MINISTRY OF EDUCATION AND TRAINING
HANOI UNIVERSITY OF MINING AND GEOLOGY

TRINH THI HOAI THU

RESEARCH ON THE IMPACT OF PERI-URBANIZATION PROCESS ON AGRICULTURAL LAND USE IN DONGANH – HANOI

Specialty : Geodesy and Mapping

Code : 62520503

SUMMARY OF THE DOCTOR THESIS

HANOI - 2015

The thesis has been completed at Department of Mine surveying, Faculty of Surveying and Mapping, Hanoi University of Mining and Geology

Scientific Supervisors:

1. Prof. Dr. Vo Chi My

Vietnam Association of Geodesy, Cartography and Remote Sensing

2. Prof. Dr. Pham Van Cu

Hanoi university of science

Examiner 1: Prof. Dr. Truong Quang Hai

Hanoi university of science

Examiner 2: Dr. Tran Dinh Luat

Vietnam natural resource and environment corporation

Examiner 3: Dr. Vu Kim Chi

Institute of Vietnamese Studies and Development Sciences

The thesis will be defended at the University examination Council at the Hanoi University of Mining and Geology, at….. h, …… 2015.

This dissertation is submitted to National library and library of Hanoi university of Mining and Geology

INTRODUCTION

1. The importance of the thesis

The process of peri- urbanization in Vietnam has caused many negative changes to recourse and enviroment in which land resources are directly and strongly affected.

Donganh is a peri-urban area that is a typical of spatial pattern of land use changes due to urbanization.

Quantitative linking between urbanization and spatial pattern of land use changes over time is necessary for decision-makers such as sustainable land use planning and spatial planning.

2. Research Objectives

a/ Objectives:

Determining the relationship between urbanization process and the structure of land use change through spatial modeling, integration of remote sensing and statistical information.

b/ Tasks:

To achieve the above objectives, the research must perform the following tasks:

- Overview of documents related to the thesis.
- Object-oriented fuzzy classification using extracted information from remote sensing data, statistical methods, and spatial modeling assessing the degree of urbanization and the effects of urbanization to structure of agricultural land use changes in the study area.

3. Scope and limitation of the thesis

The study had some limitations:

a/ Objective

The method to determine the relationship urbanization, change use and impact urbanization on land use change.

b/ Scope of the thesis

- Donganh, Hanoi.

- The study evaluated a number of major types of agricultural land use extracting from the Landsat images

4. Methodologies used in the thesis

The thesis using the following methodologies:

Synthesis methods; analysis methods; statistical methods; induction methods; modeling and experimental method.

5. Protected Arguments

Argument 1: Comparison between membership functions samples in object-oriented fuzzy classification of remote sensing data in peri-urban areas to determine appropriate indicators results in high accuracy of land use classification.

Argument 2: Integrated multinomial regression and neural network model determines quantitative of the impact of peri-urbanization on agricultural land use in the study area.

Argument 3: Characteristic factors of peri-urbanization including natural and socio-economic conditions are related the change of agricultural land use in Donganh, Hanoi.

6. New contributions of the thesis

- Comparison of the membership functions samples determines the appropriate indicators to extract information of land use.
- To develop method of assessing peri-urbanization through the application of PCA
- To make an interdisciplinary approach to socio-economic and remote sensing data integration in impact research of peri-urbanization on land use.

7. Scientific and practical meaning of thesis

a/ Scientific meaning*:*

The results of the thesis confirmed that the advantages of Object-oriented fuzzy classification to give high accuracy.

The interdisciplinary approach in the study using by regression analysis and spatial modeling to evaluate driving forces of land use change.

b/ Practical meaning*:*

The research results are documented support for the planning agencies, regulatory authorities, and build additional mechanisms policy orientation regulations.

8. Organization of the thesis

The document includes three main chapters with introduction, conclusion, and a reference.

CHAPTER ONE
LITERATURE REWIEW

1.1. Literature review of the impact of urbanization due to land use changes

1.1.1. Concept of urbanization, urban and peri - urban

Urbanization is a phenomenon of multiple and diverse range of economic, social and environmental development manifested in industry, industrial, commercial production, division of labor, housing and conversion work, etc [9]. The process of urbanization is not just happening in the areas are considered to be urban, in fact, the presence of urbanization also appear in the areas are not urban spaces, but also in the region rural and peri-urban areas [118, 152, 170].

Peri-urbanization thus indicates the phenomenon of sprawl and absorption of surrounding rural areas by cities. Urbanization in peri-urban areas emphasizes the spread formation and development of urban properties alternative to the characteristic agriculture, which is outstanding before urbanization [7].

1.1.2. Cause of land use change

Land use change is the result of the interaction of biophysical factors and socio-economic factors (Fig 1.2) [28, 69, 94]. The socio-economic factors can play a major role in the impacts on land use changes [105, 167].

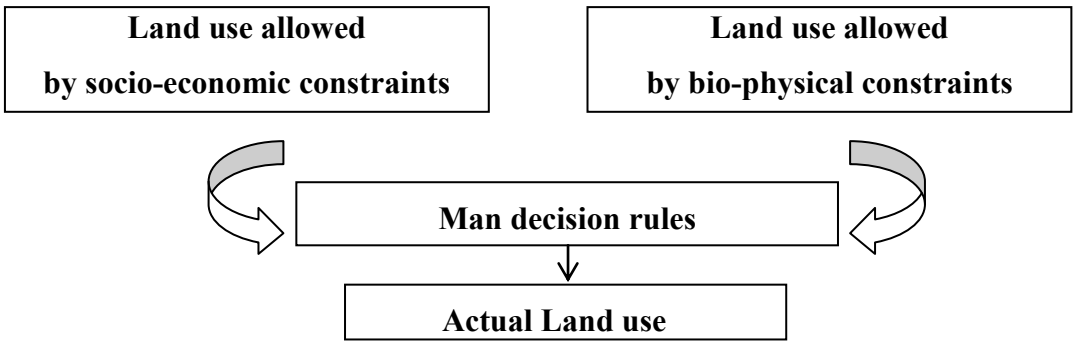


Figure 1. 2. Actual land use [48]

1.1.3. Methods of impact assessment factors to land use changes

Impact of urbanization on land use change is important aspects for many scientific researches. There are studies based on spatial data and these studies using satellite data and GIS mostly identify urban sprawl on agricultural land by assessing land use changes in the studies of the authors, such as Do Dinh Duan, Nguyen Dinh Duong, Pham Minh Hai, Hualou Long, Qingshui Lu, Somporn Sangawonge, Tran Thi Van, Xiao, S. Liu Y [45, 47, 62, 104 , 110, 137, 160, 175, 178]. There are also studies based on spatial restructuring of land-use patterns, through landscape metrics change to quantify characteristic of urban development over time and space as in the study of Xian about the evaluation urbanization and the loss of agricultural land in Yixing China [124], or in the study of Xiuhong, Pham Minh Hai, Pham Van Cu [28, 35, 63, 167]. There are studies assessing the impact of urbanization on agricultural land based on statistics and data collection in social surveys, the impact of urbanization on land use, such as in the study by Van den berg [158], Tran Duc Vien [165], Brody Lee [92] on agricultural land changes in Hanoi. These studies are assessing the overall progress the relationship between urbanization and

land use in two ways, firstly the geographical perspective and secondly the viewpoint of society. Two independent and individual approaches have exposed some limitations. Social researchers link the issue of urbanization and land use by deductive and inductive approaches [7, 158, 165], to analyze and consider degree of urbanization more generally including both economic and social considerations. Their drawback is based solely on statistical data evaluation without quantitative spatial relationship between urbanization and land use. Whereas geographers using both remote sensing and geographic information systems (GIS) only consider that the urbanization similarities urban sprawl to quantify land use changes [35, 45, 47, and 110]. However, the above factors affecting land use change shows that land use change is a complex phenomena, the complexity lies not only in land use but also and the factors. Thus, a branch or a single method is not enough to analyze them in a comprehensive and insightful way [26, 164]. When studying the issue of land use, the necessity of combining methods towards interdisciplinary approach [24, 26, 89, 105, 131, 154, and 162] through integrated multi-dimensional data which including economic and social data and spatial data as remote sensing images are typical one. Research linking remote sensing science and social science is a modern trend, this trend has been the scientists mentioned in his "People and pixels: Linking remote sensing and social science", the authors give some potential applications and advantages of the integrated data in the study of issues related to social sciences and natural sciences [73].

The integration of remote sensing data and eco-social data to analyze has created a breakthrough for the study of land use related to social issues more rigorously and comprehensively [24, 59, 104]. The statistical methods and multi - dimensional modeling is used for data analysis in researching the relationship between land use change and impact factors. The methods are considered effective when in solving space correlation problems in geographical researches that analyzes based on traditional

statistics cannot be achieved [28]. In Vietnam, there have been studies in this trend as the studies of Müller [122], Vu Kim Chi [30], Pham Thi Thanh Hien[67]. However, these studies focused only to assess the relationship between land use and natural conditions, but the assessment of land use and social economic conditions have limited.

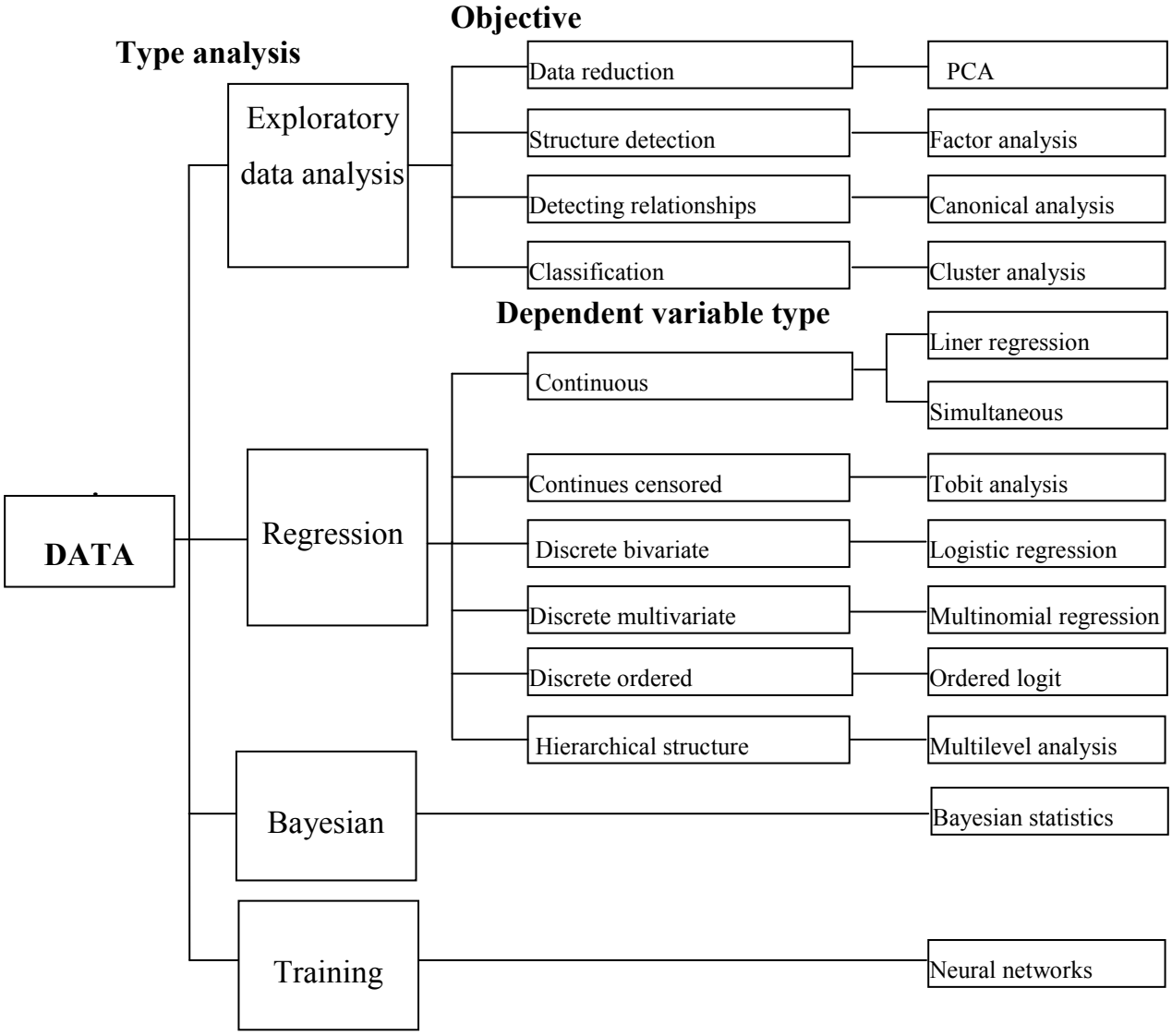


Figure 1. 1. Classification of analysis techniques base on objective and data structure [96]

There are two approaches of the analysis: liking data at point level or at polygon level [73]. There are four groups of analysis techniques for analyzing of spatial patterns: exploratory data analysis, regression, Bayesian, and training as shown in figure 1.3. The diversity in data structures, research questions, and case study-specific conditions make a

careful analysis of the requirements of the method necessary for each specific case study.

1.2. Literature review of land use/ land cover

1.2.1. Concept of land cover and land use

Land cover is the biophysical state of the earth's surface and immediate subsurface [154]. while, land use concerns the function or purpose for which the land is used by the local human population and can be defined as the human activities which are directly related to land, making use of its resources or having an impact on them [49].

Land cover and land use are closely related and they depend largely on the characteristics of land: vegetation structure, location, etc. In fact, land cover can stand individually but generally, land use is not unrelated to land cover in various ways and affects it with various implications [108, 131].

1.2.2. Classification methods for remote sensing data

When it comes to image classification, there are two issues to be addressed: the method of classification and scale of classification.

There are many methods including unsupervised and supervised classifications or statistical and nonstatistical techniques and three scales in classification including pixel, sub-pixel and object.

Currently, the trend is approaching object and non-statistical techniques for extracting land use and land cover information in which the algorithms using fuzzy theory for classification results with higher accuracy. Fuzzy classification method used to extract land-use information from images with high resolution such as Quickbird, IKONOS to the medium resolution like that Landsat [10, 98, 112, 136, 180]. When applying fuzzy theory to classify Landsat scientists have confirmed that this method for better accuracy approaches traditional statistical techniques [10, 13, 21, 27, 82, 98, 111, 112, 136, 168, 180].

1.2.3. Change detection techniques

Change detection techniques using satellite data are divided into two main groups of methods: Pre-classification and Post-classification [133]. Change detection techniques are compiled from documents by Ashbinhdu Singh (1989), David A. Mouat (1993), Ross S. Lunetta (1993), Vittorio Castelli et, al (1993), D. Lu et, al (2004) và Gong Jiannya (2008) [39, 78, 109, 133, 141] showing in figure 1.4.

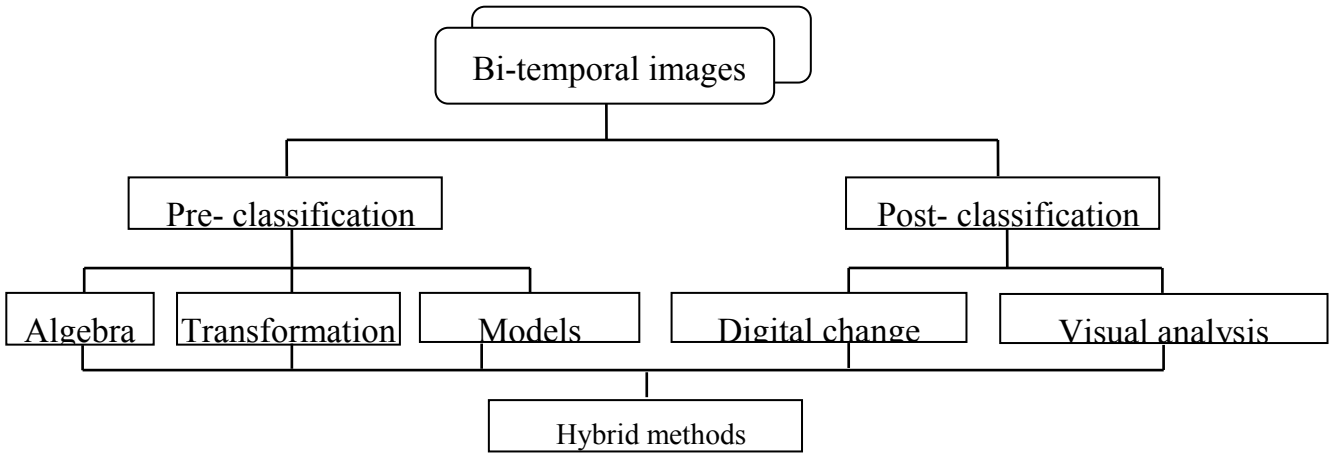


Figure 1. 2. Classification of change detection algorithms

1.3. Approaches and methodologies of the thesis

1.3.1. Approaches of thesis

1. Urbanization

In peri-urban areas, urbanization is a process in which rural areas located on the outskirts of established cities become more urban in character, in physical, economic, and social terms, often in piecemeal fashion. Therefore, it is synonymous with urban sprawl, the views expressed through the following set of criteria:

- The economic structure
- The structure of labor
- Increasing population density, standard of living
- Land use information

2. Cause of land use change

The study will assess urbanization and its effects on land use change from master to detail. In addition to the criteria of urbanization in peri-

urban areas mentioned above, the natural factors such as distance to the center, distance to roads, or the distance to residential areas, as well as the factors related to land use change should be considered.

3. Land use information:

Remote sensing data have recognized as powerful and effective one and widely applied in detecting the spatio-temporal dynamics of land-use and land-cover. Using satellite data extracted land use information base on the close relationship between land use and land cover.

4. The impact of urbanization to agricultural lands

The thesis using interdisciplinary approach assesses the impact of urbanization on land use. The interdisciplinary approach is through statistical analysis and multi - dimensional modeling.

1.3.2. Methodologies

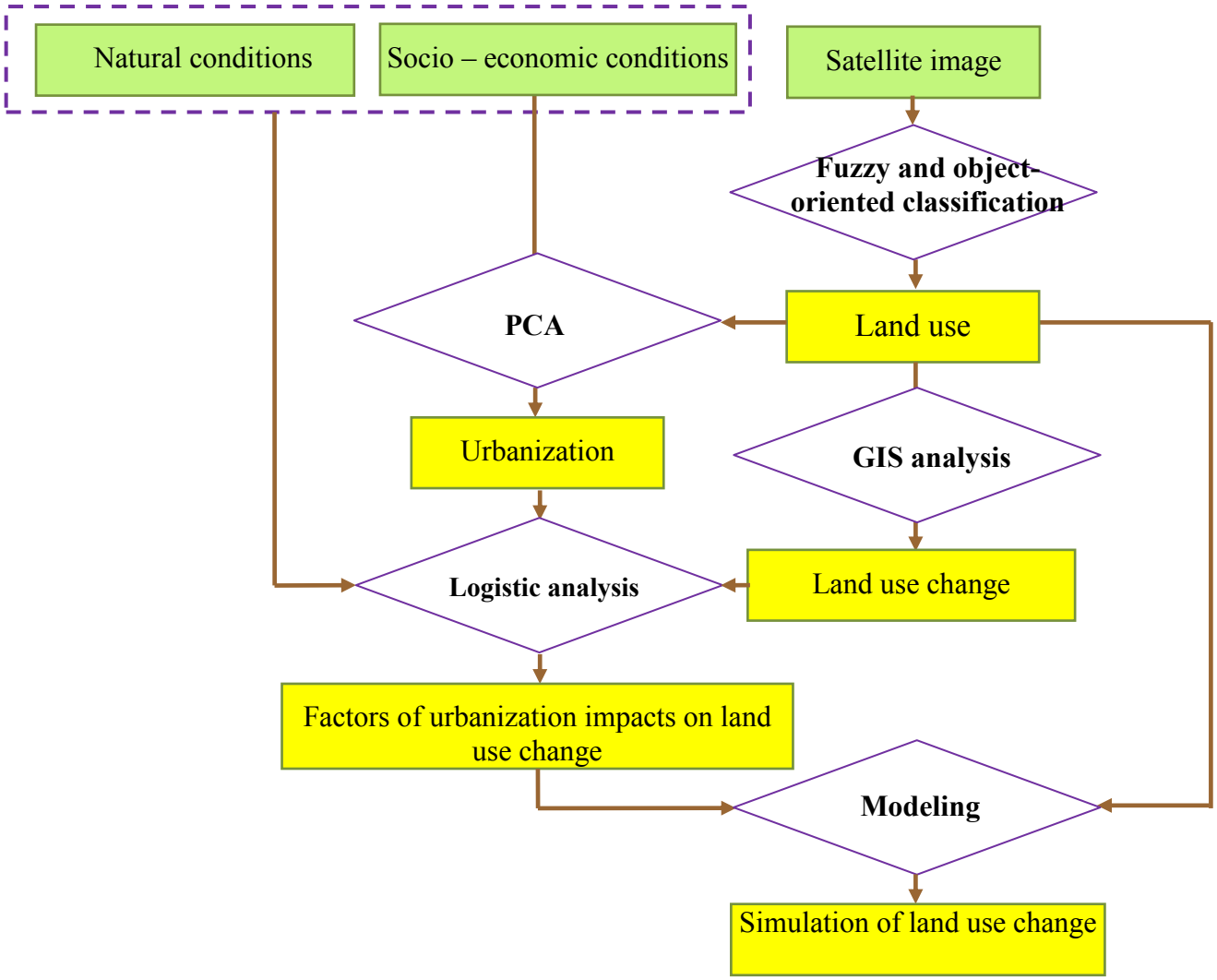


Figure 1. 3. Flow chart and methodology

CHAPTER TWO
OBJECT _ ORIENTED FUZZY CLASSIFICATION IN DONGANH, HANOI

2.1. Description of the study area

2.2. Data acquisition

2.3. Land use categories in Donganh

2.4. Object_ oriented fuzzy classification

2.4.1. Fuzzy logic

Fuzzy Logic is a form of multi-valued logic and the basic idea is to replace the two results "false", "true" of Boolean logic as a sequence of continuous [0, ..., 1] with 0 being the value "false" and 1 being the value of "right" and all values between 0 and 1 will represent the transition between wrong and right [148]. A fuzzy subset G of S is determined by a membership function μ_G, which assigns a membership grade within the interval [0, 1] to each element s. The membership grade can be expressed by:

$$\mu_G : s \rightarrow [0,1] \qquad (2.\ 1)$$

Basic architecture of the fuzzy system includes fuzzification, inference and defuzzification [153, 157] show in figure 2.5.

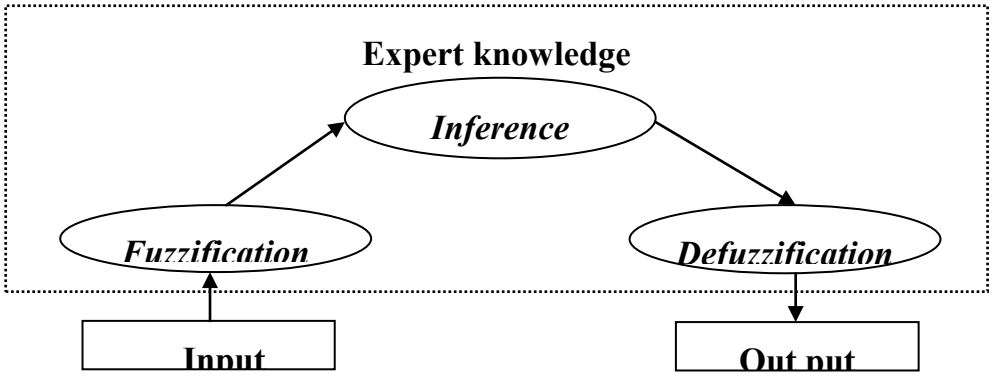


Figure 2. 5. Basic architecture of the fuzzy system

Fuzzification:

The purpose of fuzzification is to partition the feature space into fuzzy subspaces and generate rules for each fuzzy subspace.

To carry out the process of fuzzification, one must first define membership functions.

Inference:

The inference stage computes the strength contributed by the triggered rules and aggregates those triggered rules. The simplest fuzzy rule sets depend on only a« If – Then » function (If-clause as conditional clause and then – clause expressing the consequence clause).

1. The fully function

If all membership grades are equal to one (i.e., the rule condition is fully satisfied) the then clause in the rule should be fully adopted. The function is expressed blow :

If x is A then Y is C

2. The function with strength

If the rule condition is only partially satisfied, the then clause should be partially weighted:

If x is A then Y is C with strength w

Defuzzfication:

The final step, defuzzification, combines all triggered rules and generates a nonfuzzy outcome. The most popular methods of defuzzification are the center-of-gravity and the mean-of maximum methods [153]:

$$center - of - gravity = \frac{\sum_{s=1}^{n} s \times \mu(s)}{\sum_{s=1}^{n} \mu(s)} \qquad (2.\ 9)$$

$$mean - of - maximum = \sum_{s=1}^{m} \frac{s}{m} \qquad (2.\ 10)$$

2.4.2. Process classification

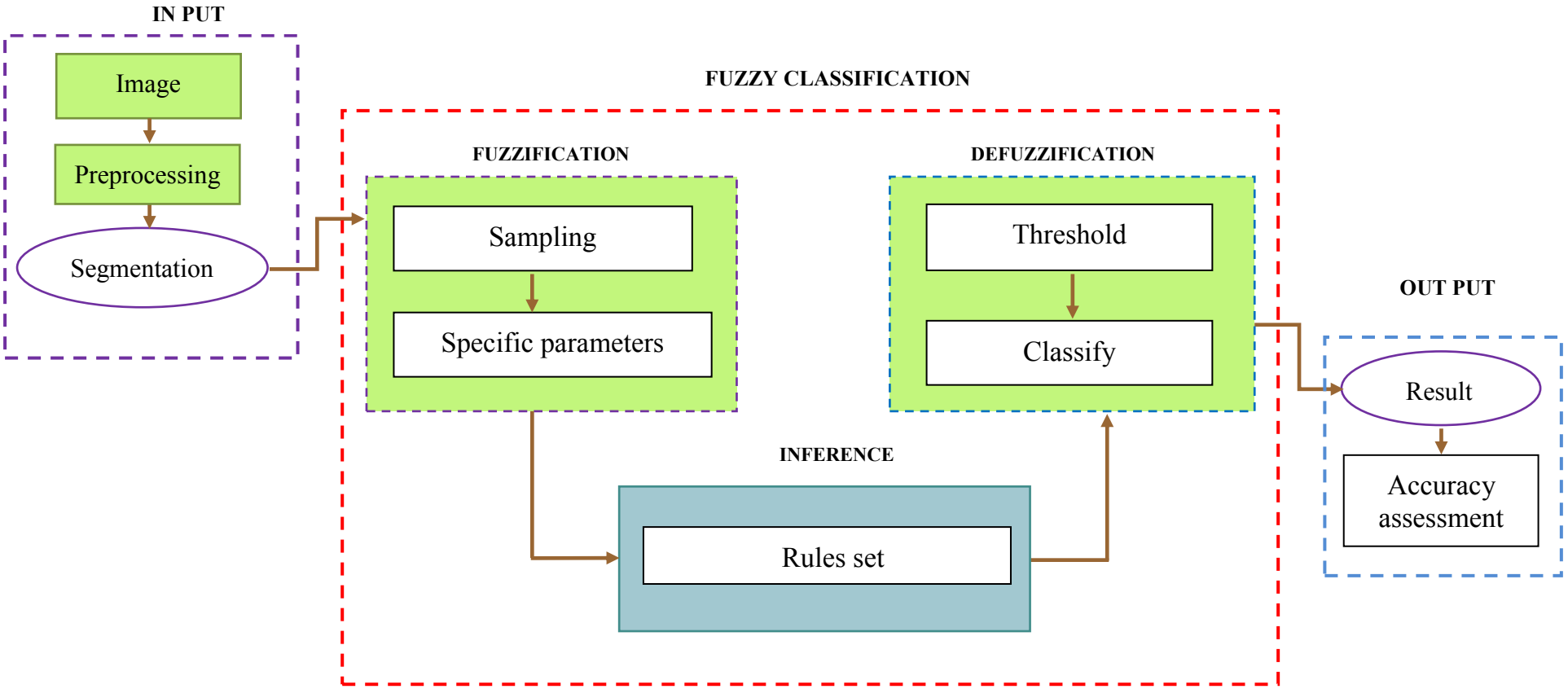


Figure 2.7. The flow chart of image process

CHAPTER THREE
ASSESSING DRIVING FORCES BEHIND STRUCTURE OF AGRICULTURAL LAND USE CHANGE IN DONGANH, HANOI

3.1. Assessing urbanization in Donganh, Hanoi

3.1.1. Criteria for determining urbanization in Donganh, Hanoi

The criteria groups include land use, labor and household income, and population and standard of living.

Table 3.1. The group of land use

Criteria	Categories	Code
Land use	- The density of build up land	XD
	- The density of paddy land	Lua
	- The density of maize land	Mau
	- The density of vegetable and paddy land	Mlua

Table 3. 2. The group of labor and household income

Criteria	Categories	Code
People of laboring age	- The density of unemployment	P_umploy
	- The density of employment in agriculture sector	P_NLTS
	- The density of employment in construction industry sector	P_CNXD
	- The density of employment in service sector	P_TNDV
Household income	- The density of household having major income from agriculture sector	I_NLTS
	- The density of household having major income from construction industry sector	I_CNXD
	- The density of household having major income from service sector	I_TNDV

Table 3.3. The group of population and standard of living

Criteria	Categories	Code
Population	- Population density	PP
Standard of living	- The density of household having a computer	PC
	- The density of household using major source of fresh water	W
	- The density of household using gas fuel for cooking	G

3.1.2. The method of assessing urbanization

Principal component analysis (PCA) is used for assessing the degree of urbanization. The main aim of principal components analysis (PCA) is to replace metrical correlated variables by a much smaller number of uncorrelated variables, which contain most of the information in the original set. PCA finds the new axes with maximum variances where the data is most spread.

3.1.3. Degree of Urbanization in Donganh, Hanoi

The resuls of PCA analysis in 2001, 2006, and 2011 are showed in figures from 3.2 to 3.4

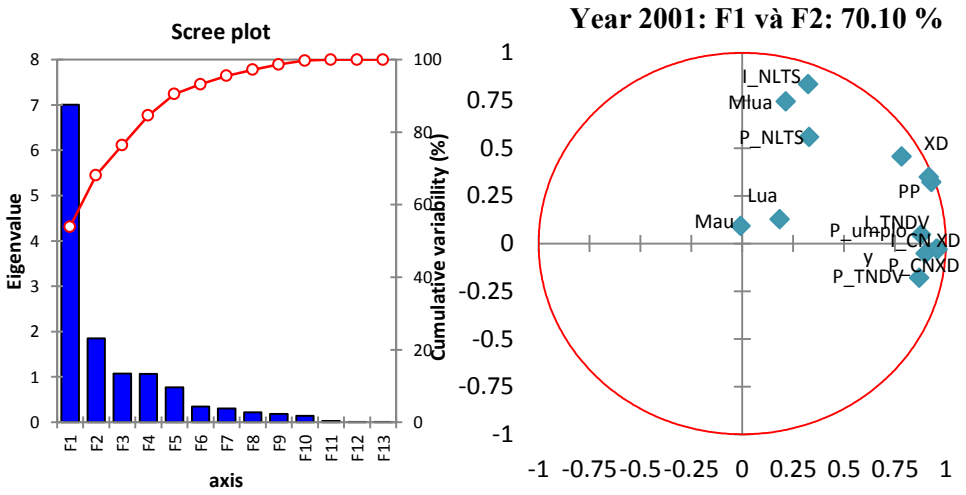


Figure 3. 2. Axes F and

The distribution of criteria in two principal components F1 and F2 in 2001

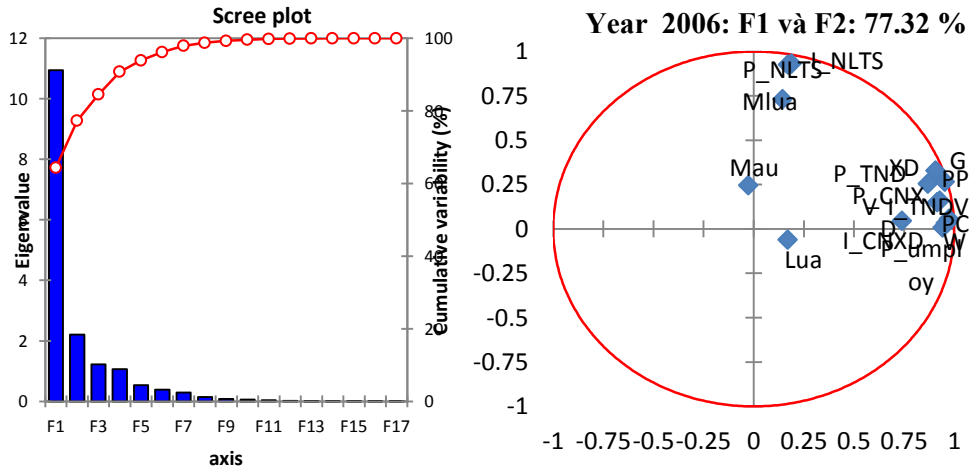


Figure 3. 3. Axes F and

The distribution of criteria in two principal components F1 and F2 in 2006

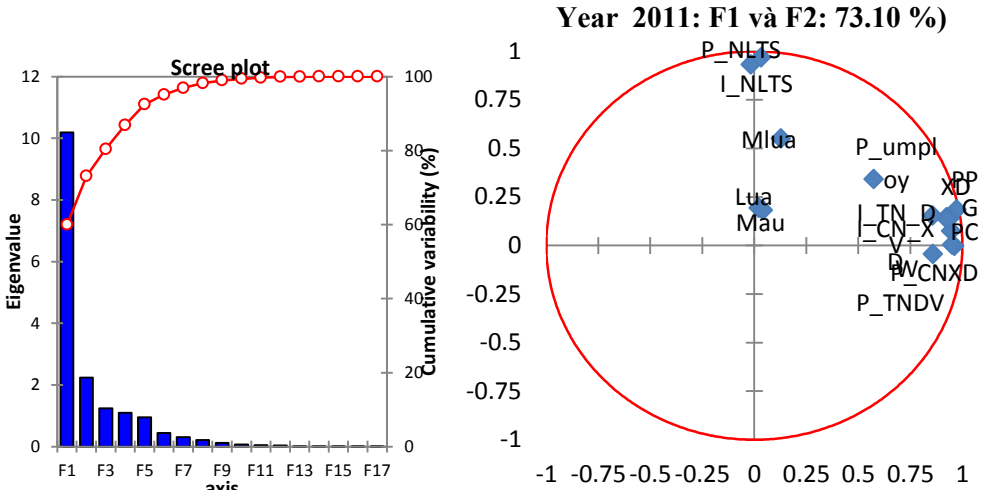


Figure 3. 4. Axes F and

The distribution of criteria in two principal components F1 and F2 in 2011

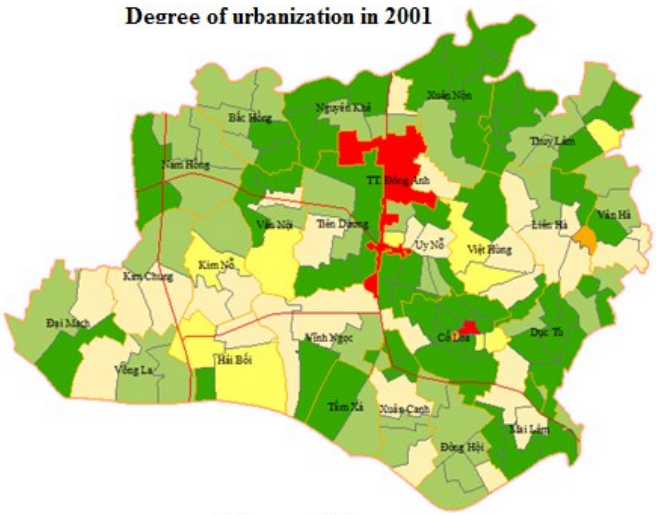


Figure 3.5

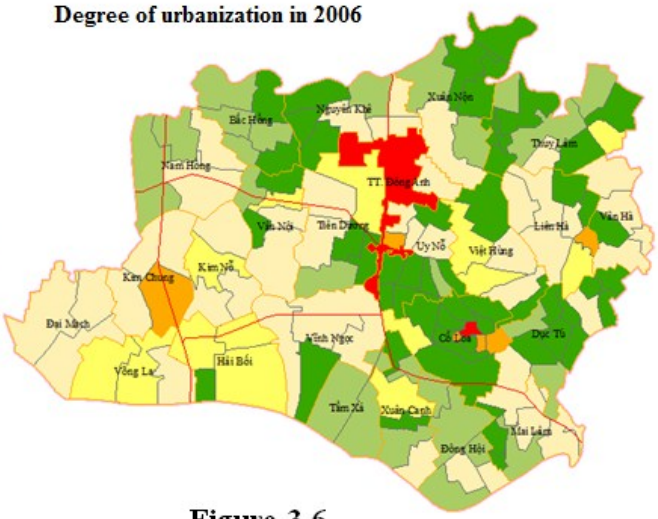


Figure 3.6

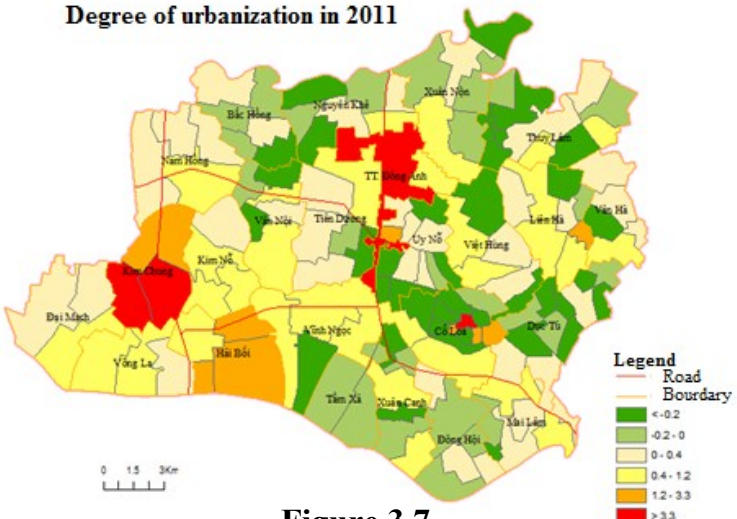


Figure 3.7

• First component (F1): Interpretation of the principal components is base on finding which variables (criteria) are strongly correlated with the components. The first component is strongly with six of the original variables, which are most variables in construction industry sector and service sector. Therefore, F1 expresses the degree of urbanization and is the new criterion represented a set of above criteria.

• Second component (F2): The second component is strongly with criteria in agriculture sector. This suggests that, F2 shows degree of activity in the agricultural field.

Figure 3.5; 3.6 and 3.7 show degree of urbanization of Donganh, Hanoi in maps that base on criteria of F1 and divided into the range of the degree increased according to cold colors to warm colors.

3.2. Agricultural land use change in Donganh, Hanoi

3.2.1. Land use change by the time

• **The 2001- 2006 period**

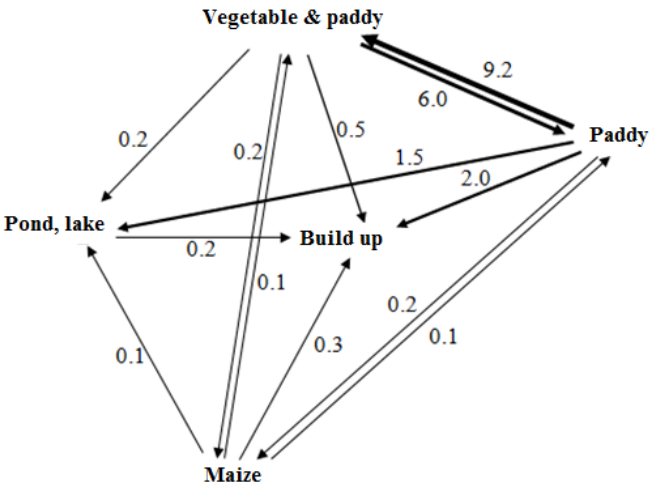


Figure 3.8

• **The 2006 – 2013 period**

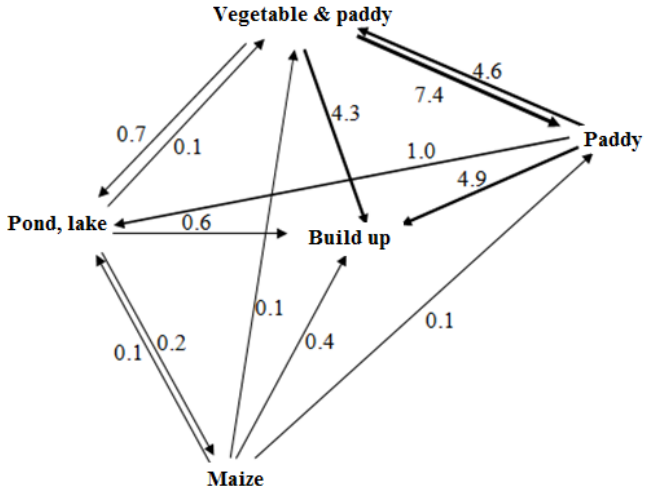


Figure 3.10

Figure 3.8 and 3.10 show major land use changes in 2001-2006 and 2006-2013 periods.

- **The 2001 – 2013 period**

Table 3.8. Land use change in period 2001- 2013

Land use	Year 2001 (ha)	Year 2013 (ha)	Land use change	
			Area (ha)	Percent (%)
Build up	2620.09	5081.08	2460.99	13.3
Paddy	7721.27	5911.00	-1810.27	-9.8
Maize	1373.77	1169.36	-204.41	-1.1
Veg & Pad	4780.93	3807.58	-973.35	-5.3
Pond, lake	1206.00	1671.99	465.99	2.5
Bare soil	80.00	179.38	99.38	0.5
River	716.24	677.92	-38.32	-0.2

3.2.2. Land use change by the dimensions

3.3. Drivers of agricultural land use change in the Donganh, Hanoi

3.3.1. Multinomial logistic regression

Multinomial logistic regression is a simple extension of binary logistic regression [70].

The binary logistic regression with the dependent variable taking only two values, 0 and 1. We have a equation:

$$g(x) = \log\left[\frac{P}{1-P}\right] = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \quad (3.1)$$

The probability of the dependent variable Y being 1 mean that the probability of an event can be estimated with the following logistic regression model:

$$P(Y=1) = \frac{e^{(\alpha+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}}{1+e^{(\alpha+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} = \frac{e^{g(x)}}{1+e^{g(x)}} \quad (3.2)$$

Multinomial logistic regression is a generalization of the binomial to more than two categories of the dependent or outcome variable. Suppose that there are k categories of outcome variable Y, denoted by 0,1,2,...,k. Therefore, there will been k-1 the regression equations. The regression of

each categories compared with the reference category, and the first category of the nominal variable is selected to be the reference category.

Using a random sampling approach with 4984 sample points in the entire study area.

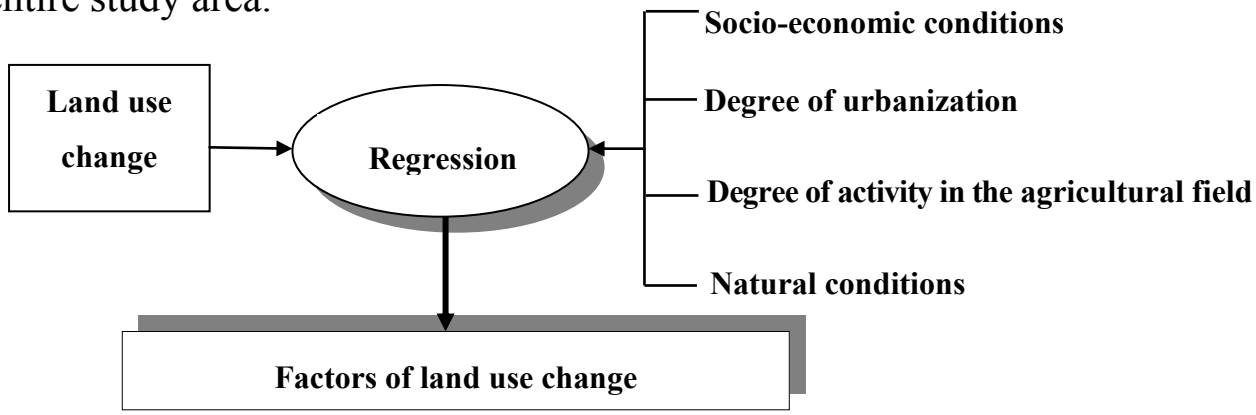


Figure 3.15. Determining factors of agricultural land use changes

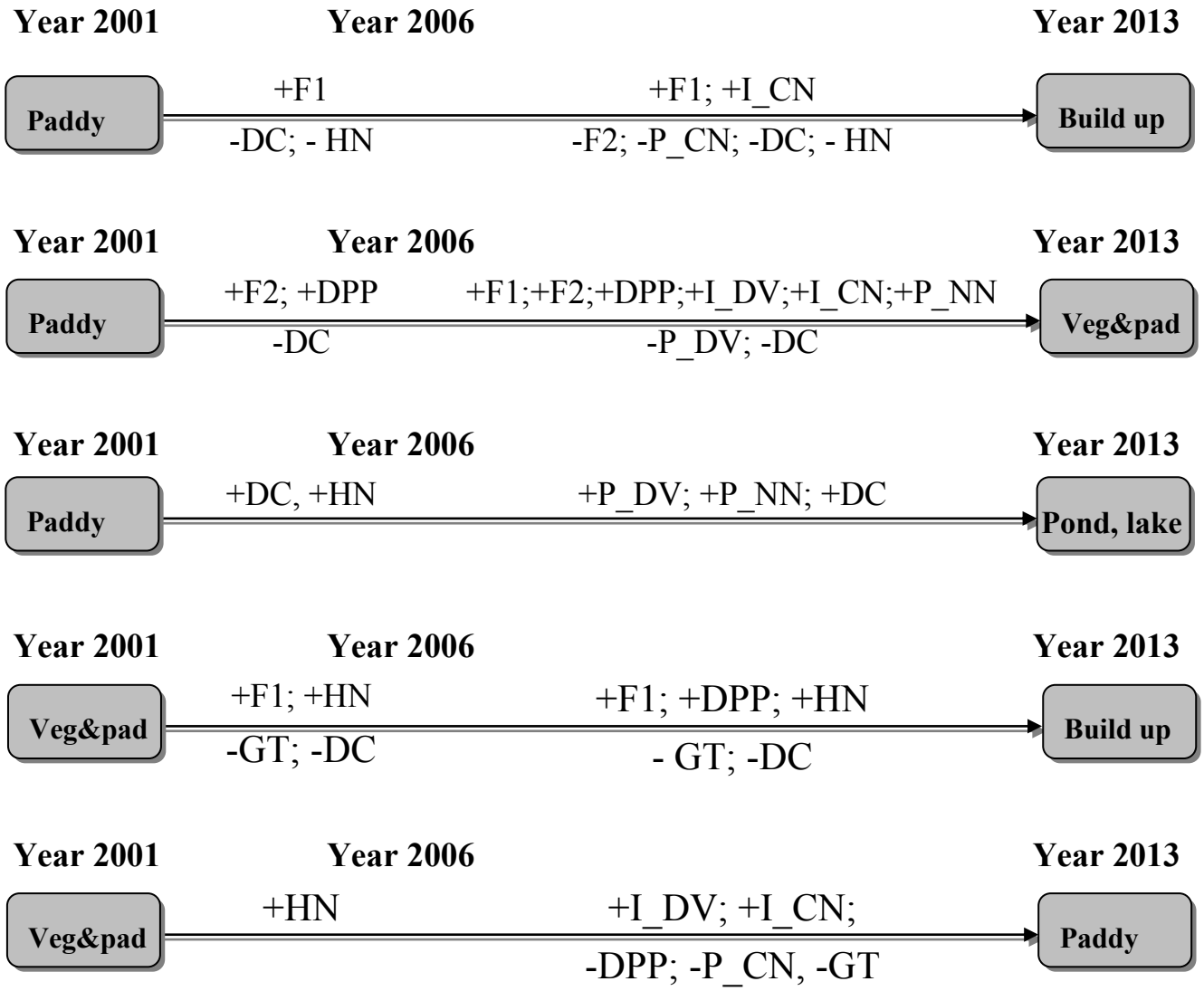
3.3.2. GIS Database of model

Table 3.13. List of variables

Variables		**Type**
Dependent variables		
None change		0
Conversion from paddy to build up land		1
Conversion from paddy to vegetable and paddy land		2
Conversion from paddy to pound land		3
Conversion from vegetable and paddy to build up land		4
Conversion from vegetable and paddy to paddy land		5
Independent variables		
Degree of urbanization (F1)		Discrete
Degree of activity in the agricultural field (F2)		Discrete
Natural conditions	Distance to Hanoi	Continuous
	Distance to major roads	Continuous
	Distance to residential areas	Continuous
Socio-	Labor sector structure	Discrete

Variables		Type
economic conditions	Income structure	Discrete
	Population density	Discrete

3.3.3. Drivers of agricultural land use changes in the Donganh, Hanoi



The results show that the natural and social– economic conditions related conversions of land use in the study area in more than 10 years. "Distance to residential areas" (DC) was an important natural factor, it influenced to almost all conversions of land use. The explanatory variables "Degree of urbanization" (F1) and "Population density" (DPP) were significant for land use conversions. The estimated coefficients for

the explanatory variable "Degree of urbanization" (F1) are positive that control the spatial pattern of agricultural land and non-agricultural land conversion for during the time. The independent variables "Population density" (DPP) were controlling positive factors in all time periods, excluding that the regression coefficient was negative from vegetable and paddy to paddy land.

3.4. Spatial modelling of agricultural land use changes

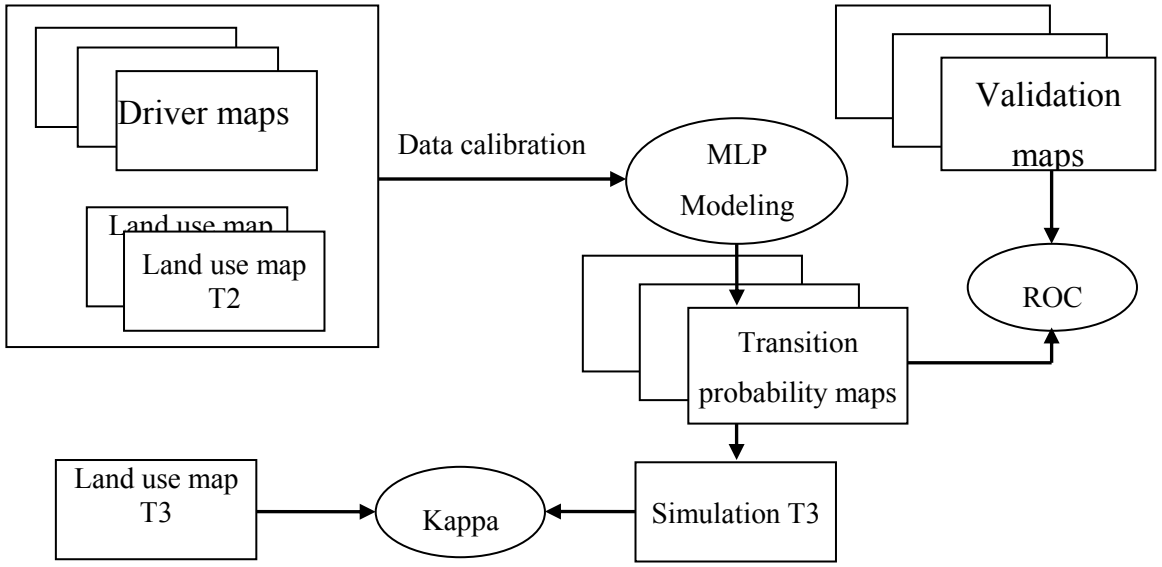
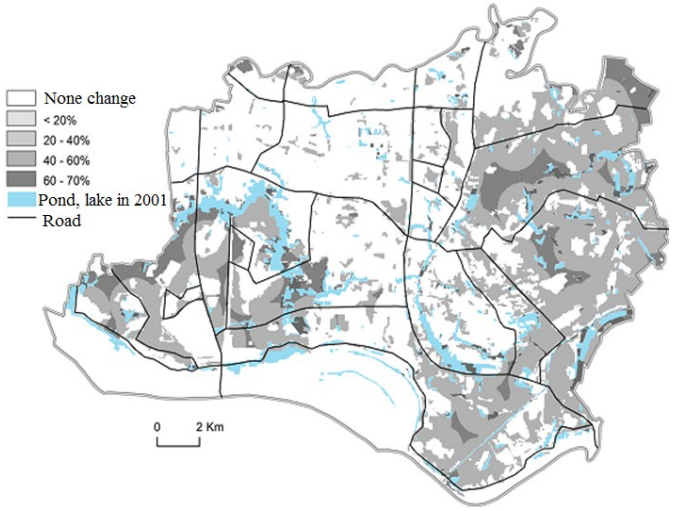


Figure 3.18. Spatial modeling of agricultural land use change

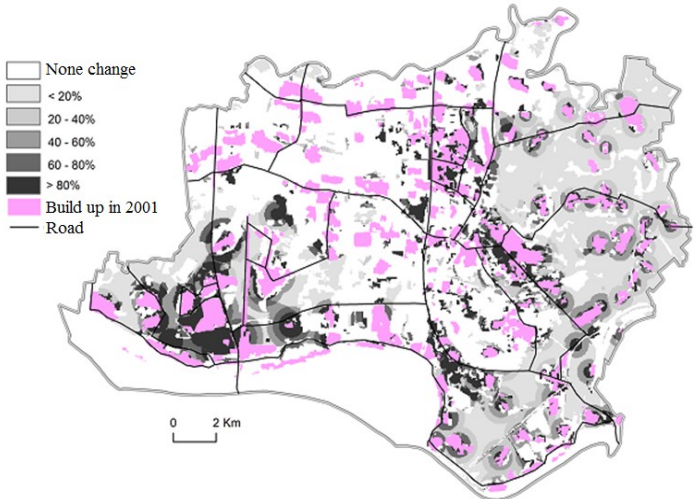
3.4.1. In put data calibration for modeling in two periods 2001-2006 and 2006-2013

3.4.2. Probability maps of the conversions

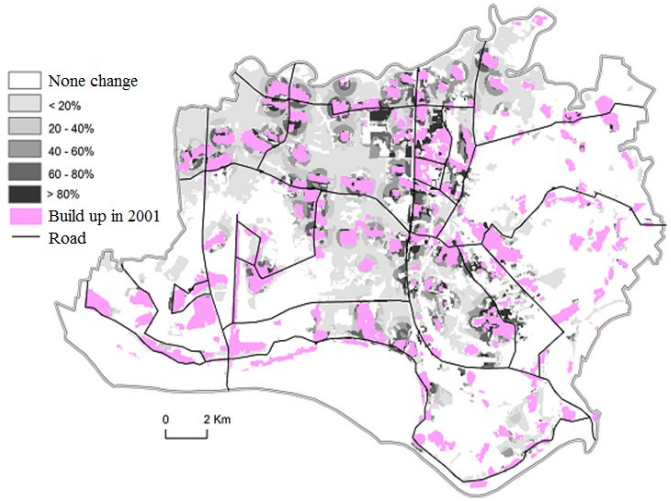
The MLP model were used to calculate for each pixel a transition probability for all conversions. MLP is a feedforward artificial neural network model. A MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. MLP utilizes a supervised learning technique called backpropagation for training the network. There will be a paticular MLP model for each of conversions.



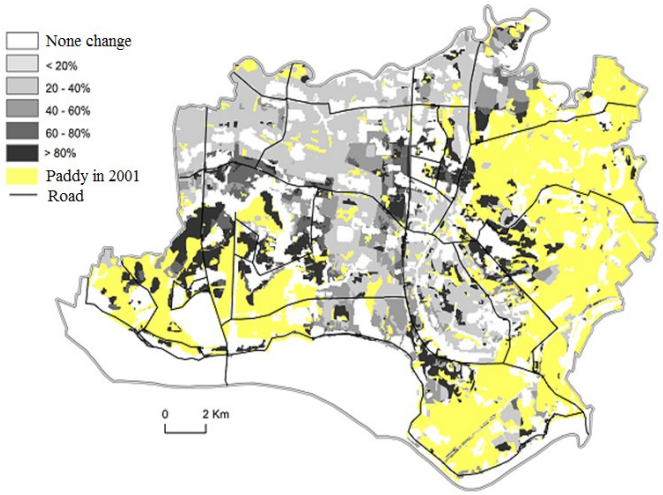
Transition probability map of the conversion from paddy to pond, lake land



Transition probability map of the conversion from paddy to build up land



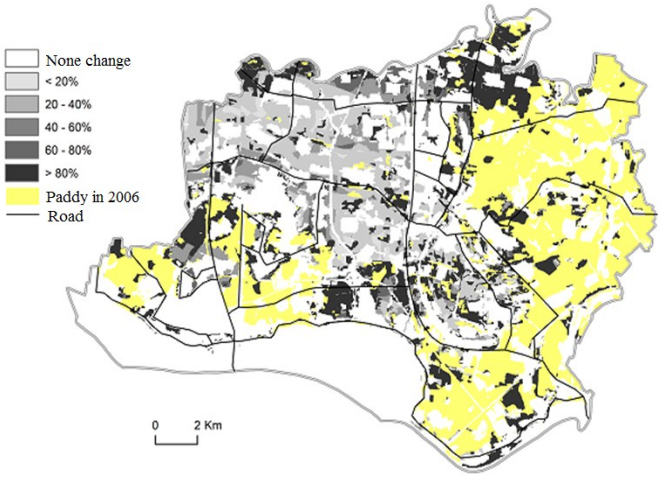
Transition probability map of the conversion from vegetable & paddy to build up land



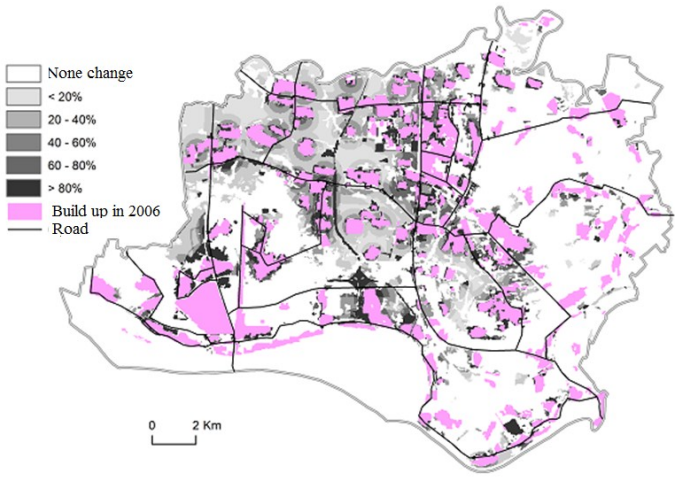
Transition probability map of the conversion from vegetable & paddy to paddy land

Transition probability map of the conversion from paddy to vegetable & paddy land

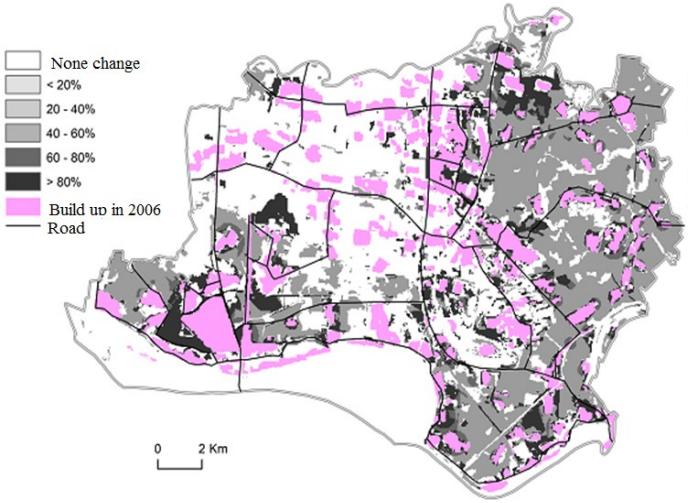
Figure 3. 20. Transition probability maps of the conversions after 2001



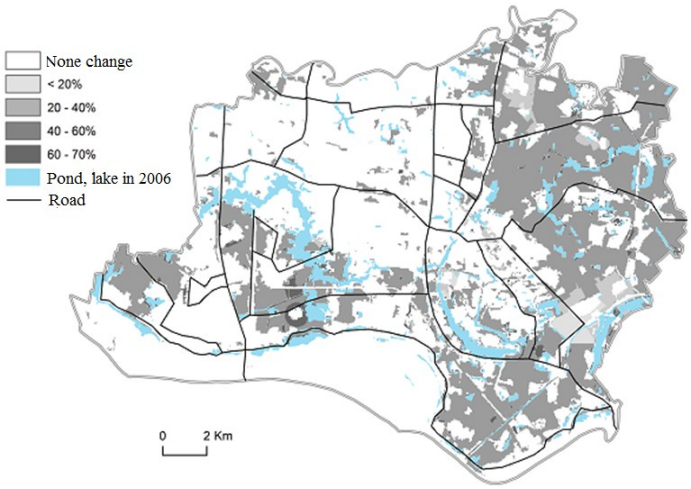
Transition probability map of the conversion from vegetable & paddy to paddy land



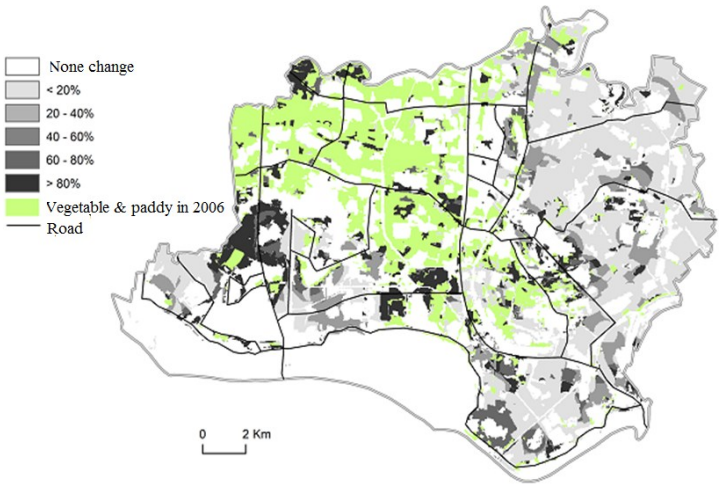
Transition probability map of the conversion from vegetable & paddy to build up land



Transition probability map of the conversion from paddy to build up land



Transition probability map of the conversion from paddy to pond, lake land



Transition probability map of the conversion from paddy to vegetable & paddy land

Figure 3. 21. Transition probability maps of the conversions after 2006

3.4.3. Validation

The validation techniques that allow to measure the model performance of land use change models. The two validation techniques are kappa analysis [134] and ROC curve [56].

The ROC curve was adroped to validate all conversions, for each conversion types will have a corresponding ROC curves. In the 2001 - 2006 period, the values of area under the curves (AUC) are determined in the 0.506 - 0.737 range. And in the 2006 - 2013 period, the value range of area under the curves (AUC) is from 0.529 to 0.730.

The kappa coefficient was estimated to be 0.727 expressed the model performs relatively well.

CONCLUSIONS

Based on theory and research results, the research results have been drawn the conclusions and recommendations of the following:

Conclusion

1. Extracting land use information: using object-oriented fuzzy classification by comparation between membership functions of samples to determine relevant paramaters for each land use categories with overlap as 0 that results high accuracy.
2. Assessment of degree of urbanization in study area: the degree of urbanization in 2001, 2006, and 2011 were detemined by using multi-criteria PCA analysis with the group of natural and economic criteria and the group of land use.
3. Agricultural land use changes: this thesis examines the five major arable land conversions during the period, which are the conversion from paddy to build up land, paddy - vegetable and paddy land conversions in both directions, the coversion from paddy to pond, lake land and the coversion from vegetable and paddy to build up land.

4. The statistical methods and spatial modeling have quantified the impact of urbanization on land use changes. The study exploring the major driving forces of agricultural land use conversions in the 2001-2013 expressed that the natural and socio – economic factors induce land use conversions in both promoted way and limited way.

Recommendations

1. Quanlifying the impact of land use change on urbanization.
2. Using object –oriented fuzzy classification for extracting land use in peri urban and intergrating statistical analysis and spatial modeling assess the interaction between natural and social econmic conditions and land coversions.

LIST OF PUBLICATIONS RELATED TO THE THESIS

1. Trinh Thi Hoai Thu (2009), "Monitoring changes in land use", 7th FIG Regional Conference, Hanoi Vietnam.
2. Vo Chi My, Trinh Thi Hoai Thu, Pham Thi Lan, Le Thi Thu Ha (2010), "Using vegetation index in monitoring exploited land cover", Proceedings of International Mining Conference on "Advanced Mining for Sustainable Development" Halong Vietnam, pp 250 – 246.
3. Trinh Thi Hoai Thu, Le Thi Thu Ha, Pham Thi Lan (2012), Comparision pixel based classification method and objectoriented method to extract land cover from high resolution, *Journal of Engineering Science Mining - Geology, 39, pp 59-65.*
4. Trinh Thi Hoai Thu (2013), "Object-oriented fuzzy approach for land use/ land cover classification in Donganh", *Journal of Geodesy and Cartography,* 30 (4) PC, *pp* 545 -554.
5. Trinh Thi Hoai Thu, Pham Thi Lan, Tong Huyen Ai (2013), "Rule set of object-oriented classification using Landsat imagery in Donganh, Hanoi, Vietnam", *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography,* Vol. 31, No. 6-2, pp 521-527.
6. Trinh Thi Hoai Thu, Le Thi Thanh Huong (2014), "Predicting land use change in Donganh, Hanoi", *Journal of Geodesy and Cartography,* 21, pp 60-64.