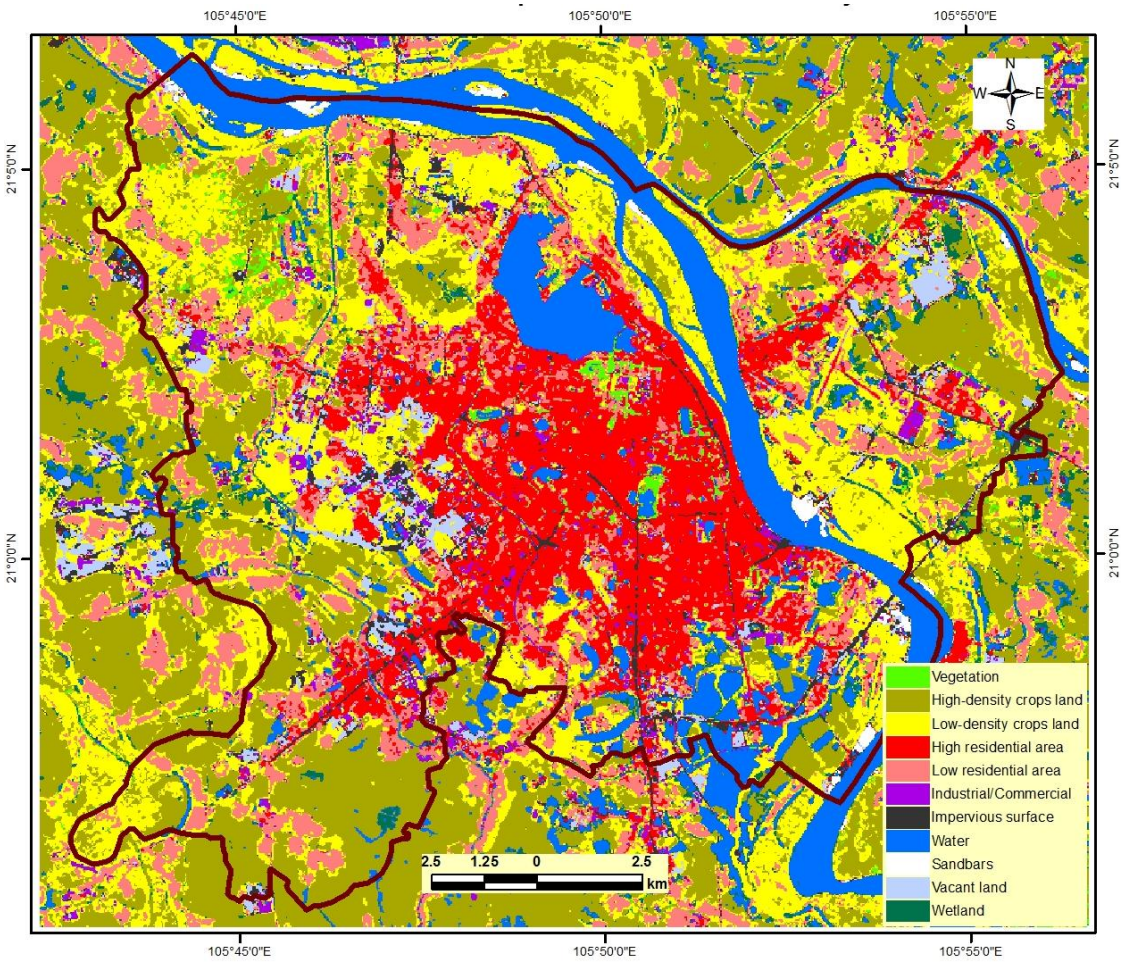


Figure 3: Land use land cover map of Hanoi inner city in 2007

The map in 2007 showed a crucial change in the LULC pattern. From the urban center in 2003, the residential area expanded mainly to the west. Some new industrial/commercial area appeared along the main roads. At that time, the water level was high, and occupied a large part of sandbars. Significantly, the map showed the appearance of some new vacant land areas next to the road and residential areas. These were different from the natural bare soil, which was often located in the field. Due to its official distribution, vacant land in this case is often transformed to urban rather than to other LULC types. In 2007, agriculture land was still the largest LULC type accounting for 12707.65 hectares (43.82 %). Urban land was the second largest LULC type occupying 11158.62 hectares (38.47% in the total land use structure). Water occupied 3555.43 hectares (12.26%), and vacant land extended over 976.8 hectares (3.37%). Vegetation has the same area and proportion as it had in 2003.

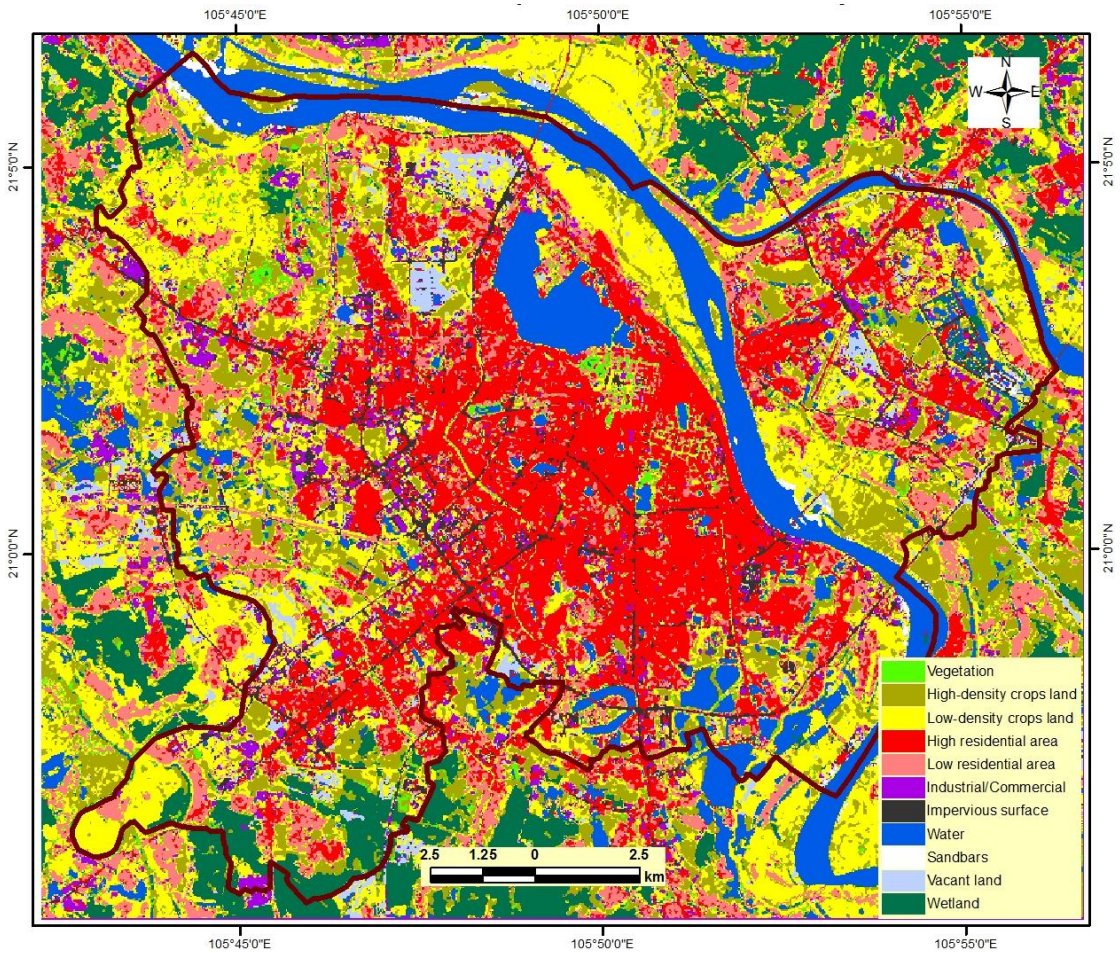


Figure 4: Land use land cover map of Hanoi inner city in 2015

The LULC map in 2015 demonstrated a huge urban land occupation (see the red color distribution in the map, as compared to others). The map also showed a very high LULC fragmentation that can be seen by the existence of many different LULC types in a small area. Besides that, the map illustrated a large wetland region located in the northeast and southwest parts of the study area. This change was not a big issue because it was simply an agricultural practice of farmers. In that case, they pumped water from river/stream to the farm to prepare for new crops. In 2015, urban class became the largest land use type, which covered 13938.87 hectares and occupied 48.07% of the total land use structure. The gain area from this class came from the increase in high residential area, industrial/commercial, and impervious surface. In contrast, agricultural area was smaller than urban area, covering 10241.21 hectares and accounting for 35.31% of the total land area. Water bodies which occupied 10.74%, increasing a small proportion. Vegetation increased about 0.5% to reach 2%. However, the increase in vegetation mostly came from outside the urban center where new urban residents are. Vacant land increased its occupied area to 3.49% and a large

part of this is preparing for new construction area. Figure 5-7 and Table 3 illustrate the change of all LULC types from 2003 to 2015.

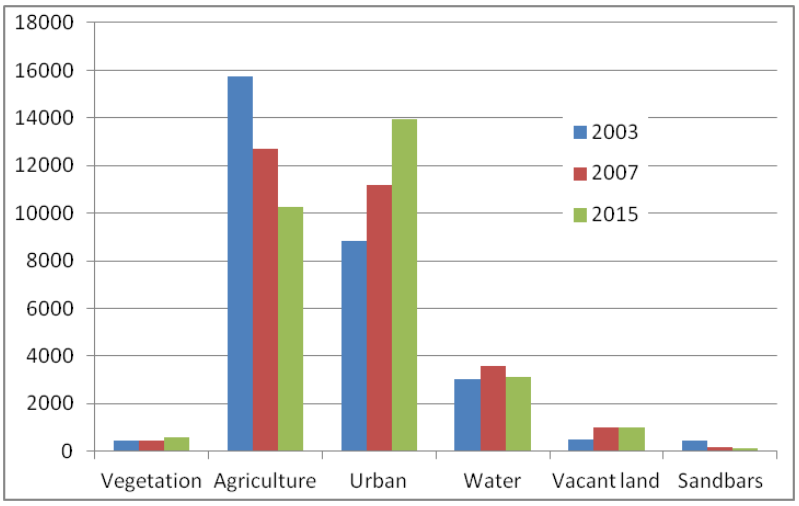


Figure 5: Land use and land cover areas in 2003, 2007, and 2015

The graph in Figure 5 demonstrates that agriculture and urban classes are the main LULC types, vacant land occupies a significant area, water and vegetation account for a small proportion of the total area and sandbars is not significant. Both agriculture and urban class are changing significantly. However, the changes in urban and agriculture classes have an opposite trend. Agriculture tends to decrease whereas urban area shows the increment through time.

Table 3: Land use and land cover change in 2003 - 2015 in Hanoi inner city

LULC types	2003 - 2007		2007 - 2015		2003-2015	
	Hectares	%	Hectares	%	Hectares	%
Vegetation	13.80	0.05	126.53	0.44	140.33	0.49
Agriculture	-3055.03	-10.54	-2466.44	-8.50	-5521.47	-19.04
High-density crops land	-702.34	-2.42	-2001.12	-6.90	-2703.46	-9.32
Low- density crops land	-2666.43	-9.20	-760.26	-2.62	-3426.69	-11.82
Wetland	313.73	1.09	294.94	1.01	608.67	2.10
Urban	2327.09	8.02	2780.26	9.59	5107.35	17.61
High residential area	215.95	0.74	2097.57	7.24	2313.52	7.98
Low residential area	1812.50	6.25	-203.31	-0.70	1609.19	5.55
Industrial/Commercial	301.59	1.04	413.59	1.43	715.18	2.47
Impervious surface	-2.95	-0.01	472.40	1.63	469.45	1.62
Water	548.53	1.89	-440.19	-1.52	108.34	0.37
Vacant land	467.76	1.61	34.78	0.12	502.54	1.73
Sandbars	-293.55	-1.01	-43.54	-0.15	-337.09	-1.16

In the first period (2003-2007) the land use structure is characterized by the decrease of agricultural land, which changed from 15762.68 hectares (54.36 %) to 12707.65 hectares (43.82 %). Urban land increased in 2327.09 hectares (8.02% in the total land use structure), from 8831.53 hectares (30.45) to 11158.62 (38.47%). Compared to 2003, vacant land gained 548.53 hectares (1.89%), and water increased 467.76 hectares (1.61%). Vegetation kept a similar proportion to the one in 2003.

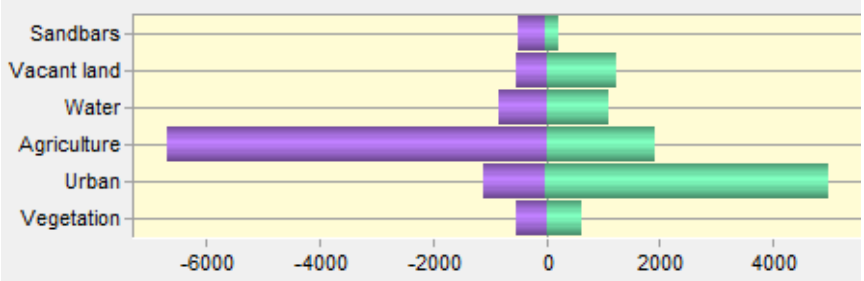


Figure 6: LULC changes between 2003 and 2007

In the second period (2007-2015), agriculture and urban classes kept the same changing trend as compared to the first period. Specifically, the agriculture and urban areas net change were the loss of 2466.44 hectares (8.5%) and the gain of 2780.26 hectares (9.59%) respectively. Other LULC types including water, vacant land and vegetation got some expansion as well.

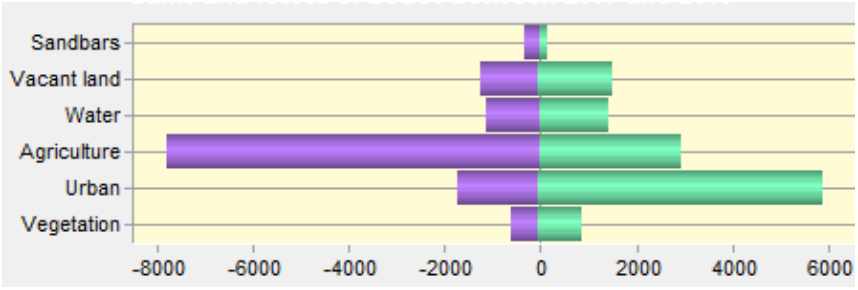


Figure 7: LULC changes between 2007 and 2015

In the LULC change of study area, we need to distinguish between the temporary and permanent transformation in LULC types. The change in many land use land cover types as high and low-density crops land, sandbars, or wetland are of a temporal type. For example, crops land could be transferred to wet land if the farmers pump the water from the river or stream to the farm to prepare for a new crop season. Sandbar also decreases the area once water level in the river grows up by flooding and vice versa. The surface temperature changed due to this is short and unpredicted time change. The permanent change such as from agriculture or vacant land to urban area

in urbanization is more significant. By analyzing such process, we will have a better idea on how urban and urbanization impact the surface temperature pattern.

Table 4 shows the transformation from other LULCs to urban class in Hanoi inner city from 2003 to 2015. In the first period, agriculture change was the biggest contribution to urban change (2639.68 hectares, equal 9.1% of total land). The conversion from vacant land, water, and vegetation were 340.36 hectares (1.17 %), 135.03 hectares (0.47%), and 68.14 hectares (0.23%) respectively. In the second period, a larger area of other LULCs changed to urban land. Urban received 2826.19 hectares (9.75%) from agriculture, 691.6 hectares (2.38%) from vacant land, 116.14 hectares (0.4%) from vegetation, and 115.21 hectares (0.4%) from water. Overall, the expansion of urban class in the 2003-2015 period came from taking 17.45% from agriculture, 1.37% from vacant land, 0.72% from water, and 0.47% from vegetation.

Table 4: Area and percent change from other LULC types to urban in 2003-2015

	Urban 2007			Urban 2015			Urban 2015	
LU 2003	Area	%	LU 2007	Area	%	LU 2003	Area	%
Vegetation	68.14	0.23	Vegetation	116.14	0.40	Vegetation	136.87	0.47
Agriculture	2639.68	9.10	Agriculture	2826.19	9.75	Agriculture	5060.85	17.45
Water	135.03	0.47	Water	115.21	0.40	Water	211.38	0.73
Vacant land	340.36	1.17	Vacant land	691.60	2.38	Vacant land	396.32	1.37

From these analyses, we can also conclude that the urbanization in the first period (2003-2007) is faster than in the second period (2007-2015) as it took only four years to gain more or less the same area to that gained in the 2007-2015 period. This could be explained by the impact of economic crisis in the real estate market in 2010.

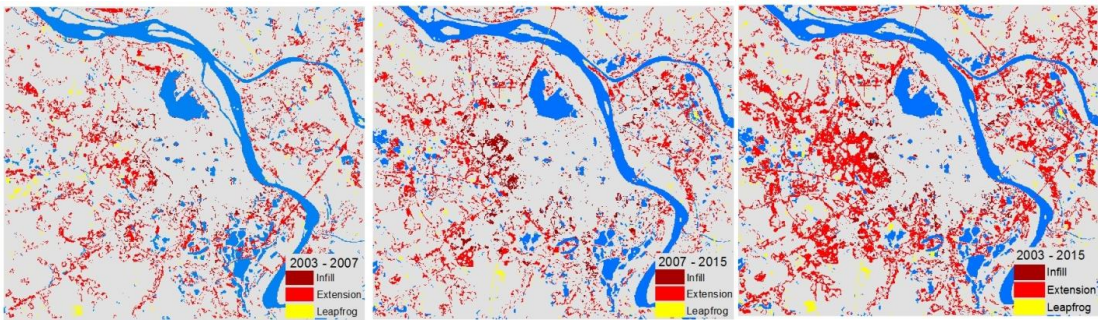


Figure 8: Urban development types from 2003 to 2015

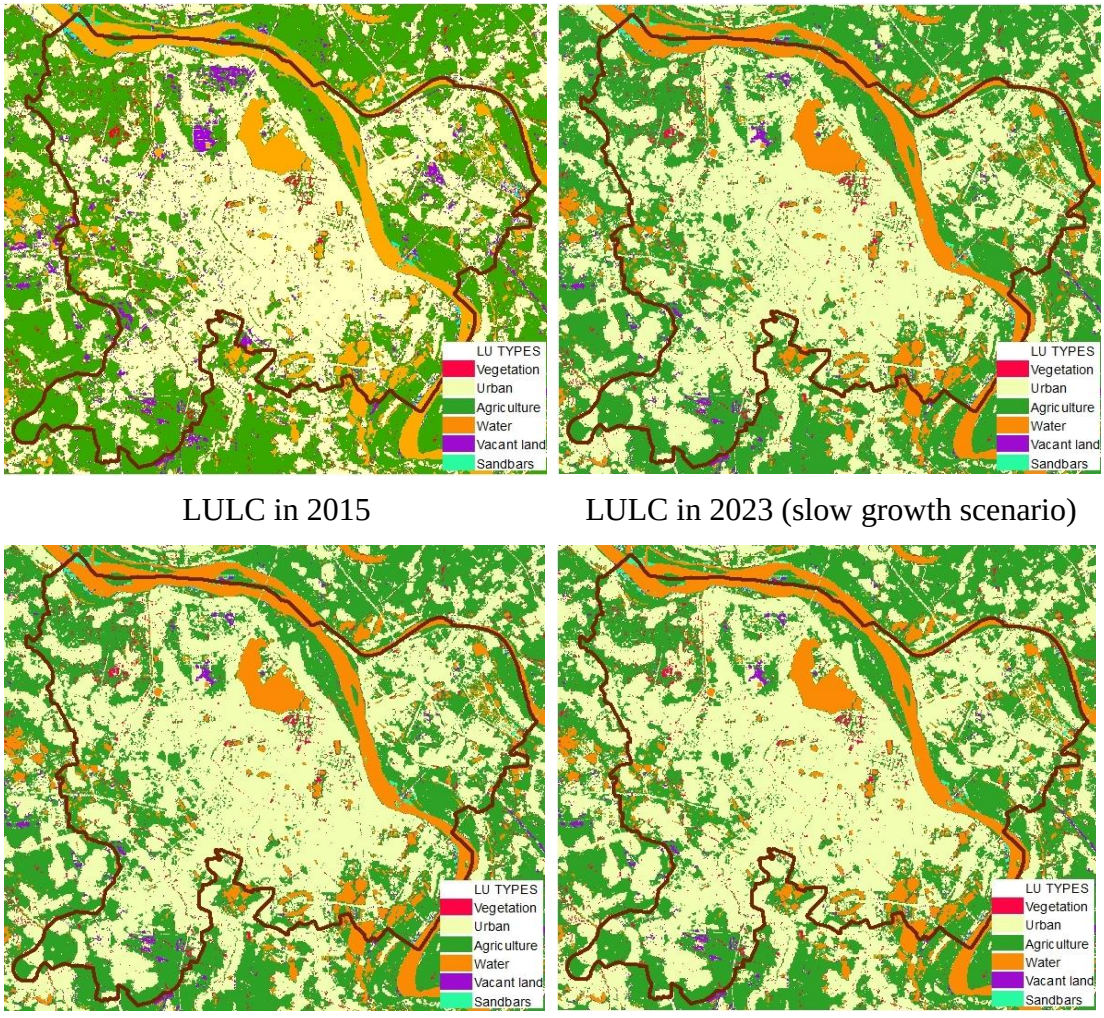
Figure 8 demonstrates urban fragmentation in terms of development type in the Hanoi inner city. Urban growth divided into three types of development: infill, expansion, and leapfrog. The general trend in urban development includes extending

from core urban and the development of new urban area. In both of these developments, the road systems play a key role as main urban expansion driving factor.

4.2. Future LULC simulation

Using the Land change modeler software provided by Clark labs (https://clarklabs.org/), we predict LULC change in the study area based on three scenarios: slow, medium, and fast growth scenarios (Figure 9). From the map in 2023, we can clearly see the expansion of the new urban area. Urban grew up significantly in the southeast part of the city due to the concentration of more advantage conditions such as land resources and transportation.

Figure 9: Current and predicted land use land cover in and around the study area



LULC in 2015 LULC in 2023 (slow growth scenario)

LULC in 2023 (medium growth scenario) LULC in 2023 (fast growth scenario)

Tables 5 and 6 demonstrate the LULC change in the study area in the 2015-2023 period under three different scenarios.

Table 5: Area and proportion of LULC types in 2015 and 2023 (whole study area)

	2015		2023 (slow growth)		2023 (medium growth)		2023 (fast growth)	
LULC types	Hectares	%	Hectares	%	Hectares	%	Hectares	%
Vegetation	1024.65	2.01	1024.65	2.01	1024.65	2.01	1024.65	2.01
Urban	19276.47	37.77	22226.40	43.56	23253.39	45.57	24210.18	47.44
Agriculture	23409.27	45.87	21510.72	42.15	20504.52	40.18	19560.69	38.33
Water	5500.89	10.78	5500.89	10.78	5500.89	10.78	5500.89	10.78
Vacant land	1640.34	3.21	588.96	1.15	568.17	1.11	555.21	1.09
Sandbars	178.38	0.35	178.38	0.35	178.38	0.35	178.38	0.35

Table 6: Area and proportion of LULC types in 2015 and 2023 (inner city)

	2015		2023 (slow growth)		2023 (medium growth)		2023 (fast growth)	
LULC types	Hectares	%	Hectares	%	Hectares	%	Hectares	%
Vegetation	579.24	2.00	579.24	2.00	579.24	2.00	579.24	2.00
Urban	13935.78	48.06	15597.45	53.79	16138.35	55.66	16651.08	57.43
Agriculture	10239.21	35.31	9304.29	32.09	8774.55	30.26	8268.93	28.52
Water	3116.61	10.75	3116.61	10.75	3116.61	10.75	3116.61	10.75
Vacant land	1011.42	3.49	284.67	0.98	273.51	0.94	266.40	0.92
Sandbars	112.32	0.39	112.32	0.39	112.32	0.39	112.32	0.39

In 2023, urban land will reach 15597.45 hectares at the slow growth scenario, 16138.35 hectares at the medium growth scenario, and 16651.08 hectares at the fast growth scenario. Urban area will occupy 53.79 %, 55.66 %, and 57.43 % respectively following these scenarios. Consequently, agricultural land, the LULC type that has highest potential to be transformed into urban land, will decrease. Agricultural land will cover 9304.29 hectares, 8774.55 hectares, and 8268.93 hectares in the slow, medium, and fast growth scenario.

4.3. Land surface temperature, NDVI, and NDBI patterns

4.3.1. Land surface temperature

The maps in Figure 10 represent the LST pattern of the study area in 2003, 2007, and 2015. LST ranged from 21.7 ^{0}C to 29.2 ^{0}C across the city in 2003, 18.9^{0}C to 23.9 ^{0}C in 2007, and 25.5 ^{0}C to 34.2 ^{0}C in 2015. Mean LST in the region was 23.7, 20.8, and 28.4 ^{0}C respectively. From the maps, we can clearly see the differences in the LST patterns. The high LST area (heavy yellow and red color) appears as a big island surrounded by low LST region (represented by blue color tone). This core

exists in the center where of highly urbanized zone is. Between these two LST pattern is a small transition LST area marked by light yellow color exists. In addition, there are some small cold islands inside the hot area.

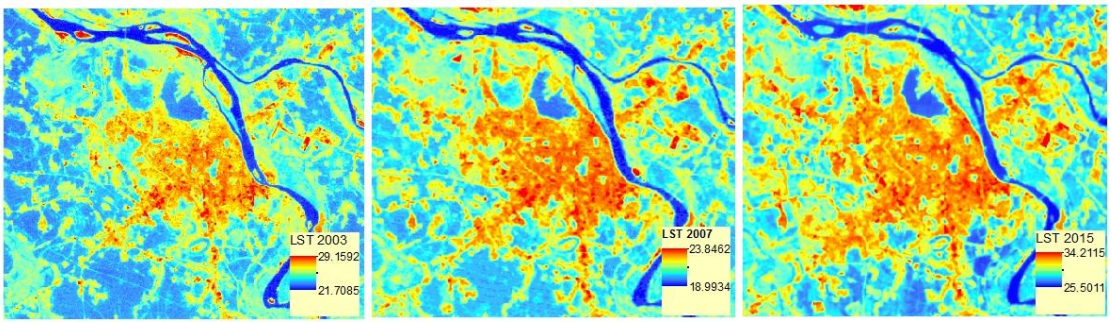


Figure 10: Mean surface temperature of different LULC types

Mean LST within each LULC type will give a clearer understanding about LST change through the study area (Figure 11). In all three dates, water got the lowest LST mean value (22.6^0C in 2003, 19.8^0C in 2007, and 26.5^0C in 2015). Urban land was the hottest area where all of its LULC types had very high mean LST compared to others. Within urban class, high residential area gave the highest mean LST value (24.6^0C, 21.6^0C, and 29.5^0C), even higher than impervious surface (24.4^0C, 21.2^0C, and 29.2^0C) and industrial/commercial area (24.4^0C, 21.4^0C, and 29.1^0C). With the concentration of more vegetation and open area, low residential area was the coldest urban LULC type, which received the mean LST value of 24.1^0C, 21.2^0C, and 28.9^0C. The mean LST of vegetation (23.3^0C, 20.7^0C, 27.9^0C) and high-density crops land (22.8^0C, 20.1^0C, and 27.4^0C) were quite low, only higher than water. The mean LST in low-density crops land was quite low (23.6^0C, 20.6^0C, and 28^0C) compared to the urban type. Vacant land got quite similar value in relation to low residential area, which was 24.3^0C, 21.3^0C, and 28.6^0C. Sandbars had very high mean LST value (24.6^0C, 21.3^0C, and 28.5^0C) compared to other non-artificial surfaces (vegetation, agriculture land, and water). The mean LST from these three dates reveal that the parts with high LST (the temperature is higher than the average LST of the whole area) correspond to urban area, sandbars, and vacant land. The colder part includes water bodies, vegetation, and high-density crops land recording lower LST value than the average value.

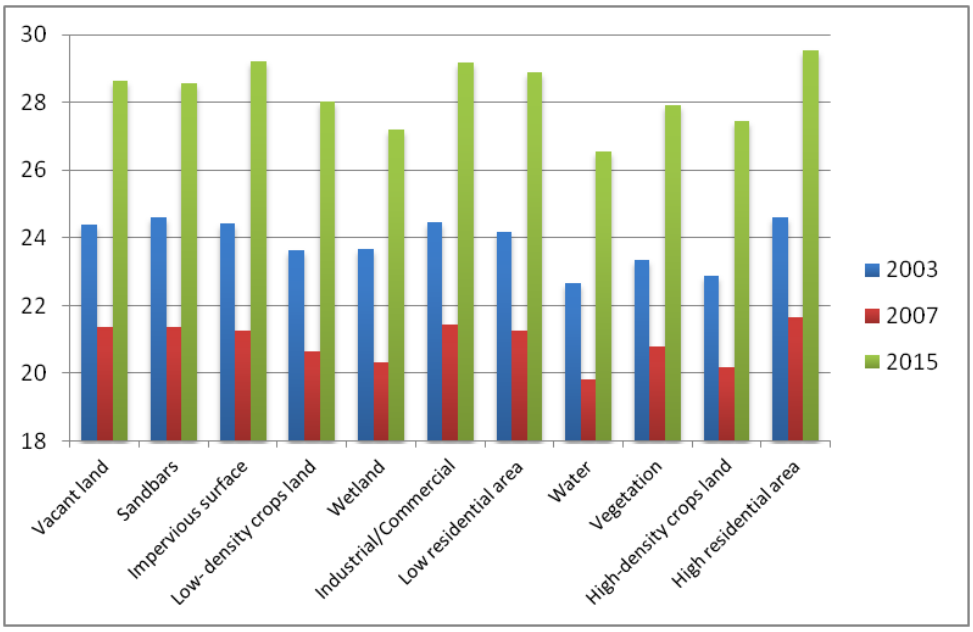


Figure 11: Mean LST by different LULCs

In order to have a better understanding about the difference in LST value between each LULC type, we compared their mean LST with water mean LST (Table 7).

Table 7: The difference of mean LST between other LULC types and water

	2003	2007	2015
Vacant land-water	1.73	1.55	2.08
Sandbars-water	1.95	1.55	2.03
Impervious surface-water	1.78	1.42	2.66
Low- density crops land-water	0.97	0.81	1.49
Wetland-water	1.02	0.49	0.67
Industrial/Commercial-water	1.80	1.60	2.62
Low residential area-water	1.53	1.44	2.36
Vegetation-water	0.71	0.96	1.37
High-density crops land-water	0.23	0.34	0.90
High residential area-water	1.97	1.85	2.99

Table 7 reveals that all of the urban areas represent higher difference LST, ranging from 1.5-3 Celsius degrees. The difference between high-density crops land and vegetation (except vegetation in 2015) is less than 1 degree.

4.3.2. Normalized difference vegetation index (NDVI)

The NDVI spatial distribution is shown in Figure 12. The areas with the lowest vegetation levels appeared over the Red river and the West Lake where the main water bodies of the city were. Low NDVI values concentrated in the center

corresponding to the urban area. High NDVI values were primarily distributed around the urban area, which was occupied by agriculture land. Some small islands of high NDVI values located inside the urban area corresponded to green urban areas including parks and open spaces.

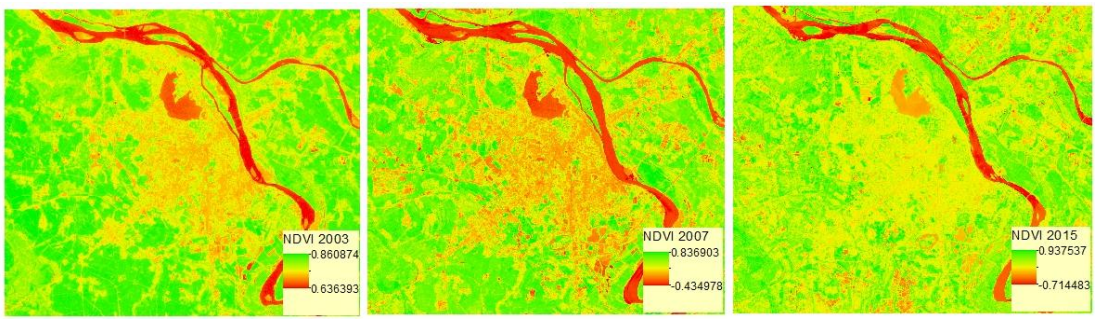


Figure 12: Normalized difference vegetation index (NDVI) in 2003, 2007, and 2015

Bars plot in Figure 13 demonstrates the mean NDVI within each LULC type. The LULC types with highest mean NDVI value were high-density crops land (0.64-0.73), vegetation (0.57 - 0.7), and low-density crops land (0.53 - 0.58) due to the dominance of vegetated cover.

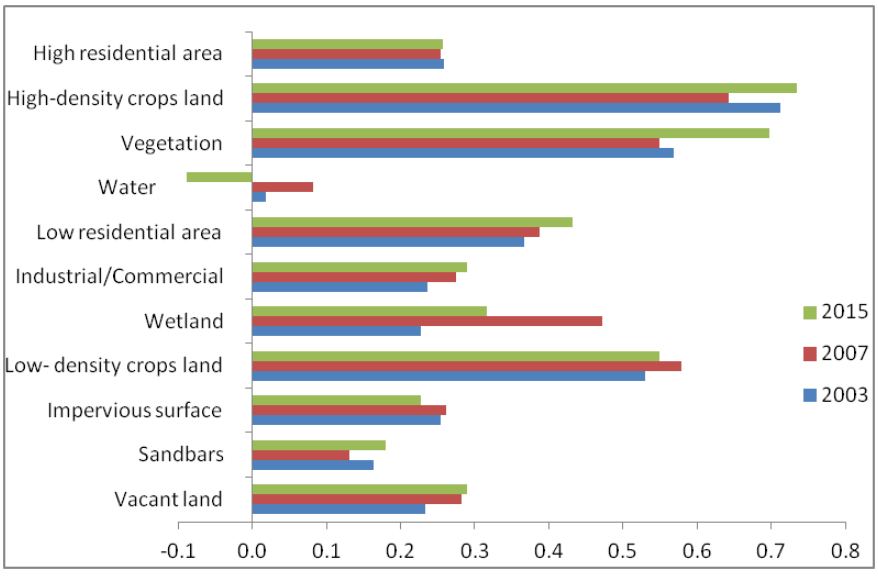


Figure 13: Mean NDVI value within each LULC types in 2003, 2007, and 2015

With the lack of vegetation, water showed the lowest NDVI value (around -0.1 - 0.01). Sandbars had low NDVI value, which was below 0.2. High residential area, vacant land, and industrial/commercial area had quite similar mean NDVI value (0.23 - 0.29). By having more vegetation and open land with grass coverage, low residential area had higher NDVI value than other urban LULC types. Except for the case of

water, other LULCs demonstrated that lower NDVI would have higher LST. Water had the lowest NDVI value but also got lowest LST due to its special characteristic. From this problem, we took the water out in the regression analysis between NDVI and LST to have a better result.

4.3.3. Normalized difference built-up index (NDBI)

The NDBI maps represent an opposite pattern as compared to the NDVI maps (Figure 14). Urban center, which was a low-NDVI area, got high NDBI value. Around the urban core (especially the southwest area) showed low NDBI value. Water bodies represent the lowest NDBI.

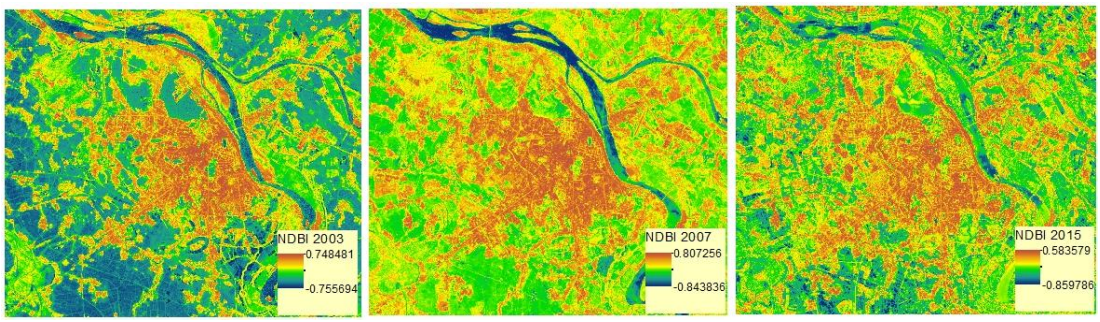


Figure 14: NDBI within each LULC type in 2003, 2007, and 2015

A graph of mean NDBI within each LULC type for 2003, 2007, and 2015 is shown in figure 15. The mean NDBI value was very small in all the LULC types.

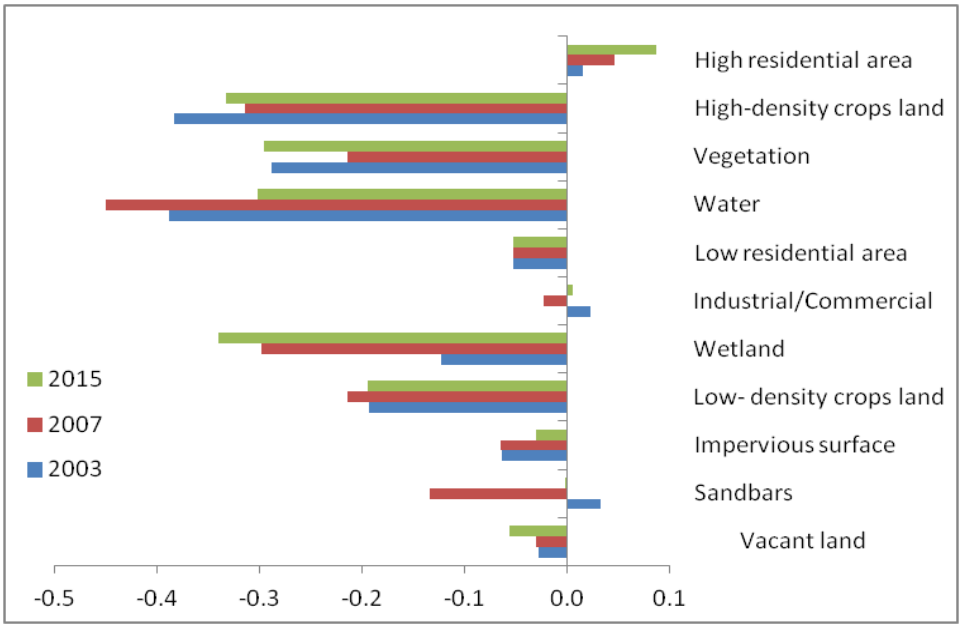


Figure 15: Mean NDBI value within each LULC type in 2003, 2007, and 2015

Water had the lowest mean NDBI (from -0.4 to -0.3) value and the same happened to NDVI. The trend in NDBI value with other LULC types was the opposite to that of NDVI as lower NDBI values would represent more vegetated LULC types. After water, high-density crops land and vegetation demonstrated very low NDBI values (with a mean value ranging from -0.3 to -0.2). Compared to these densely vegetated LULCs, low-density crops land, which had less vegetation concentration, got higher NDBI value (between −0.2 and -0.1). In urban group, high residential area had highest NDBI (between 0.01 - 0.08). Mean NDBI value of low residential was -0.05, industrial/commercial varied from -0.02 to 0.02, and impervious surface differed from -0.06 to -0.03.

4.4. Relationship between LST and LULC

4.4.1. Correlation between LST, and NDVI, NDBI

First, we investigate the correlation between LST, NDVI, and NDBI for all LULC types in the study area. To assess this relationship, we applied a multiple regression analysis between LST, NDVI and NDBI for all the pixels in study area (water area was excluded due to the problem mentioned above). The multiple regression model developed in the study is defined as:

LST = 24.67 + 2.93NDBI - 0.93NDVI (2003)

LST = 21.48 + 3.30NDBI - 0.36NDVI (2007)

LST = 29.05 + 4.81NDBI - 0.33NDVI (2015)

where:

LST is the land surface temperature (Celsius degree);
NDBI is Normalized Difference Built-up Index
NDVI is Normalized Difference Vegetation Index

The coefficients, standard error, t statistic, and p-values are shown in Table 8. The results indicate a significant negative correlation between LST and NDVI and positive correlation between LST and NDBI. The negative NDVI coefficient and positive NDBI coefficient from the three dates demonstrate that within the study area, vegetation contributes to a decrease in the LST while built-up area is related to the increase in LST. In other words, LULCs with higher NDVI will generally get lower LST value whereas LULCs with higher NDBI will obtain higher LST. The regression

model from all three dates in 2003, 2007, and 2015 shows a strong linear relationship (R^2 reached values from 0.7 to 0.8). In 2003, the increase in built-up index raised land surface temperature by 2.93°C. In an inverse direction, an increase in the vegetation index contributed to lower surface temperature by 0.93°C. In 2007 and 2015, the contribution of NDBI to surface temperature was 3.3°C and 4.81°C respectively whereas NDVI lowered the surface temperature by 0.36°C and 0.33°C respectively.

Table 8: Regression analysis parameters in 2003, 2007, and 2015

May 5, 2003					
	Coefficient	Std.Error	t-Statistic	Probability	R^2
Constant	24.67	0.002	12180.35	0.000	0.8
NDVI	-0.93	0.006	- 143.14	0.000	
NDBI	2.93	0.007	379.58	0.000	
May 24, 2007					
	Coefficient	Std.Error	t-Statistic	Probability	R^2
Constant	21.48	0.002	9107.08	0.000	0.8
NDVI	-0.36	0.007	-50.62	0.000	
NDBI	3.3	0.008	398.76	0.000	
July 1, 2015					
	Coefficient	Std.Error	t-Statistic	Probability	R^2
Constant	29.07	0.003	9608.95	0.000	0.7
NDVI	-0.39	0.008	-40.32	0.000	
NDBI	4.15	0.009	512.73	0.000	

The results also imply that built-up might be a more effective indicator than the vegetation index. In addition, NDBI impact tends to be stronger. As we saw in the 2003-2015 time interval, the NDBI coefficient raised from 2.93 (2003) to 3.3 (2007), and to 4.81 (2015) whereas NDVI coefficient decreased from 0.93 (2003) to 0.36 (2007) and 0.33 (2015). This can be explained by analyzing the trend in the LULC change in such period. Urban area, the high NDBI and low NDVI LULC type has increased and occupied part of other low NDBI and high NDVI LULC types such as crops land, water, and vegetation.

To have a deeper understanding about the detailed relationship between LST and LULC, we built the regression model for each LULC type (Tables 9 - 11). The correlation between NDVI, NDBI and LST within each LULC types showed some important results. First, we saw a general trend in urban LULC in the sense that all urban LULC types had negative correlation with NDVI and positive correlation with

NDBI. This brings us an important conclusion about the role vegetation has in the UHI effect mitigation.

Table 9: Correlation between LST, and NDVI, NDBI within each LULCs in 2003

LULC types	Regression function	R^2	Std.Error	Probability
Vegetation	2.38NDBI - 1.01NDVI + 24.58	0.33	0.022	0.000
High-density crops land	2.43NDBI - 0.78NDVI + 24.36	0.50	0.004	0.000
Low- density crops land	2.47NDBI - 0.45NDVI + 24.31	0.48	0.003	0.000
Wetland	2.32NDBI - 0.21NDVI + 23.99	0.28	0.019	0.000
High residential area	2.26NDBI - 1.64NDVI + 24.99	0.38	0.011	0.000
Low residential area	2.29NDBI - 1.10NDVI + 24.67	0.35	0.010	0.000
Industrial/Commercial	2.38NDBI - 1.33NDVI + 24.66	0.46	0.020	0.000
Impervious surface	4.01NDBI - 0.97NDVI + 24.93	0.59	0.016	0.000
Water	1.14NDBI + 1.88NDVI + 23.04	0.67	0.005	0.000
Vacant land	1.66NDBI - 1.60NDVI + 24.78	0.34	0.018	0.000

Table 10: Correlation between LST, and NDVI, NDBI within each LULCs in 2007

LULC types	Regression function	R^2	Std.Error	Probability
Vegetation	2.72NDBI - 0.92NDVI + 21.75	0.38	0.021	0.000
High-density crops land	2.88NDBI + 0.87NDVI + 20.48	0.29	0.005	0.000
Low- density crops land	1.96NDBI - 0.42NDVI + 21.26	0.27	0.006	0.000
Wetland	1.33NDBI + 0.17NDVI + 20.58	0.14	0.010	0.000
High residential area	1.74NDBI - 0.64NDVI + 21.73	0.17	0.008	0.000
Low residential area	1.74NDBI - 0.36NDVI + 21.44	0.18	0.009	0.000
Industrial/Commercial	1.40NDBI - 1.27NDVI + 21.71	0.31	0.012	0.000
Impervious surface	2.04NDBI - 1.40NDVI + 21.68	0.28	0.013	0.000
Water	1.68NDBI + 1.12NDVI + 20.43	0.61	0.004	0.000
Vacant land	3.04NDBI - 0.17NDVI + 21.32	0.14	0.013	0.000

Table 11: Correlation between LST, and NDVI, NDBI within each LULCs in 2015

LULC types	Regression function	R^2	Std.Error	Probability
Vegetation	5.63NDBI + 1.04NDVI + 28.70	0.32	0.049	0.000
High-density crops land	3.16NDBI + 1.52NDVI + 27.32	0.09	0.014	0.000
Low- density crops land	2.81NDBI - 0.05NDVI + 28.43	0.11	0.009	0.000
Wetland	0.77NDBI + 0.40NDVI + 27.27	0.12	0.006	0.000
High residential area	3.27NDBI - 0.77NDVI + 29.34	0.17	0.008	0.000
Low residential area	2.84NDBI - 0.02NDVI + 28.93	0.16	0.015	0.000
Industrial/Commercial	3.16NDBI - 0.88NDVI + 29.28	0.30	0.011	0.000
Impervious surface	4.13NDBI - 0.29NDVI + 29.32	0.18	0.015	0.000
Water	0.40NDBI + 1.56NDVI + 26.76	0.44	0.005	0.000
Vacant land	3.40NDBI - 0.27NDVI + 28.74	0.18	0.014	0.000

On the other hand, water had positive effect in relation to both NDBI and NDVI. That means water with more vegetation and built-up materials (sand, gravel, soil) will have higher LST. However, the results area a little bit surprising in relation to the different trend observed considering the same LULC type. For example, NDVI affected negatively to LST in vegetation area in 2003 and 2007 but its effect was the opposite in 2015. Wetland also had the same problem, i.e., the correlation was different between 2003 versus 2007 and 2015.

A possible answer for this problem might be that vegetation in the urban area had higher mean LST than another one that was located outside the urban area. This was strongly reflected in 2015 when urban class significantly expanded (Figure 16). In this case, vegetation mean LST value inside the urban area reached 28.4^{0}C whereas outside it was 27.4^{0}C. From this, we may conclude that vegetation in urban area with high NDVI and high LST value generated the counterintuitive inverse trend in the model. We also observed that the correlation between each LULC type and LST in 2007 and 2015 is more similar than between 2003 and 2015. For instance, NDVI and LST in wetland had a negative relationship with LST in 2003 whereas this LULC types in 2007 and 2015 got a positive correlation. High-density crops fields got positive correlation between NDVI and LST in 2007 and 2015 but got negative correlation in 2003.

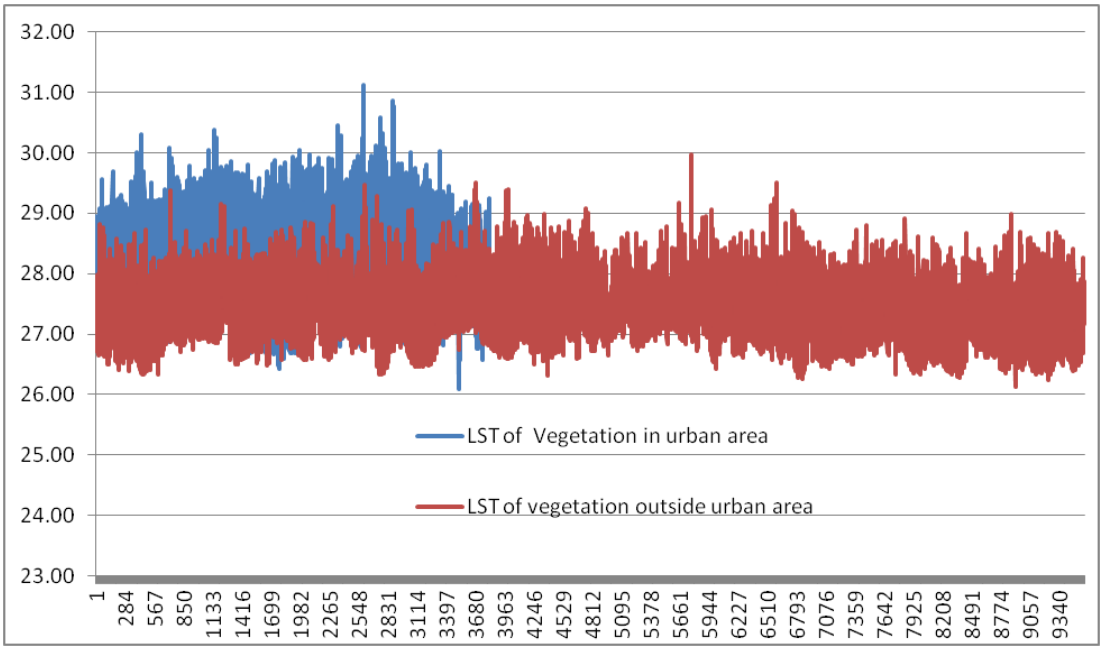


Figure 16: LST of vegetation within and outside urban area in 2015 (Celsius degrees)

We can therefore conclude that the relationship between LULC and LST is varied between each LULC type and changes taking into account the geographic location/pattern. To have a better insight into the LULC-LST correlation, we need to explain these local differences. We suggest that the linear regression model can be useful in discovering the general trend as well as providing the big picture of the relationship between LST and LULC. However, a linear relationship is not always the best choice when considering a more local level analysis. On the other hand, we might also bear in mind the NDVI and NDBI behaviours in relation to the temperature. In order to better solve those problems, we decided to apply a non-linear regression model (Kernel Ridge Regression).

4.4.2. The impacts of LULC change and urbanization in urban warming effect

To gain insight into the impact of LULC change on the urban heat island effect, we determined how the urban heat island pattern changed with the LULC change. First, we applied Getis-Ord Gi* statistics to determine where LST with either high or low values clustered spatially. Later, we linked the change in hotspot pattern to land use land cover change through time. This method gives a better demonstration of the urban heat island effect, rather than focusing only on the high or low LST absolute value separately. By applying Getis-Ord Gi* statistics, we created hotspot maps of Hanoi inner city in three different dates (Figure 17). This may give us a better understanding about the city urban heat island effect. The identification of hot spot or cold spot by such method does not depend on whether the mean surface temperature is high or low. This implies that the effect of different LST values through time is reduced, so we can compare the results in a more confident way.

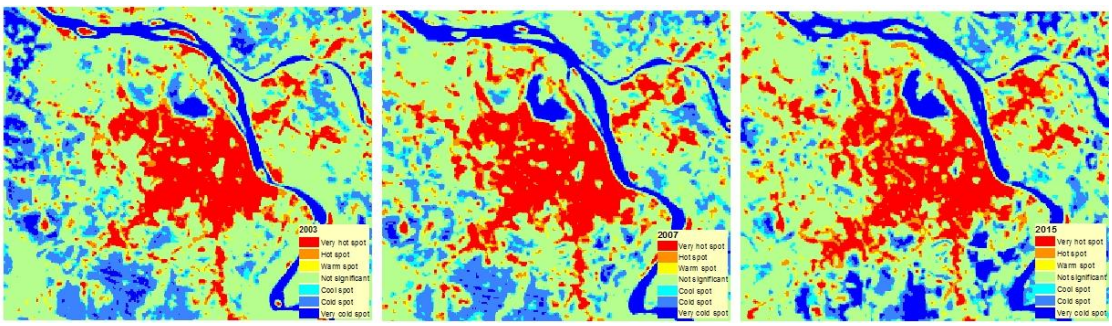


Figure 17: Hotspot analysis in 2003, 2007, and 2015

The maps show that hot regions (all the hot spots whose significance is higher than 95%) are highly clustered in the urban center, along main roads and industrial

zones. The cold regions correspond to vegetation high concentration areas, rivers and lakes. On the other hand, hot spot regions tend to expand through time. It is also recognized that when LST increases, some "not significant" areas tend to become a cold spot. This is represented by the transformation to cold spot of many lakes and parks inside the urban center in 2015.

Tables 12 - 14 give a detailed view of the contribution of LULCs to LST clustering in Hanoi inner city. In general, hot spot occupies larger area than cold spot does. More than 20% of the city is always warmer whereas less than 10% of the city is always colder than the mean zonal LST. The hottest land cover type is urban while the coldest one is water. Water bodies consist of the largest proportion of cold spots, which contribute more than 50% to the total cold spot area. Agriculture plays a significant role in supplying cold spots besides water bodies. Vegetation has a small contribution to the number of cold spots.

Table 12: Proportion of hot and cold area in different LULCs in 2003 (%)

	Thermal pattern						
LULC types	Cold spot (99% sig)	Cold spot (95% sig)	Cold spot (90% sig)	Not sig	Hot spot (90% sig)	Hot spot (95% sig)	Hot spot (99% sig)
Vegetation	0.001	0.030	0.055	1.244	0.037	0.077	0.071
Urban	0.003	0.022	0.022	6.578	1.870	3.939	18.021
Agriculture	0.544	7.709	3.802	37.200	1.457	1.897	1.748
Water	4.276	1.118	0.587	4.242	0.061	0.049	0.037
Vacant land	0.000	0.002	0.002	0.451	0.168	0.397	0.736
Sandbars	0.001	0.001	0.007	0.434	0.079	0.214	0.811
Total	4.826	8.882	4.474	50.150	3.673	6.573	21.423

Table 13: Proportion of hot and cold area in different LULCs in 2007 (%)

	Thermal pattern						
LULC types	Cold spot (99% sig)	Cold spot (95% sig)	Cold spot (90% sig)	Not sig	Hot spot (90% sig)	Hot spot (95% sig)	Hot spot (99% sig)
Vegetation	0.00	0.02	0.02	1.06	0.07	0.13	0.25
Urban	0.00	0.01	0.02	8.69	2.54	5.53	21.68
Agriculture	0.51	6.25	3.37	30.84	0.94	1.11	0.76
Water	6.12	1.30	0.75	3.78	0.08	0.10	0.14
Vacant land	0.00	0.00	0.00	1.05	0.32	0.70	1.29
Sandbars	0.00	0.01	0.01	0.24	0.04	0.08	0.17
Total	6.64	7.59	4.17	45.66	4.00	7.65	24.29

Table 14: Proportion of hot and cold area in different LULCs in 2015 (%)

	Thermal pattern						
LULC types	Cold spot (99% sig)	Cold spot (95% sig)	Cold spot (90% sig)	Not sig	Hot spot (90% sig)	Hot spot (95% sig)	Hot spot (99% sig)
Vegetation	0.093	0.124	0.104	1.364	0.068	0.122	0.124
Urban	0.022	0.076	0.088	13.352	3.399	8.242	22.891
Agriculture	1.374	2.669	2.794	25.566	0.788	1.187	0.939
Water	6.849	1.054	0.455	2.267	0.038	0.043	0.038
Vacant land	0.011	0.035	0.045	2.230	0.229	0.419	0.520
Sandbars	0.022	0.016	0.015	0.253	0.021	0.042	0.016
Total	8.369	3.974	3.501	45.032	4.543	10.054	24.527

Hot spots tend to increase through time (from 29.7% in 2003 to 34.6% in 2015) but these are different trends between LULC types. Figure 18 reveals that urban class accounts for the largest hot spot area. It has continuously increased from 2003 to 2015. Agriculture has decreased the hot spot area from 1000 hectares in 2003 to 500 hectares in 2015. The change in other land use types including vegetation and vacant land has not become stable.

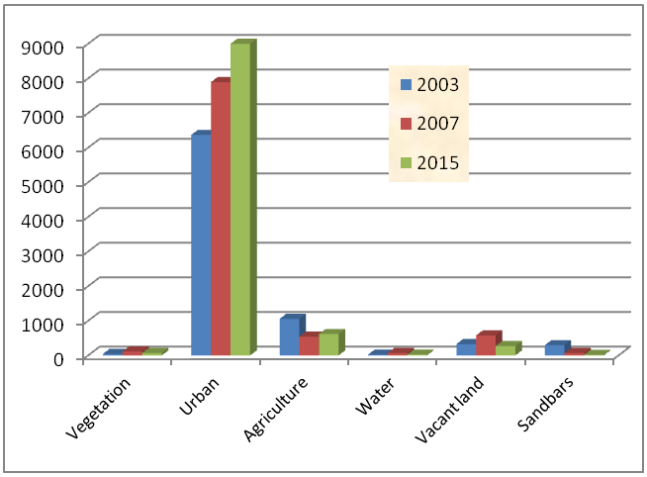


Figure 18: Change in hot spots total area (in hectares) for different LULC types in 2003, 2007, and 2015 (considering significance ≥ 95 %)

Land use land cover change was forced by urbanization made by the transformation between other land use land cover types to urban land use. Therefore, it is important to relate the urbanization and its fragmentation to the UHI effect. We can conclude (see Table 15) that urban built-up is the hottest area type whereas rural open land is the coldest area. The LST spatial pattern can be seen as directly correlated to the transition of different zones. LST decreases from the hottest area-

urban core to suburban area, rural built-up, and reaches the lowest mean LST at the rural open land. Within the urban zone, the urbanized open land has lower LST compared to urban and suburban built-up area. This shows the important role of open area such as parks, public space, and other non-built-up areas in reducing the UHI effect.

Table 15: Mean LST by urban landscape fragmentation (Celsius degrees)

Urban landscape	2003	2007	2015
Urban built-up	24.7	21.6	29.4
Suburban built-up	24.2	21.2	28.7
Rural built-up	23.9	20.8	28.0
Urbanized open land	23.9	20.9	28.3
Rural open land	23.2	20.3	27.5

Related to urbanization, we determined the mean LST in different urban development types. The LST statistics in each urban development type was shown in Table 16.

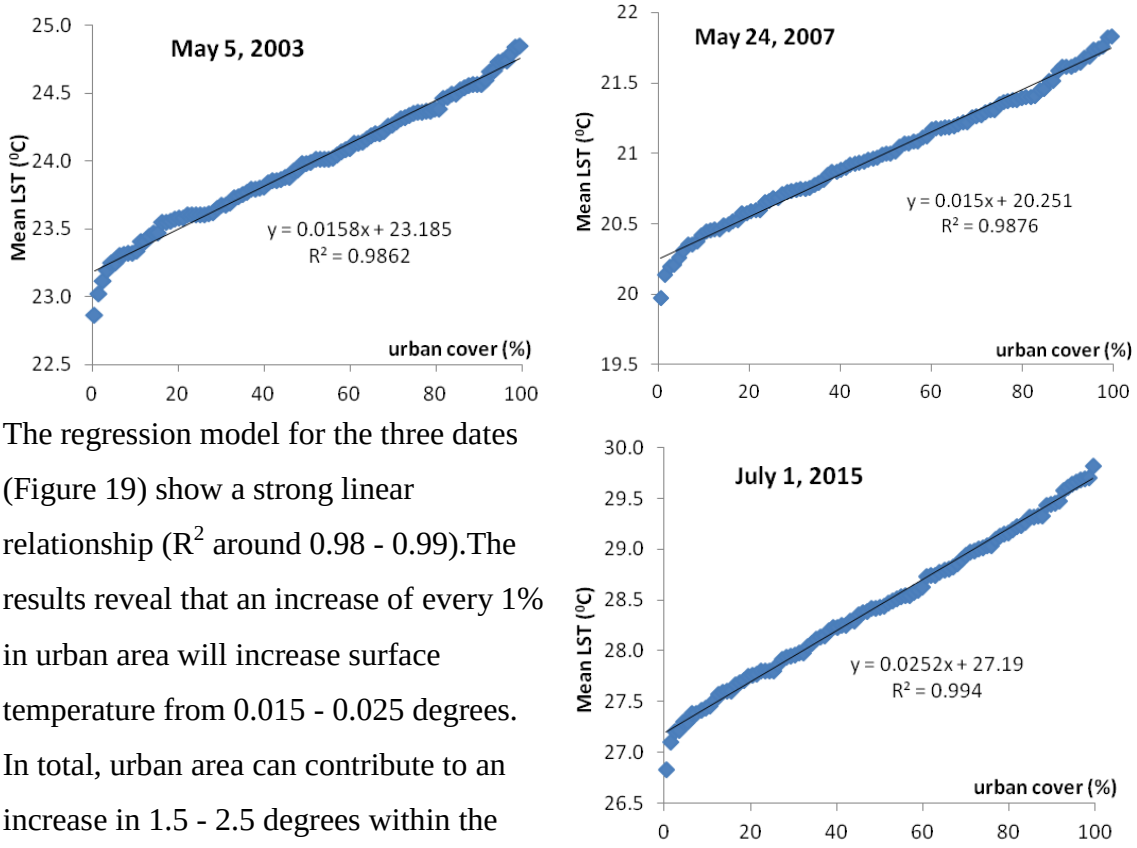
Table 16: LST at different urban development types (Celsius degrees)

	2003 - 2007	*2007-2015*	*2003 - 2015*
Infill	25.98	29.21	29.35
Extension	25.01	28.78	28.96
Leapfrog	24.43	28.17	28.35

Table 16 shows that infill development has the highest mean temperature and leapfrog development represents the lowest one. This can be explained by two main reasons. First, surrounding the infill urban area there is high LST land use type such as built-up area, which positively influences the infill area's LST. In contrast, surrounding leapfrog urban area there are other land use types with lower LST such as agriculture land. This helps reducing the leapfrog area mean LST. In addition, leapfrog urban area often have good planning policies with an appropriate percentage of public areas like parks and lakes. This good LULC structure contributes to the LST decrease. Urban area under expansion often has higher LST than "leapfrog" and lower LST than "infill" because its surrounding areas include both urban and other LULC types.

To analyse the relationship between urban cover and temperature, we used zonal statistics tool in ArcGIS to estimate the mean LST at each percentage of urban area (from 0 to 100%). We then applied linear regression to explore the correlation between LST and urban cover (Figure 19).

Figure 19: Linear regression analysis between urban rate and land surface temperature in 2003, 2007, and 2015



The regression model for the three dates (Figure 19) show a strong linear relationship (R^2 around 0.98 - 0.99).The results reveal that an increase of every 1% in urban area will increase surface temperature from 0.015 - 0.025 degrees. In total, urban area can contribute to an increase in 1.5 - 2.5 degrees within the study area.

4.5. Land surface temperature prediction and future LST pattern

We applied Kernel Ridge Regression (Saunders, Gammerman & Vovk, 1998) to predict LST in 2015. KRR was trained using the data corresponding to 2003 and 2007. The maps below show the predicted LST for 2015 using window sizes of 5x5, 10x10, and 20x20 (Figure 20 - 25).

At 5x5 window size (150m resolution), we can see a very clear difference between predicted and test LST using data in 2003 as the training set. The differences spread out all the parts of the city. The predicted LST using 2007 data has better match with test data. The differences also distribute at the same pattern but the range of differences has significantly reduced. Urban and water area predicted LST is better than other LULC types. Sandbars and agriculture class (especially wetland class) show the biggest difference.

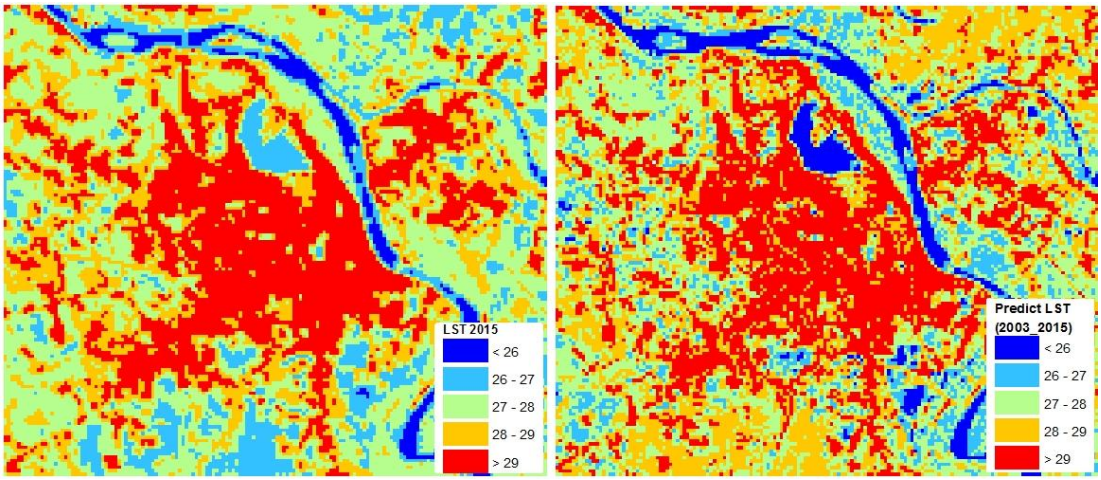


Figure 20: Real and predicted LST using data in 2003 as the training set (5x5 window size)

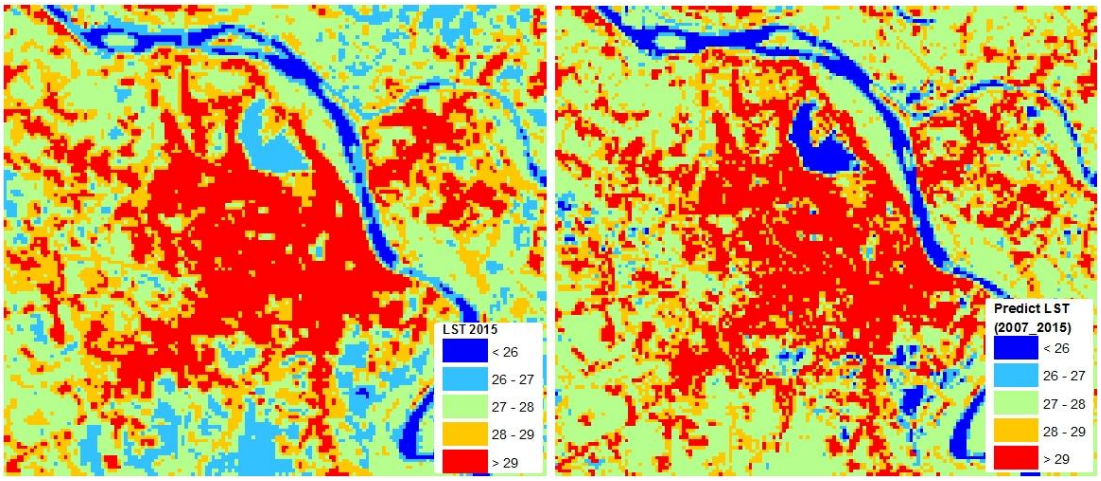


Figure 21: Real and predicted LST using data in 2007 as the training set (5x5 window size)

Using a 10x10 window size (300m resolution), the result has improved both for 2003 and 2007 data. However, the prediction using data in 2003 still shows a big difference as compared to the real LST values at the northeast and southwest parts. The predicted LST using data in 2007 has better result that the difference of LST in the northeast and southwest parts is closer to the real LST.

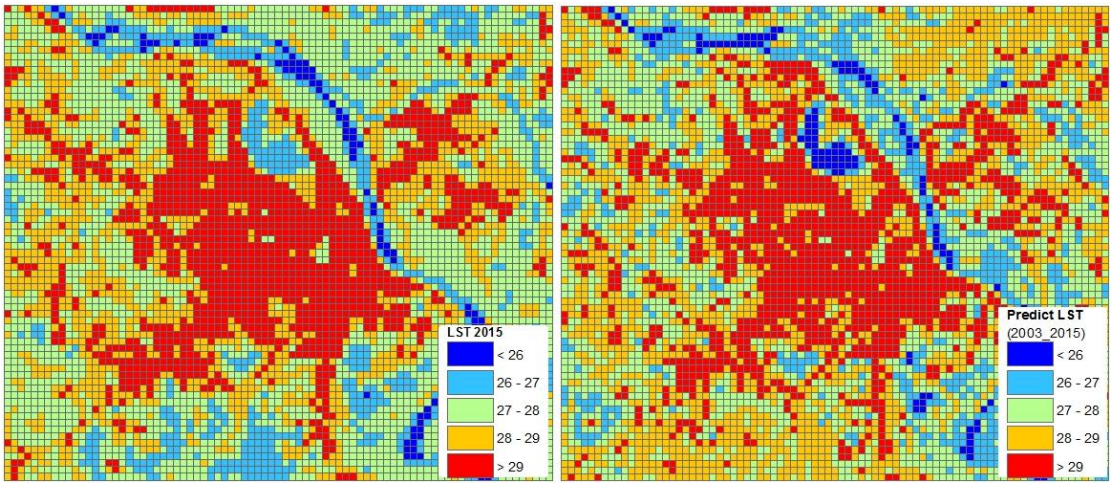


Figure 22: Real and predicted LST using data in 2003 as the training set (10x10 window size)

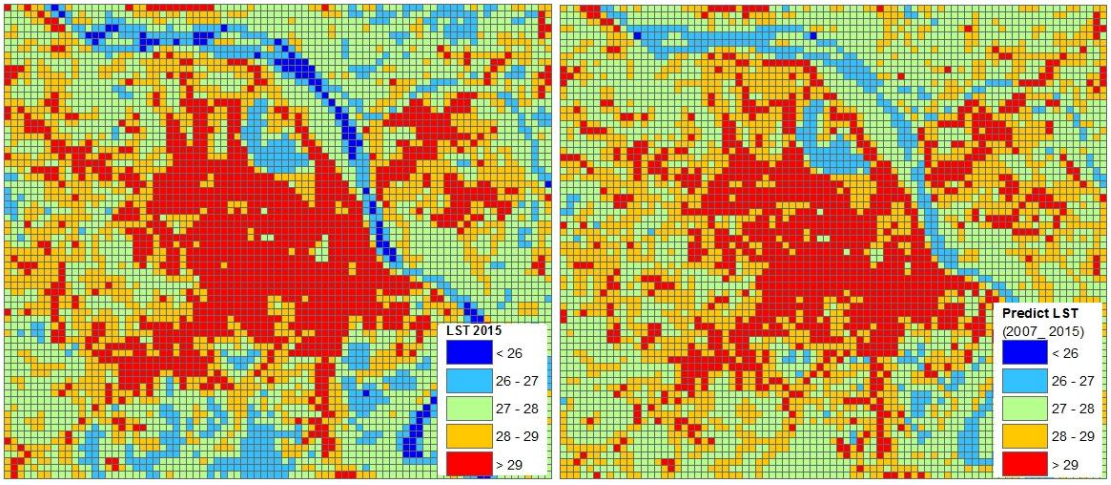


Figure 23: Real LST and predicted using data in 2007 as the training set (10x10 window size)

The result from a 20x20 window size (600m resolution) using the 2003 and 2007 data showed us a better prediction compared to using other window sizes. From the maps (Figures 24 and 25) we can clearly see that the predicted LST using 600 meters resolution data gave a very good forecast in which most of the area had the same LST pattern and value. It also demonstrates that using 2007 data gives a better prediction than applying 2003 data.

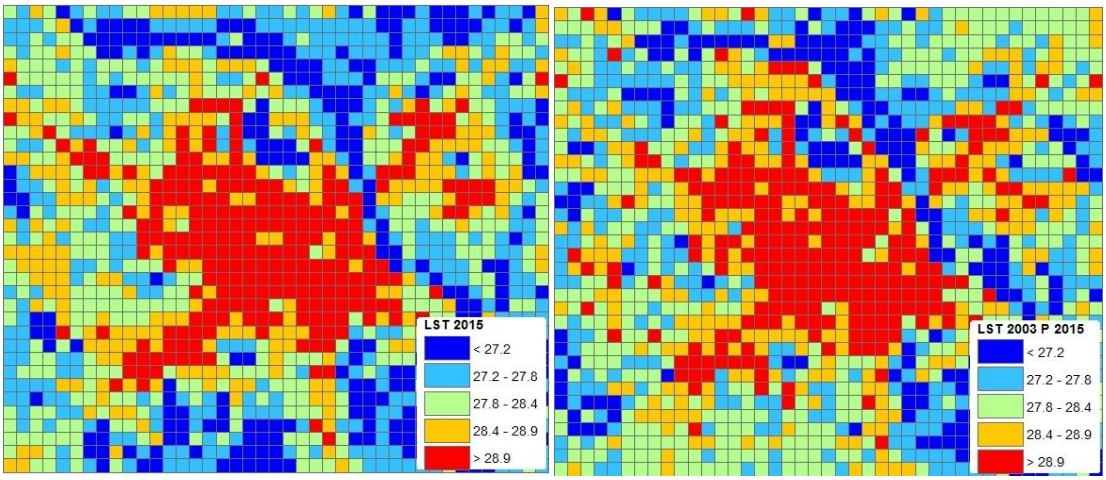


Figure 24: Real and predicted LST for 2015 using data in 2003 as the training set (20x20 window size)

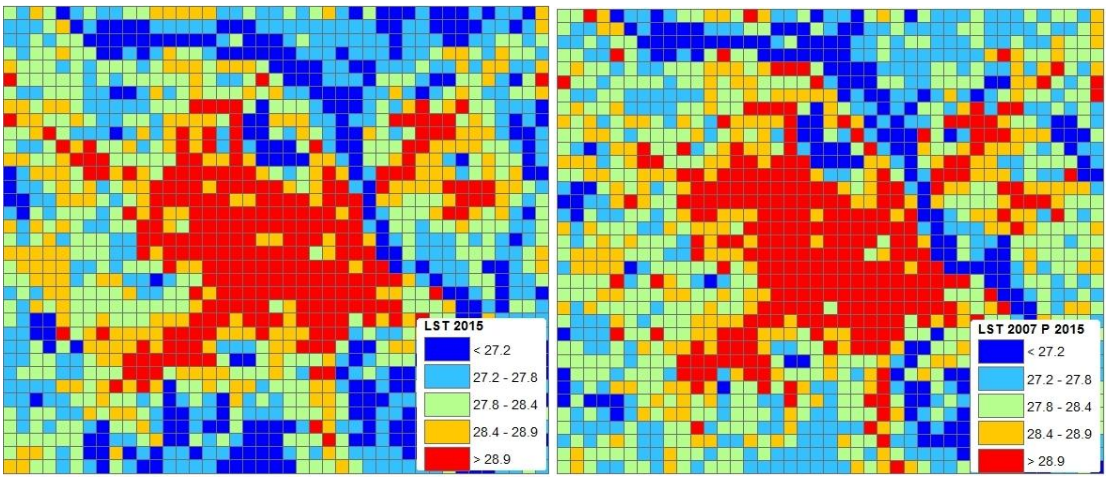


Figure 25: Real and predicted LST for 2015 using data in 2007 as the training set (20x20 window size)

To have a better prediction quality assessment, we compared the level and percentage of matched LST pixels between predicted value and test value in the study area (Table 17).

Table 17: Assessment of the difference between predicted LST and test LST (area percentage)

	2003 predict 2015			2007 predict 2015		
Window size	$< 0.5^0C$	$0.5^0C - 1^0C$	$> 1^0C$	$< 0.5^0C$	$0.5^0C - 1^0C$	$> 1^0C$
5x5	70.33	24.59	5.08	75.31	21.14	3.55
10x10	78.59	18.11	3.30	83.45	15.03	1.52
20x20	85.87	12.23	1.90	90.56	8.82	0.62

Table 17 shows that using 2007 as the training data and a 20x20 window size had a better performance, generating a large area with a < 0.5 degrees difference.

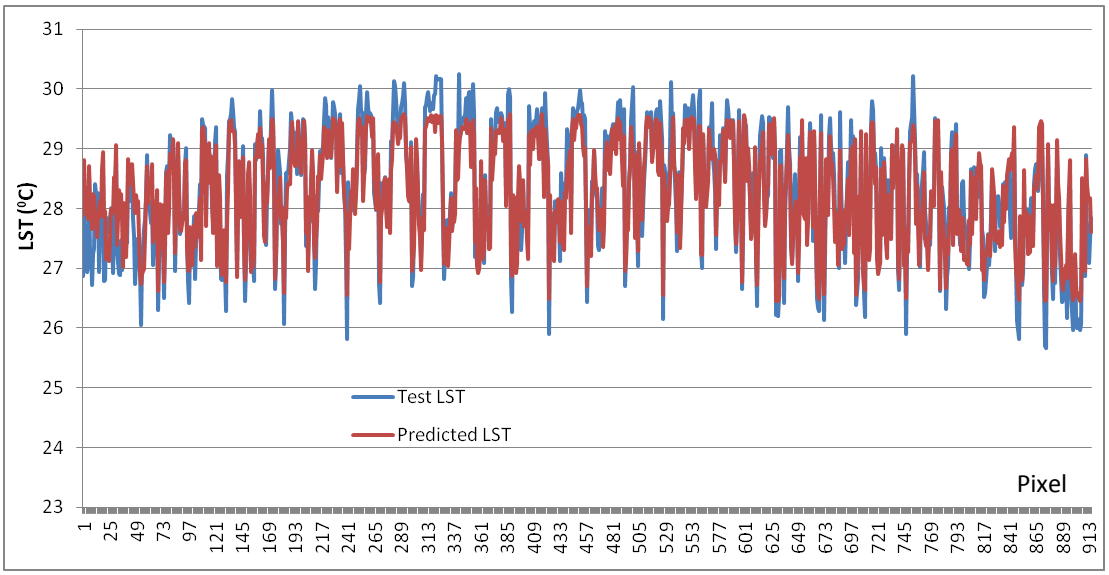


Figure 26: Variations of the test and predicted LST in 2015 using data in 2003

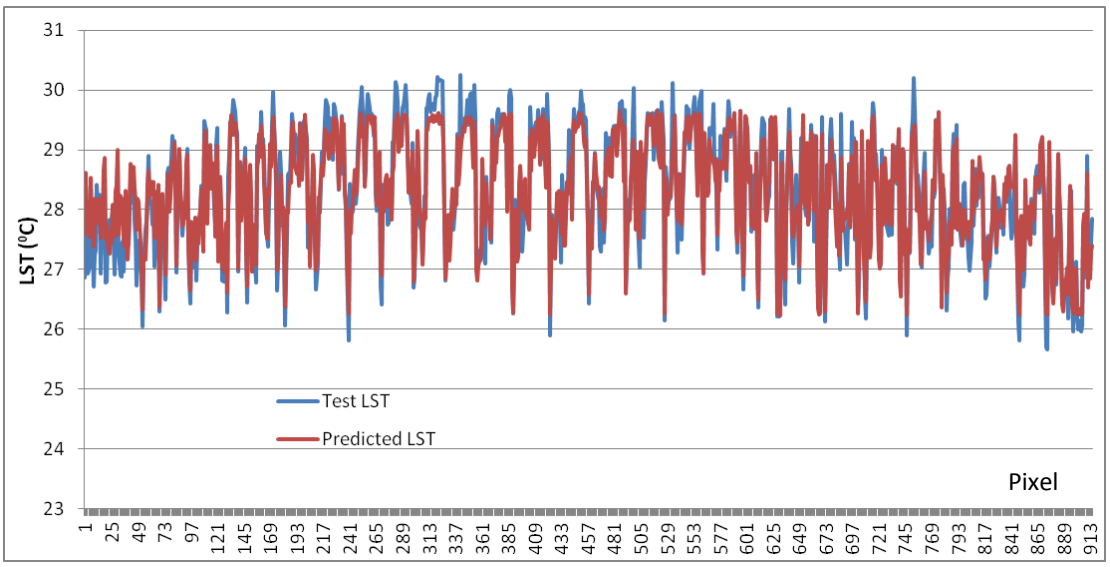


Figure 27: Variations of the test and predicted LST in 2015 using data in 2007

The results from LST prediction indicate that input LULC data and the window size significantly affects on the predicted values. When we investigated the relationship between LULC and LST using NDBI and NDVI indices in section 4.4, we concluded that 2007 and 2015 LULLC data had more similar trend in its relationship with LST. This is the reason of the better prediction capability when using data in 2007 as the training set. At three different window sizes, with an appropriate LULC input, the resolution of 600m (20x20) gives the best prediction. At

this level, both absolute LST value and LST pattern of predicted results are very close to the real values and patterns. From these findings, we performed the future LST prediction for 2023 using the 20x20 window size (600m resolution) and 2007 data as the training set. The map in Figure 28 shows the LST pattern in 2015 and 2023.

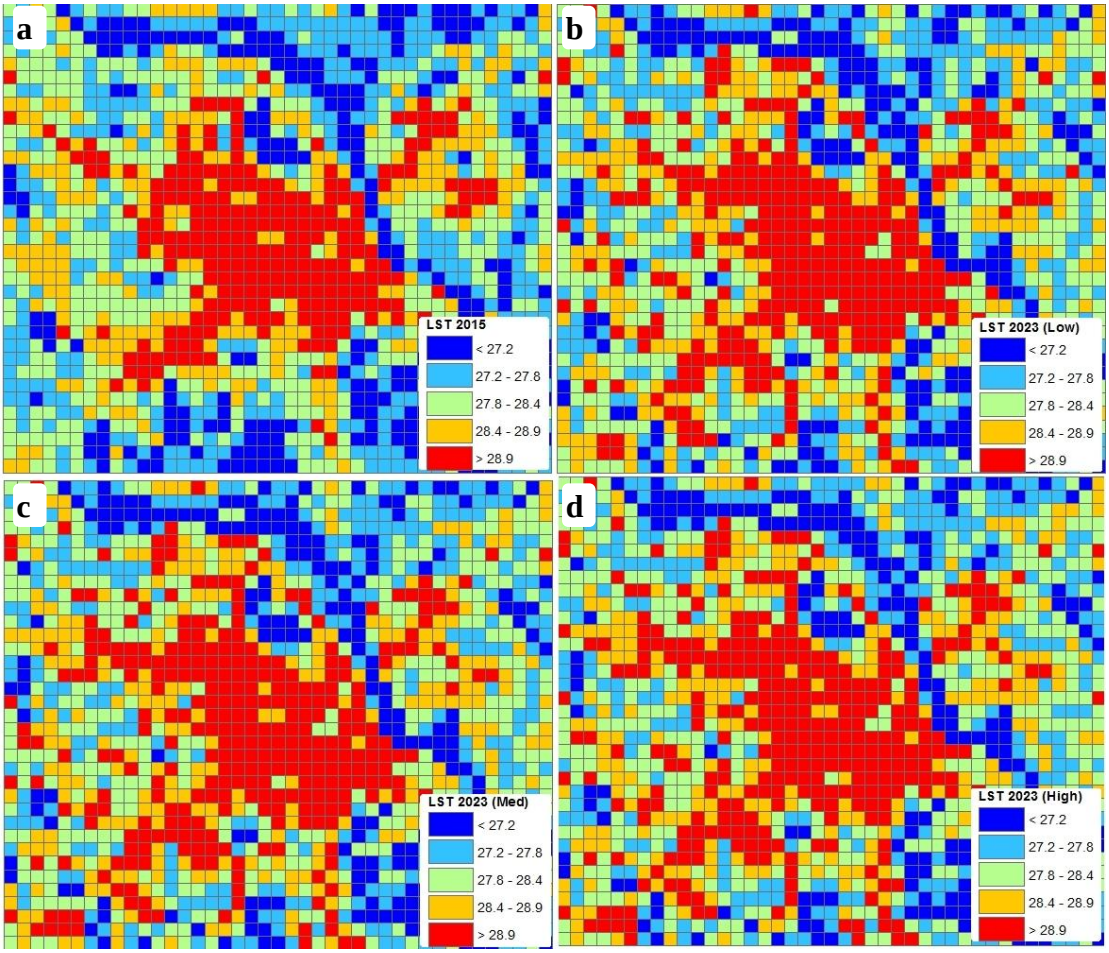


Figure 28: Map of LST in 2015 and 2023 at different simulation scenarios

(a: LST in 2015, b: LST in low growth scenario, c: LST in 2023 in medium growth scenario, d: LST in 2023 in high growth scenario).

All these maps demonstrate a general LST pattern having a high hot area concentration in the center, and cold area along the river, northeast, and southeast parts. However, the hot pattern tends to expand crucially on the southwest part of the city. Another trend is the increase of the hot pattern in the northwest area, continuing from the existed urban area. These changes have the same trend in the relation to the urban expansion. Tables 18 and 19 give a more detailed picture about the LST pattern change. In general, the LST pattern in the period of 2015-2023 is characterized by the decrease of

cold and cool LST areas and the increase of hot and very hot areas. Warm LST area does not change very much as it increases or decreases in a reduced number of zones.

At the low growth scenario, cold LST area changes from 8784 hectares (17%) to 7704 hectares (14.98%). Cool area losses 2232 hectares (4.32%), decreasing from 13068 hectares (25.3%) to 10548 hectares (20.98%). The opposite trend appears from hot and very hot LST areas. Specifically, the very hot area increases 936 hectares (1.81%) from 10296 hectares (19.93%) in 2015 to 11484 hectares (21.74%) in 2023. Hot area gains 2412 hectares (4.67%), increasing from 7128 hectares (13.8%) to 9504 hectares (18.47%).

Table 18: Area and proportion of LST pattern in 2015 and 2023 at different growth scenarios

	2015		2023- Low		2023-Med		2023-High	
LST pattern	Hectares	%	Hectares	%	Hectares	%	Hectares	%
Cold	8784	17.00	7704	14.98	7740	14.91	7812	15.12
Cool	13068	25.30	10548	20.98	10836	20.42	10224	19.79
Warm	12384	23.97	12420	23.83	12312	24.04	12456	24.11
Hot	7128	13.80	9504	18.47	9540	18.40	9504	18.40
Very hot	10296	19.93	11484	21.74	11232	22.23	11664	22.58

Table 19: Change in LST pattern in 2015-2023 at different urban growth scenarios

	2023-Low		2023- Med		2023-High	
LST pattern	Hectares	%	Hectares	%	Hectares	%
Cold	-1044	-2.02	-1080.00	-2.09	-972.00	-1.88
Cool	-2232	-4.32	-2520.00	-4.88	-2844.00	-5.51
Warm	-72	-0.14	+36.00	+0.07	+72.00	+0.14
Hot	+2412	+4.67	+2376.00	+4.60	+2376.00	+4.60
Very hot	+936	+1.81	+1188.00	+2.30	+1368.00	+2.65

At the medium growth scenario, cold LST area decreases from 8784 hectares (17%) to 7740 hectares (14.91%), losing an area of 1080 hectares (2.09%). Cool area losses 2520 hectares (4.88%), decreasing from 13068 hectares (25.3%) to 10836 hectares (20.42%). Very hot area increases 1188 hectares (2.3%), from 10296 hectares (19.93%) in 2015 to 11232 hectares (22.23%) in 2023. Hot area gains 2376 hectares (4.6%), increasing from 7128 hectares (13.8%) to 9540 hectares (18.4%).

At the fast growth scenario, the change in LST pattern is highlighted by the large increase in the very hot area. In the 2015-2023 period, very hot area increases 1368 hectares (2.65%) from 10296 hectares (19.93%) in 2015 to 11664 hectares

(22.58%). Compared to the other two scenarios, high urban growth makes the largest contribution to the increase in the very hot LST area.

LST pattern change through urbanization assessment in the study area shows that urban growth creates a severe impact on UHI. This process reduces the agriculture land and non-built-up space. Consequently, urbanization leads to the loss of cold and cool area and the gain of hot and very hot area. Hence, it is necessary to consider an appropriate proportion of green space and open land in the city to mitigate the UHI effect.

5. CONCLUSION

Our study explored the LULC change and urbanization to investigate the relationship between these processes and UHI effect in and around Hanoi inner city. We discovered urban and agriculture areas are two main LULC types in the LULC structure of the city. Water also plays an important role as it occupies more than 10% of total land use structure. The study results showed that LULC in study area has changed significantly in the period of 2003-2015. The main driver of LULC change is the transformation from agriculture land to urban area due to urbanization and industrialization. From the analysis of LULC change and characteristics, we recognized some irrational issue emerging in the entire area. First, the expansion of urban area included a large area of "infill" development that filled up the available open space existed in the city. Besides, many industrial areas were located inside the urban area. This might be the consequence of the fast urbanization and lack of urban growth control policy. In addition, water area including lakes and pools inside urban area was decreased due to the urban construction. All these issues would create the negative effects on the urban population and environment.

We also explored the LST characteristic to understand the thermal pattern of the city in general and in a more detail within each LULC type in summer time. We found that urban area got very high LST compared to other areas. Within urban area, LULC type that has more vegetation and public space will have lower LST. Regarding to the impact of urbanization on UHI, we conclude that LST is affected by both urban density and type of urban development. We recommend that "infill" urban development is a bad behaviour, which we should avoid in urban development. Such process of filling up the open land in the city brings many negative impacts. It crucially increases the urban warming effect within the city. Besides, it reduces the living conditions of people through decreasing the public space.

We applied the linear regression method, using NDBI and NDVI indices to investigate the correlation between LULC and LST. The result gave the same trend in many studies that have done in which NDBI showed the direct relationship and NDVI indicated the inverse correlation with LST. We performed a deeper investigation by analyzing the correlation between NDBI, NDVI, and LST within each LULC type. This exploration gave the conclusion that different LULC types would have different correlation. It also revealed a very interesting point in which the relationship in the

same LULC type might vary from geographical distribution as well as temporal change. Composition and pattern of each LULC type may strongly affect to the correlation between LULC and LST.

We performed the non-linear regression model using KRR algorithm to predict future LST based on future LULC in different scenarios. Results indicated that the resolution of 600m (20x20 window size) was most suitable for investigating the relationships between LULC and LST. More important, we recommend that choosing the LULC input is very important step to archive a better prediction. After examining the effect of LULC characteristic and its change on LST from the past to present and future, we confirmed that UHI has been affected by both LULC structure and LULC change. In addition, UHI will be influenced by speed and type of urban development. It is an important recommendation to urban planners and policy decision makers. They have to consider not only the rate of urbanization but also its fragmentation in city planning. They need to control both the urbanization speed and urban landscape structure to reduce the urban warming effect.

Our study was mainly based on daytime LST and using three different dates in summer time. We believe that the findings will be more comprehensive if we use both daytime and nighttime LST and more data instead of only three years. Future study using higher resolution imageries for LULC classification may give a better explanation of LULC composition-LST correlation, as well as result that is more practical for urban planners. We need to test other regression methods to predict LST prediction as well as apply other possible window size to have better result.

REFERENCES

Adams, M. P., & Smith, P. L. (2014). A systematic approach to model the influence of the type and density of vegetation cover on urban heat using remote sensing. *Landscape and Urban Planning, 132, 47-54.*

Angel, S., Parent, J., & Civco, D. L. (2012). The fragmentation of urban landscapes: global evidence of a key attribute of the spatial structure of cities, 1990–2000. *Environment and Urbanization, 24(1), 249-283.*

Andhang, R., Han, S.L., Phuong, T.T.T., Tetsu, K., Takahiro, T.K.M.. 2014. The Cooling Effect of Green Strategies Proposed in the Hanoi Master Plan for Mitigation of Urban Heat Island. *30th International Plea Conference, 16-18 December, CEPT University, Ahmedabad, India.*

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

Barsi, J., Barker, J. L., & Schott, J. R. (2003). An atmospheric correction parameter calculator for a single thermal band earth-sensing instrument. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International* (Vol. 5, pp. 3014-3016). IEEE.

Burchfield, M., Overman, H. G., Puga, D., & Turner, M. A. (2006). Causes of sprawl: A portrait from space. *The Quarterly Journal of Economics, 587-633.*

Chun, B., & Guldmann, J. M. (2014). Spatial statistical analysis and simulation of the urban heat island in high-density central cities. *Landscape and urban planning, 125, 76-88.*

Coseo, P., & Larsen, L. (2014). How factors of land use/land cover, building configuration, and adjacent heat sources and sinks explain Urban Heat Islands in Chicago. *Landscape and Urban Planning, 125, 117-129.*

Coseo, P., & Larsen, L. (2015). Cooling the Heat Island in Compact Urban Environments: The Effectiveness of Chicago's Green Alley Program. *Procedia Engineering, 118, 691-710.*

Guo, G., Wu, Z., Xiao, R., Chen, Y., Liu, X., & Zhang, X. (2015). Impacts of urban biophysical composition on land surface temperature in urban heat island clusters. *Landscape and Urban Planning, 135, 1-10.*

Grimmond, S. (2007). Urbanization and global environmental change: local effects of urban warming. *The Geographical Journal, 173(1), 83-88.*

Juer, S., Shihong, D., Xin, F., Luo, G. (2014). The relationships between landscape compositions and land surfacetemperature: Quantifying their resolution sensitivity with spatial regression models. *Landscape and Urban Planning, 123, 145– 157.*

Junxiang, L., Conghe, S., Lu, C., Feige, Z., Xianlei, M., Jianguo, W. (2011). Impacts of landscape structure on surface urban heat islands: A case study of Shanghai, China. *Remote Sensing of Environment, 115, 3249–3263.*

Kim, J. P., & Guldmann, J. M. (2014). Land-use planning and the urban heat island. *Environment and Planning B: Planning and Design, 41, 1077-1099.*

Kikegawa, Y., Genchi, Y., Yoshikado, H., & Kondo, H. (2003). Development of a numerical simulation system toward comprehensive assessments of urban warming countermeasures including their impacts upon the urban buildings' energy-demands. *Applied Energy, 76(4), 449-466.*

Kourtidis, K., Georgoulias, A. K., Rapsomanikis, S., Amiridis, V., Keramitsoglou, I., Hooyberghs, H., ... & Melas, D. (2015). A study of the hourly variability of the urban heat island effect in the Greater Athens Area during summer. *Science of The Total Environment, 517, 162-177.*

Kumar, K. S., Bhaskar, P. U., & Padmakumari, K. (2012). Estimation of land surface temperature to study urban heat island effect using LANDSAT ETM+ image. *International journal of Engineering Science and technology, 4(2), 771-778.*

Luyssaert, S., Jammet, M., Stoy, P. C., Estel, S., Pongratz, J., Ceschia, E., & Dolman, A. J. (2014). Land management and land-cover change have impacts of similar magnitude on surface temperature. *Nature Climate Change,4(5), 389-393.*

Meineke, E. K., Dunn, R. R., & Frank, S. D. (2014). Early pest development and loss of biological control are associated with urban warming. *Biology letters, 10(11), 20140586.*

Mirzaei, P. A. (2015). Recent challenges in modeling of urban heat island. *Sustainable Cities and Society, 19, 200-206.*

Myint, S. W., Wentz, E. A., Brazel, A. J., & Quattrochi, D. A. (2013). The impact of distinct anthropogenic and vegetation features on urban warming. *Landscape Ecology, 28(5),959-978.*

Nam, T.H.H., Kubota, T., Trihamdani, A.R. (2015) Impact of urban heat island under the Hanoi Master Plan 2030 on cooling loads in residential buildings. *International Journal of Built Environment and Sustainability, 2(1), pp.48-61.*

Niem, T.H., Wen, L., Renee, L. N., Darlene, O.G., Duy,T.X. Southeast Asia 2050-2100: Imagining New Lifeways/Lifestyles. In After Climate Change and Culture-shift: Imagining a World, ed. J. Norwine. *New York, NY: Springer, 2013.*

Oke, T. R. (1982). The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society, 108(455), 1-24.*

Palit, A. K., & Popovic, D. (2006). Computational intelligence in time series forecasting: theory and engineering applications. *Springer Science & Business Media.*

Plocoste, T., Jacoby-Koaly, S., Molinié, J., & Petit, R. H. (2014). Evidence of the effect of an urban heat island on air quality near a landfill. *Urban Climate,10, 745-757.*

Quattrochi, D. A., & Luvall, J. C. (1999). Thermal infrared remote sensing for analysis of landscape ecological processes: methods and applications. *Landscape ecology, 14(6), 577-598.*

Radhi, H., Fikry, F., & Sharples, S. (2013). Impacts of urbanisation on the thermal behaviour of new built up environments: A scoping study of the urban heat island in Bahrain. *Landscape and Urban Planning, 113, 47-61.*

Rizwan, A. M., Dennis, L. Y., & Chunho, L. I. U. (2008). A review on the generation, determination and mitigation of Urban Heat Island. *Journal of Environmental Sciences, 20(1), 120-128.*

Rouse Jr, J., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA special publication, 351, 309.*

Rotem-Mindali, O., Michael, Y., Helman, D., & Lensky, I. M. (2015). The role of local land-use on the urban heat island effect of Tel Aviv as assessed from satellite remote sensing. *Applied Geography, 56, 145-153.*

Saunders, C., Gammerman, A., & Vovk, V. (1998). Ridge regression learning algorithm in dual variables. In *(ICML-1998) Proceedings of the 15th International Conference on Machine Learning (pp. 515-521).*

Sobrino, J. A., Jiménez-Muñoz, J. C., & Paolini, L. (2004). Land surface temperature retrieval from LANDSAT TM 5. *Remote Sensing of environment, 90(4), 434-440.*

Taha, H. (1997). Urban climates and heat islands: albedo, evapotranspiration, and anthropogenic heat. *Energy and buildings, 25(2), 99-103.*

Thuy, T.D., Duy, T.X., Duc, T.T., Luong, D.N.H. (2014). Using Landsat data to study the relationship between land surface temperature and land use land cover of Hanoi city. *Proceedings of National Geography Association Conference, Vol. 6 (2014), 102 - 110, Ho Chi Minh city, Vietnam.*

Walawender, J. P., Szymanowski, M., Hajto, M. J., & Bokwa, A. (2014). Land surface temperature patterns in the urban agglomeration of Krakow (Poland) derived from Landsat-7/ETM+ data. *Pure and Applied Geophysics, 171(6), 913-940.*

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing, 27(14), 3025-3033.*

Yonezawa, G. 2009. Generation of DEM for Urban Transformation of Hanoi, Vietnam. *Kyoto Working Papers on Area Studies 62(60), 1-10.*

Yuan, F., & Bauer, M. E. (2007). Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in Landsat imagery. *Remote Sensing of Environment, 106(3), 375-386.*

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing, 24(3), 583-594.*

Zhou, W., Qian, Y., Li, X., Li, W., & Han, L. (2014). Relationships between land cover and the surface urban heat island: seasonal variability and effects of spatial and thematic resolution of land cover data on predicting land surface temperatures. *Landscape ecology, 29(1), 153-167.*

Klaus Tempfli, Norman Kerle, Gerrit C. Huurneman, Lucas L. F. Janssen, 2009. *Principles of Remote Sensing. An introductory textbook* (ITC Educational Textbook Series 2) (ITC, Enschede).

Cong giao tiep dien tu thanh pho Hanoi (Hanoi portal, 2015). http://www.hanoi.gov.vn.

ESRI, 2016. *How Hot Spot Analysis (Getis-Ord Gi*) works?* http://pro.arcgis.com/en/pro-app/tool-reference/spatial-statistics/h-how-hot-spot-analysis-getis-ord-gi-spatial-stati.htm.Accessed 16 February 2016.

Landsat, N. A. S. A. (7). *Science Data Users Handbook. 2011-03-11*. http://landsathandbook. gsfc. nasa. gov/inst_cal/prog_sect8_2. html. Accessed 18 October 2015.

Landsat, N. A. S. A. (8). *Science Data Users Handbook*. 2015-june. http://landsat.usgs.gov/l8handbook.php. Accessed 23 September 2015.

World bank. (2011). *Vietnam Urbanization Review: Technical Assistance Report*.

Masters Program in Geospatial Technologies

EXPLORING LAND USE LAND COVER CHANGE TO UNDERSTAND URBAN WARMING EFFECT IN HANOI INNER CITY, VIETNAM

Tran Xuan Duy

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

Masters Program in **Geospatial Technologies**

Supported by:
Education and Culture
ERASMUS MUNDUS