Modeling Urban Growth in Southern Thailand

Potjamas Chuangchang

A Thesis Submitted in Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Research Methodology

Prince of Songkla University

2016

Thesis Title Modeling Urban Growth in Southern Thailand

Author Miss Potjamas Chuangchang

Major Program Research Methodology

Major Advisor	**Examining Committee:**
....	....Chairperson
(Asst. Prof. Dr. Phattrawan Tongkumchum)	(Asst. Prof. Dr. Chokchai Lueangthuwapranit)
	....
	(Asst. Prof. Dr. Phattrawan Tongkumchum)
Co-advisors	
....	....
(Asst. Prof. Dr. Chamnein Choonpradub)	(Asst. Prof. Dr. Chamnein Choonpradub)
....	....
(Emeritus Prof. Dr. Don McNeil)	(Emeritus Prof. Dr. Don McNeil)
	....
	(Prof. Dr. Khairil Anwar Notodiputro)

The Graduate School, Prince of Songkla University, has approved this thesis as fulfillment of the requirements for the Doctor of Philosophy Degree in Research Methodology.

....

(Assoc. Prof. Dr. Teerapol Srichana)

Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgements have been made of any assistance received.

…….………………………………. Signature

(Assist. Prof. Dr. Phattrawan Tongkumchum)

Major Advisor

…….………………………………. Signature

(Miss Potjamas Chuangchang)

Candidate

I hereby certify that this work has not been accepted in substance for any degree, and is not being currently submitted in candidature for any degree.

..………………………………. Signature

(Miss Potjamas Chuangchang)

Candidate

ชื่อวิทยานิพนธ์ โมเดลการพัฒนาเป็นเมืองในภาคใต้ของไทย

ผู้เขียน นางสาวพจมาศ เชื่องช้าง

สาขาวิชา วิธีวิทยาการวิจัย

ปีการศึกษา 2558

บทคัดย่อ

การจัดการข้อมูลการใช้ประโยชน์ที่ดินมีสำคัญที่จะต้องใช้วิธีการที่เหมาะสม วิทยานิพนธ์นี้ศึกษาเกี่ยวกับวิธีการใหม่ที่มีความสำคัญเนื่องจากสามารถบูรณาการกับข้อมูลการใช้ประโยชน์ที่ดินจากการสำรวจระยะไกลจากดาวเทียมได้ วิธีการดังกล่าวอาศัยการเปลี่ยนโครงสร้างของข้อมูลจากอนาลอคเป็นดิจิตอล ทำให้สามารถนำข้อมูลไปวิเคราะห์โดยใช้วิธีการทางสถิติเพื่อแสดงการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินระหว่างสองช่วงเวลาได้ โดยการใช้ประโยชน์ที่ดินในปีที่สนใจเป็นตัวแปรตาม และการใช้ประโยชน์ที่ดินในปีก่อนหน้าเป็นตัวแปรอิสระ กรณีที่ตัวแปรตามเป็นแบบกลุ่มที่มีสองระดับ สามารถใช้ตัวแบบการการถดถอยแบบลอจิสติกวิเคราะห์ข้อมูลได้ แต่เนื่องจากข้อมูลแต่ละกริดไม่เป็นอิสระต่อกันจึงใช้วิธี Variance inflation factor จัดการความสัมพันธ์ของข้อมูล วิทยานิพนธ์เล่มนี้แบ่งการศึกษาออกเป็นสามตอนดังนี้

ตอนที่หนึ่งมีวัตถุประสงค์เพื่ออธิบายการเปลี่ยนโครงสร้างของข้อมูลจากอนาลอคเป็นดิจิตอล โดยจากเดิมบันทึกข้อมูลเป็นโพลิกอนเปลี่ยนเป็นบันทึกแบบกริดที่มีขนาด 1 เฮกตาร์และใช้ข้อมูลการใช้ประโยชน์ที่ดินของจังหวัดภูเก็ตจากปี พ.ศ. 2510-2552 กรมพัฒนาที่ดินแบ่งการใช้ประโยชน์ที่ดินเป็น 5 ประเภท ได้แก่ พื้นที่ป่าไม้ พื้นที่เกษตรกรรม พื้นที่ชุมชนและสิ่งปลูกสร้าง พื้นที่น้ำ และพื้นที่เบ็ดเตล็ด

ในที่นี้จัดประเภทการใช้ประโยชน์ที่ดินเป็น 2 กลุ่ม ได้แก่พื้นที่ชุมชนและสิ่งปลูกสร้าง ซึ่งหมายถึงพื้นที่เมือง และพื้นที่อื่นๆ จากนั้นจึงสร้างตัวแบบการถดถอยแบบลอจิสติกของความเป็นเมืองโดยมีการใช้ประโยชน์ที่ดินในปีก่อนหน้าที่แยกเป็นพื้นที่ทางตอนเหนือหรือใต้เป็นตัวแปรอิสระ โดยใช้วิธี variance inflation factors โดยคำนวณค่าคลาดเคลื่อนมาตราฐานของสัดส่วนความเป็นเมือง เนื่องจากข้อมูลมีความสัมพันธ์กันในแต่ละกริด ผลการศึกษาพบว่าการเปลี่ยนโครงสร้างของข้อมูลจากอนาลอคเป็นดิจิตอลมีความเหมาะสมสำหรับการวิเคราะห์การเปลี่ยนแปลงการใช้ประโยชน์ที่ดินด้วยการสร้างตัวแบบ และการประยุกต์ใช้วิธีการดังกล่าวกับข้อมูลการใช้ประโยชน์ที่ดินในจังหวัดภูเก็ตพบความเปลี่ยนแปลงความเป็นเมืองส่วนใหญ่เกิดขึ้นในทางตอนใต้ของภูเก็ต

ตอนที่สองศึกษาความเป็นเมืองในพื้นที่บริเวณทางหลวงแผ่นดินหมายเลข 4 ระหว่างจังหวัดพัทลุงถึงหาดใหญ่จากปี พ.ศ. 2534-2552 โดยแบ่งพื้นที่ศึกษาออกเป็นพื้นที่ทางตอนเหนือ ทางตอนกลาง และทางตอนใต้ โดยประยุกต์ใช้วิธีการเปลี่ยนโครงสร้างของข้อมูลจากอนาลอคเป็นดิจิตอล และการวิเคราะห์การถดถอยแบบลอจิสติกของความเป็นเมือง โดยตัวแปรอิสระคือการใช้ประโยชน์ที่ดินกับพื้นที่ที่ศึกษา ผลการศึกษาพบว่าพื้นที่ความเป็นเมืองเฉลี่ยเพิ่มขึ้นโดย เกิดขึ้นร้อยละ 3 ในปี พ.ศ. 2543 และร้อยละ 5 ในปี พ.ศ. 2552 โดยความเป็นเมืองส่วนใหญ่เกิดขึ้นทางตอนเหนือ

ตอนที่สามเป็นการสร้างตัวแบบความเป็นเมืองใน 17 ตำบลของจังหวัดภูเก็ตในปี พ.ศ. 2543-2552 โดยใช้การวิเคราะห์การถดถอยแบบลอจิสติก ที่มีการใช้ประโยชน์ที่ดินในปี พ.ศ. 2543 และตำบลเป็นตัวแปรอิสระ และใช้พื้นที่ใต้โค้ง ROC เปรียบเทียบกับตัวแบบที่มีเฉพาะตัวแปรการใช้ประโยชน์ที่ดินในปี พ.ศ. 2543 ผลการศึกษาพบว่า พื้นที่ของความเป็นเมืองเพิ่มขึ้น 4,557 เฮคแตร์ โดยส่วนใหญ่เปลี่ยน

มาจากพื้นที่เกษตรกรรม และพื้นที่ความเป็นเมืองส่วนใหญ่พบในตำบลฉลองและตำบลตลาดเหนือ ตัวแบบที่มีตัวแปรการใช้ประโยชน์ที่ดินในปี พ.ศ. 2543 และตำบลให้ค่าพื้นที่ใต้โค้ง ROC เท่ากับ 0.83 แสดงว่าตัวแบบมีความเหมาะสม

โดยสรุปการศึกษาครั้งนี้ให้ข้อมูลเกี่ยวกับแนวโน้มของการเปลี่ยนแปลงใช้ประโยชน์ที่ดินสำหรับนักวางแผนและนักพัฒนาในการจัดการการใช้ประโยชน์ที่ดินในอนาคต วิธีการเปลี่ยนโครงสร้างของข้อมูลจากอนาลอคเป็นดิจิตอลเหมาะสำหรับการพัฒนาตัวแบบทางสถิติเพื่ออธิบายการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินระหว่างสองช่วงเวลาโดยใช้โปรแกรม R ซึ่งเป็นฟรีซอฟแวร์

Thesis Title Modeling Urban Growth in Southern Thailand

Author Miss Potjamas Chuangchang

Major Program Research Methodology

Academic Year 2015

ABSTRACT

It is important to have a method that makes appropriate use of land-use data. This thesis focuses on methods, which we believe are new and important, particularly because they can be integrated with remote sensing data on land cover from Earth-orbiting satellites. The method is basically analog-to-digital conversion. The method allows us to analyze land-use development and to show how land-use has changed between two time periods using statistical methods. The land-use of each grid at the later time point is an outcome and the land-use at the previous time point is determinant. Logistic regression model was used when outcome is binary (developed land/other). Since data at neighboring grid points are not independent. Variance inflation factor was used to handling spatial correlation. The model based on weighted sum contrasts was used.

The first part aims to propose method for predicting land-use change. This method is based on an analog-to-digital conversion which replaces polygonal shapes by coded grid points. The method is applied to data from a survey of Phuket province from 1967-2009 where land-use was classified broadly as forest, agriculture, urban, water bodies and miscellaneous land. Logistic regression was used to predict a binary land-use outcome (urban/other), and location combined with land-use at a previous survey was a determinant. To account for correlation in land-use amongst nearby plots of land, variance inflation factors were used to compute standard errors of proportions of urban growth. This study shows that analog-to-digital conversion methods are useful approaches to develop appropriate statistical models for land-use change. The greater urbanization was observed in the southern parts of Phuket during the period of study.

The second part aims to investigate the change of developed land in three different locations along Highway 4 Road from Phattalung to HatYai. The method involves creating a digitized grid of geographical coordinates covering the study area. The land-use codes and plot identifiers were recorded in database tables indexed by grid coordinates. Logistic regression of land development adjusted for spatial correlation was used to model its change over a 9-year period using land-use at the previous survey combined with location as a determinant. The results show increasing average percentages of developed land (3% in 2000 and 5% in 2009). Land development occurred mostly in the northern location along the Pattalung to HatYai road.

The third part aims to model land development in the 17 sub-districts in Phuket province of Thailand from 2000 to 2009. Logistic regression was used to monitor changes in land-use over this period and predict future changes. The ROC curve was used to measure the performance of the model. Land-use from a previous survey in 2000 and sub-district identity were included as determinants. The area of developed land increased by 4,557 ha over the study period. Agricultural land in 2000 was more likely to become developed in

2009 and developed land was more likely to remain developed land in 2009. Land development occurred mostly in Chalong and Talat Nua sub-districts. The area under the ROC curve was 0.83 indicating a reasonably good fit of the model.

In conclusion, this study provides useful information on the trend of land-use development for government planners and developers to manage land-use change in the future. The methods are useful approaches to develop an appropriate statistical model for land-use change between two periods using freely available software R program.

Acknowledgements

I would like to express my sincere appreciation to Asst. Prof. Dr. Phattrawan Tongkumchum, my supervisor, whose guidance, support, motivation, vast experience, and untiring help during the course of my PhD thesis. Without her guidance, endless patience and encouragement, this thesis would not have been possible to accomplish. Besides, I have also acquired valuable insights through her instructions, not only in academic studies but also enthusiasm and vigor in life.

My sincere thanks also go to my co-supervisors, Prof. Dr. Don McNeil and Asst. Prof. Dr. Chamnein Choonpradub, for their guidance, encouragement, support and comments. They also giving me so many wonderful opportunities and helped me gain confidence and orientation in the right direction which enabled me to successfully complete this thesis.

I would like to acknowledge the help and support to all the people of the Department of mathematics and computer science, starting from Asst. Prof. Dr. Apiradee Lim, Asst. Prof. Dr. Nittaya McNeil, Asst. Prof. Dr. Metta Kuning and all Ph.D. students.

I would like to thank Prof. Dr. Alfrado Huete for giving me the chance to start my experience at the Plant Functional Biology and Climate Change Cluster, School of the Environment of the University of Technology, Sydney (UTS). My appreciation to Dr. Hilary Green for giving me the opportunity and took care of me when I did the research in Sydney. I have very fond memories of my time there. Thank go to Greig Rundle for English correction.

I would like to extend my gratitude to Prince of Songkla University for granting me the scholarship to complete my PhD study and bring this research a success.

Last of all, I would like to convey my special thanks to my lovely parents and my supportive sister for all their encouragement and support and to all others who directly or indirectly contributed to the success of the study.

Potjamas Chuangchang

Contents

Contents (cont.)

Contents (cont.)

List of Tables

List of Figures

List of Figures (cont.)

Page

List of Figures (cont.)

Chapter 1

Introduction

This thesis investigates and compares scientific methods for analyzing and graphing land-use data collected from regular surveys undertaken by government authorities and for forecasting land-use changes over time. It uses samples collected by the Department of Land Development in Thailand from Phuket, Phattalung and Songkhla provinces to identify patterns of change in developed land in these provinces. The analysis focuses on urban growth. Urban growth is generally considered to be a land-use change from a non-urban category to urban category.

1.1 Background and rationale

Many scientific studies, particularly in environmental science and management, are concerned about land-use (Zeng *et al*. 2008; Kamyab and Mahiny, 2013; Achmad *et al*., 2015). Land development and urban growth has resulted in land-use change in many areas, especially in developing countries where urbanization rates are high, impacting on the environment, the social structure and the economy of the region. Recent studies have shown that urbanized land tends to replace either agricultural land or forest land (Alsaaideh *et al*., 2012); Alsharif and Pradhan, 2014; Jiang *et al*., 2015).

Several studies in recent decades have developed new and improved models of changes in land-use based on remote sensing (RS) and geographical information systems (GIS) to explain and forecast future development (Sudhira *et al*., 2004; Liu and Zhou, 2005; Huang *et al*., 2008; Eyoh *et al*., 2012). It shows scientific knowledge of land-use change is well established in research areas of science including geography, computer science, image processing, and remote sensing, whereas statistical methodologies, in particular time series modeling, have not yet been applied to the same extent.

Statistical models of land-use change and urban growth have become important tools for city planners, economists, ecologists and resource managers, facilitates timely and effective action for sustainable growth of urban regions (Herold *et al.*, 2001). These models can provide insight into the dynamics of the urbanization system and can be used to forecast future development trends. Different modeling methods have been adopted in studies of land-use change from the perspective of their utility for predicting changes in land-use intensification (Lambin *et al.*, 2000). Logistic regression has been used to model urban growth (Allen and Lu, 2003; Cheng and Masser, 2003; Hu and Lo, 2007; Nong and Du, 2011; Eyoh *et al.*, 2012; Alsharif and Pradhan, 2014; Tayyebi *et al.*, 2014). However, these studies do not have predictors comprising land-use type in a previous year, and spatial correlation, which causes violation of the independent assumption of errors, is difficult to handle (Hu and Lo, 2007). Moreover, conventional survey and mapping techniques used in land-use change studies are expensive and time consuming for the expansion of developed land. Such information is not readily available in most developing countries.

1.2 Methodology for Land-Use Research

Land-use research methodology is relatively new and rapidly developing, due to the explosion of data available from remote sensing, the continuing rapid increase in computational power, and the widespread availability of free and open-source geographical information systems software. Before outlining the methodology we use, it is instructive to briefly outline the historical development of scientific research methodology from which it has arisen.

Until 400 years ago scientific development was largely influenced by Aristotle, who lived in the 4th century BC and argued that one could deduce all the laws governing the universe by pure thought. In this philosophy there was no place for scientific experiments, and, largely due to belief in revelation by religious authorities, it held sway in Europe for nearly 2000 years. During this long period many scientific fallacies prevailed, such as the idea that the Earth was the centre of the universe and that in the absence of air resistance a heavier body fell to ground faster than a lighter one.

The scholar now credited with countering the Aristotelian deductive science was Francis Bacon (1561-1626), who is regarded as the father of scientific research methodology. He disparaged Aristotle's deductive logic by contrasting a spider with a bee, arguing that the bee is superior because it pollinates flowers and takes natural products to produce nutritious honey, whereas the predator spider draws its web from its own entrails. Bacon thus pioneered what is now regarded as the "scientific method" based on induction from empirical data as well as purely deductive logic.

A more recent pioneer in the development of scientific research methodology is Karl Popper (1902-1994), who developed the idea of evidence-based scientific decision-making. In this method, a scientific theory can never be proven correct. It can only be proven incorrect. Popper supported the idea of falsification, where a hypothesis must be potentially disprovable for it to be regarded as scientific, and believed that theological, epistemological and metaphysical questions were not falsifiable, and therefore not scientific. Popper firmly believed that such concepts should not be given credence by science, saying that how the hypothesis was derived was an unimportant part of the scientific method. He believed that the only important issue was that the hypothesis was testable and falsifiable. He felt that an alternative to proving theories, scientists should instead try to falsify them, a belief still held by many scientists.

Popper's approach finds a parallel in statistical hypothesis testing, where a sufficiently small *p-value* validly constructed using data from a scientific experiment justifies a decision to reject the hypothesis.

P-values were developed by Ronald Fisher (1890-1962), the geneticist and statistician who developed modern statistical science. In this Fisherian formulation, a p-value is the probability of obtaining an effect at least as extreme as that observed, assuming the validity of a *null* hypothesis (the hypothesis that a specified scientific claim is not true).

P-values prevailed for most of the 20^{th} century, almost to the point that a sufficiently small p-value summarizing the result of a scientific study was a necessary requirement for publication in many journals (leading to a proliferation of errors known as *publication bias*). In his book Exploratory Data Analysis (1977), John Tukey (1915-2000) argued for p-values to be buttressed by appropriate graphical summaries of data, including box plots and confidence intervals. He described *data analysis* as a name for what applied statisticians did, differentiating this term from formal statistical inference. He argued that Data Analysis was a new scientific discipline, not just a branch of Mathematics. Although he was a Professor of

Mathematics at Princeton University, he argued that Mathematics is not strictly a scientific discipline, because its ultimate standard of validity is agreed-upon logical consistency and provability, rather than reliance on the test of experience as the ultimate standard of validity.

Tukey is now regarded by many as the pioneer of a new and rapidly developing field of science known as Data Science, described by Donoho (2015) as "learning from data", and encompassing much of the theory and methods used by statisticians as well as machine learning concepts.

The scientific basic for our thesis is that the research methodology needed to undertake effective research in land-use is to be found in Data Science. These methods involve developing procedures or models using specified observed variables known as *determinants* to predict or forecast unknown *outcome* variables, based on representative samples in which both determinants and outcomes are known. Traditional statisticians prefer to use regression models to do this, whereas machine learning advocates argue in favour of black-box procedures including random forests. These methods are described in Ripley (1996) and Breiman (2001). Xu *et al*. (2014) provide an example involving trends in vegetation in British Columbia using recent remote sensing data from NASA Earth-orbiting satellites to compare these approaches.

Human understanding of the world depends on designing experiments to test theories and carrying out appropriate analyses of the resulting data (Hawking, 1988). Such data analysis is greatly assisted by graphical methods that allow the data to tell their story. This is especially true for land-use data, which vary with respect to time of year and period of time, and regional location.

In our thesis we use regression modeling rather than machine learning methods because this approach provides more interpretable conclusions for our data. The methods are described in detail in Chapter 2.

1.3 Research objectives

Land-use in general has been considered as local environmental issue, but it is becoming a force of global importance. Many of the problems that threaten mankind's survival on the planet result from the increased human use of land area. Knowledge about land-use has become important to manage and direct changes in the area. The objectives of this research were thus as follows:

1. To display land-use change using graphical methods and maps.
2. To model land-use data to detect priorities for future land-use change.
3. To develop and validate a model to predict urban growth using logistic regression.

1.4 Literature review

1.4.1 Land-use data

Land-use research currently is of scientific interest due to the availability of data from remote sensing, widely use of global positioning systems (GPS), and the readily available GIS software. These data comprise hundreds of land-use plots. The boundaries of the plots create the polygons that can be stored in database tables. Then, the polygonal data structure can be created. GIS information such as imagery and land properties can be extracted using GIS software.

Most land-use studies have used commercial software based on polygonal data structure. Data storage and analysis require high capacity computer facilities, creating difficulties for researchers with limited facilities. Land-use changes based on the polygonal data structure are difficult to measure. Polygonal data structures can provide thematic maps to display patterns of land-use type but the data are difficult to analyse due to changes in the polygons with time. The polygons are replaced when the land-use changes over time. Some polygons disappear, others appear, and existing ones change their shape. Polygonal data thus have different plot boundaries at different periods, and need matching to a common set of boundaries. Statistical methods to measure change over time cannot be simply used because statistical methods for comparing time-varying outcomes require that these outcomes are defined on the same sample space at each time period.

GIS analysis of land-used is based on not only polygons but also pixels. Options vary on the appropriate unit for assessing land-use change it has been argued that and an ideal spatial unit does not exist (Stehman and Wickham, 2011). Knowledge gaps here are how land-use is defined for suitable for statistical analysis and what appropriate unit of analysis can be used to simply detect change and measure its accuracy.

The limitation of polygonal data structure for statistical analysis has been documented (Thinnukool *et al.*, 2014; Chuangchang and Tongkumchum, 2014). The data structure can be improved by conversion to digital structure. The data can be recorded as points on a grid, for which land-use change is easily measured because the grid stays put while only the data change. Digitized data structure can be used directly for statistical analysis of land-use change. It is thus preferable to polygonal data structure for spatial data analysis.

1.4.2 Land-use model

In recent years, a lot of improvements have been made in measuring the change in land-use, understanding the causes of land-use changes and developing predictive models for such change. Such predictive models support analyses made on the causes and effects of land-use changes to obtain more insight on the land-use system and to support land-use policy and planning. The models are helpful in simplifying the complicated nature of socio-economic and biophysical forces that effect the rate and spatial pattern of change in land-use and for estimating the effects of land-use changes. Furthermore, predictive models can explore future changes in land-use under different conditions. Summarizing, land-use models are important and duplicable tools, adding to our existing intellectual abilities to analyze the change in land-use and to make more informed decisions (Costanza and Ruth, 1998).

1.4.3 Urban growth modeling

Many modeling approaches have been used to analyse and predict urban growth trends and some of these approaches are artificial neural networks (ANN), cellular automata (CA), empirical statistical models, etc.

1. Artificial neural networks (ANN) - ANN are the tool for pattern recognition that can describe the complex and non-linearity of environmental processes. They are a tools that use a machine learning algorithm in order to model complex behaviour.

ANN is independent of functional relationships, makes no assumptions with regards to the data distributional properties and requires no prior knowledge on the relationship of variable.

Triantakonstantis and Stathakis (2015) used an ANN to predict urban growth in Athens, Greece. An ANN for model urban growth was developed to simulate the urban changes. Urban changes from 1990 and 2000 were used for model simulation. Prediction for 2006 was achieved and the results were validated using a reference map of 2006. This research showed that ANNs are independent of the relationship between input data, hence, assumptions about spatial autocorrelation and multi-collinearity should be taken into account.

Pijanowskia *et al*. (2002) integrated ANN and geographical information system (GIS) to predict the change in land-use, where GIS is used to develop the spatial predictor variables. First, they designed the network and inputs from historical data using subset of the inputs to train the network, and tested the neural network with full data sets of the inputs and finally predicted changes based on the information from the neural network.

2. Cellular automata (CA) - CA models are made up of simulation environment that is represented by a space grid (raster), in which the attributes of each cell are determined by a set of transition rules, taking in accounts the attributes of neighboring cells.

Al-shalabi *et al*. (2013) used cellular automata with GIS to simulate and forecast the urban growth and change in land-use for Sana'a (Yemen) from 2004-2020. The growth pattern of CA model presented a compact and high density development. The limitations of the model are highly dependent on the data quality. High quality data give more accurate outputs. Other weaknesses affecting the models are growth in multiple directions and also their inability to detect new urban development that sprang up away from existing urban areas.

Furthermore, these applications emphasize on the simulation of spatial pattern rather than on the interpretation or comprehension of the spatio-temporal process of urban growth.

3. Empirical statistical models - Logistic regression is one of empirical-statistical methods widely use to model urban. It is an appropriate model for analysis of binary data (Hosmer and Lemshow, 2004). The model provides estimated effects of independent variables and their precisions. Urban growth modelling aims to explain the dynamic processes related to it, and therefore interpreting the models becomes very important for gaining knowledge about the processes that drive the spatial pattern changes.

Existing logistic regression models usually ignore spatial correlation in land-use data, which affect the accuracy of fitting of land-use modeling and the statistical methodology for considering correlation in such model is not well developed for logistic regression models. It is therefore essential to account for spatial correlation in land-use change models (Pontius *et al*., 2001; Hu and Lo, 2007; Zeng *et al*., 2008).

1.4.4 Spatial Correlation

The problem of using conventional statistical method in spatial land-use analysis is the assumption that data are independent. However, for land-use data, this assumption does not hold, because neighboring observations on land-use are typically spatially correlated. The conventional method does not consider the spatial autocorrelation that exists in the spatial data. The standard errors provided by this method could be underestimated. Ignoring spatial autocorrelation means that importance of variables which might not be relevant to the outcome variables might be overestimated (Overmars *et al*., 2003). Various spatial statistical methods have been used to take spatial correlation into account. They can be summarized as follows.

1. Mixed model - Spatial structure were detected and then the spatial autoregressive model with imposing restrictions were used. The imposing restriction depends on spatial structure defined as weight when fitting regression. However, this method is not suitable for our land-use data because this approach can be relatively slow for large data sets (sample size is greater than 10,000)

2. Autologistic regression - An autologistic regression model was developed by combining logistic regression model with autocorrelation effects. It accounted for spatial autocorrelation by adding an autocovariance variable, which calculated for a specific neighbourhood sizes. Then, significance of predictors were used to determine which neighbourhood size give the most parsimonious autologistic regression model. Size of neighbourhood and type of weighting function are possible sensitive parameters, which can be optimised through trial and error.

3. Generalised estimating equations (GEE) - GEE divide the data into smaller clusters before also modelling the variance-covariance relationship. Zeger and Liang (1986) developed this approach which is an extension of generalised linear models (GLMs). Measuring the outcomes repeatedly through time or space, the GEE method take correlations within clusters of sampling units into account by means of a parameterised correlation matrix, while assuming correlations between clusters to be zero. In a spatial context the clusters can be interpreted as geographical regions, if distances between different regions are large enough. GEE require high storage capacity for solving the GEE score equation without clustering as we used it in our fixed model. Application of the fixed model will therefore be limited for models on data with larger sample size.

1.4.5 Land-use change

Land-use change is not necessarily a change in the use of land but a modification in the purpose of the land which includes intensity and management.

Land-use change has been accepted as an important driver of environmental changes on all spatial and temporal scales. Earth-atmosphere interactions, forest fragmentation, and biodiversity loss are significantly affected by land-use change. Land-use change also related to monitoring and management of natural resource. Land-use change is a critical issue due to its great influence in global warming, loss of biodiversity, and impact on human life.

A wide-range of changes in land are rapidly occurring across several regions of the world are currently undergoing rapid, wide-ranging changes in land. Thailand has seen fast urbanization and tremendous economic growth in the last few decades. Economic development activity is focused in and around many provinces. These changes have quickly transformed Thailand from a subsistent agrarian economy into a rapidly industrialized country. There is pressure on change in land-use pattern due to growing urbanization. Studies on land-use change have been reported in Thailand (see, for example, Muttitanon and Tripathi, 2005; Prabnarong and Thongkao, 2006; Swangjang and Iamaram, 2011). A study in coastal areas of Ban Don Bay, Surat Thani using time series Landsat imagery from 1990, 1993, 1996 and 1999 (Muttitanon and Tripathi, 2005). Overlaying GIS technique was used to find how much growth occurred between different periods, and cross-tabulation was used to determine the change of earlier land-use and the current land-use. This study showed an increasing trend of shrimp farms, forest/mangrove and urban areas with a decreasing trend of agricultural and wasteland areas.

Prabnarong and Thongkao (2006) investigated land-use changes around Pak Panang Bay, Nakhon Si Thammarat Province in 1974, 1991 and 2003 by using aerial photographs with geographic information system techniques. Mangroves, paddy fields and other type of land-uses changed to shrimp farm. The expansion of shrimp farm areas into the mangrove areas directly affected the eco-system of Pak Panang Bay, which was declared a Ramsar Site.

Swangjang and Iamaram (2011) described land-use change near Suvarnabhumi International Airport in 1994 and 2002. They focused on the patterns of land-use before and during airport development. They showed that urbanization of the airport location area increased between 10.07% and 15.57%. These changes of land-use complied with the integrated town and country planning.

Few empirical studies have focused on urban growth. A basic reason for this lack of research is inconsistency of data structure, which make the creation of appropriate model very complicated. Urban growth is one of the most important types of land-use changes currently affecting Thailand.

1.5 Structure of the Thesis

This introductory first chapter introduced of the background and rationale, study objectives and literature review. The next chapter focuses on the methodology, highlighting the procedures and techniques used for analyze and model land-use change, including data, land-use categories, data structure, measure of land-use change, method to display the change and statistical methods. The third chapter presents the preliminary results of land-use change using polygonal to grid-point conversion method and modeling result. The last chapter presents conclusions and recommendation for future work.

Chapter 2

Methodology

The aim of this study was to develop and improved method for analyzing land-use change over time. This chapter describes the research methodology including study area, data source and data management, land-use categories, land-use change, graphical methods and statistical methods.

2.1 Data

Land-use data were obtained from Thailand's Land Development Department. The land-use database comprises polygonal "shape files" of land-use plots recorded at regular surveys of every province. The data can be read and displayed using GIS software, and subsequently restructured into relational database tables and analyzed using a general purpose program such as R.

According to Thailand's Land Development Department, land-use is generally classified into five categories comprising urban and built up land (U), agricultural land (A), forest land (F), water bodies (W) and miscellaneous land (M). Each land-use category is further divided into sub-categories (level 2 and level 3). Some classification levels are shown in Table 2.1. For example, urban and build up land (U) of level 1 is classified into six types of level 2 comprising city, town, commercial (U1), village (U2), institutional land (U3), transportation, commercial and utility (U4), industrial land (U5) and other (U6). U2 of level 2 is further divided into abandoned village (U200), village (U201), and hill tribe village (U202).

Table 2.1 Examples of land-use classifications in Thailand

Level 1	Level 2	Level 3
U: Urban and built up land	U1: City, Town, Commercial	U1: City, Town, Commercial
	U2: Village	U200: Abandoned village U201: Village U202: Hill tribe village
	U3: Institutional land	U3: Institutional land
	U4: Transportation, Communication and Utility	U401: Airport U402: Railway station U403: Bus station U404: Harbour U405: Road
	U5: Industrial land	U500: Abandoned factory U501: Industrial estate U502: Factory U503: Agricultural product trading centers
	U6: Other	U600: Abandoned area U601: Recreation area U602: Golf course U603: Cemetery U604: Refugee camp
A: Agricultural land	A1: Paddy field	A100: Abandoned paddy field A101: Rice paddy
	A2: Field crop	A200: Abandoned field crop
	: :	: :
F: Forest land	F1: Evergreen forest	F100: Disturbed evergreen forest F101: Dense evergreen forest
	: :	: :
W: Water bodies	W1: Natural water body	W101: River, Canal W102: Natural water resource
	: :	: :
M: Miscellaneous land	M1: Rangeland	M101: Grass M102: Scrub M103: Bamboo
	: :	: :

Universal Transverse Mercator (UTM) coordinates were used as an input data file to create the land-use thematic map. The UTM coordinates are based on meter unit of measure, which incorporates the simplicity of the decimal system and it is easy to comprehend units of ten and always have the same two directional designators and never carry negative values.

2.2 Steps of analysis process

We have described the methodology developed for our study, detailing the steps covered for the final output as show in Figure 2.1

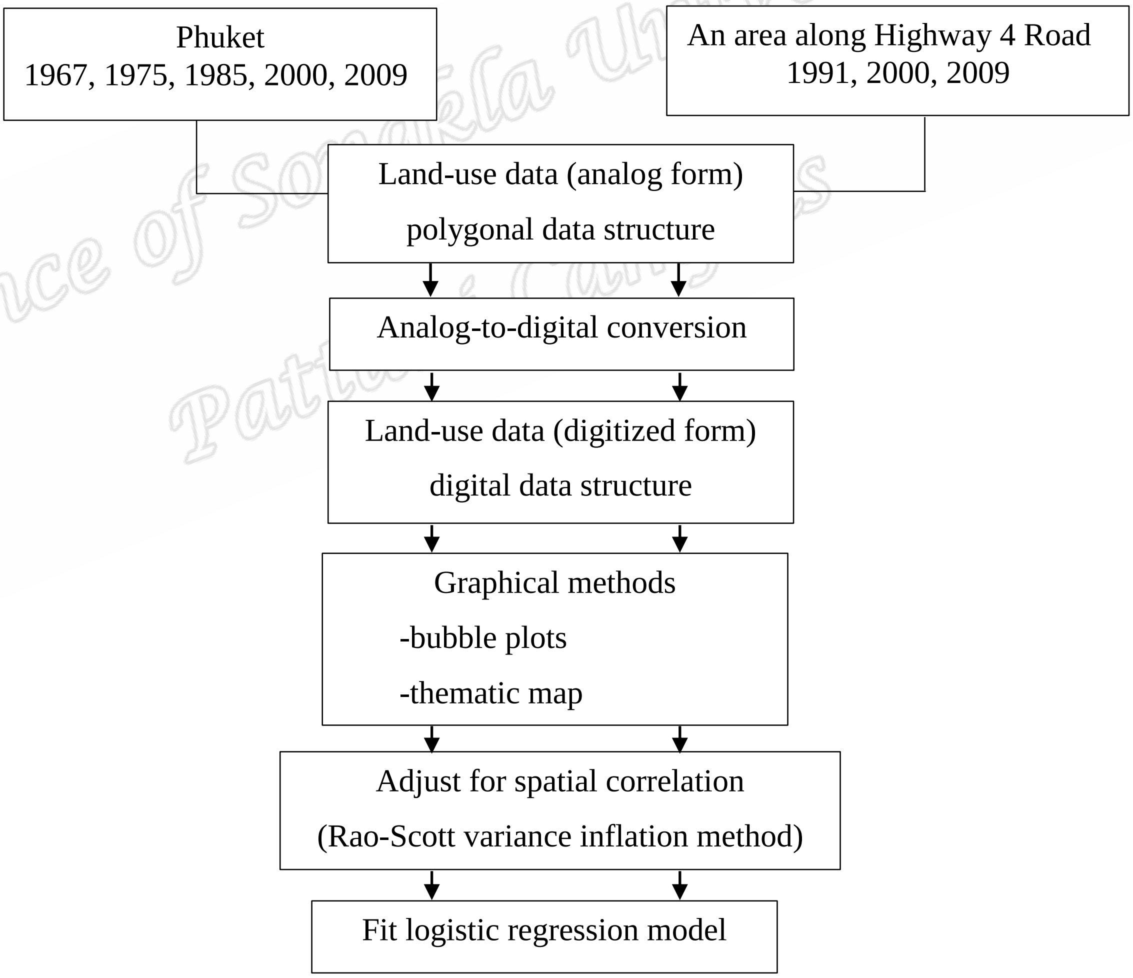


Figure 2.1 Steps of analysis process

2.3 Study area

2.3.1 Phuket

Phuket Province has an area of about 53,900 hectare (539 square kilometers) and is the second-smallest province of Thailand. Furthermore, Phuket is the only island big enough to be a province. The province is separated from the mainland by a narrow channel and connected by Sarasin Bridge. Phuket Province is divided into three districts which are further subdivided into 17 sub-districts. It formerly derived its wealth from tin and rubber

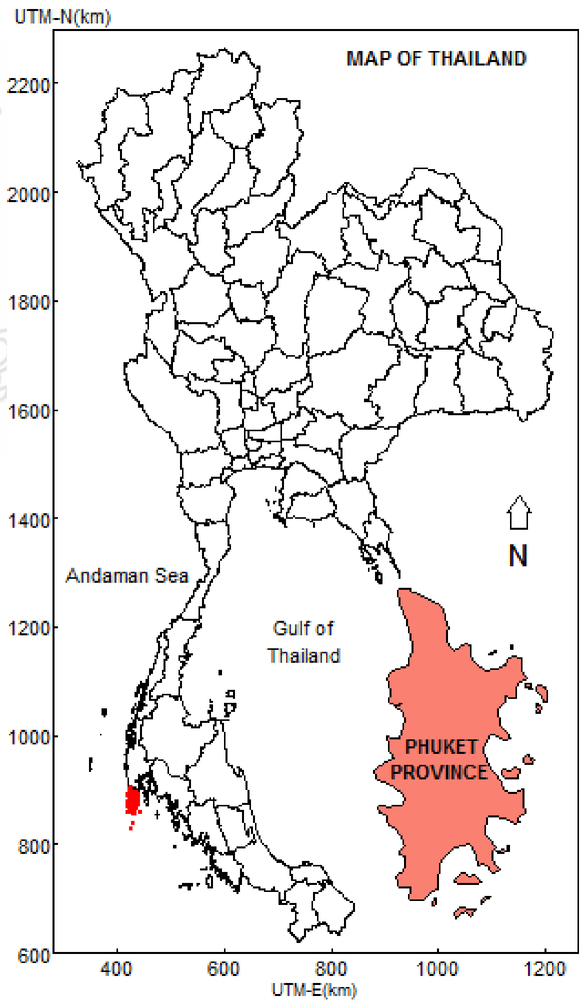


Figure 2.2 Phuket province of Thailand

2.1.2 An area along Highway 4 Road between Phattalung and HatYai

An area along Highway 4 Road between Phattalung and HatYai in Southern Thailand is shown in Figure 2.3. This part of the road is approximately 95 kilometers long and it goes through five districts (Meaung Phatthalung, Khao Chison, Bang Kaeo, Tamot and Pa Bon) of Phattaung and four districts (Rattaphum, Khuan Niang, Bang Klam and HatYai) of Songkhla. The study area comprises three locations. The area for each location is about 700 km^2 (25 km×28 km).

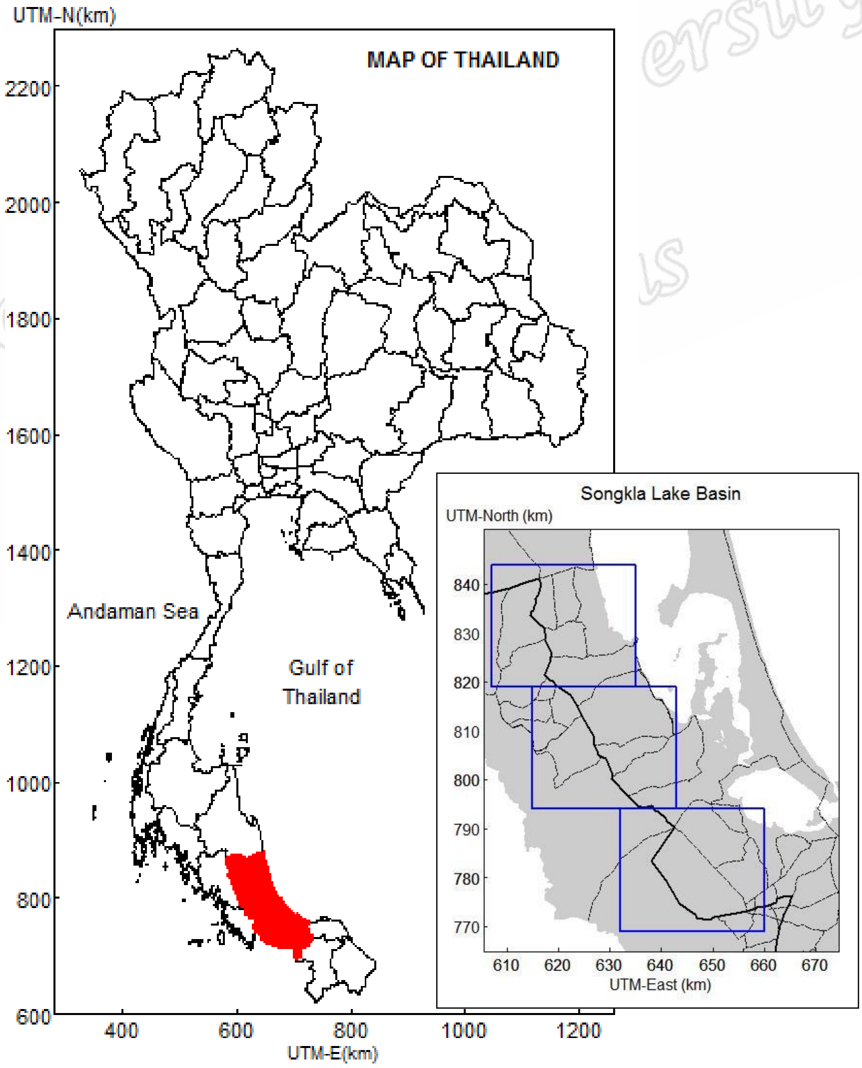


Figure 2.3 Location of area along Highway 4 Road between Phattalung and HatYai

To illustrate the methods for conversion analog to digitized grid we focus on data from Phuket province. Thematic maps of land-use categories for Phuket province in 2009 are shown in Figure 2.4. The left panel shows five land-use categories of level 1 comprising agriculture (A), forest (F) miscellaneous land (M), urban (U) and water (W), respectively. The right panel shows land-use sub-categories for a sub-region of level 2 and 3. For agricultural land (A), it is further classified into paddy field (A1) in level 2, and further classified into abandoned paddy fields (A101) and rice paddies (A102).

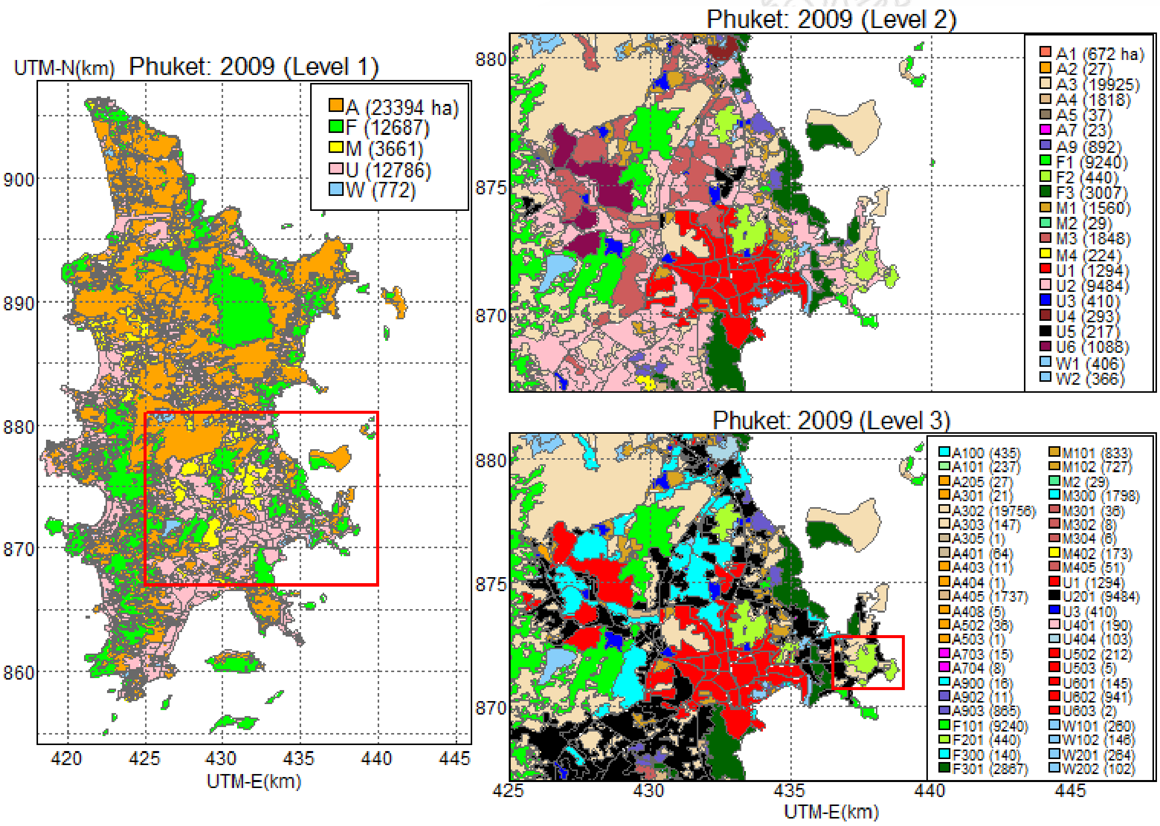


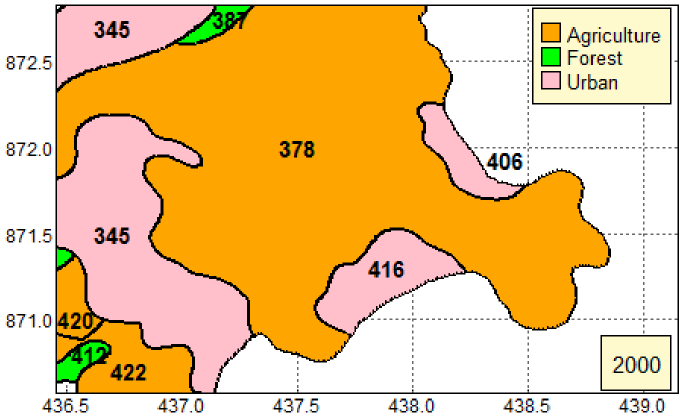
Figure 2.4 Classifications categories of Phuket province land-use in 2009

2.4 Data structure

2.4.1 Polygonal data structure

Land-use plots as polygons can be stored in database tables. Figure 2.5 shows an example of polygonal data structure of land-use in an inset map of the bottom right panel in Figure 2.4. Polygonal data structure stored polygonal plot identified as plotID, land-use codes (luCode), and area of each polygon (for the whole province) in one table. Points of coordinates x and y for drawing the polygon identified as PointID are stored in another table. PlotID field is a primary key linking between these two tables. It comprises eight polygonal plots identified as PlotIDs 345, 378, 387, 406, 412, 416, 420 and 422 comprising of three categories of land-use recorded as forest, agricultural land and urban. PlotID 345 is urban (U) whereas the plotID 422 is agricultural land (A). For each plotID the right panel shows pointID comprising coordinates x and y for drawing the polygon. The pointID field determines the order in which the boundary points (x,y) are connected to obtain a closed polygon for each land-use plot. For example, there are 447 points connected to obtain a closed polygon from plotID 422. The whole of Phuket province in 2000 consists of 571 such polygons.

plotID	luCode	area	desEng
1	F	1881776	Forest
2	U	532235.9	Urban
3	F	16876	Forest
4	F	50776.61	Forest
:	:	:	:
345	U	23098570	Urban
378	A	4191809	Agriculture
387	F	338592.3	Forest
406	U	89158.15	Urban
412	F	740800.6	Forest
416	U	187198.6	Urban
420	A	85411.61	Agriculture
:	:	:	:
422	A	293820.8	Agriculture
570	A	36782.09	Agriculture
571	A	195943.2	Agriculture



plotID	pointID	x	y
345	1	433.20	878.28
345	2	433.21	878.28
:	:	:	:
345	11570	433.20	878.28
378	1	437.48	873.46
378	2	437.49	873.45
:	:	:	:
378	1698	437.48	873.46
387	1	437.29	873.78
387	2	437.30	873.78
:	:	:	:
387	429	437.29	873.78
406	1	438.49	871.78
406	2	438.49	871.77
:	:	:	:
406	191	438.49	871.78
412	1	436.03	871.86
412	2	436.04	871.86
:	:	:	:
412	773	436.03	871.86
416	1	437.73	870.90
416	2	437.71	870.92
:	:	:	:
416	228	437.73	870.90
420	1	436.53	871.37
420	2	436.54	871.36
:	:	:	:
420	202	436.53	871.37
422	1	436.43	870.92
422	2	436.44	870.92
:	:	:	:
422	447	436.43	870.92

Figure 2.5 Polygonal data structure; the right panel lists polygonal data for creating map on the left panel

The polygons are replaced when the land-use changes with time. Some polygons disappear, others appear, and existing ones change their shape. Moreover, coordinates shift from year to year due to change in GPS setting. Therefore, land-use changes based on the polygonal data structure are difficult to measure. A limitation of polygonal data structure for statistical analysis has been documented (Thinnukool *et al*., 2014; Chuangchang and Tongkumchum, 2014). The data structure can be improved by conversion to data on a fixed grid. The new method is basically analog-to-digital conversion, replacing polygonal shapes by coded grid points. This analog-to-digital conversion method is explained by Thinnukool *et al*. (2014). The steps of data conversion are summarized in Figure 2.6.

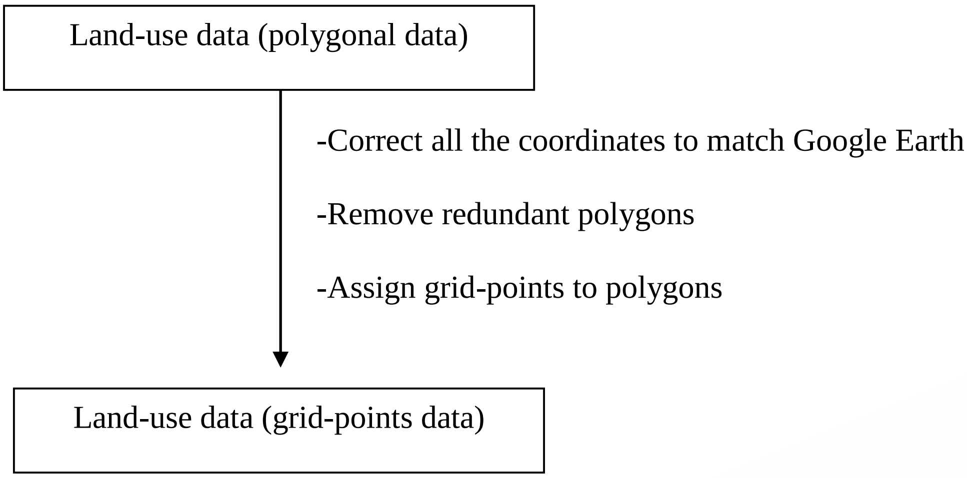


Figure 2.6 Steps of data conversion

According to the standards for measuring, land-use data have changed in recent decades giving incorrect geometric positions, leading to geometric distortion. These coordinate shifts are quite substantial and complicate the accurate measurement of land-use change. Assuming Google Earth coordinates are correct and that these locations have not changed over recent decades, it is desirable to convert all land-use coordinates to be consistent with Google Earth coordinates based on a bilinear transformation method (Thinnukool *et al.*, 2014). This method we use for this conversion takes the form

$$u = a_1 + b_1x + c_1y + d_1xy \qquad (1)$$

$$v = a_2 + b_2x + c_2y + d_2xy \qquad (2)$$

In these equations (x, y) are coordinates in the files that need to be corrected and (u, v) are their corresponding corrected values that agree with Google coordinates.

The parameters (a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2) in equations (1) and (2) are determined by using the data for the coordinate shifts (dx, dy) at the four locations of the specified rectangle as described in detail by Thinnukool *et al*. (2014). These equations can be written in matrix form as

$$\boldsymbol{g} = \boldsymbol{F}\,\boldsymbol{h} \qquad (3)$$

where $\boldsymbol{g}$ is the column vector (u_1, v_1, u_2, v_2, u_3, v_3, u_4, v_4), $\boldsymbol{h}$ is the column vector (a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2) and $\boldsymbol{F}$ is an eight dimensional square matrix given by:

$$\begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \\ u_4 \\ v_4 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & x_1y_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x_1 & y_1 & x_1y_1 \\ 1 & x_2 & y_2 & x_2y_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x_2 & y_2 & x_2y_2 \\ 1 & x_3 & y_3 & x_3y_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x_3 & y_3 & x_3y_3 \\ 1 & x_4 & y_4 & x_4y_4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x_4 & y_4 & x_4y_4 \end{bmatrix} \times \begin{bmatrix} a_1 \\ b_1 \\ c_2 \\ d_2 \\ a_3 \\ b_3 \\ c_4 \\ d_4 \end{bmatrix}$$

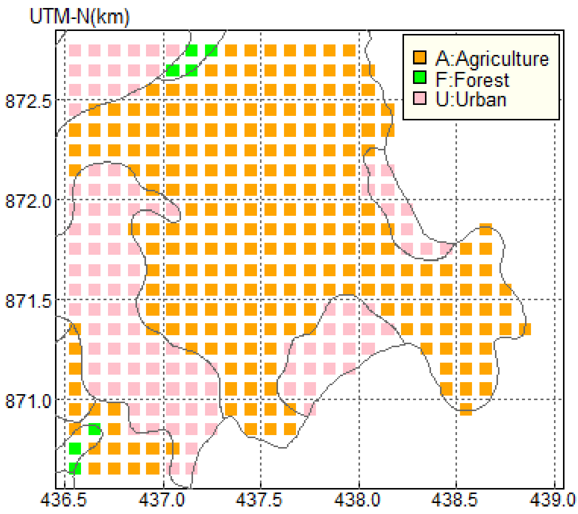
All coordinates are corrected to match Google Earth. Polygonal structure is converted to a digitized data structure using a computer program. The computation comprises of assigning grid points to polygons. This uses the function *point.in.polygon* in the spatial (sp) library of the R program.

When larger polygonal land-use plots contain smaller plots such as farm ponds it creates duplicated polygons and needed to removes holes from polygons. Once their holes have been removed, it is necessary to order the polygons by increasing area, because once the grid-points within a polygon containing a hole, they cannot be overwritten.

2.4.2 Digital data structure

Digital data structure involves creating a digitized grid of geographical coordinates covering the whole of study area and storing the land-use codes and plot identifiers as fields in database tables indexed by the grid coordinates. The digitized land-use data are recorded at points on a square grid of specified dimension. Altitude data in Thailand are gridded at 111 meter intervals (1^o of latitude = 111 km (Polovina *et al.*, 2008)), so that one degree of latitude contains 100 grid points. In this study we selected 100 meter intervals digitizing land-use data (100 by 100 meter). So the digitised land-use data are recorded at points on a square grid with and area of one hectare.

The digital data structure facilitates measurement of land-use change because the land-use at each year is defined on a constant grid. Figure 2.7 shows the digital data structure of the same area as Figure 2.5. The data table on the right panel contains plotID, pointID and luCode. The plotID refers to polygon, and pointID with x, and y coordinates here refers to grid point. For example, the plotID 387 contains four grid boxes. The area of each land-use type can be calculated simply as the numbers of grid in the polygon. The area of plotID 387 is four ha because it comprises four grids.



plotID	pointID	x	y	luCode
345	1	436.55	871.45	U
345	2	436.55	871.55	U
:	:	:	:	:
345	82	437.25	871.15	U
378	1	436.55	871.15	A
378	2	436.55	871.25	A
:	:	:	:	:
378	239	438.85	871.35	A
387	1	437.05	872.65	F
387	2	437.15	872.65	F
:	:	:	:	:
387	4	437.25	872.75	F
406	1	438.05	872.05	U
406	2	438.05	872.15	U
:	:	:	:	:
406	10	438.45	871.75	U
412	1	436.55	870.65	F
412	2	436.55	870.75	F
412	3	436.65	870.85	F
416	1	437.65	871.05	U
416	2	437.65	871.15	U
:	:	:	:	:
416	19	438.15	871.35	U
420	1	436.55	870.95	A
420	2	436.55	871.05	A
:	:	:	:	:
420	5	436.55	871.35	A
422	1	436.55	870.85	A
422	2	436.65	870.65	A
:	:	:	:	:
422	13	437.05	870.75	A

Figure 2.7 Digital data structure; the right panel lists grid-point data for creating the left panel

2.5 Measure land-use change

The digital structure is appropriate for analysis of land-use change, which can be measured in terms of loss and gain from two periods of time as shown in Figure 2.8 The land-use change in area from an inset map of the bottom right panel in Figure 2.4, which is in Rasada sub-district of Phuket province. Where land-use in 2000 and 2009 are shown in the left panels, loss in 2000 and gain in 2009 are shown in right panels.

The total area of each panel is 550 ha. In 2000, land area is 375 ha and sea (Z) area is 175 ha. In 2009, land area is 369 ha and sea (Z) area is 181 ha. The area of losses in 2000 (or gains in 2009) is 198 ha. The top right panel shows only the land-use in 2000 that was lost, mainly from agricultural land; orange colour indicates areas of agriculture loss. The bottom right panel shows that changes in 2009 were mostly due to gains in forest and urban areas. The major change is that agricultural land in 2000 changed to forest in 2009.

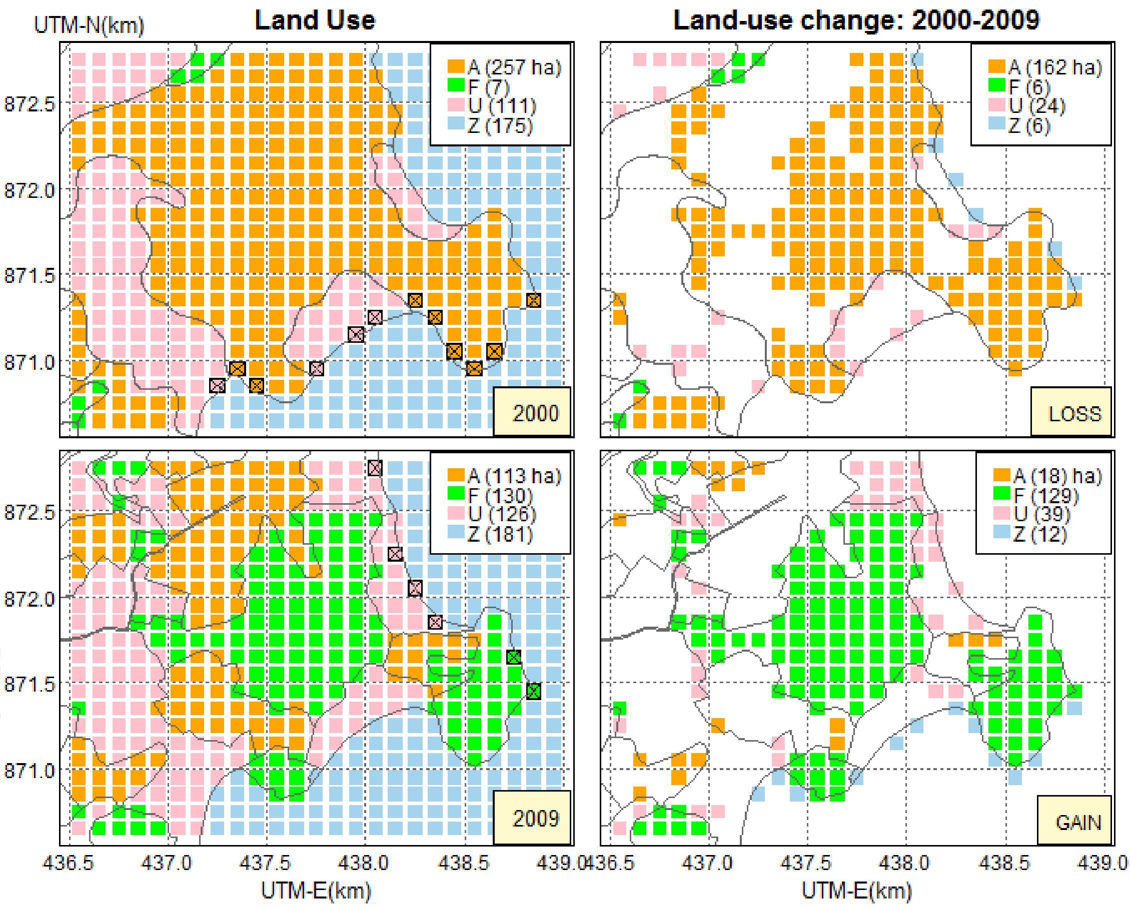


Figure 2.8 Land-use change in area from 2000-2009 with losses from 2000 (upper right panel) and gains to 2009 (lower right panel)

A digital data structure provides a simple calculation for land-use change. Focusing only on land, it is 375 ha in 2000 and 369 ha in 2009. Note that the shape of the study area changed over the 9-year period, which complicates the measurement of land-use change. This problem can be avoided by considering change with respect only to land present in both 2000 and 2009. This common area is 363 ha. The area of 12 ha in

2000 and 6 ha in 2009 marked as (⊠) in the map of Figure 2.8 are omitted. Therefore, common area was used to detect change between periods of time.

2.6 Method to display change

Measure of change can be displayed using cross-tabulation, bubble plots and thematic map.

2.6.1 Cross-tabulation and Bubble plots

Table 2.2 shows a corresponding cross-tabulation of three categories of land-use in 2000 and 2009. It displays area of change in land-use and area remaining the same. The area of urban land increased from 107 to 122 ha over the period, forest land increased from 7 to 128 ha and agricultural land decreased from 249 to 113 ha. In 2009, 154 ha of agricultural land were lost: 120 ha to forest land and 34 ha to urban land.

Table 2.2 Land-use change in area: 2000-2009

		2009			
		Urban (U)	Agriculture (A)	Forest (F)	total
2000	Forest (F)	1	5	1	7
	Agriculture (A)	34	95	120	249
	Urban (U)	87	13	7	107
	total	122	113	128	363

Bubble plots are effective methods to summarize change in land-use. They display percentages of land-use categories from one period to the next. Land-use for a given year can be graphed as a thematic map using a separate colour for each category, as shown in Figure 2.4. Graphing land-use change is more complicated because many additional colours are needed to show data from a cross-tabulation such as Table 2.2.

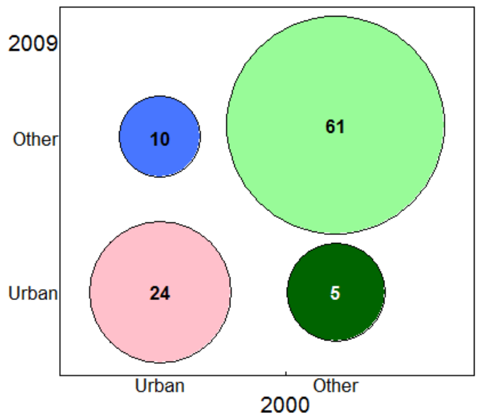
To simplify this task we combined land-use type into two categories (U and A+F). As a result, the cross-tabulation is reduced to a 2×2 table requiring only four colours, as shown in Figure 2.9 using a bubble plot.

Change in land-use is effectively summarized in a cross-tabulation giving area in hectares (Figure 2.9a) of land-use categories from one period to the next. These numbers can be converted to percentage and displayed as a bubbleplot matrix (Figure 2.9b).

The benefit of bubble plot is easy to understand because the size of bubble represents the percentages of land-use categories from 2000 to 2009. The colour represents categories change.

		2009		
		Urban(U)	Other(A+F)	total
2000	Other(A+F)	35	221	256
	Urban(U)	87	20	107
	total	122	241	363

(a)



(b)

Figure 2.9 Land-use change: (a) change of land-use in 2000-2009 (ha) and (b) percentage of land-use categories

2.6.2 Thematic map

Thematic map can be used to display land-use in each year (as in Figure 2.4). It can also be used to show the distributions of land-use change and where the changes occur.

Figure 2.10 shows percentages of land-use change in an inset map of Figure 2.4 from 1967-2009 using a thematic map of change corresponding to the bubbleplot matrix in Figure 2.9. The thematic map divides land-use change into four colors, pink for urban land that remained urban (U-U), blue for other land that changed to urban land (O-U), dark green for urban that change to other land (U-O), and green for other land that remained other land (O-O).

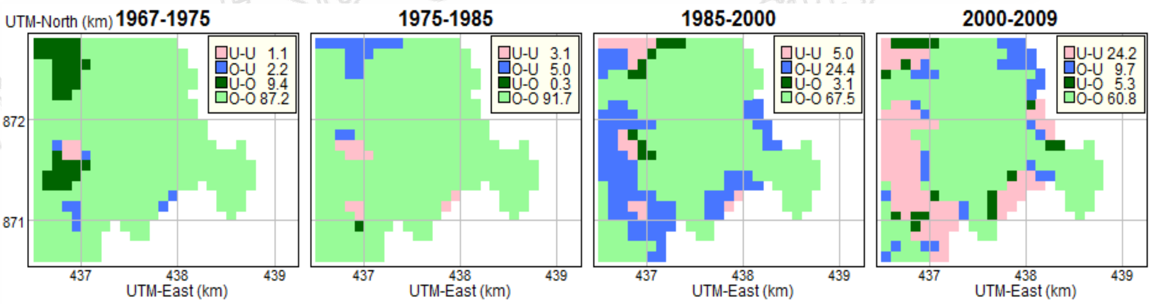


Figure 2.10 Land-use percentage changes: 1967-2009

2.7 Statistical methods

Statistical methods of land-use are complicated by changing boundaries of polygonal land-use plots. The data structure is improved by gridding, in which the polygons that vary in shape and size are replaced by a rectangular and unchanging grid of points on which the land-use is defined.

For simplicity in statistical analysis, the outcome is current land-use in 1 hectare plot classified into two categories (urban/other). The determinants are location and land-use in previous year. Figure 2.11 shows path diagram.

Determinants		**Outcomes**
1: Location 2: Land-use in previous year	→	Current land-use in 1 hectare plot (Urban land/Other)

Figure 2.11 path diagram

These data can be analyzed by logistic regression model. The model provides estimates of urban land in current year in each category of the determinant together with measures of precision of the estimates. The model also gives the p-value for testing association between determinant and outcome.

2.7.1 Logistic regression model

Logistic regression is a powerful empirical method appropriate when the outcome is binary (Hosmer and Lemeshow, 2000).

2.7.1.1 Simple logistic regression

A simple logistic regression is a model with single determinant. Location and land-use at a previous year were combined into a location-land-use factor. Since the effects of location and land-use from a previous survey as determinants of land-use change might not be additive, there is some advantages in combining them to form a single factor corresponding to all location-land-use group combinations.

The model is fitted to a digitized grid data with location-land-use factor at a previous year as determinant. This model formulates the logit of the probability p_i of urban outcome in the year of interest as a function of location-land-use factor determinant. The simple model of urban land thus takes the form

$$\text{logit}(p_i) = \log\left[\frac{p_i}{1-p_i}\right] = \mu + \alpha_i \qquad (2.1)$$

where μ is a constant and the term α_i, refers to location-land-use group at a previous year. The equation may be inverted to give an expression for the probability p_i as

$$p_i = \frac{1}{1+\exp(-(\mu+\alpha_i))} \qquad (2.2)$$

This simple model in equation 2.1 was applied to land-use data of Phuket province and land-use data along highway 4 road from Phattalung to HatYai.

For Phuket province, land-use data was classified into two categories comprising urban land (U) and other land (O). Urban land comprises urban and built up land and mines. The other land comprises agricultural land, forest, water body and miscellaneous land except mines. Phuket province was divided into two locations (north: above UTM-North 880 km and south). Location and land-use at a previous year were combined into four levels of a location-land-use factor with two land-use groups for two locations. Four set of analysis were conducted. Firstly, urban land in 1975 is an outcome and location-land-use group in 1967 is a determinant. Secondly, urban land in 1985 is an outcome and location-land-use group in 1975 is a determinant. Thirdly, urban land in 2000 is an outcome and location-land-use group in

1985 is a determinant. Finally, urban land in 2009 is taken as an outcome and location-land-use group in 2000 as a determinant.

For area along highway 4 road from Phattalung to HatYai, land-use data in 1991, 2000 and 2009 was classified into four main groups namely undeveloped land (UD), paddy field and other agriculture (PF+), rubber plantation (RP) and developed land (Dev). In land-use for level 1, undeveloped land comprises forest, water bodies and miscellaneous land while paddy field and other agriculture is the agricultural land except for rubber plantation. Developed land is the urban and built up land. Detailed descriptions of land-use categories are shown in Table 2.1. The determinants are location and land-use categories nine years earlier. Locations and land-use categories were thus combined into 12 levels of a categorical variable. These 12 levels comprise combinations of four land-use groups, and three locations along the Phattalung to HatYai road. To measure land-use change between two periods of time, two sets of analysis were conducted. First, developed land in 2000 was taken as the outcome and location by land-use group from preceding surveys in 1991 as the determinant. Second, developed land in 2009 was taken as the outcome and location by land-use group in 2000 as a determinant.

2.7.1.2 Multiple logistic regression

Multiple logistic regression was used for further analysis of land-use data in Phuket province from 2000 to 2009. The objective is to model land development at sub-districts level. Two factors were considered as determinants of developed of land, namely land-use group in 2000 and sub-district identity. Land-use group in 2000 has four levels: (1) undeveloped land, (2) other agriculture, (3) rubber plantation and (4) developed land. The sub-district factor has 17 levels, one for each sub-district.

Simple and multiple logistic regression model of developed land (Dev) were fitted. The simple model included only one factor (land-use group in 2000), while the full model included two factors. The full model formulates the logit of the probability p_{ij} of developed land (Dev) in terms of the two determinant factors as follows:

$$\log\left[\frac{p_{ij}}{1-p_{ij}}\right] = \mu + \alpha_i + \beta_j \tag{2.3}$$

In this model μ is a constant and the terms α_i and β_j refer to land-use group in 2000 and sub-district, respectively.

Conventional statistical analysis such as logistic regression assumes that data samples are independent. However, this assumption clearly does not hold for data defined at grid-points just 100 meters apart. Data from neighbouring plots are likely to be correlated, violating the independence assumption, giving incorrect standard errors. Therefore, method for handling spatial correlation is needed.

2.7.2 Variance Inflation Factors (VIF)

In other studies with geographical data (see, for example, McNeil and Chooprateep, 2014) such spatial correlation is handled by aggregating data into larger regions with acceptably small correlation between adjoining regions, or by using factor analysis and multivariate regression to adjust for spatial correlations. For our land-use data, the generalized estimating equations (GEE) (Zeger and Liang, 1986) method could be used by dividing the region into groups of plots (clusters) and estimating common fixed correlations between plots in the same group. However, this method has two difficulties. They are (a) large cluster size requiring excessive computation even on the most powerful available computer, and (b) it assumes a common correlation within land-use polygons.

Another option is a conventional method widely used in survey sampling, based on variation inflation factors (Rao and Scott, 1992). Standard errors of proportions can be corrected using variance inflation factors (VIF). The VIF is specified by how much the sample size of a cluster needs to be increased to compensate for the clustering. For the simplest case, when the correlations between binary outcomes in clusters of size m have equal correlation ρ, the VIF is $1 + (m - 1)\rho$, and the standard error of the log odds ratio is increased by the square root of this factor. This formula shows that even small correlations between outcomes can have a substantial effects in large clusters (McNeil, 2014). This method avoids the problem in the GEE method by computing effective sample sizes for each land-use plot based on their sample variances giving a set of VIFs, from which standard errors are applied to fitted values from a logistic model to compute confidence intervals.

2.7.3 Weighted sum contrasts

The adjusted percentages of urban land/developed land for each determinant are thus presented using graphs of confidence intervals. The conventional treatment contrasts method gives different graphs based on which level in the factor is selected as the reference group, and provides wider confidence intervals when this reference group has smaller sample size. To compare sub-district and land-use group effects with overall mean, rather than with an arbitrary sub-district, the standard errors for the estimated parameters in the model were based on weighted sum contrasts, as described by Tongkumchum and McNeil (2009). Using this method, the confidence intervals for proportions may be classified into three levels; (1) totally above the mean, (2) crossing the mean, and (3) totally below the mean.

2.7.4 Receiver Operating Characteristic (ROC) curve

A receiver operating characteristic (ROC) is useful indicator of how well the model is able to perform classification. Generally for evaluating the accuracy of a logistic model that predicts a binary outcome. This is a plot of the true positive rate (sensitivity) against the false positive rate (1-specificity) for different possible thresholds of the model. A graph displays the predictive accuracy of the logistic model using the area under the ROC curve (AUC). The AUC is particularly important for evaluating how well the method can discriminate the outcome values. An ideal model would have an AUC of 1. Denoting the predicted outcome as 1 (developed land) if $p \geq c$ or 0 (other) if $p < c$, the ROC curve plots the proportion of positive outcomes correctly predicted by the model) against the false positive rate (proportion of all outcomes incorrectly predicted), as *c* varies. Choosing *c* to match numbers of predicted and observed outcomes means that equal weights are assigned to false positive and false negative prediction errors. The AUC measures the performance of a model and represents model accuracy (Sakar and Midi, 2010; Takahashi *et al*., 2006). It shows how well a model predicts a binary outcome, which varies from 0.5 to 1. An AUC close to 1 signifies that the model has almost perfect discrimination while an AUC close to 0.5 indicates poor discrimination (Hanley and McNeil, 1982).

All data analysis and graphical displays are carried out using R (R Development Core Team, 2012).

Chapter 3

Results

This chapter presents results from preliminary analysis and statistical modeling of land-use change. The fundamental output of this research are the thematic maps, bar charts, bubble plots and modeling result. In section 3.1 (paper I), illustrated polygonal to grid-point conversion method and result for using this method for land-use changes in the Phuket province from 1967 to 2009. In section 3. 2 (paper II), we report land-use change along highway 4 in southern Thailand from 2000 to 2009. In section 3.3 (paper III), we report land-use change in Phuket province by sub-district from 2000 to 2009.

3.1 Land-use change in Phuket province from 1967-2009

(**Paper I:** Modelling Urban Growth over Time using Grid-digitized Method with Variance Inflation Factors applied to Spatial Correlation (Chuangchang *et al*., 2016a))

3.1.1 Preliminary results

This part focused on land-use conversion from other land to urban land and aimed to model urban growth in Phuket province.

Figure 3.1 shows percentages of land-use change in Phuket province from 1967- 2009 using a bubbleplot matrix. Bubbles with lighter colours along the diagonal line denote no land-use change from each period whereas bubbles with darker colours off the diagonal line denote land-use changes from one period to the next.

For example, the top right panel shows how land-use changed from 2000-2009 in the north of Phuket province. In 2000, 11% was urban land and 89% was other land. In 2009, 14% was urban land and 86% was other land. On the diagonal line, the pink bubble shows that 7% was urban land in 2000 and that this remained urban land in 2009. The light green bubble shows that 82% was other land in 2000 and that this remained other land in 2009.

The small green bubble shows that 4% was urban land in 2000 and this was changed to be other land in 2009. The blue bubble shows that 7% was other land in 2000 and this was changed to be urban land in 2009.

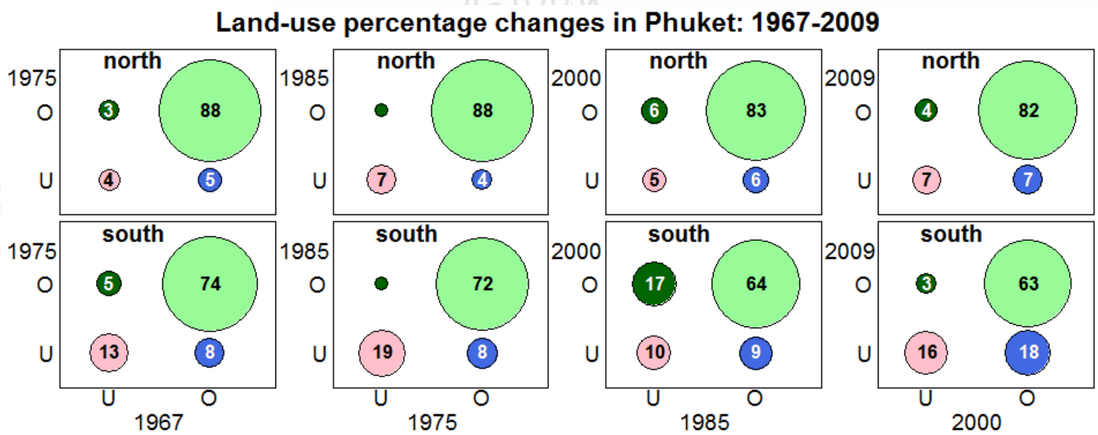


Figure 3.1 Land-use percentage change in Phuket province from 1967-2009. The top row shows the change of land-use in the north and the bottom row is for the south of Phuket province

Thematic maps are useful for showing where changes occur. Figure 3.2 shows percentages of land-use change in Phuket province from 1967-2009 using a thematic map corresponding to the bubbleplot matrix. The thematic map uses the same colours as in the bubbleplot matrix. This figure shows land-use changes from one period to the next for each category. For example, in the rightmost panel, light green colour denote other land in 2000 that remained other land in 2009 (O-O) 72.2% and dark green denote urban land in 2000 that changed to other land in 2009 (U-O) 3.8%. Thus, other land (all green colours) in 2009 was about 76.0% (72.2% + 3.8%) of all land.

Pink colour denotes urban land in 2000 that remained urban land in 2009 (U-U) 11.3% and blue colour denotes other land in 2000 that changed to urban land in 2009 (O-U) 12.7%. Thus, urban land (pink and blue colours) in 2009 comprised about 24.0% (11.3%+12.7%) of all land.

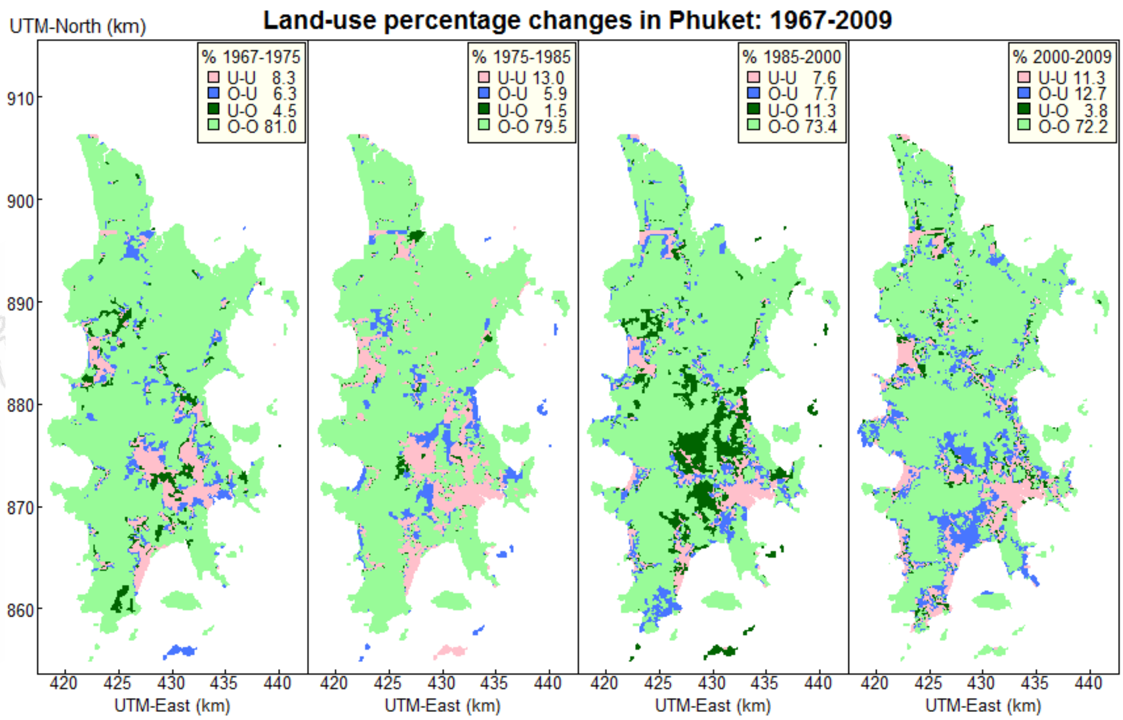


Figure 3.2 Thematic map of land-use change in Phuket province from 1967-2000. The maps show percent of land-use change 1967-1975, 1975-1985, 1985-2000 and 2000-2009, respectively

3.1.2 Modelling results

Logistic regression models were used for predicting the percentages of urban land corresponding to each land-use category. The results from the logistic regression model are displayed using 95% confidence intervals for the percentage of urban land in a given year in each category of location-land-use group in a previous survey superimposed on the bar chart. A confidence interval completely above or below the mean line indicates that the factor is significantly higher or lower than expected. Hence, location-land-use group in the previous period was statistically associated with urban land in the year of interest.

Figure 3.3 shows bar charts of land-use change for each group and model-based 95% confidence intervals of percentage change to the following period by location and land-use from preceding surveys carried out in 1967, 1975, 1985, 2000 and 2009. The red horizontal lines represent overall percentages of urban land in each year.

The bar charts show percentages of urban land from one period to the next. For example, in southern Phuket province, 70% of urban in 1967 remained urban in 1975, while 9% of other land became urban land. Similarly, more than 90% of urban in 1975 remained urban in 1985, while 10% of other land became urban land. For the next period, more than 35% of urban land in 1985 remained urban in 2000, while 12% of other land became urban. Finally, more than 80% of urban land in 2000 remained urban in 2009, while 22% of other land became urban land. Surprisingly, the proportion of urban land in the south of Phuket remained relatively low in the period from 1985-2000.

The overall percentages of urban land in 1975, 1985, 2000 and 2009 were 15%, 19%, 16% and 24%, respectively (red line). The blue vertical line segments denote 95% confidence intervals from the logistic regression model using the variance inflation method to account for correlations within land-use plots.

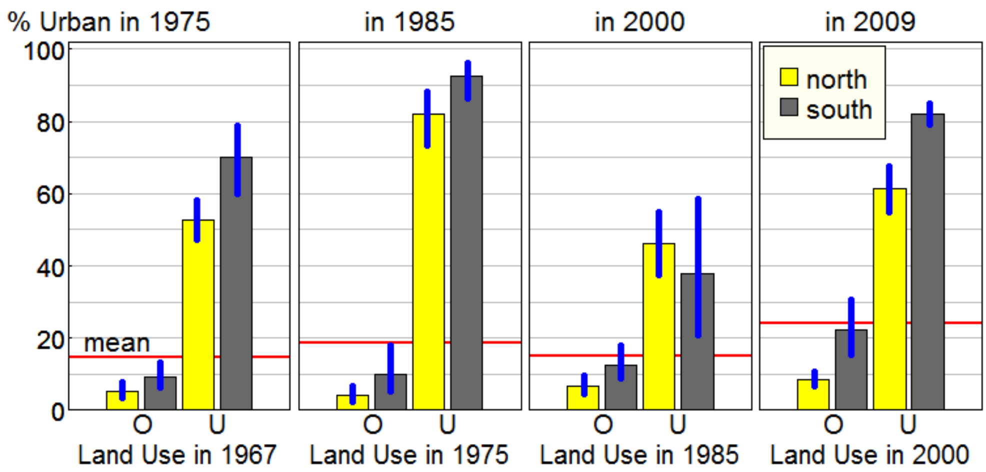


Figure 3.3 Bar charts of land-use change for each group and 95% confidence intervals of percentage change by location and land-use from preceding surveys 8-15 year ago in Phuket: 1967-2009

3.2 Land-use change along Highway 4 from 1991 to 2009

(**Paper II:** Modeling land development along highway 4 in Southern Thailand (Chuangchang *et al.*, 2014))

3.2.1 Preliminary results

The thematic map (Figure 3.4) in this section show the land-use types along highway 4 for 1991, 2000 and 2009 for three locations. The top row shows land-use in the northern location, middle row shows land-use in the central and the bottom row shows land-use in the southern location.

Most of the land in the northern part was used for paddy fields and other agriculture (PF+). Rubber plantation (RP) was the largest category of land-use in the central and southern locations.

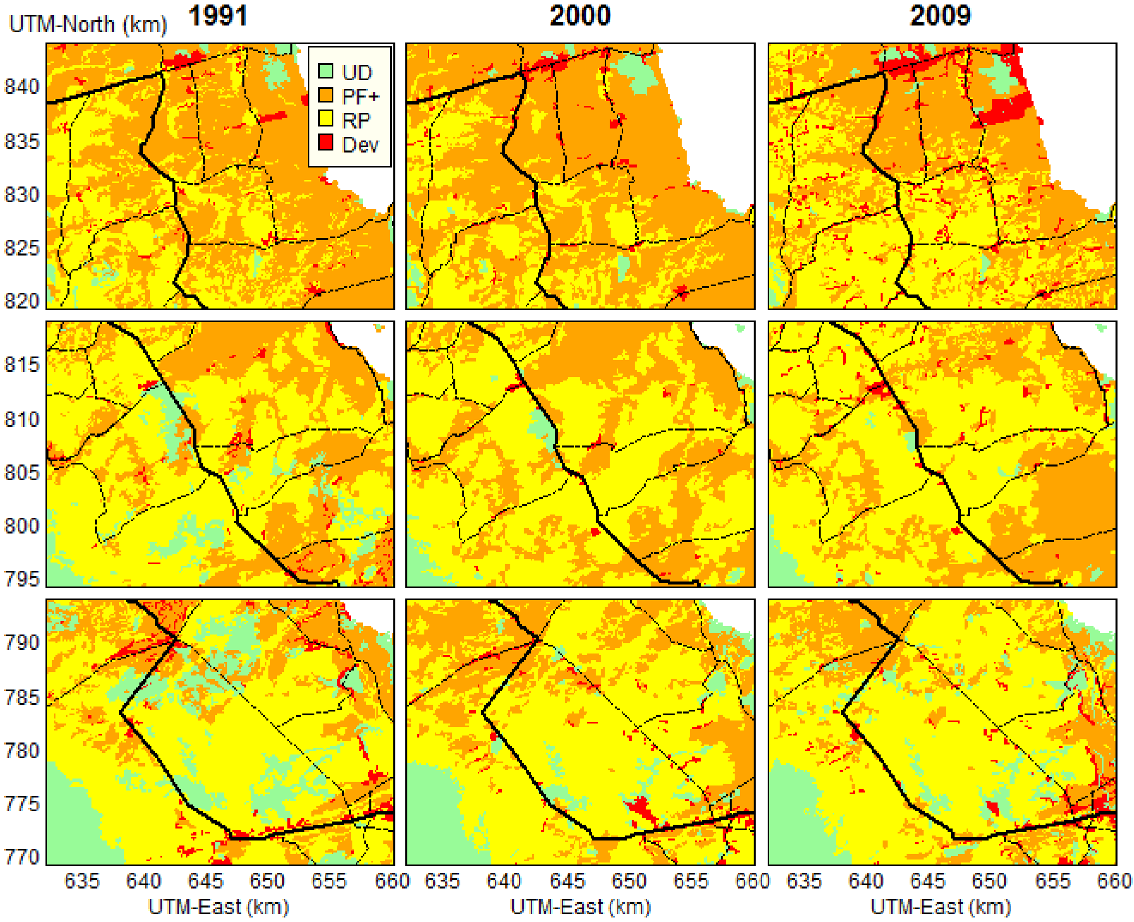


Figure 3.4 Land-use maps along the Phattalung to HatYai road in 1991, 2000 and 2009 for the three locations

The bar charts in the Figure 3.5 show the area (in ha) of undeveloped land (UD), paddy field and other agriculture (PF+), rubber plantation (RP), and developed land (Dev).

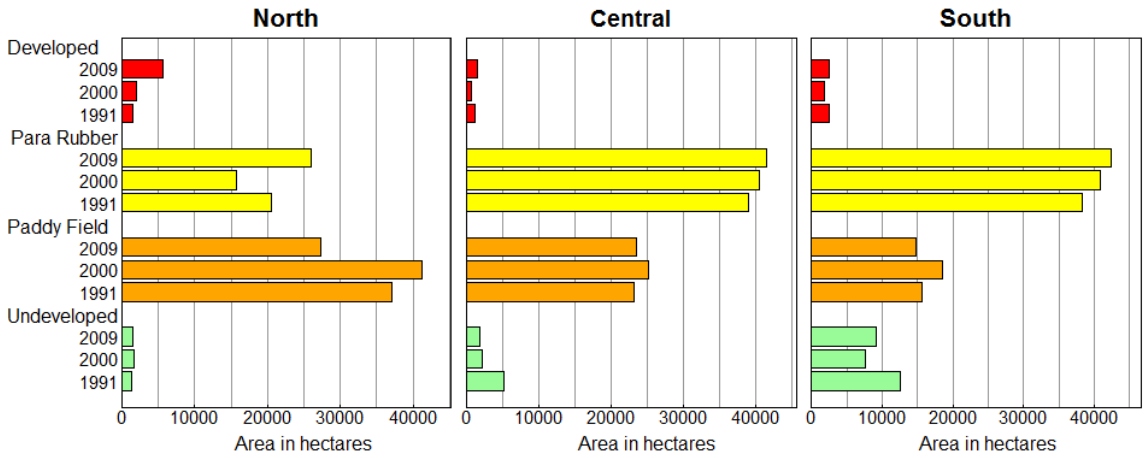


Figure 3.5 Bar charts of land-use along the Phattalung to HatYai road in 1991, 2000 and 2009 for three locations

Figure 3.6 shows a bubble plot matrix of percentages of land-use category change in three locations. The top row shows the change of land-use from 1991 to 2000 and the bottom row shows the change from 2000 to 2009. Lighter grey or lighter colors (pink, yellow, orange and green) along the diagonal denote no change in land-use. Darker gray or darker colors (dark green, grey and red) off the diagonal denote land-use changes from one period to the next.

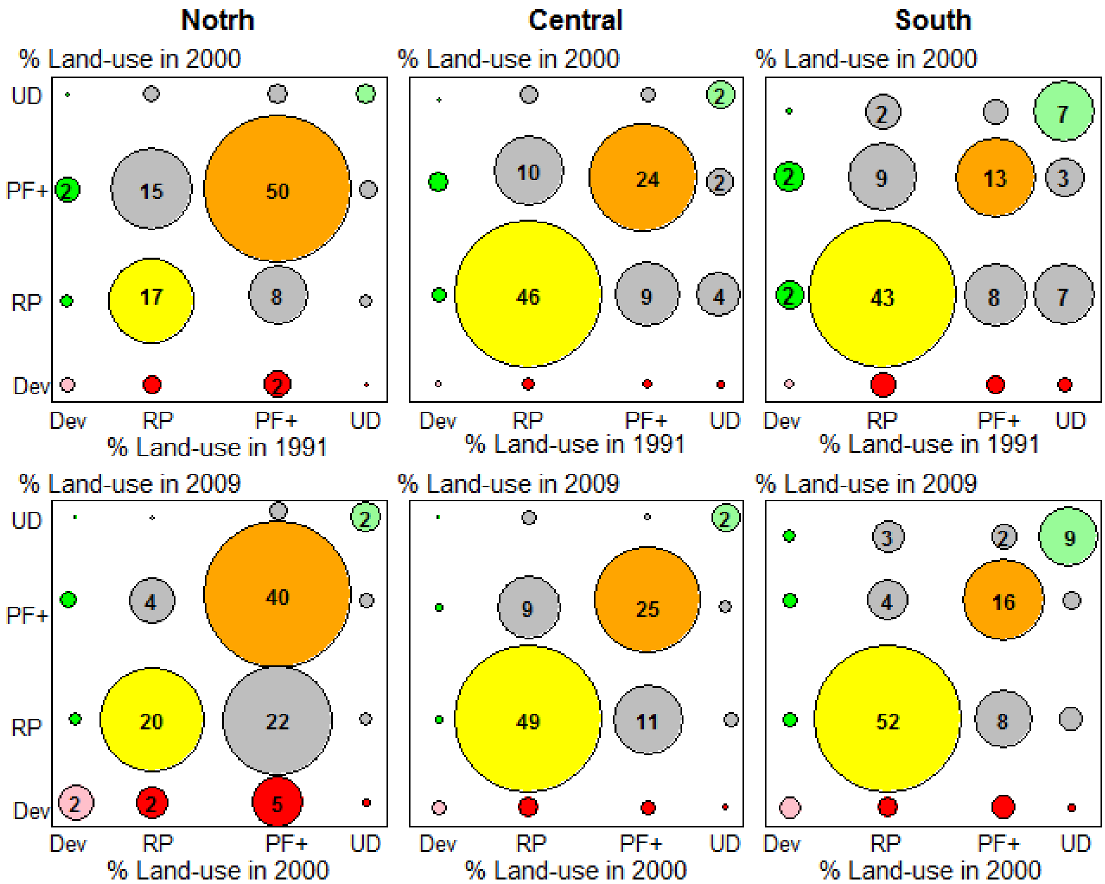


Figure 3.6 Land-use percentage changes along the Phattalung to HatYai road are the periods 1991-2000 and 2000-2009 for the northern, central and southern locations

For example, the top right panel shows the change in land-use from 1991-2000 in the southern location. In 1991, 4.28% of the land was developed land, 55.52% was rubber plantation, 22.70% was paddy field and other agriculture and 17.50% was undeveloped land. In 2000, 2.92% was developed land, 60% was rubber plantation, 27% was paddy field and other agriculture and 10.08% was undeveloped land.

On the diagonal, the pink bubble shows that less than 1% of the land that was developed land in 1991 remained developed in 2000. The yellow bubble shows that 43% of the land that was rubber plantation in 1991 remained so in 2000. The orange bubble shows that 13% of the land that was paddy field and other agricultural land in

1991 remained so in 2000. The light green bubble shows that 7% of the land that was undeveloped in 1991 remained so in 2000.

Off the diagonal line in the first column, the green bubbles, show that out of 4% of the land that was developed in 1991, 2% changed to paddy field and other agriculture and the other 2% changed to rubber plantation in 2000. In the second column, the grey bubbles show that out of the 11% that was rubber plantation in 1991, 2% became undeveloped and the other 9% changed to paddy field and other agriculture in 2000. The red bubble shows that 1% of the land that was rubber plantation in 1991 became developed in 2000. In the third column, the grey bubbles show that out of the 9% of land that was paddy field and other agriculture in 1991, 1% became undeveloped and the other 8% changed to rubber plantation in 2000. The red bubble shows that less than 1% of the land that was paddy field and other agriculture in 1991 became developed in 2000. In the fourth column, the grey bubbles show that the 10% of the land that was undeveloped in 1991, 3% changed to paddy field and other agriculture and the other 7% changed to rubber plantation in 2000. The small red bubble shows that less than 1% of the land that was undeveloped in 1991 became developed in 2000.

3.2.2 Modeling results

The logistic regression model gave estimates of the percentages of the change in land-use to developed land and their corresponding standard errors. Standard errors were used to construct 95% confidence intervals for comparing the percentages of change to developed land for each of location by land-use type with their average percentages.

Figure 3.7 shows bar charts of the crude percentages of the change in land-use to developed land by location and by land-use group with 95% confidence intervals superimposed. The average percentage is shown by the horizontal red line with 3% for 2000 and 5% for 2009. Confidence intervals above the average line reflect the groups that were more likely to become developed land. The percentages of developed land that remained developed were higher in the 2000-2009 period than those in the 1991-2000 period. Greater land development occurred in the north, and the percentage of paddy field and other agriculture (PF+) that became developed was higher in 2000-2009 than in 1991-2000.

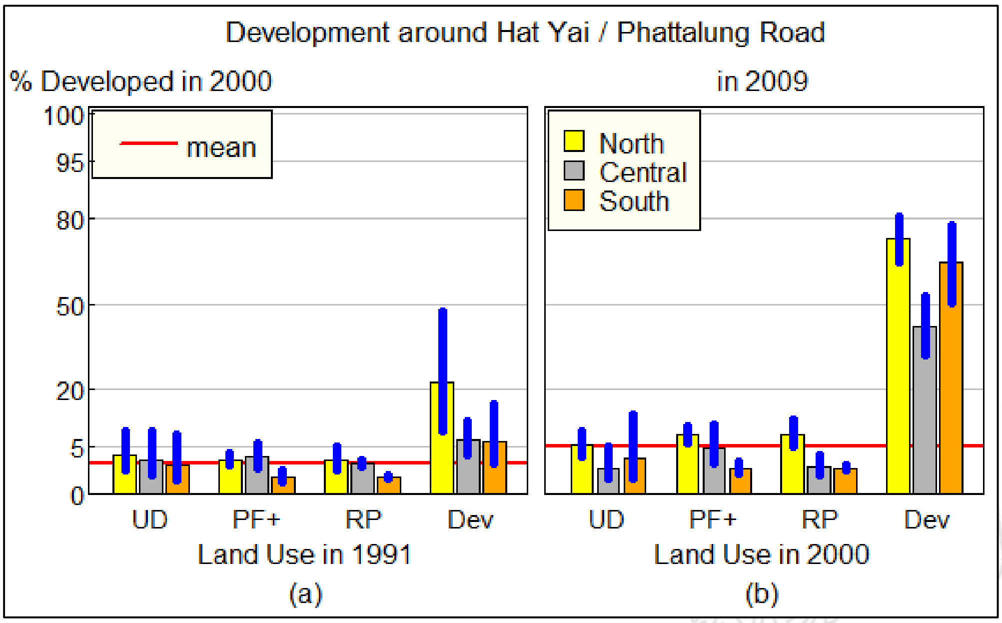


Figure 3.7 Bar charts of percentages of land that changed to developed land along the Phattalung to HatYai road from (a) 1991-2000 and (b) 2000-2009 by previous land-use and location

In Figure 3.7a, the developed land in 1991 (22.26% in the northern location, 6.23% in the central location, and 6.09% in the southern location) remained developed in 2000. Less than 4% of undeveloped land, paddy field and other agriculture and rubber plantation became developed. The 95% confidence interval for developed land in the north is substantially above the mean and it is marginally higher than the average in the central location whereas it is not different from the average in the south. Thus, the developed land in the north and the central location were more likely to remain developed land. All of the confidence intervals for the other land-use groups (UD, PF+ and RP) were lower or across the mean. Thus, these groups were less likely to become developed land, especially PF+ and RP in the south.

In Figure 3.7b, the developed lands in 2000 (73.59% in the northern, 41.86% in the central location and 65.61% in the southern) remained developed in 2009. Less than

8% of undeveloped land, paddy field and other agriculture and rubber plantation became developed. The percentages of developed land in three locations are substantially above the mean for all groups. Thus, they were more likely to remain developed. The results for other land-use groups were similar to the earlier period.

3.3 Modeling developed land in Phuket province of Thailand: 2000-2009

(**Paper III:** Modeling developed land in Phuket province of Thailand: 2000-2009 (Chuangchang *et al*., 2016b))

3.3.1 Preliminary results

The objectives of this study are to detect and evaluate land-use change and to identify the pattern of developed land that occurred in Phuket province from 2000 to 2009. The land-use categories were classified into four main groups comprising undeveloped land (UD), rubber plantation (RP), other agriculture (OA), and developed land (Dev).

The bar chart in Figure 3.8 shows the area (ha) of land-use categories in Phuket province from 2000 and 2009. Most land was used for rubber plantation, with an area of 25,250 ha in 2000 and 19,791 ha in 2009. Both developed and undeveloped land areas have been increasing, with areas of 7,834 ha in 2000, 12,391 ha in 2009, 13,280 ha in 2000 and 16,152 ha in 2009, respectively.

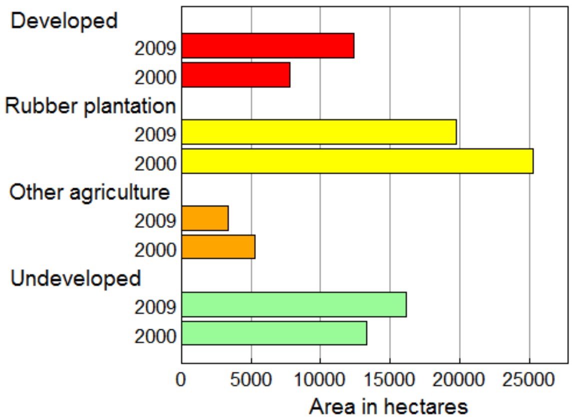


Figure 3.8 Bar chart of land-use in 2000 and 2009

The map in Figure 3.9 illustrates land-use change from 2000 to 2009. The left panel demonstrates land-use in 2000 whereas the right panel shows land-use in 2009 and the middle panel shows loss and gain of land. Undeveloped and developed land gained 4,557 ha and 2,872 ha, respectively. Other agriculture and rubber plantation lost 1,972 ha and 5,459 ha to other categories.

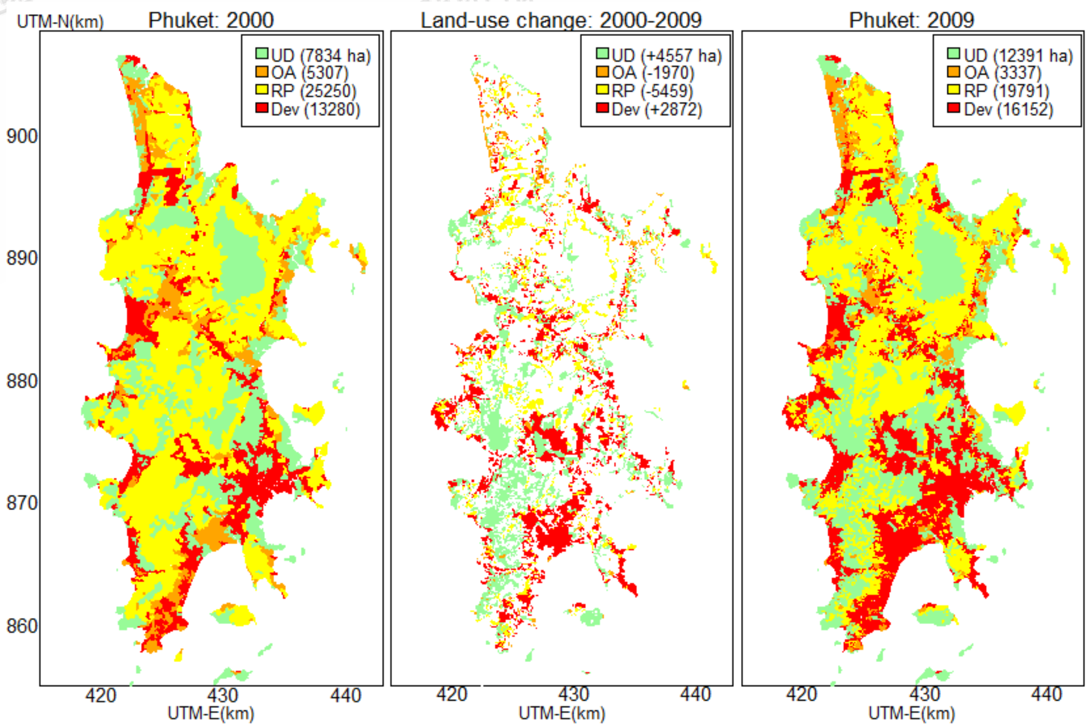


Figure 3.9 Land-use maps of Phuket province in 2000-2009

3.3.2 Modeling result

Figure 3.10 shows crude percentages of developed land classify by the two determinant factors (land-use in 2000 and sub-district) as red circles, together with their corresponding 95% confidence intervals for differences between these percentages and the average percentage shown as the horizontal red line. The average percentage of developed land in 2009 was 23.98%. The boxes show 95% confidence intervals coloured according to their location above (cyan), across (yellow) or below (pink) the mean. For each factor, the confidence intervals are adjusted for the effect of the other factor, showing a result that adjusts for any correlation between determinants. The vertical bars denote relative sample sizes. The sample sizes is number of grids that represent area in each category. The plotted values above the average line reflect the groups that were more likely to change to developed land.

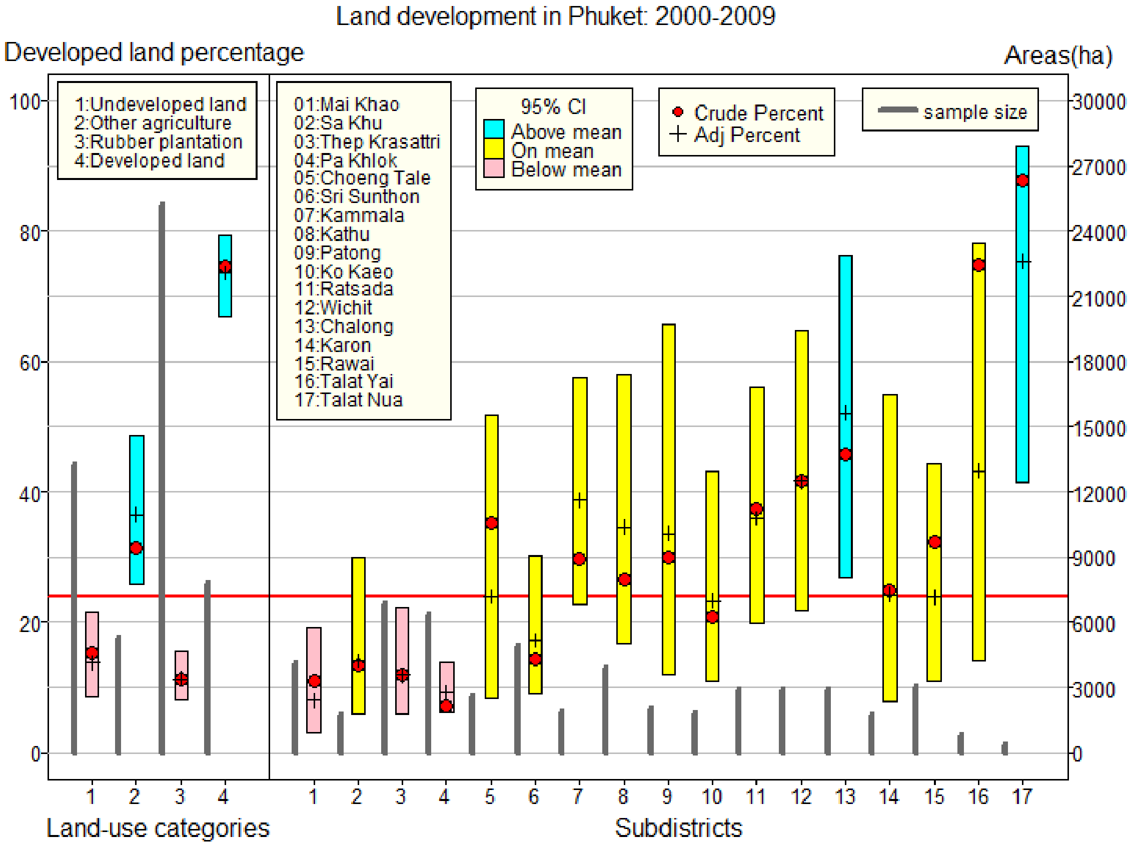


Figure 3.10 Adjusted percentages of land that changed to developed land in Phuket province from 2000-2009 with rectangles denoting 95% confidence intervals of percentage change by land-use categories and sub-district identity. Vertical bars denote sub-district areas

The 95% confidence interval for both other agriculture and developed land in 2000 are higher than the mean whereas that for rubber plantation is lower. Thus, other agriculture and developed land in 2000 were significantly more likely to change to developed land. The percentages of developed land in Chalong and Talat Nua were all substantially above the mean. Thus, these two sub-districts were more likely to have increased levels of developed land.

The full model was assessed using the ROC curve and compared with a simple model. Figure 3.11 shows the ROC curve for the simple model with just land-use group in

2000 as a single factor, together with the ROC curve for the full model with both factors included.

The cut-off point in the curve, where the observed and the predicted number of developed land in 2009 were the same, was used to report sensitivity and specificity of the model. The simple model represents AUC of 0.751, 60.6% sensitivity and 85.7% specificity whereas the full model represents AUC of 0.833, 63.1% sensitivity and 88.3% specificity. When we only compared an area above the diagonal line, the simple model had AUC of 75.1% and the full model had AUC of 83.3%. In other words, the simple model had an error of 24.9% whereas the full model had an error of 16.7%. This means our model reduced the error 8.2%. It is clear that the full model has the ability to predict number of developed land better than the simple model.

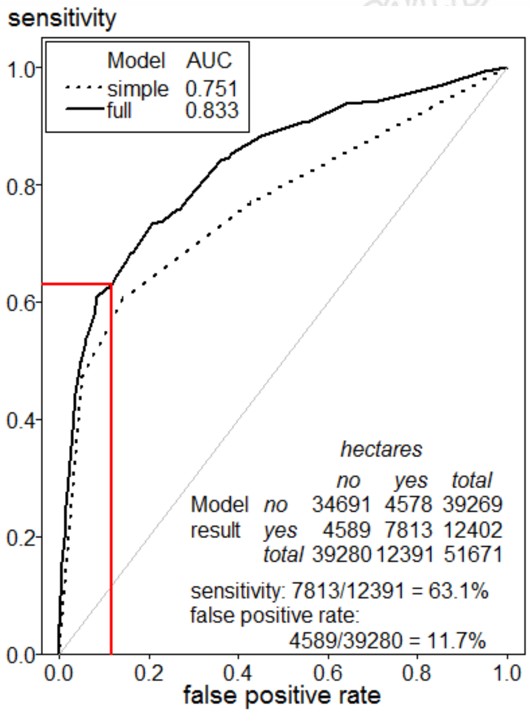


Figure 3.11 ROC curve for logistic regression models with results from full model

The map in Figure 3.12 illustrates the results from the logistic regression model. A false positive rate is when indicates that the grid was developed land in 2009 when did not (the result is *false*), while a false negative is when indicates that developed land was not occurred (the result is *negative*), but it occurred in 2009 (the result is *false*).

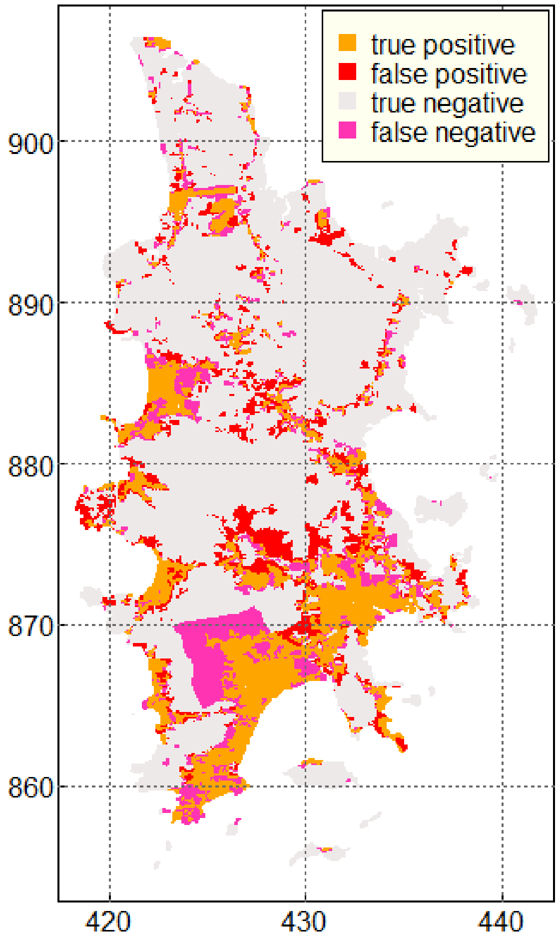


Figure 3.12 Results from the logistic regression model

Chapter 4

Summary

This thesis has described methods for analyzing land-use data from Thailand's Department of Land Development. The data are stored as polygonal “shape” files which have different plot boundaries at different points in time, and need matching to a common set of boundaries. The polygons are replaced when the land-use changes over time. Some polygons disappear, others appear, and existing ones change their shape. Therefore, land-use changes based on the polygonal data structure are difficult to measure. Statistical methods cannot be simply used because conventional statistical methods for comparing time-varying outcomes require that these outcomes are defined on the same sample space at each time period. Moreover, the polygonal data storage and analysis require high capacity computer facilities, creating difficulties for researchers with limited facilities.

The method for converting data was used in this study. It is basically an analog-to-digital conversion, replacing polygonal shapes by coded grid points. The grid size of 1 hectare is chosen. The unit of analysis is the digitized grid and only the data is changed. The digital data structure facilitates measurement of land-use change because the land-use at each year is defined on a constant grid.

The digitized data structure allows us to create map to show where and how and land-use has changed between time periods. Size of the change is easily computed by counting the grids and simply displayed using bar chart, cross-tabulation, and bubble plot. The land-use of each grid at the later time point is considered as an outcome and the land-use at the previous time point as determinant. Since data at neighboring grid points are not independent. Then, analysis of categorical data taking into account spatial correlation can be used. For a binary outcome (developed land or other), logistic regression was used.

Although logistic regression has been used in other studies of land-use research, it has not been used when a predictor is land-use in a previous year. Moreover, the method we have used to correct for the substantial spatial correlation that exists in land-use is largely new, and extends further the "variance inflation factor" method.

It is important to have a method that makes appropriate use of historical land-use data such as those available from Thailand's Department of Land Development. These data contain a lot of valuable information about history and culture development that is not available elsewhere and tends to get forgotten and lost. Such information is also valuable to planners and developers.

4.1 Findings from study samples

Findings on land development in Phuket province and Highway 4 along HatYai-Phatthalung draw from three publications.

4.1.1 Modelling Urban Growth over Time using Grid-digitized Method with Variance Inflation Factors applied to Spatial Correlation (Chuangchang *et al*., 2016a)

Greater urbanization was observed in the southern parts of Phuket province. The percentages of southern urban land in each three periods of time 1967-1975, 1975-1985 and 2000-2009 were higher than those in northern Phuket province, where 70-

90% of urban land remained urban land. The major land-use changes occurring in the area are conversion of forest land and agricultural land to urban land. But for 1985-2000 there was more urban land in the north (47% compared to 37%, approximately), a situation affected by reduced mining activity. Tin mining industry in the past has been converted to other land-use types. Agricultural land in Phuket province is mostly distributed across all areas, but with limited success since Phuket's environment was not conducive to large-scale commercial agriculture. The government has started to encourage tourism and some tin-mining entrepreneurs transferred their interest to the new industry (Na ranong, 2007). Since then, investment in the tourist industry in Phuket grew at a remarkable rate. As a result, Phuket tourism has expanded rapidly and become the main economic activity.

4.1.2 Modeling land development along highway 4 in Southern Thailand

(Chuangchang *et al*., 2014)

Average percentages of developed land increase from 2000 to 2009 (3% in 2000 and 5% in 2009). Land development occurred mostly in the north. This may be due to the proximity to Phatthalung City. Moreover, the north is also located at the intersection of two highways (Highway 4 and 41). Road networks can influence the conversion of land to developed land. That developed land occurs closer to road networks was previously reported for Lop Buri Province, Central Thailand (Patarasuk and Binford, 2012). Urban growth and conversion of agricultural land to urban area have been found to occur closer to road networks in disparate locations, including the Kansas City Metropolitan area in the central U.S.A. (Underhill, 2004), Puerto Rico (del Mar Lopez *et al*., 2001) and in Beijing, PR China (Zhang *et al*., 2002).

The other land-use categories were less likely to change to developed land, especially paddy field and other agriculture and rubber plantation in the south. This can be seen by noting the higher percentage of rubber plantations. Due to a higher expected income from rubber plantations farmers have tended to convert land to rubber plantations rather than develop their land. It is reported that in the Phatthalung watershed, a quarter of the paddy field area has been converted into rubber plantation due to the higher incomes from rubber production (Pensuk and Shrestha, 2008).

The undeveloped land was more likely to change to other categories. This has occurred elsewhere, for example, forest lands were transformed to agricultural land, particularly for shrimp farms in Pak Panang Bay (Prabnarong and Thongkao, 2006) and Ban Don Bay (Muttitanon and Tripathi, 2005). However, various factors affected the extent of land-use change at each location.

4.1.3 Modeling developed land in Phuket province of Thailand: 2000-2009

(Chuangchang *et al*., 2016b)

In this study, logistic regression is used to model land development. It shows where land development would have probably occurred in 2009. The result indicates that agricultural land and developed land were more likely to change to developed land in 2009. Our findings show a similar pattern of land-use change, especially from agricultural land to developed land, as those found elsewhere in the literature. A recent study on Banir Dar, Ethiopia by Haregeweyn *et al.* (2012) reported that the built-up areas increased from 80 ha in 1957 to 155 ha in 1944, primarily converted from the agricultural land. For Delhi, during period 1997-2008, Mohan *et al.* (2011) found that built up land increased by 17% mainly due to conversion from agricultural and waste land. For central Jordan over the period 1987-2005, Alsaaideh *et al.* (2012) found that a high percentage of agricultural lands were converted to urban areas. A similar study in Freetown, Sierra Leone by Forkuor and Cofie (2011) showed that 27% of agricultural land in 1986 was converted to residential purposes in 2000. In the Al Gharbiya governorate of Egypt from the 1972 to 2005 by Belal and Moghanm (2011) reported that urban areas increased by 7.2-5.8%, causing loss of productive agricultural lands. Another study in Puerto Rico city by Lopez *et al.* (2001) showed that rapid losses of agricultural lands occurred as a result of urban expansion since 1950.

In our study of Phuket, land development mostly occurred in Chalong and Talat Nua sub-districts. These sub-districts are located in Mueang Phuket (the capital). Chalong is located in the central southern part of Phuket, where most visitors to islands south of Phuket depart from Chalong pier. Tourists number up to 3,000 per day. The economy of the sub-district is growing. Moreover, Chalong is also located at the intersection of four roads (4021, 4022, 4024 and 4028). Road networks can influence the conversion of land to developed land. A recent study in Lop Buri province of Thailand by Patarasuk and Binford (2012) reported developed land occurs closer to road networks.

To confirm the model capability, the ROC technique is used in land-use change modeling studies (Pontius and Schneider, 2001; Hu and Lo, 2007; Wang and Mountrakis, 2011; Arsanjani *et al*., 2013). ROC curves give the proportion of positive outcomes correctly and incorrectly predicted by the model. AUC is currently considered to be the standard method to assess the accuracy of models. The AUC is well known in public health of research and there are several scales for the AUC value interpretation but, in general ROC curves with AUC below 0.75 are not clinically useful.

4.2 Conclusions

Increasing urban growth through the world has aroused concerns over the degradation of our environment. Modeling and simulation are required to understand the dynamics of complex urban systems and to evaluate the impacts of urban growth on environment.

This study provides useful information on the trend of land-use development for government planners and developers to manage land-use change in the future. The methods are useful approaches to develop an appropriate statistical model for land-use change. The thematic map, bubble and confidence intervals plots not only provide sufficient information on the land-use change but also provide quantitative information about urban change by simple geographical from freely available software R program.

The results from these regional scale assessments have provided interesting insights into the future of the region. Logistic regression could be an appropriate model for regional assessments of urban land-use change, the results of which could be used to guide more localized modeling efforts. The visualization of potential land-use change has proven to be a powerful tool for raising public awareness and facilitating discussion. The results for the unmanaged trends scenario are especially important to public discussion since they demonstrate the potential losses in resource lands that could occur if the observed rates of land-use change were to continue into the future. The result on change and patterns of change are major steps towards filling in of the information gap and creation of database for monitoring land. This effort would facilitate decision making on mitigating the impacts of land-use dynamics on resources as well as provide a basis for future research.

Location by land-use group is just one factor that contributes to land-use change. Land-use change is usually the result of a combination of multiple factors including economic, biophysical, social, and political drivers, such as income, rainfall and population dynamics (Geist and Lambin, 2002). This study used locations by land-use group as a determinant of the change to developed land.

Thailand's land-use database is updated every few years, and can provide a rich data resource for historians, property investors, environmental scientists, and planning agencies concerned with sustainable development of natural resources.

The statistical model can be easily extended to situation with several predictors. For example, land-use at a previous survey time, location, ownership, accessibility or proximity to roads and transport hubs, climate, and population density may be incorporated into a model to predict future land-use at each grid-point.

This thesis focuses on methods, which we believe are new and important, particularly because they can be integrated with remote sensing data on land cover from Earth-orbiting satellites.

4.3 Limitations

Although Thai land-use data contain a lot of valuable information, the data are not readily available later than 2009. Limited availability of recent land-use data lead to less beneficial use in policy planning. This is common a limitation in land-use research using survey data because land-use surveys are expensive and time consuming. To understand land development phenomena it needs long land-use data records. The MODIS polar-orbiting satellites data on vegetation index (NDVI) could be in cooperated.

4.4 Recommendations for further studies

The MODIS polar-orbiting satellites data are now available for NDVI. Further research should focus on using MODIS NDVI data. It can be used for further investigation of land-use change. Further investigation on forest should also be investigated for ecological conservation.

References

Achmad, A., Hasyim, S., Dahlan, B. and Aulia, D. N. 2015. Modeling of urban growth in tsunami-prone city using logistic regression: Analysis of Banda Aceh, Indonesia. Applied Geography. 62, 237-246.

Allen, J. and Lu, K. 2003. Modeling and prediction of future urban growth in the Charleston region of South Carolina: a GIS-based Integrated Approach. Conservation Ecology. 8(2), 2.

Alsaaideh, B., Al-Hanbali, A., and Tateishi, R. 2012. Assessment of land use/cover change and urban expansion of the central part of Jordan using remote sensing and GIS. Asian Journal of Geoinformatics, 11(3).

Al-shalabi, M., Pradhan, B., Billa, L., Mansor, S. and Althuwaynee, O.F. 2013. Manifestation of remote sensing data in modeling urban sprawl using the SLEUTH model and brute force calibration: a case study of Sana'a city, Yemen. Journal of the Indian Society of Remote Sensing. 41(2), 405-416.

Alsharif, A.A., and Pradhan, B. 2014. Urban sprawl analysis of Tripoli Metropolitan city (Libya) using remote sensing data and multivariate logistic regression model. Journal of the Indian Society of Remote Sensing. 42(1), 149-163.

Arsanjani, J.J., Helbich, M., Kainz, W. and Boloorani, A.D. 2013. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. International Journal of Applied Earth Observation and Geoinformation. 21, 265-275.

Belal, A.A. and Moghanm, F.S. 2011. Detecting urban growth using remote sensing and GIS techniques in Al Gharbiya governorate, Egypt. The Egyptian Journal of Remote Sensing and Space Science. 14(2), 73-79.

Breiman, L. 2001. Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statistical Science. 16(3), 199-231.

Cheng, J. and Masser, I. 2003. Urban growth pattern modeling: a case study of Wuhan city, PR China. Landscape and urban planning. 62(4), 199-217.

Chuangchang, P. and Tongkumchum, P. 2014. Modeling land development along highway 4 in Southern Thailand. Songklanakarin Journal of Science & Technology. 36(6), 719-725.

Chuangchang, P., Thinnukool, O., T. and Tongkumchum, P. 2016a. Modelling Urban Growth over Time using Grid-digitized Method with Variance Inflation Factors applied to Spatial Correlation. Arabian Journal of Geosciences. (In process). Accepted in February 2016.

Chuangchang, P., Sangkhaduang, T. and Tongkumchum, P. 2016b. Modeling Developed Land in Phuket Province of Thailand: 2000-2009. Journal of Social Sciences & Humanities. 24(2), (In process). Accepted in June 2015.

Costanza, R. and Ruth, M. 1998. Using dynamic modeling to scope environmental problems and build consensus. Environmental management. 22(2), 183-195.

del Mar Lopez, T., Aide, T.M. and Thomlinson, J.R. 2001. Urban expansion and the loss of prime agricultural lands in Puerto Rico. Ambio: a Journal of the Human environment. 30(1), 49-54.

Donoho, D. 2015. 50 years of Data Science. Based on a presentation at the Tukey Centennial workshop, Princeton NJ, September 18, 2015, 1-41.

Eyoh, A., Olayinka, D.N., Nwilo, P., Okwuashi, O., Isong, M. and Udoudo, D. 2012. Modelling and predicting future urban expansion of Lagos, Nigeria from remote sensing data using logistic regression and GIS. International Journal of Applied Science and Technology. 2(5), 116.

Forkuor, G. and Cofie, O. 2011. Dynamics of land-use/land-cover change in Freetown Sierra Leone and its effects on urban and peri-urban agriculture- A remote sensing approach. International Journal of Remote Sensing. 32(4), 1017-1037.

Geist, H.J. and Lambin, E.F. 2002. Proximate Causes and Underlying Driving Forces of Tropical Deforestation Tropical forests are disappearing as the result of many pressures, both local and regional, acting in various combinations in different geographical locations. BioScience. 52(2), 143-150.

Hanley, J.A. and McNeil, B.J. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 143, 29-36.

Haregeweyn, N., Fikadu, G., Tsunekawa, A., Tsubo, M. and Meshesha, D.T. 2012. The dynamics of urban expansion and its impacts on land use/land cover change and small-scale farmers living near the urban fringe: A case study of Bahir Dar, Ethiopia. Landscape and urban planning. 106(2), 149-157.

Herold, M., Menz, G. and Clarke, K.C. 2001. Remote sensing and urban growth models-demands and perspectives. In Symposium on remote sensing of urban areas, 35.

Helmer, E.H. 2004. Forest Conservation and land development in Puerto Rico. Landscape Ecology. 19, 29-40.

Hosmer, D.W. and Lemeshow, S. 2000. *Applied logistic regression*. John Wiley & Sons.

Hu, Z. and Lo, C.P. 2007. Modeling urban growth in Atlanta using logistic regression. Computers, Environment and Urban Systems. 31(6), 667-688.

Huang, W., Liu, H., Luan, Q., Jiang, Q., Liu, J. and Liu, H. 2008. Detection and prediction of land use change in Beijing based on remote sensing and GIS. The International Archives of the Photogrammetry. Remote Sensing and Spatial Information Sciences. 37, 75-82.

Jiang, Y., Fu, P. and Weng, Q. 2015. Assessing the impacts of urbanization-associated land use/cover change on land surface temperature and surface moisture: a case study in the Midwestern United States. Remote Sensing. 7(4), 4880-4898.

Kamyab, H. R. and Salman Mahiny, A. 2013. Modeling urban growth effects on landscape structure in Gorgan city area. Ecopersia. 1(2), 99-112.

Lambin, E.F., Rounsevell, M.D.A., Geist, H.J. 2000. Are agricultural land-use models able to predict changes in land-use intensity? Agriculture, Ecosystems and Environment. 82, 321-331.

Liu, H. and Zhou, Q. 2005. Developing urban growth predictions from spatial indicators based on multi-temporal images. Computers, Environment and Urban Systems. 29(5), 580-594.

McNeil, D. 2014. Epidemiological Statistics-Further Issues. Wiley StatsRef: Statistics Reference Online. DOI: 10.1002/9781118445112.stat00321.pub2

McNeil, N. and Chooprateep, S. 2014. Modeling sea surface temperatures of the North Atlantic Ocean. Theoretical and applied climatology. 116(1-2), 11-17.

Mohan, M., Pathan, S.K., Narendrareddy, K., Kandya, A. and Pandey, S. 2011. Dynamics of urbanization and its impact on land use/land cover: A case study of Megacity Delhi. Journal of Environmental Protection. 2, 1274-1283.

Muttitanon, W. and Tripathi, N.K. 2005. Land use/land cover changes in the coastal zone of Ban Don Bay, Thailand using Landsat 5 TM data. International Journal of Remote Sensing. 26(11), 2311-2323.

Na ranong, A. 2007. Economic changes from tin mining to tourist industry in Phuket 1941-1987. Master of Art thesis. Department of History, Graduate School, Silpakorn University.

Nong, Y., and Du, Q. 2011. Urban growth pattern modeling using logistic regression. Geo-Spatial Information Science. 14(1), 62-67.

Overmars, K.P., De Koning, G.H.J. and Veldkamp, A. 2003) Spatial autocorrelation in multi-scale land use models. Ecological modelling. 164(2), 257-270.

Patarasuk, R. and Binford, M.W. 2012. Longitudinal analysis of the road network development and land-cover change in Lop Buri province, Thailand, 1989-2006. Applied Geography. 32(2), 228-239.

Pijanowski, B.C., Brown, D.G., Shellito, B.A. and Manik, G.A. 2002. Using neural networks and GIS to forecast land use changes: a land transformation model. Computers, Environment and Urban systems. 26(6), 553-575.

Polovina, J.J., Howell, E.A. and Abecassis, M. 2008. Ocean's least productive waters are expanding. Geophysical Research Letters. 35(3).

Pontius, R.G. and Schneider, L.C. 2001. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems and Environment. 85, 239-248.

Prabnarong, P. and Thongkao, S. 2006. Land Use Changes in Pak Panang Bay using Arial Photographs and Geographic Information System. Walailak Journal of Science and Technology (WJST). 3(1), 93-104.

R Development Core Team. 2012. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for statistical Computing. Available online: http://www.r-project.org [September 6, 2012].

Rao, J.N.L. and Scott, A.J. 1992. A sample method for the analysis of clustered binary data. Biometrics. 48(2), 577-585.

Ripley, B.D. 1996. *Pattern recognition and neural networks*. Cambridge university press.

Sakar, S. and Midi, H. 2010. Importance of assessing the model adequacy of binary logistic regression. Journal of Applied Sciences. 10(6), 479-486.

Stehman, S.V. and Wickham, J.D. 2011. Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. Remote Sensing of Environment. 115(12), 3044-3055.

Sudhira, H.S., Ramachandra, T.V. and Jagadish, K.S. 2004. Urban sprawl: metrics, dynamics and modelling using GIS. International Journal of Applied Earth Observation and Geoinformation. 5(1), 29-39.

Swangjang, K. and Iamaram, V. 2011. Change of Land Use Patterns in the Areas Close to the Airport Development Area and Some Implicating Factors. Sustainability. 3, 1517-1530.

Takahashi, K. Uchiyama, H. Yanagisawa, S. Kamae, I. 2006. The logistic regression and ROC analysis of group-based screening for predicting Diabetes incidence in four years. Kobe Journal of Medical Sciences. 52(6), 171-180.

Tayyebi, A., Perry, P.C. and Tayyebi, A.H. 2014. Predicting the expansion of an urban boundary using spatial logistic regression and hybrid raster-vector routines with remote sensing and GIS. International Journal of Geographical Information Science. 28(4), 639-659.

Thinnukool, O., Chuangchang, P., and Kongkraphan, K. 2014. Analyzing land-use change using grid-digitized method. Songklanakarin Journal of Science and Technology. 36 (2), 235-240.

Tongkumchum, P. and McNeil, D. 2009. Confidence intervals using contrasts for regression model. Songklanakarin Journal of Science and Technology. 31(2), 151-156.

Triantakonstantis, D. and Stathakis, D. 2015. Urban Growth Prediction in Athens, Greece, Using Artificial Neural Networks. International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering. 9(3), 193-197.

Tukey, J.W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Underhill, K.K. 2004. Land cover dynamics and urban sprawl indicators: a casestudy in the Kansas City Metropolitan Area. Unpublished Ph.D. Dissertation,University of Missouri e Kansas City, Kansas City.

Wang, J. and Mountrakis, G. 2011. Developing a multi-network urbanization model: a case study of urban growth in Denver, Colorado. International Journal of Geographical Information Science. 25, 229-253.

Xu, Y., Knudby, A. and Ho, H.C. 2014. Estimating daily maximum air temperature from MODIS in British Columbia, Canada. International Journal of Remote Sensing. 35(24), 8108-8121.

Zeger, S. and Liang, K.Y. 1986. Longitudinal data analysis for discrete and continuous outcomes. Biometrics. 42, 121-130

Zeng, Y.N., Wu, G.P., Zhan, F.B. and Zhang, H.H. 2008. Modeling spatial land use pattern using autologistic regression. The international archives of the photogrammetry remote sensing and spatial information sciences. 115-118.

Zhang, Q., Wang, J., Peng, X., Gong, P., and Shi, P. 2002. Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data. International Journal of Remote Sensing. 23(15), 3057-3078.

Appendix I

Modelling Urban Growth over Time using Grid-digitized Method with Variance Inflation Factors applied to Spatial Correlation

Potjamas Chuangchang, Orawit Thinnukool and Phattrawan Tongkumchum*

**Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Mueang, 94000 Thailand*

Abstract

Analysis of land-use change over time is useful information to support urban planning and management policies. Most land use modelling studies have used polygonal data structure. The main limitation of polygonal data structure is that it is difficult to measure changes in land-use. This study proposes another method for predicting land-use change. This method is based on an analog-to-digital conversion which replaces polygonal shapes by coded grid points. The method is applied to data from a survey of Phuket province from 1967 to 2009 where land-use was classified broadly as forest, agriculture, urban, water bodies and miscellaneous land. Logistic regression was used to predict a binary land-use outcome (urban/other), and location combined with land-use at a previous survey was a determinant. To account for correlation in land-use amongst nearby plots of land, variance inflation factors were used to compute standard errors of proportions of urban growth. The result of the present study discloses that greater urbanization was observed in the southern parts of Phuket during the period of study and surprisingly, that reforestation occurred in 1985-2000. This study shows that analog-to-digital conversion methods are useful approaches to develop appropriate statistical models for land-use change.

Keywords: Urban growth, Digitization, Logistic regression, Land-use, Phuket province

Introduction

Land-use data contain rich and valuable information about historical and cultural developments. Such data provide essential information relating to land development, a subject of intensive research in remote sensing and geographic information systems (GIS) as well as in economics (see for example, Bach et al.(2006); Lewis (2010); Stehman and Wickham (2011); Guo et al.(2013)).

It is important to have methods that make appropriate use of land-use data. Improved knowledge of historical land-use not only improves our ability to manage the legacy of this history but also provides an insight into, or confirmation of, the likely impacts of current and future urbanization on land-use, environment and ecosystem (Sheffield and Morse-McNabb 2012). The quantity and the location of land-use changes are the main issues to be addressed by city planners and decision makers, especially in a rapidly changing environment. Thus, the main objective of the modelling process is to understand and to predict future urban growth. This paper focuses on methods for analysing and graphing land-use data from regular surveys undertaken by the Thai Department of Lands, and to use this method to explore land-use change over time.

Land-use research is of current scientific interest due to the availability of data from remote sensing, widespread use of global positioning systems (GPS) and the availability of GIS software. These data comprise hundreds of land-use plots. The boundaries of the plots create the polygons that can be stored in database tables. Then, the polygonal data structure can be created. GIS information such as imagery and land properties can be extracted using GIS software.

Most land-use studies have used commercial software based on polygonal data structure. Data storage and analysis require high-capacity computer facilities, creating difficulties for researchers with limited facilities. Land-use changes based on the polygonal data structure are difficult to measure.

Polygonal data structures can provide thematic maps to display patterns of land-use type, but the data are difficult to analyse because the polygons change. The polygons are replaced when the land-use changes over time. Some polygons disappear, others appear, and existing ones change their shape. Polygonal data thus have different plot boundaries at different periods and need matching to a common set of boundaries. Statistical methods to measure change over time cannot be simply used because statistical methods for comparing time-varying outcomes require that these outcomes are defined on the same sample space at each time period.

GIS analysis of land-use is based on not only polygons but also pixels. It has been argued that options vary on the appropriate unit for assessing land-use change, and an ideal spatial unit does not exist (Stehman and Wickham, 2011). Knowledge gaps here are how land-use is defined for its suitability for statistical analysis and what appropriate unit of analysis can be used to simply detect change and measure its accuracy.

The limitation of polygonal data structure for statistical analysis has been documented (Thinnukul et al. 2014; Chuangchang and Tongkumchum 2014). The data structure can be improved by conversion to digital structure. The data can be recorded as points on a grid, for which land-use change is easily measured because the grid stays put while only the data change. The grid data structure can be used directly for statistical analysis of land-use change. Digitized data structure is thus preferable to polygonal data structure for spatial data analysis.

Recent studies have developed new and improved models of changes in a specific category such as urban growth based on remote sensing and GIS to explain future development (Sudhira et al. 2004; Liu and Zhou 2005; Eyoh et al. 2012; Al-sharif and Pradhan 2015). Many modelling approaches have been used to analyse and predict urban growth trends, and some of these approaches are artificial neural networks (Triantakonstantis and Stathakis 2015), cellular automata (Li et al. 2014; Feng et al. 2015) and logistic regression (Alsharif and Pradhan 2014; Tayyebi et al. 2014; Achmad et al. 2015). It shows that scientific knowledge of land-use change is well developed in research areas of geoscience, computing, image processing and remote sensing, whereas statistical methods, particular time series modeling, have not yet been applied to the same extent.

Modelling of urban growth is regarded as an efficient way to understand the mechanisms of urban dynamics, to evaluate current urban systems and to provide planning support in urban growth management, where land-use models can help to build future growth scenarios and to assess possible environmental impacts (Lambin et al. 2006). Land-use change and urban growth models are important tools that support planning and development of sustainable urban areas. These models provide a mechanism for exploratory analysis of detailed case studies. They can help to identify the key driving variable behind land-use changes and be readily developed for a particular geographic place and time. Typically, regression methods are used to describe spatial and temporal aspects of land-use change.

Logistic regression has been used in urban growth modelling (Wu and Yeh 1997; Landis and Zhang 1998; Allen and Lu 2003; Cheng and Masser 2003; Hu and Lo 2007; Nong and Du 2011; Eyoh et al. 2012; Alsharif and Pradhan 2014; Tayyebi et al. 2014). Logistic regression is an appropriate model for analysis of binary data (Hosmer and Lemshow 2004). The model provides estimated effects of independent variables and their precisions. However, to our knowledge, it has not been used when a predictor is land-use type in previous year. Urban growth modelling aims to explain the dynamic processes related to it, and therefore interpretability of models is becoming crucial for gaining knowledge of the processes driving the change of spatial patterns. However, spatial correlation, which causes violation of the assumption of independent residuals, is often difficult to handle (Hu and Lo 2007). Violation of the independence assumption results in incorrect standard errors of the estimates. It is therefore essential to have a method to handle spatial correlation.

In other studies with geographical data (see, for example, McNeil and Chooprateep, 2014) such spatial correlation is handled by aggregating data into larger regions with acceptably small correlation between adjoining regions, or by using factor analysis and multivariate regression to adjust for spatial correlations. The generalized estimating equations (GEE) (Zeger and Liang, 1986) method could be used by dividing the region into groups of plots (clusters) and estimating common fixed correlations between plots in the same group. However, this method has difficulties due to large cluster size and it assumes a common correlation within land-use polygons.

Another option is a conventional method widely used in survey sampling, based on variation inflation factors (Rao and Scott, 1992). It is a simple method for analysis of clustered binary data based on the concept of design effect and effective sample size widely used in sample surveys. It assumes no specific models for intra-cluster correlation. This method avoids the problem in the GEE by computing effective sample sizes for each land-use plot based on their sample variances, from which standard errors are applied to fitted values from a logistic model to compute confidence intervals.

This study offers a method for predicting land-use change over time. This method is based on analog-to-digital conversion, replacing polygonal shapes by coded grid points. Logistic regression can thus be used to analyse the data, and a method based on variance inflation factors can also be used to handle the problems arising from spatial correlations. The method is illustrated using urban growth in Phuket province of Thailand. Understanding land-use change patterns helps in planning for effective natural resource utilization and provision of infrastructure facilities.

Materials and methods

Data and study area

Land-use data of 1967, 1975, 1985, 2000 and 2009 for Phuket province were obtained from Thailand's Land Development Department. The land-use database comprises polygonal "shape files" of land-use plots recorded at regular surveys of every province. The data can be read and displayed using GIS software. Alternatively, the data files can be restructured into relational database tables and used with a general purpose program such as R.

The study area is Phuket, one of the urban tourism expansion cities in the south of Thailand. The province derives much of its income from tourism.

Phuket province includes the largest island in Thailand plus nearly 30 small islets. It is located in the tropical zone of the west coast of the southern part of Thailand in the Andaman Sea (Fig. 1). The main island itself has an area of about 53,900 ha (539 km^2). The province comprises about 55% undulating-rolling topography and 36% mountainous, mostly located in the south.

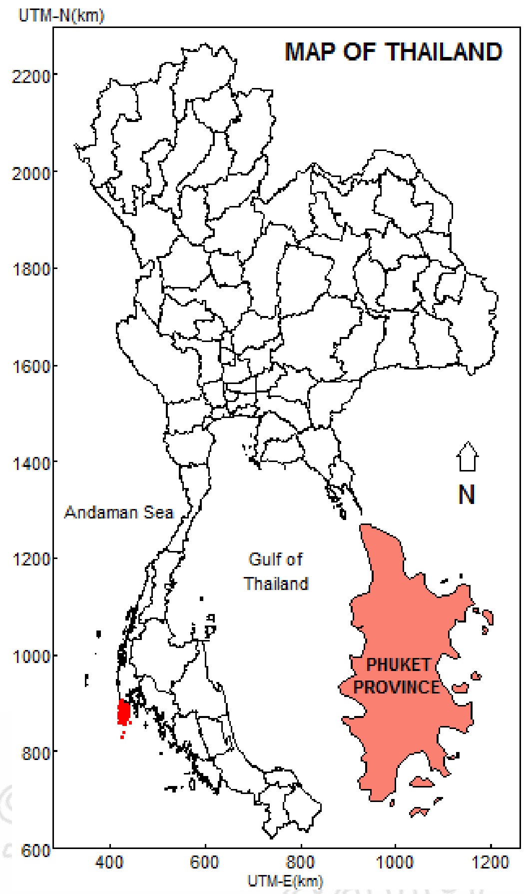


Fig. 1 Phuket province (red) of Thailand.

Land-use categories

According to Thailand's Land Development Department, land-use is generally classified into five categories (level 1) comprising urban and built-up land (U), agricultural land (A), forest land (F), water bodies (W) and miscellaneous land (M). Each land-use category is further divided into subcategories (level 2 and level 3). Some classification levels are shown an example in Table 1. For example, urban and build-up land (U) of level 1 is classified into six types of land-use of level 2 comprising city, town, commercial (U1), village (U2), institutional land (U3), transportation, communication and utility (U4), industrial land (U5) and other (U6). Some of level 2 such as U2, U4, U5 and U6 are further divided into subcategories of level 3. For example, U2 comprises abandoned village (U200), village (U201), and hill tribe village (U202).

Table 1 Examples of land-use classifications in Thailand.

Level 1	Level 2	Level 3
U: Urban and built up land	U1: City, Town, Commercial	U1: City, Town, Commercial
	U2: Village	U200: Abandoned village U201: Village U202: Hill tribe village
	U3: Institutional land	U3: Institutional land
	U4: Transportation, Communication and Utility	U401: Airport U402: Railway station U403: Bus station U404: Harbour U405: Road
	U5: Industrial land	U500: Abandoned factory U501: Industrial estate U502: Factory U503: Agricultural product trading centers
	U6: Other	U600: Abandoned area U601: Recreation area U602: Golf course U603: Cemetery U604: Refugee camp
A: Agricultural land	A1: Paddy field	A100: Abandoned paddy field A101: Rice paddy
	A2: Field crop	A200: Abandoned field crop
	⋮	⋮
	A9: Aquacultural land	A905: Crocodile farm
F: Forest land	F1: Evergreen forest	F100: Disturbed evergreen forest F101: Dense evergreen forest
	⋮	⋮
	F6 : Agro - forestry	F6 : Agro - forestry
W: Water bodies	W1: Natural water body	W101: River, Canal W102: Natural water resource
	⋮	⋮
	W2: Reservoir (Built-up)	W203: Irrigation canal
M: Miscellaneous land	M1: Rangeland	M101: Grass M102: Scrub M103: Bamboo
	⋮	⋮
	M4: Other	M405: Landfill

Universal Transverse Mercator (UTM) coordinates were used as an input data file to create the land-use thematic map. The UTM coordinates are based on meters as unit of measure, which incorporates the simplicity of the decimal system, and it is easy to comprehend units of ten and always have the same two directional designators that never carry negative values. Thematic maps of land-use categories for Phuket province in 2009 are shown in Fig. 2. This map is created from a digitized grid data structure. The total area is 53,300 ha. This is the common area after fixing the coordinate shift from year to year due to changes in GPS settings. The left panel shows five categories of level 1. In this level 23,394 ha is agriculture (A), 12,687 ha is forest (F) while 3,661, 12,786 and 772 ha are miscellaneous land (M), urban (U) and water (W), respectively. The right panel shows land-use categories for a sub-region of level 2 and 3. For agricultural land, it is 672 ha of paddy field (A1). Paddy fields comprise 435 ha of abandoned paddy fields (A100) and 237 ha of rice paddies (A101). The map based on the digitized grid not only provides the location of each land-use type but also provides the area of land.

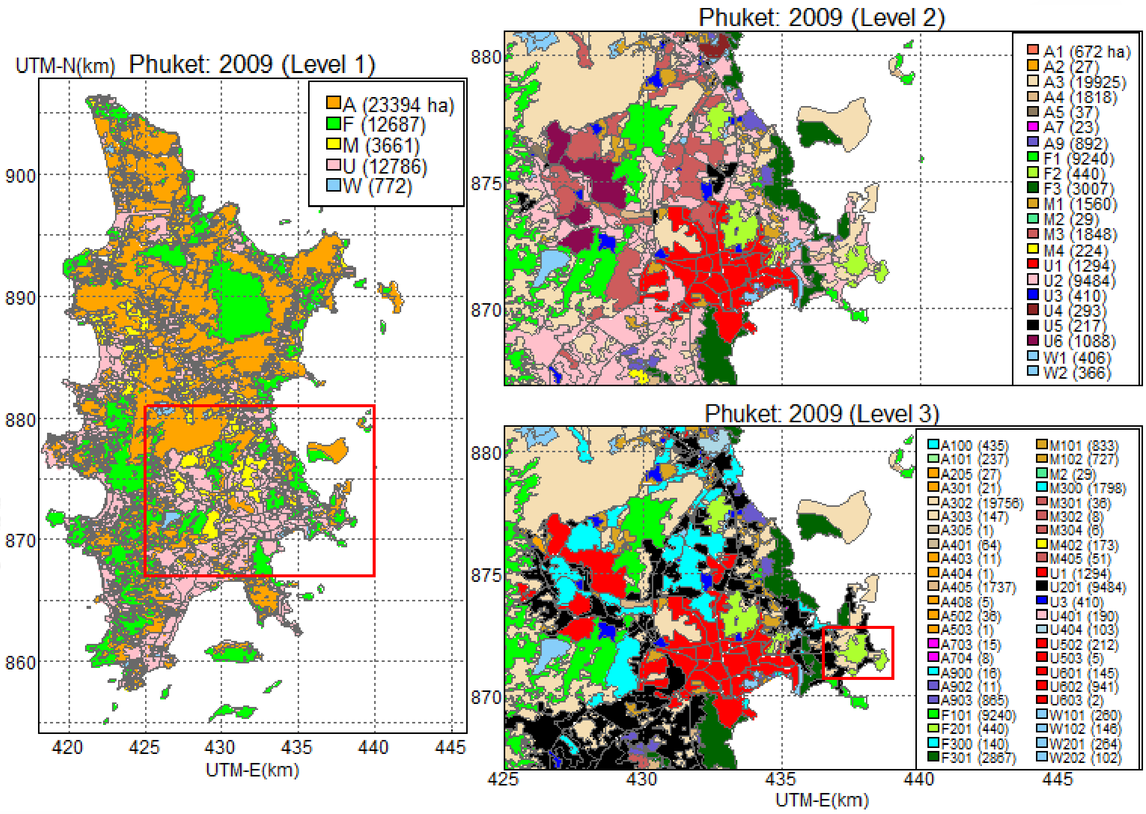


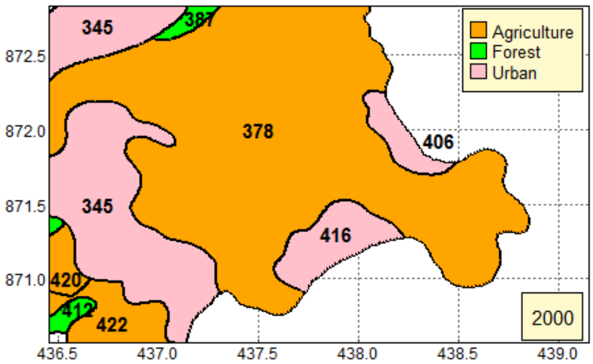
Fig. 2 Classification of land-use in Phuket province in 2009: urban land (U), agricultural land (A), forest land (F), water (W) and miscellaneous land (M). Left panel shows five categories of level 1. Right panel shows land-use categories for a sub-region of level 2 and 3.

Data structure

Polygonal data structure

Land-use plots as polygons can be stored in database tables. Fig. 3 shows an example of polygonal data structure of land-use as an inset map in the bottom right panel in Fig. 2. Polygonal data structure stores the polygonal plot identified as plotID, land-use codes (luCode), and area of each polygon in one table. Points of coordinates x and y for drawing the polygon identified as pointID are stored in another table. PlotID field is a primary key linking these two tables. It comprises eight polygonal plots identified as PlotIDs 345, 378, 387, 406, 412, 416, 420 and 422 comprising three categories of land-use recorded as forest, agricultural land and urban. PlotID 345 is urban (U), whereas plotID 422 is agricultural land (A). For each plotID, the right panel shows pointID comprising coordinates x and y for drawing the polygon. The pointID field determines the order in which the boundary points (x,y) are connected to obtain a closed polygon for each land-use plot. For example, there are 447 points connected to obtain a closed polygon from plotID 422. The whole of Phuket province consists of 571 such polygons.

plotID	luCode	area	desEng
1	F	1881776	Forest
2	U	532235.9	Urban
3	F	16876	Forest
4	F	50776.61	Forest
:	:	:	:
345	U	23098570	Urban
378	A	4191809	Agriculture
387	F	338592.3	Forest
406	U	89158.15	Urban
412	F	740800.6	Forest
416	U	187198.6	Urban
420	A	85411.61	Agriculture
:	:	:	:
422	A	293820.8	Agriculture
570	A	36782.09	Agriculture
571	A	195943.2	Agriculture



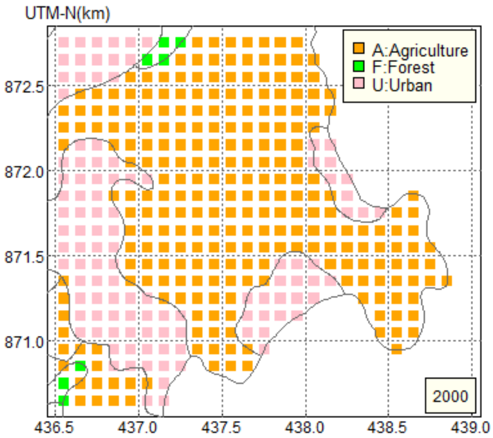
plotID	pointID	x	y
345	1	433.20	878.28
345	2	433.21	878.28
:	:	:	:
345	11570	433.20	878.28
378	1	437.48	873.46
378	2	437.49	873.45
:	:	:	:
378	1698	437.48	873.46
387	1	437.29	873.78
387	2	437.30	873.78
:	:	:	:
387	429	437.29	873.78
406	1	438.49	871.78
406	2	438.49	871.77
:	:	:	:
406	191	438.49	871.78
412	1	436.03	871.86
412	2	436.04	871.86
:	:	:	:
412	773	436.03	871.86
416	1	437.73	870.90
416	2	437.71	870.92
:	:	:	:
416	228	437.73	870.90
420	1	436.53	871.37
420	2	436.54	871.36
:	:	:	:
420	202	436.53	871.37
422	1	436.43	870.92
422	2	436.44	870.92
:	:	:	:
422	447	436.43	870.92

Fig. 3 Polygonal data structure; the right panel lists polygonal data for creating the left panel

Digital data structure

Digital data structure involves creating a digitized grid of geographical coordinates covering the whole of the study area and storing the land-use codes and plot identifiers as fields in database tables indexed by the grid coordinates. The digitized land-use data are recorded at points on a square grid of specified dimension. We selected 100 meter intervals digitizing land-use data (100 by 100 meter). Altitude data are gridded at 111 meter intervals (1° of latitude = 111 km (Polovina et al. 2008)), so that one degree of latitude contains 100 grid points. Having such a database overcomes many difficulties land-use researchers face when measuring and predicting land-use change.

The digital data structure facilitates measurement of land-use change because the land-use at each year is defined on a constant grid. Fig. 4 shows the digital data structure of the same area in Fig. 3. The data table on the right panel contains plotID, pointID and luCode. The plotID refers to polygon, pointID with x and y coordinates here refers to grid point and luCode is land-use categories. For example, plotID 387 contains four grid boxes. So, the digitized land-use data are recorded at points on a square grid with an area of 1 ha. The area of each land-use type can be calculated simply as the number of grid points in the polygon. The area of plotID 387 is 4 ha because it comprises four grid points.



plotID	pointID	x	y	luCode
345	1	436.55	871.45	U
345	2	436.55	871.55	U
⋮	⋮	⋮	⋮	⋮
345	82	437.25	871.15	U
378	1	436.55	871.15	A
378	2	436.55	871.25	A
⋮	⋮	⋮	⋮	⋮
378	239	438.85	871.35	A
387	1	437.05	872.65	F
387	2	437.15	872.65	F
⋮	⋮	⋮	⋮	⋮
387	4	437.25	872.75	F
406	1	438.05	872.05	U
406	2	438.05	872.15	U
⋮	⋮	⋮	⋮	⋮
406	10	438.45	871.75	U
412	1	436.55	870.65	F
412	2	436.55	870.75	F
412	3	436.65	870.85	F
416	1	437.65	871.05	U
416	2	437.65	871.15	U
⋮	⋮	⋮	⋮	⋮
416	19	438.15	871.35	U
420	1	436.55	870.95	A
420	2	436.55	871.05	A
⋮	⋮	⋮	⋮	⋮
420	5	436.55	871.35	A
422	1	436.55	870.85	A
422	2	436.65	870.65	A
⋮	⋮	⋮	⋮	⋮
422	13	437.05	870.75	A

Fig. 4 Example of digital data structure; the right panel lists grid-point data for creating the left panel.

The digital structure is appropriate for analysis of land-use change, which can be measured in terms of loss and gain from two periods of time as shown in Fig. 5. Where land-use in 2000 and 2009 are shown in the left panels, losses from 2000 and gains in 2009 are shown in the right panels. The total area of each panel is 550 ha. In 2000, land area is 375 ha and sea (Z) area is 175 ha. In 2009, land area is 369 ha and sea (Z) area is 181 ha. The area of losses in 2000 and gains in 2009 is 198 ha. The top right panel shows only the land-use in 2000 that was lost, mainly from agricultural land; orange colour indicates areas of agriculture loss. The bottom right panel shows that changes in 2009 were mostly due to gains in forest and urban areas.

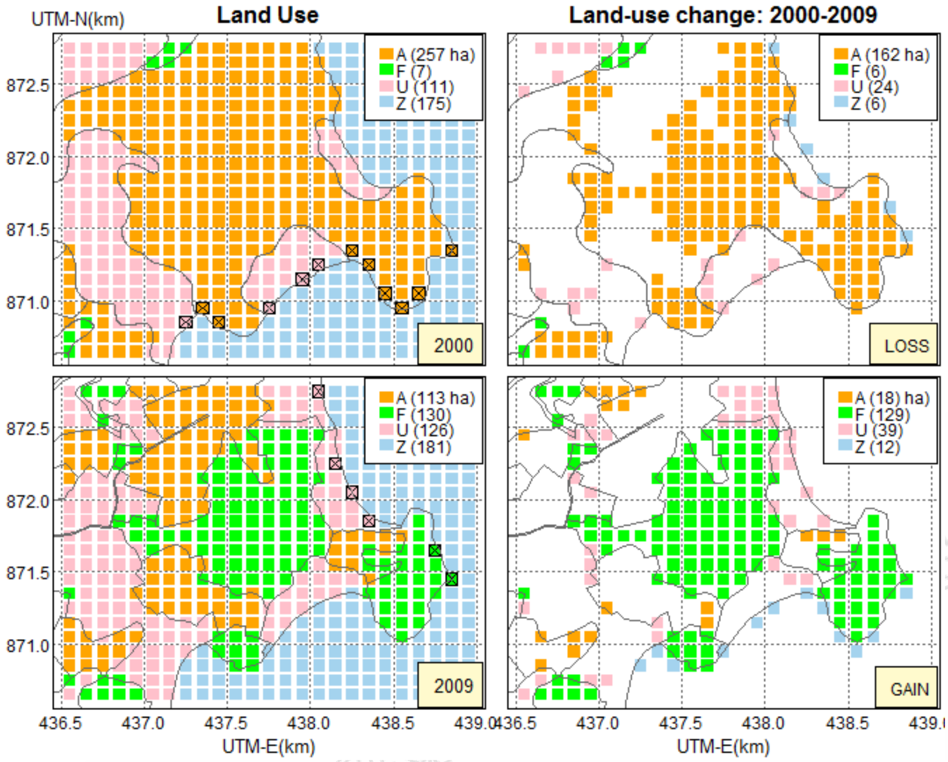


Fig. 5 Land-use change in area from 2000-2009 with losses from 2000 (upper right panel) and gains to 2009 (lower right panel).

A digital data structure provides a simple calculation for land-use change. Focusing only on land, it is 375 ha in 2000 and 369 in 2009. Note that the shape of the study area changed over the 9-year period, which complicates the measurement of land-use change. This problem can be avoided by considering change with respect only to the land present in both 2000 and 2009. This common area is 363 ha. The area of 12 ha in 2000 and 6 ha in 2009 marked as ⊠ in the maps of Fig. 5 are omitted. Table 2 shows a corresponding cross-tabulation of three categories of land-use in 2000 and 2009. It displays area of changes in land-use and area remaining the same. The area of urban land increased from 107 to 122 ha over the period, and that of forest increased from 7 to 128 ha, while that of agricultural land decreased from 249 to 113 ha. In 2000, 154 ha of agricultural land was lost: 120 ha to forest and 34 ha to urban.

Table 2 Land-use change in area: 2000-2009.

		2009			
		Agriculture	Forest	Urban	total
	Agriculture	95	120	34	249
2000	Forest	5	1	1	7
	Urban	13	7	87	107
	total	113	128	122	363

Graphical methods

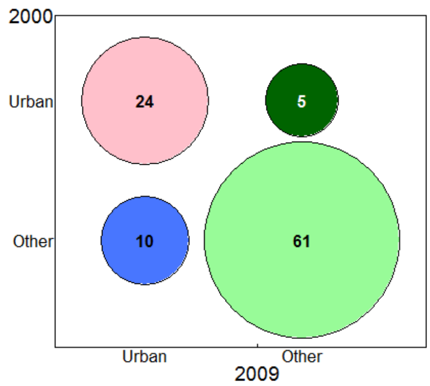
Bubble plots

Bubble plots and cross-tabulation are effective methods to summarize change in land-use. They give area or percentages of land-use categories from one period to the next. While land-use for a given year can be graphed as a thematic map using a separate colour for each category, as shown in Fig. 2, graphing land-use change is more complicated because many additional colours are needed to show data from a cross-tabulation such as Table 2. To simplify this task, we combined land-use type into two categories (U and A+F). As a result, the cross-tabulation is reduced to a 2×2 table requiring only four colours, as shown in Fig. 6 using a bubble plot.

Change in land-use is effectively summarized in a cross-tabulation giving area in hectares (Fig. 6a) or percentages of land-use categories from one period to the next. These numbers can be displayed as a bubble plot matrix (Fig. 6b). In our study, we show area in hectares using cross-tabulation and percentages of changes using the bubble plot.

		2009		
		Urban(U)	Other(A+F)	total
2000	Urban(U)	87	20	107
	Other(A+F)	35	221	256
	total	122	241	363

(a)



(b)

Fig. 6 Table and bubble plot showing land-use change: (a) change of land-use in 2000-2009 (ha) and (b) percentage of land-use categories.

Thematic map

Land-use data can be displayed using thematic maps. These maps can be used to represent the distributions of land-use change and highlight variations. Thematic maps can show not only the distribution of land-use but also where changes occur. Fig.7 shows percentages of land-use change in the study area from 1967 to 2009 using a thematic map corresponding to the bubble plot in Fig. 6. The thematic map divides land-use change into four colours, pink for urban land that remained urban (U-U), blue for other land that changed to urban land (O-U), dark green for urban that changed to other land (U-O) and green for other land that remained other land (O-O). Although the bubble shows similar information as the thematic map, it can give a quick comparison of sizes of changes.

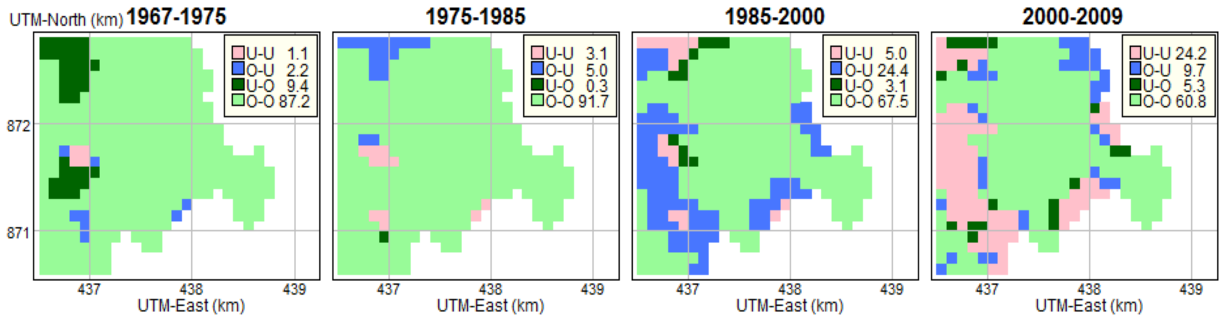


Fig. 7 Land-use percentage changes: 1967-2009.

Statistical methods

A statistical model provides estimates of land use change together with measures of precision of the estimates. The statistical significance of the change is also provided.

Statistical modelling of land-use is complicated by changing boundaries of polygonal land-use plots. The data structure is improved by gridding, in which the polygons that vary in shape and size are replaced by an unchanging grid of points on which the land-use is defined.

The outcome variable is land-use type (urban or other) at each grid-point in the year of interest. Land-use data are thus analysed using logistic regression because the outcome at each grid-point is binary. Location and land-use at a previous survey are combined into four levels of a location-land-use factor with two land-use groups (urban or other) for two locations of Phuket province (north: above UTM-North 880 km and south). Since the effects of location and land-use from a previous survey as determinants of land-use change might not be additive, there is some advantage in combining them to form a single factor corresponding to all location-land-use group combinations. Therefore, the only determinant variable in the model is a location- land-use factor.

The model is fitted to a digitized grid data for Phuket province with location-land-use at a previous survey as determinant. This model formulates the logit of the probability p_i of urban outcome in the year of interest as a function of a location- land-use factor determinant. The model thus takes the form

$$\text{logit}(p_i) = \log\left[\frac{p_i}{1-p_i}\right] = \mu + \alpha_i \qquad (1)$$

where μ is a constant and α_i refers to location-land-use at a previous survey. This equation may be inverted to give an expression for the probability p_i as

$$p_i = \frac{1}{1+\exp(-(\mu+\alpha_i))} \qquad (2)$$

Four sets of analysis were conducted. Firstly, urban land in 1975 is an outcome and location-land-use group in 1967 is a determinant. Secondly, urban land in 1985 is an outcome and location-land-use group in 1975 is a determinant. Thirdly, urban land in 2000 is an outcome and location-land-use group in 1985 is a determinant. Finally, urban land in 2009 is taken an outcome and location-land-use group in 2000 as a determinant.

Conventional statistical analysis such as logistic regression assumes that data samples are independent. However, this assumption clearly does not hold for data defined at grid-points just 100 meters apart. Data from neighbouring plots are likely to be correlated, violating the independence assumption, giving incorrect standard errors.

Spatial clustering occurs when subjects are sampled from villages or families sharing particular attributes, or when repeated measurements are taken on the same subjects. Clustering also occurs in time owing to seasonal effects. In other studies of geographical data (see, for example, McNeil and Chooprateep (2014)) such spatial correlation is handled by aggregating data into larger regions with acceptably small correlation between adjoining region, or by using factor analysis and multivariate regression to adjust for spatial correlations. For our land-use data, (GEE) (Zeger and Liang 1986) method could be used by dividing the region into groups of plots (clusters) and estimating common fixed correlations between plots in the same group. However, this method has two difficulties: (a) the region has a large cluster size requiring excessive computation even on the most powerful available computer, and (b) this method assumes a common correlation within land-use polygons.

We use a conventional method widely used in survey sampling, based on variance inflation factors (VIF). VIF provide a simple method for comparing clustered binary data outcomes with group-specific covariates. This method is based on the concepts of design effect and effective sample size (Rao and Scott 1992). When the correlations between binary outcomes in clusters of size m have equal correlation ρ, the VIF is $1 + (m - 1)\rho$, and standard error of the log odds ratio is increased by the square root of this factor. This formula shows that even small correlations between outcomes can have a substantial effects in large clusters. For example, if $\rho = 0.1$ and $m = 31$, the standard error is doubled, and the sample size would need to be quadrupled to compensate for the clustering (McNeil, 2014). This method avoids the problem in the GEE method by computing effective sample sizes for each land-use plot based on their sample variances giving a set of VIFs, from which standard errors are applied to fitted values from a logistic model to compute confidence intervals.

All data analysis and graphical displays were carried out using R (R Development Core Team, 2012).

Results

The digital data structure allows us to explore land-use change from one period to the next by using bubble plots and thematic maps. The unit of analysis is a grid box with an area 1 ha. The percentages of change were simply calculated from counting numbers of grid boxes. Logistic regression shows where urban land was expected to occur. It also shows percentage of urban land and its precision in terms of 95% confidence intervals. Moreover, the benefit of using a logistic regression model to analyse land-use data is that the model can be extended when more than one predictor is needed.

Land-use change in Phuket province

Fig. 8 shows percentages of land-use change in Phuket province from 1967 to 2009 using a bubbleplot matrix. Bubbles with lighter colours along the main diagonal line with positive slope denote no change from each period, whereas bubbles with darker colours off the diagonal line denote land-use changes from one period to the next.

For example, the top right panel shows how land-use changed from 2000 to 2009 in the north of Phuket province. In 2000, 11% was urban and 89% was other land. In 2009, 14% was urban and 86% was other land. On the diagonal line, the light green bubble shows that 82% of the land was other land in 2000 and that this remained as other land in 2009. The pink bubble shows that 7% of the land was urban in 2000 and that this remained as urban in 2009.

The blue bubble shows that 7% of other land in 2000 changed to urban land in 2009. The small green bubble shows that 4% of urban land in 2000 changed to other land in 2009.

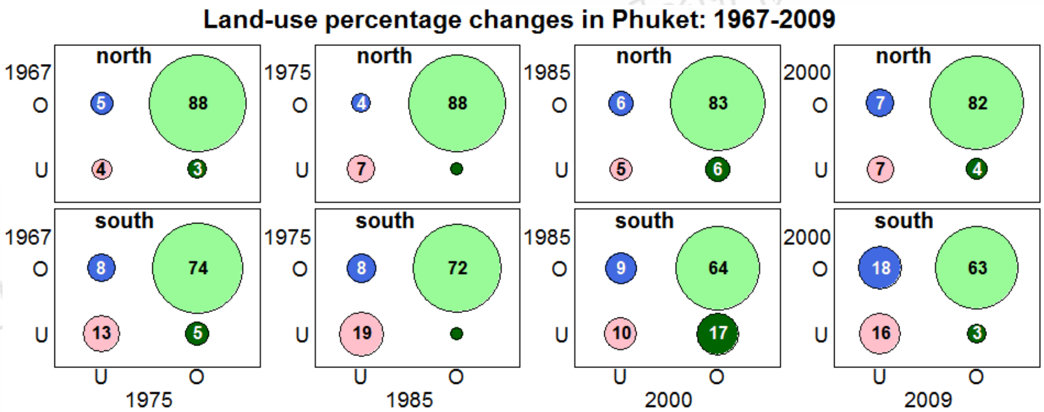


Fig. 8 Land-use percentage change in Phuket province from 1967-2009. The top row shows the change of land-use in the north and the bottom row is for the south of Phuket province.

Thematic maps are useful for showing where changes occur. Fig. 9 shows percentages of land-use change in Phuket province from 1967 to 2009 using a thematic map corresponding to the bubbleplot matrix in Fig 8. The thematic map uses the same colours as in the bubbleplot matrix. This figure shows land-use changes for each category from one period to the next and overall land-use changes for each category. For example, in the rightmost panel, light green colour denotes other land in 2000 that remained other land in 2009 (O-O) at 72.2% and dark green denotes urban land in 2000 that changed to other land in 2009 (U-O) at 3.8%. Thus, other land (all green colours) in 2009 was about 76.0% (72.2 + 3.8%) of all land.

Pink colour denotes urban in 2000 that remained urban in 2009 (U-U) at 11.3% and blue colour denotes other land in 2000 that changed to urban in 2009 (O-U) at 12.7%. Thus, urban land (pink and blue colours) in 2009 comprised about 24.0% (11.3+12.7%) of all land.

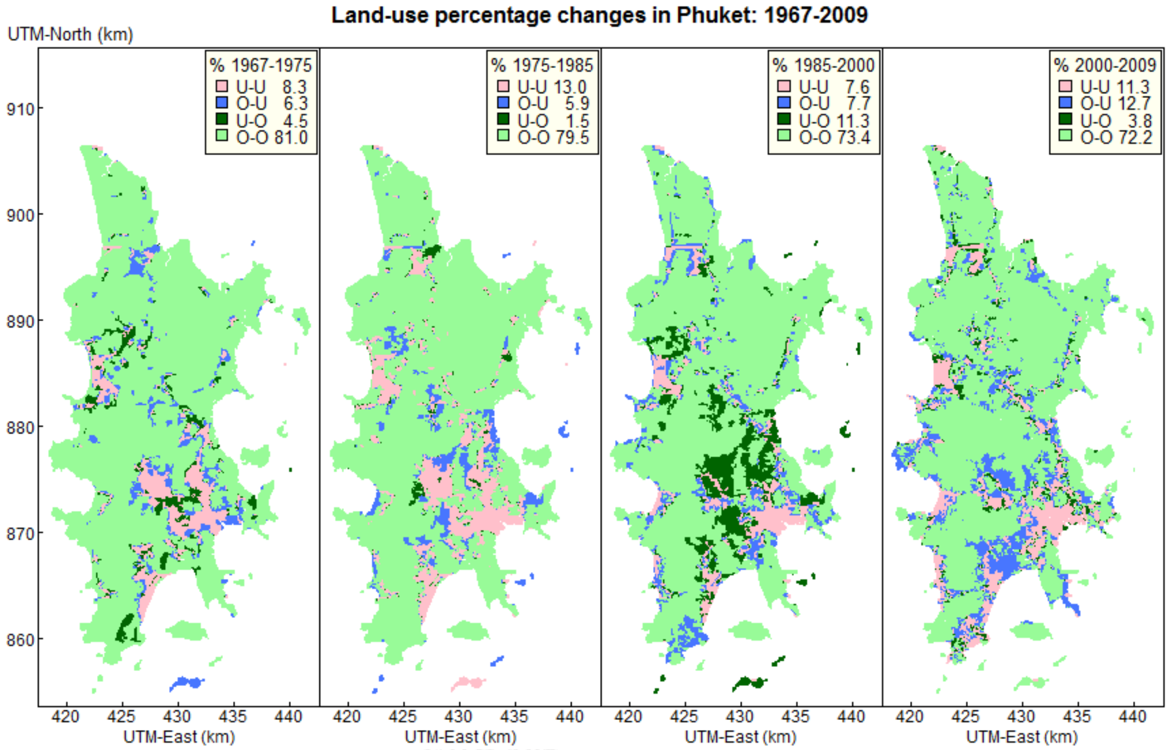


Fig. 9 Thematic map of land-use change in Phuket province from 1967-2000. The maps show percent of land-use change 1967-1975, 1975-1985, 1985-2000 and 2000- 2009, respectively.

Modelling results

Logistic regression models were used for predicting the percentages of urban land corresponding to each land-use category. The results from the logistic regression model are displayed using 95% confidence intervals for the percentage of urban land in a given year in each category of location-land-use group in a previous survey superimposed on the bar chart. A confidence interval completely above or below the mean line indicates that the factor is significantly higher or lower than expected. Hence, location-land-use group in the previous period was statistically associated with urban land in the year of interest.

Fig. 10 shows bar charts of land-use change for each group and model-based 95% confidence intervals of percentage change to the following period by location and land-use from preceding surveys carried out in 1967, 1975, 1985, 2000 and 2009. The red horizontal lines represent overall percentages of urban land in each year.

The bar charts show crude percentages of urban land from one period to the next. For example, in southern Phuket province, 70% of urban in 1967 remained urban in 1975, while 9% of other land became urban. Similarly, more than 90% of urban in 1975 remained urban in 1985, while 10% of other land became urban. For the next period, more than 35% of urban land in 1985 remained urban in 2000, while 12% of other land became urban. Finally, more than 80% of urban land in 2000 remained urban in 2009, while 22% of other land became urban. Surprisingly, the proportion of urban land in the south of Phuket remained relatively low in the period from 1985 to 2000.

The overall percentages of urban land in 1975, 1985, 2000 and 2009 were 15, 19, 16 and 24%, respectively. The blue vertical line segments denote 95% confidence intervals from the logistic regression model using the variance inflation method to account for correlations within land-use plots.

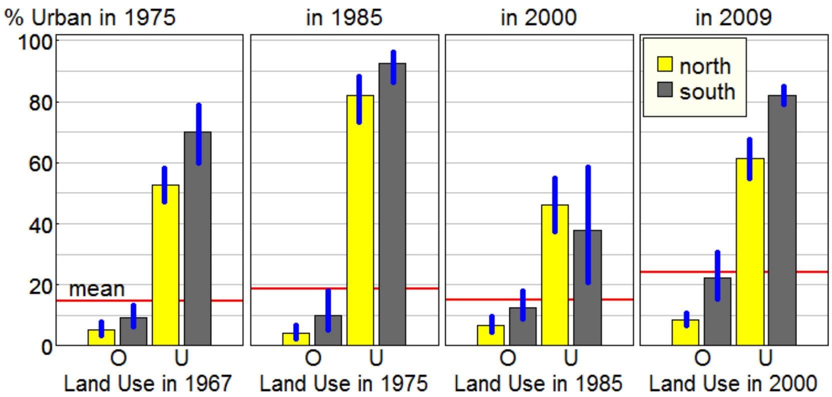


Fig. 10 Bar charts of land-use change for each group and 95% confidence intervals of percentage change by location and land-use from preceding surveys 8-15 year ago in Phuket: 1967-2009.

Discussion

In summary, this paper has introduced methods for analysing and graphing land-use data. It emphasizes exploring land-use change over time. The method used is analog-to-digital conversion, replacing polygonal shapes by digital grid points. It allows researchers to map land-use development and to show how land-use changes between two periods using statistical methods. The land-use grid at the later time point is considered as an outcome and the land-use type at the previous time point as the determinant. Then, logistic regression is used to develop the models. Since data at neighbouring grid points are not independent, a method for adjusting confidence intervals based on VIF is used to account for spatial correlation.

Polygonal data structure has some limitations in statistical analysis. Although it can provide thematic maps for displaying pattern of land-use for a given year, the data are difficult to analyse because the polygons change over time. Previous studies (Hun et al. 2011; Guo et al. 2013) describe the use of polygon to construct accurate maps. They also address shape change. Cross-tabulation, bubble plots and thematic maps are methods for preliminary land-use data analysis. Logistic regression is a statistical technique to determine the significance and strength of predictors when the outcome is binary. It needs to account for spatial correlation.

The study results show that the average percentages of urban land increased in three periods, by 15% in 1967-1975, 19% in 1975-1975 and 24% in 2000-2009. The major land-use changes were conversion of forest and agricultural land to urban. Our result is similar to those found in the city of

Hangzhou (Jin-Song et al. 2009), in central Jordan over the period 1987-2005 (Alsaaideh et al. 2012) and in Banir Dar, Ethiopia (Haregeweyn et al. 2012). Greater urbanization was observed in the southern part of Phuket province. The percentages of urban land in each of three periods of time 1967-1975, 1975-1985 and 2000-2009 were higher than those in northern Phuket province, where 70-90% of urban land remained urban. This may be due to the facts that the main city of the province and tourism activities are located in the south. There has been similar evidence that tourism developments in the form of resorts and service buildings are replacing agricultural lands and natural areas of forest, wetland or scrublands (Gossling 2002; Burak et al. 2004; Symeonakis et al. 2007; Rico-Amoros et al. 2009). Urban growth is not randomly distributed and tends to occur on forest and agricultural land. There are more likely to be developed because of their location near existing urban centers and the main highway of the island (Patarasuk and Binford 2012). Almost all the land-use changes that occurred in the study areas were because of changing human influences. Natural regeneration of forests will become increasingly difficult. Forest cover in the province was found mainly on high elevations and in protected areas. Agricultural land in Phuket province is mostly distributed across all areas, but with limited success since Phuket’s environment is not conducive to large-scale commercial agriculture.

Surprisingly, the urban land in the south of Phuket was relatively low in 1985-2000 compared to other periods. There was more urban land in the north (47% compared to 37%, approximately), a situation affected by reduced mining activity. Tin mines have been converted to forest land. The tin mining industry has gradually failed to generate economic growth in Phuket, especially after 1985 when the price of tin fell by half. However, with its natural resources, Phuket later emerged as a tourist destination with great potential. The government started to encourage tourism and some tin-mining entrepreneurs transferred their interest to the new industry. Since then, investment in the tourist industry in Phuket has grown at a remarkable rate. As a result, tourism in Phuket has expanded rapidly and has become the main economic activity in the region. Over the past four decades, Phuket has developed into one of Asia's foremost tourist destinations with more than four million visitors a year. The money brought into the island by tourism has made Phuket the second wealthiest province in Thailand after Bangkok. The most important land-use changes from 1985-2000 involve reforestation. The increase of the forest land is largely due to abandonment of mining activities that have returned to forest land. In recent times the rate of deforestation around the globe has increased dramatically. The world’s forests are being cut down and converted to other land-use every year (FAO 2006). At the same time, planting of trees has resulted in new forests being established while in other areas forests have expanded on to abandoned land through natural regeneration, thus reducing the net loss of total forest area. This partly explains the lower urban land of Phuket in 1985-2000. It is thus interesting to find out why and how this situation occurred in further studies.

Conclusion

This study has introduced a method for analysing land-use data. The method is basically an analog-to-digital conversion, replacing polygonal shapes by coded grid points. This method is very useful for the analysis of land-use. The bubble plot matrices and thematic maps were used to display land-use and urban growth from 1967 to 2009 in Phuket province. The thematic map, bubble plots and bar charts with confidence intervals not only provide sufficient information on the land-use change but also provide quantitative information about urban change in simple geographic areas using freely available software R program. Logistic regression models were used to fit the land-use data for predicting the percentages of urban based on the land-use in the previous survey and location. VIF were used to obtain valid confidence intervals that account for correlation between these sub-plots. It can be implemented using any standard computer program for the analysis of independent binary data after a small amount of preprocessing. The method is applied to a variety of problems involving clustered binary data. The logistic regression model can allocate and specify where urbanization will occur. Thus, different scenarios of future urbanization patterns are predicted by allocation cells that represent the estimated size of the required urban land on the predicted probability map. The prediction of future urban patterns for different years is extremely useful in understanding the temporal urban expansion process. This method provides information about future urbanization trends in Phuket province for government planners and developers to manage land-use change in the future. The methods are useful to develop appropriate statistical models for land-use change. The statistical model can be extended. Various determinants, such as land-use at a previous survey time, location, ownership, accessibility or proximity to roads and transport hubs, climate, and population density may be incorporated into a model to predict future land-use at each grid-point.

Acknowledgements The authors acknowledge suggestions from Prof. Don McNeil and Greig Rundle. We also thank Thailand’s Land Development Department for providing the land-use data. This study was supported by a Prince of Songkla University scholarship.

References

Achmad A, Hasyim S, Dahlan B, Aulia DN (2015) Modeling of urban growth in tsunami-prone city using logistic regression: Analysis of Banda Aceh, Indonesia. Appl Geogr 62: 237-246

Allen J, Lu K (2003) Modeling and prediction of future urban growth in the Charleston region of South Carolina: a GIS-based integrated approach. Conserv Ecol 8:2

Alsaaideh B, Al-Hanbali A, Tateishi R (2012) Assessment of Land Use/Cover Change and Urban Expansion of the Central Part of Jordan Using Remote Sensing and GIS. *AJG* 11(3):1-9

Alsharif AA, Pradhan B (2014) Urban sprawl analysis of Tripoli Metropolitan city (Libya) using remote sensing data and multivariate logistic regression model. J Indian Soc Remote 42:149-163

Alsharif AA, Pradhan B (2015). Spatio-temporal Prediction of Urban Expansion Using Bivariate Statistical Models: Assessment of the Efficacy of Evidential Belief Functions and Frequency Ratio Models. Appl Spat Anal Policy 1-19

Bach M, Breuer L, Frede HG, Huisman JA, Ott A, Waldardt R (2006) Accuracy and congruency of three different digital land-use maps. Landscape Urban Plan 78:289-299

Burak S, Dogan E, Gazioglu C (2004) Impact of urbanization and tourism on coastal environment. Ocean Coast Manag 47(9): 515-527

Cheng J, Masser I (2003) Urban growth pattern modeling: a case study of Wuhan city, PR China. Landscape Urban Plan 62:199-217

Chuangchang P, and Tongkumchum P (2014) Modeling land development along highway 4 in Southern Thailand. Songklanakarin J Sci Technol 36 (6): 719-725

Eyoh A, Olayinka DN, Nwilo P, Okwuashi O, Isong M, Udoudo D (2012) Modelling and predicting future urban expansion of Lagos, Nigeria from remote sensing data using logistic regression and GIS. IJAST 2:116-124

Feng Y, Liu Y, Batty M (2015) Modeling urban growth with GIS based cellular automata and least squares SVM rules: a case study in Qingpu–Songjiang area of Shanghai, China. Stoch Env Res Risk A 1-14

Guo L, Du S, Haining R, Zhang L (2013) Global and local indicators of spatial association between points and polygons: A study of land use change. Int J Appl Earth Obs Geoinf 21:384-396

Gssling S (2002) Global environmental consequences of tourism. Glob Environ Change 12(4): 283-302

Haregeweyn N, Fikadu G, Tsunekawa A, Tsubo M, Meshesha DT (2012) The dynamics of urban expansion and its impacts on land use/land cover change and small-scale farmers living near the urban fringe: A case study of Bahir Dar, Ethiopia. Landscape Urban Plan 106:149-157

Hosmer DW, Lemeshow S (2000) Applied logistic regression. John Wiley & Sons.

Hu Z, Lo CP (2007) Modeling urban growth in Atlanta using logistic regression. Comput Environ Urban Syst 31:667-688

Huh Y, Yu K, Heo J (2011) Detecting conjugate-point pairs for map alignment between two polygon data sets. Comput Environ Urban Syst 35: 250-262

Jin-Song D, Wang K, Li J, Deng YH (2009) Urban land use change detection using multisensor satellite images. Pedosphere 19:96-103

Lambin EF, Geist H, Rindfuss RR (2006) Introduction: local processes with globalimpacts. In: Land-use and land-cover change. Springer Berlin Heidelberg, pp 1-8

Landis J, Zhang M (1998) The second generation of the California urban futures model. Environ Plann B 25:795-824

Lewis DJ (2010) An economic framework for forecasting land-use and ecosystem change. Resour Energy Econ 32:98-116

Li X, Liu X, Yu L (2014) A systematic sensitivity analysis of constrained cellular automata model for urban growth simulation based on different transition rules. IJGIS 28(7): 1317-1335

Liu H, Zhou Q (2005) Developing urban growth predictions from spatial indicators based on multi-temporal images. Comput Environ Urban Syst 29(5): 580-594

McNeil N, Chooprateep S (2014) Modeling sea surface temperatures of the North Atlantic Ocean. Theor Appl Climatol 116:11-17

McNeil D (2014) Epidemiological Statistics-Further Issues. Wiley StatsRef: Statistics Reference Online. DOI: 10.1002/9781118445112

Nong Y, Du Q (2011) Urban growth pattern modeling using logistic regression. P Soc Photo-Opt Ins 14:62-67

Patarasuk R, Binford MW (2012) Longitudinal analysis of the road network development and land-cover change in Lop Buri province, Thailand, 1989-2006. Appl Geogr 32(2): 228-239

Polovina JJ, Howell EA, Abecassis M (2008) Ocean's least productive waters are expanding. Geophys Res Lett 35:1-5

R Development Core Team (2012) The R project for statistical computing. http://www.r-project.org. Accessed 24 September 2012

Rao JNL, Scott, AJ (1992) A sample method for the analysis of clustered binary data. Biometrics 48:577-585

Rico-Amoros AM, Olcina-Cantos J, Sauri D (2009) Tourist land use patterns and water demand: evidence from the Western Mediterranean. Land Use Policy 26(2): 493-501

Sheffield and Morse-McNabb (2012) Historical land cover assessment: challenges and achievements. school of mathematical and geospatial sciences RMIT university. http://ceur-ws.org/Vol-1328/GSR2_Sheffield.pdf. Accessed 10 November 2015

Stehman SV, Wickham JD (2011) Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. Remote Sens Environ 115:3044-3055

Sudhira HS, Ramachandra TV, Jagadish KS (2004) Urban sprawl: metrics, dynamics and modelling using GIS. Int J Appl Earth Obs Geoinf 5(1): 29-39

Symeonakis E, Calvo-Cases A, Arnau-Rosalen E (2007) Land Use Change and Land Degradation in Southeastern Mediterranean Spain. Environ Manage 40(1): 80-94

Tayyebi A, Perry PC, Tayyebi AH (2014) Predicting the expansion of an urban boundary using spatial logistic regression and hybrid raster-vector routines with remote sensing and GIS. Int J Geogr Inf Sci 28:639-659

The Food and Agriculture Organization of the United Nations (FAO) (2006) Forest losses and gains: where do we stand?. http://www.unep.org/vitalforest/Report/VFG-02-Forest-losses-and-gains.pdf. Accessed 16 November 2015

Thinnukool O, Chuangchang P, Kongkraphan K (2014) Analyzing land-use change using grid-digitized method. Songklanakarin J Sci Technol 36:235-240

Triantakonstantis D, Stathakis D (2015) Urban Growth Prediction in Athens, Greece, Using Artificial Neural Networks. Int. J. Civil Struct. Constr Archit Eng, 9: 193-197

Wu F, Yeh AG (1997) Changing spatial distribution and determinants of land development in Chinese cities in the transition from a centrally planned economy to a socialist market economy: a case study of Guangzhou. Urban Stud 34:1851-1880

Zeger S, Liang KY (1986) Longitudinal data analysis for discrete and continuous outcomes. Biometrics 42:121-130

Songklanakarin J. Sci. Technol.
36 (6), 719-725, Nov. - Dec. 2014

SJST SONGKLANAKARIN JOURNAL OF SCIENCE AND TECHNOLOGY
http://www.sjst.psu.ac.th

Original Article

Modeling land development along highway 4 in Southern Thailand

Potjamas Chuangchang and Phattrawan Tongkumchum*

Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Mueang, 94000 Thailand.

Received: 19 July 2014; Accepted: 27 October 2014

Abstract

This study aims to investigate the change of developed land in three different locations along Highway 4 Road from Phattalung to HatYai. The method involves creating a digitized grid of geographical coordinates covering the study area. The land-use codes and plot identifiers were recorded in database tables indexed by grid coordinates. Logistic regression of land development adjusted for spatial correlation was used to model its change over a 9-year period using land-use at the previous survey combined with location as a determinant. The results show increasing average percentages of developed land (3% in 2000 and 5% in 2009). Land development occurred mostly in the northern location along the Pattalung to HatYai road.

Keywords: land-use data, grid-digitization, urban growth, land development, logistic regression model

1. Introduction

Modeling of land-use change has attracted substantial attention because it helps to explain the process of land-use change and thus assist relevant public policy-making. Recent publications on modeling the change in land-use include Lambin *et al.* (2000); Veldkamp and Lambin (2001); Verburg and Veldkamp (2001); Weng (2002); Aspinall, (2004); Veldkamp and Verburg (2004); and Heistermann *et al.* (2006). Relevant urban land development studies include those by Allen and Lu (2003); Barredo and Demicheli (2003); Cheng and Masser (2003); Barredo *et al.* (2004); Henriquez *et al.* (2006); Aguayo *et al.* (2007); Hu and Lo (2007); He *et al.* (2008); Luo and Wei (2009); Nong and Du (2011); Al-shalabi *et al.* (2013a); Al-shalabi *et al.* (2013b); and Alsharif and Pradhan (2014). Although developed land usually constitutes only a small proportion of the whole area, it is a catalyst for wider social and environmental changes. Modeling the change of land-use to developed land therefore provides valuable information to planners, developers and policy makers.

* Corresponding author.
Email address: tphattra@bunga.pn.psu.ac.th

Scientific knowledge of land-use change is well developed in research areas of science, geography, computer science, image processing, and remote sensing, whereas statistical methodologies, in particular time series modeling, have not yet been applied to the same extent. Several studies in recent decades have developed new and improved models of changes in land-use to developed land based on remote sensing (RS) and geographical information systems (GIS) to explain and forecast future development (Sudhira *et al.*, 2004; Liu and Zhou, 2005; Huang *et al.*, 2008; Eyoh *et al.*, 2012).

Modeling approaches used for analysis of land-use change have different purposes. Some dynamic, process-based simulation models (Cassel-Gintz and Petschel-Held, 2000; Stephenne and Lambin, 2001) are useful to predict changes whereas stochastic and optimization methods (Maxwell *et al.*, 2000; Fischer and Sun, 2001) are suitable for describing the decision making for land management. Empirical statistical models identify current land-use changes by applying regression models to relate historical land-use changes and other factors. Logistic regression has been widely used (Allen and Lu, 2003; Cheng and Masser, 2003; Hu and Lo, 2007; Nong and Du, 2011; Eyoh *et al.*, 2012; Alsharif and Pradhan, 2014; Tayyebi *et al.*, 2014). However, spatial correlation, which causes violation of the assumption of independent residuals, is often ignored because the statistical methodology for considering correlation is not well developed for logistic regression models (Hu and Lo, 2007).

Moreover, conventional survey and mapping techniques used in land-use change studies are expensive and time consuming for the expansion of developed land. Such information is not readily available in most developing countries. Analyses of land-use change based on sample data and inferential statistical approaches are needed.

This study uses graphical and statistical methods to display and model the extent of land development. After converting land-use data from polygons to a digitized grid, we use logistic regression with variance inflation factors to obtain confidence intervals for the extent of change in land development.

2. Methods

2.1 Study area

The study area comprises three locations along Highway 4 between Phattalung and HatYai in Southern Thailand (Figure 1). This part of the road is approximately 95 kilometers long and it goes through five districts (Mueang Phatthalung, Khao Chaison, Bang Kaeo, Tamot, and Pa Bon) of Phattaung and four districts (Rattaphum, Khuan Niang, Bang Klam and HatYai) of Songkhla. The area for each region is about 700 km^2 (25 km × 28 km).

2.2 Data

The land-use data for the three locations for 1991, 2000, and 2009 were obtained from the Thailand Department of Land Development, which has records of data from remote sensing and regular surveys in every province. The data are stored as polygonal *shape files* of land-use plots. The files can be restructured into a relational database table. The data structure for statistical analysis is based on an analog-to-digital conversion method, using a grid with dimension 100 m ×100 m, as explained by Thinnukool *et al.* (2014).

2.3 Land-use categories

Land-use was classified into four main categories comprising undeveloped land (UD), paddy field and other agriculture (PF+), rubber plantation (RF), and developed land (Dev). Descriptions of these categories are shown in Table 1.

2.4 Preliminary data analysis and variables

Land-use data can be displayed using thematic maps. These maps represent only the distributions of land-use in different years and locations and thus highlight variations, but they cannot be used directly to compare the areas of land-use categories. Bar charts are appropriate for comparing areas of land use categories by year and location. For detecting association of land-use categories between two periods of time, the percentages of land-use change are depicted using bubble plots with sizes representing magnitudes of change.

The outcome is binary, developed land (Dev) or other. The determinants are location and land-use categories nine years earlier. Since the effects of location and land-use category as determinants of land development might not be additive, there is some advantage in combining them to form a single factor corresponding to all location by land-use group combinations. Locations and land-use categories were thus combined into 12 levels of a categorical variable. These 12 levels comprise combinations of four land-use groups, namely undeveloped land (UD), paddy field and other agriculture (PF+), rubber plantation (RP) and developed land (Dev), and three locations: northern, central and southern locations along the Phattalung to HatYai road.

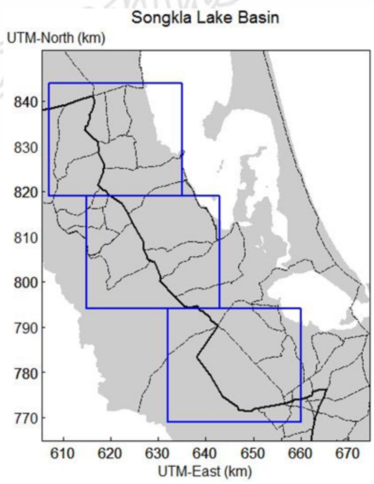


Figure 1. Location of study area.

Table 1. Land-use classification.

Land-use categories	Descriptions
Undeveloped land (UD)	Forest, grassland, water bodies, marsh and swampand miscellaneous land
Paddy field and other agriculture (PF+)	Paddy field, field crop perennial, orchard, horticulture, pasture and aquatic plant
Rubber plantation (RF)	Rubber plantation
Developed land (Dev)	City, town, commercial, village, institutional land, transportation, communication and industrial land

To measure land-use change between two periods of time, two sets of analyses were conducted. First, developed land in 2000 was taken as the outcome and location by land-use group from preceding surveys in 1991 as the determinant. Second, developed land in 2009 was taken as the outcome and location by land-use group in 2000 as a determinant.

2.5 Statistical analysis

Logistic regression was used for modeling the association between developed land (Dev) and location by land-use group. This model was fitted to the digitized grid data. It formulates the logit of the probability p_i of developed land (Dev) outcome in the year of interest and the location by land-use group nine years earlier as a determinant, and thus takes the form

$$\log\left[\frac{p_i}{1-p_i}\right]=\mu+\alpha_i, \qquad (1)$$

where μ is a constant and the term α_i refers to the location by land-use group nine years earlier.

Conventional statistical regression modeling assumes that the individual observations are uncorrelated. However, digitized grid land-use data have substantial spatial correlation and thus violate this assumption (Hu and Lo, 2007). Variation inflation factors (Rao and Scott, 1992) were used to account for this spatial correlation and thus obtain valid confidence intervals. The graphical displays and statistical analyses in this study were performed using R (R Development Core Team, 2011).

3. Results

3.1 Land-use change

Figure 2 shows a thematic map that displays the land-use patterns along the Phattalung to HatYai road in 1991, 2000, and 2009 for three locations. Most of the land in the northern part was used for paddy fields and other agriculture (PF+). Rubber plantation (RP) was the largest category of land-use in the central and southern locations.

The bar charts in the Figure 3 show the area (in ha) of undeveloped land (UD), paddy field and other agriculture (PF+), rubber plantation (RP), and developed land (Dev).

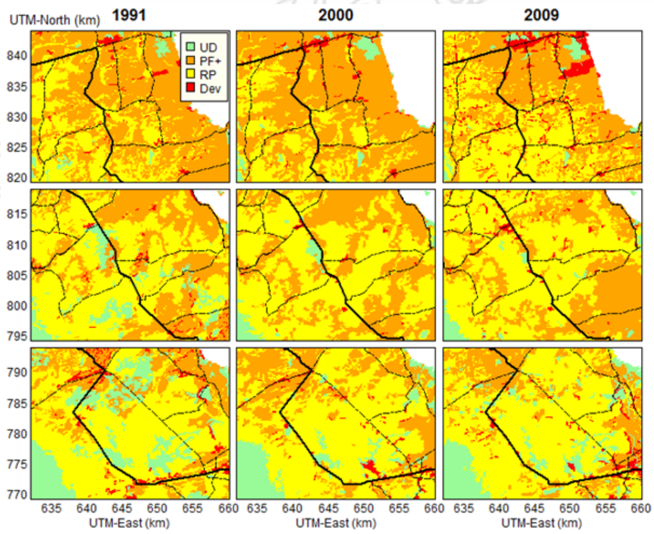


Figure 2. Land-use maps along the Phattalung to HatYai road in 1991, 2000, and 2009 for the three locations shown in Figure 1.

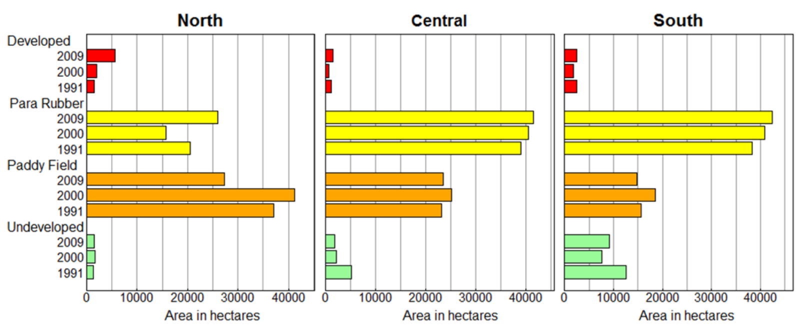


Figure 3. Bar charts of land-use along the Phattalung to HatYai road in 1991, 2000, and 2009 for three locations shown in Figure 1.

Figure 4 shows a bubble plot matrix of percentages of land-use category change in three locations. The top row shows the change of land-use from 1991 to 2000 and the bottom row shows the change from 2000 to 2009. Lighter grey or lighter colors (pink, yellow, orange and green) along the diagonal denote no change in land-use. Darker gray or darker colors (dark green, grey and red) off the diagonal denote land-use changes from one period to the next.

For example, the top right panel shows the change in land-use from 1991- 2000 in the southern location. In 1991, 4.28% of the land was developed land, 55.52% was rubber plantation, 22.70% was paddy field and other agriculture and 17.50% was undeveloped land. In 2000, 2.92% was developed land, 60% was rubber plantation, 27% was paddy field and other agriculture and 10.08% was undeveloped land.

On the diagonal, the pink bubble shows that less than 1% of the land that was developed land in 1991 remained developed in 2000. The yellow bubble shows that 43% of the land that was rubber plantation in 1991 remained so in 2000. The orange bubble shows that 13% of the land that was paddy field and other agricultural land in 1991 remained so in 2000. The light green bubble shows that 7% of the land that was undeveloped in 1991 remained so in 2000.

Off the diagonal line in the first column, the green bubbles, show that out of 4% of the land that was developed in 1991, 2% changed to paddy field and other agriculture and the other 2% changed to rubber plantation in 2000. In the second column, the grey bubbles show that out of the 11% that was rubber plantation in 1991, 2% became undeveloped and the other 9% changed to paddy field and other agriculture in 2000. The red bubble shows that 1% of the land that was rubber plantation in 1991 became developed in 2000. In the third column, the grey bubbles show that out of the 9% of land that was paddy field and other agriculture in 1991, 1% became undeveloped and the other 8% changed to rubber plantation in 2000. The red bubble shows that less than 1% of the land that was paddy field and other agriculture in 1991 became developed in 2000. In the fourth column, the grey bubbles show that the 10% of the land that was undeveloped in 1991, 3% changed to paddy field and other agriculture and the other 7% changed to rubber plantation in 2000. The small red bubble shows that less than 1% of the land that was undeveloped in 1991 became developed in 2000.

3.2 Modeling results

The logistic regression model gave estimates of the percentages of the change in land-use to developed land and their corresponding standard errors. Standard errors were used to construct 95% confidence intervals for comparing the percentages of change to developed land for each of location by land-use type with their average percentages.

Figure 5 shows bar charts of the crude percentages of the change in land-use to developed land by location and by land-use group with 95% confidence intervals superimposed.

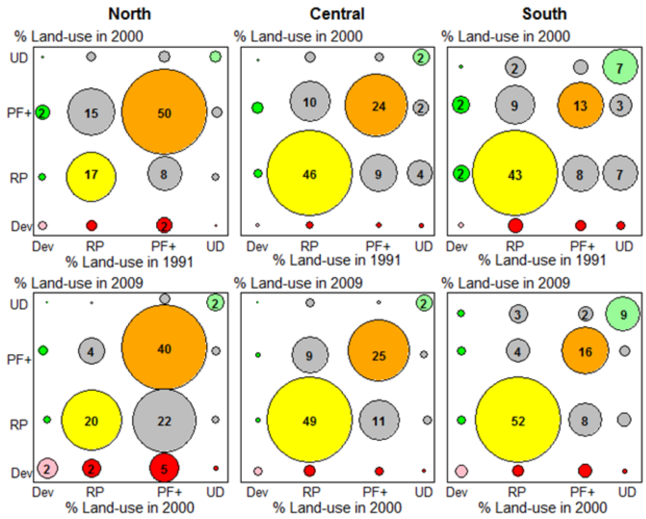


Figure 4. Land-use percentage changes along the Phattalung to Hat Yai road are the periods 1991-2000 and 2000-2009 for the northern, central, and southern locations.

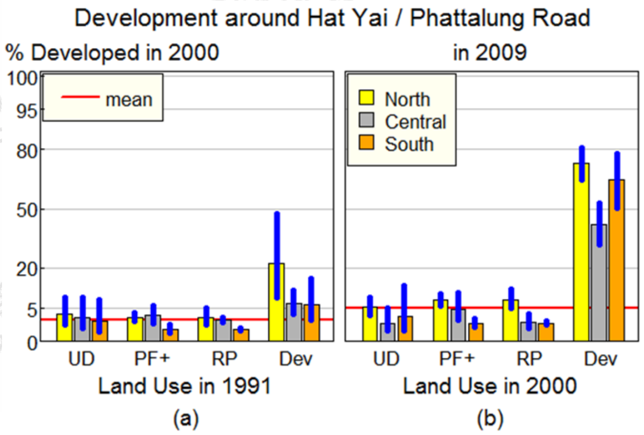


Figure 5. Bar charts of percentages of land that changed to developed land along the Phattalung to HatYai road from (a) 1991-2000 and (b) 2000-2009 by previous land-use and location.

The average percentage is shown by the horizontal red line with 3% for 2000 and 5% for 2009. Confidence intervals above the average line reflect the groups that were more likely to become developed land. The percentages of developed land that remained developed were higher in the 2000-2009 period than those in the 1991-2000 period. Greater land development occurred in the north, and the percentage of paddy field and other agriculture (PF+) that became developed was higher in 2000-2009 than in 1991-2000.

In Figure 5a, the developed land in 1991 (22.26% in the northern location, 6.23% in the central location, and 6.09% in the southern location) remained developed in 2000. Less than 4% of undeveloped land, paddy field and other agriculture and rubber plantation became developed. The 95% confidence interval for developed land in the north is substantially above the mean and it is marginally higher than the average in the central location whereas it is not different from the average in the south. Thus, the developed land in the north and the central location were more likely to remain developed land. All of the confidence intervals for the other land-use groups (UD, PF+ and RP) were lower or across the mean. Thus, these groups were less likely to become developed land, especially PF+ and RP in the south.

In Figure 5b, the developed lands in 2000 (73.59% in the northern, 41.86% in the central location and 65.61% in the southern) remained developed in 2009. Less than 8% of undeveloped land, paddy field and other agriculture and rubber plantation became developed. The percentages of developed land in three locations are substantially above the mean for all groups. Thus, they were more likely to remain developed. The results for other land-use groups were similar to the earlier period.

4. Discussion and Conclusions

Land-use change along the Phatthalung to HatYai road from 1991 to 2009 was investigated using statistical methods. The data were converted from polygons to a digitized grid. Thematic maps and bar chart were used to display the data. Bubble plot matrices were used to summarize percentages of land-use change from one period to the next. Logistic regression was used to delineate the association between developed land and location by land-use group nine years earlier.

Our results show increasing average percentages of developed land (3% in 2000 and 5% in 2009). Land development occurred mostly in the north. This may be due to the proximity to Phatthalung City. Moreover, the north is also located at the intersection of two highways (Highway 4 and 41). Road networks can influence the conversion of land to developed land. That developed land occurs closer to road networks was previously reported for Lop Buri Province, Central Thailand (Patarasuk and Binford, 2012). Urban growth and conversion of agricultural land to urban area have been found to occur closer to road networks in disparate locations, including the Kansas City Metropolitan area in the central U.S.A. (Underhill, 2004), Puerto Rico (del Mar Lopez *et al.*, 2001) and in Beijing, PR China (Zhang *et al.*, 2002).

The other land-use categories were less likely to change to developed land, especially PF+ and RP in the south. This can be seen by noting the higher percentage of rubber plantations. Due to a higher expected income from rubber plantations farmers have tended to convert land to rubber plantations rather than develop their land. It is reported that in the Phatthalung watershed, a quarter of the paddy field area has been converted into rubber plantation due to the higher incomes from rubber production (Pensuk and Shrestha, 2008).

The undeveloped land was more likely to change to other categories rather than developed land. This has occurred elsewhere, for example, forest lands were transformed to agricultural land, particularly for shrimp farms in Pak Panang Bay (Prabnarong and Thongkao, 2006) and Ban Don Bay (Muttitanon and Tripathi, 2005). However, various factors affected the extent of land-use change at each location.

Location by land-use group is just one factor that contributes to land-use change. Land-use change is usually the result of a combination of multiple factors including economic, biophysical, social, and political drivers, such as income, rainfall and population dynamics (Geist and Lambin, 2002). This study used locations by land-use group as a determinant of the change to developed land. Further research should include other variables such as ownership, accessibility or proximity to roads and transport hubs, climate, and population density.

Acknowledgements

The authors acknowledge suggestions from Prof D. McNeil. We also would like to thank Thailand's Land Development Department for providing the land use data. Our study was supported by a Prince of Songkla University scholarship.

References

Aguayo, M., Wiegand, T., Azocar, G., Wiegand, K. and Vega, C. 2007. Revealing driving forces of mid-cities urban growth patterns using spatial modeling: a case study of Los Angeles (Chile). Ecology and Society. 12(1), 13.

Allen, J. and Lu, K. 2003. Modeling and prediction of future urban growth in the charleston region of south Carolina: a GIS-based integrated approach. Conservation Ecology. 8(2), 2.

Al-shalabi, M., Billa, L., Pradhan, B., Mansor, S., and Al-Sharif, A.A. 2013. Modelling urban growth evolution and land-use changes using GIS based cellular automata and SLEUTH models: the case of Sana'a metropolitan city, Yemen. Environmental Earth Sciences. 70(1), 425-437.

Al-shalabi, M., Pradhan, B., Billa, L., Mansor, S., and Althuwaynee, O.F. 2013. Manifestation of remote sensing data in modeling urban sprawl using the SLEUTH model and brute force calibration: a case study of Sana'a city, Yemen. Journal of the Indian Society of Remote Sensing. 41(2), 405-416.

Alsharif, A.A. and Pradhan, B. 2014. Urban sprawl analysis of Tripoli Metropolitan city (Libya) using remote sensing data and multivariate logistic regression model. Journal of the Indian Society of Remote Sensing. 42(1), 149-163.

Aspinall, R. (2004). Modelling land use change with generalized linear models—a multi-model analysis of change between 1860 and 2000 in Gallatin Valley. Montana. Journal of Environmental Management. 72(1), 91-103.

Barredo, J.I. and Demicheli, L. 2003. Urban sustainability in developing countries' megacities: modelling and predicting future urban growth in Lagos. Cities. 20(5), 297-310.

Barredo, J.I., Demicheli, L., Lavalle, C., Kasanko, M. and McCormick, N. 2004. Modelling future urban scenarios in developing countries: an application case study in Lagos. Nigeria. Environment and Planning B. 31(1), 65-84.

Cassel-Gintz, M. and Petschel-Held, G. 2000. GIS-based assessment of the threat to world forests by patterns of non-sustainable civilisation nature interaction. Journal of Environmental Management. 59(4), 279-298.

Cheng, J. and Masser, I. 2003. Urban growth pattern modeling: a case study of Wuhan city, PR China. Landscape and Urban planning. 62(4), 199-217.

del Mar Lopez, T., Aide, T.M. and Thomlinson, J.R. 2001. Urban expansion and the loss of prime agricultural lands in Puerto Rico. Ambio: A Journal of the Human environment. 30(1), 49-54.

Eyoh, A., Olayinka, D.N., Nwilo, P., Okwuashi, O., Isong, M. and Udoudo, D. 2012. Modelling and predicting future urban expansion of Lagos, Nigeria from remote sensing data using logistic re-gression and GIS. International Journal of Applied Science and Technology. 2(5), 116-124.

Fischer, G. and Sun, L.X. 2001. Model based analysis of future land-use development in China. Agriculture Ecosystems and Environment. 85(1-3), 163-176.

Geist, H.J. and Lambin, E.F. 2002. Proximate causes and underlying driving forces of tropical deforestation. BioScience. 52(2), 143-150.

He, C., Okada, N., Zhang, Q., Shi, P. and Li, J. 2008. Modelling dynamic urban expansion processes incorporating a potential model with cellular automata. Landscape and Urban Planning. 86(1), 79-91.

Henríquez, C., Azócar, G. and Romero, H. 2006. Monitoring and modeling the urban growth of two mid-sized Chilean cities. Habitat International. 30(4), 945-964.

Heistermann, M., Muller, C. and Ronneberger, K. 2006. Land in sight? Achievements, deficits and potentials of continental to global scale land-use modeling. Agriculture, Ecosystems and Environment. 114(2), 141-158.

Hu, Z. and Lo, C.P. 2007. Modeling urban growth in Atlanta using logistic regression. Computers, Environment and Urban Systems. 31(6), 667-688.

Huang, W., Liu, H., Luan, Q., Jiang, Q., Liu, J. and Liu, H. 2008. Detection and prediction of land use change in Beijing based on remote sensing and GIS. The International Archives of the Photogrammetry. Remote Sensing and Spatial Information Sciences. 37, 75-82.

Lambin, E.F., Rounsevell, M.D.A. and Geist, H.J. 2000. Are agricultural land-use models able to predict changes in land-use intensity? Agriculture, Ecosystems and Environment. 82, 321-331.

Luo, J. and Wei, Y.H. 2009. Modeling spatial variations of urban growth patterns in Chinese cities: the case of Nanjing. Landscape and Urban Planning. 91(2), 51-64.

Liu, H., and Zhou, Q. 2005. Developing urban growth predictions from spatial indicators based on multi-temporal images. Computers, Environment and Urban Systems. 29(5), 580-594.

Maxwell, B.D., Johnson, J. and Montagne, C., 2000. Predicting land use change and ecosystem impacts in and around a rural community. In Spatial Information for Land Use Management, M.J. Hill and R.J. Aspinall, editors. Gordon and Breach, Reading, MA, U.S.A., pp. 183-197.

Muttitanon, W. and Tripathi, N.K. 2005. Land use/land cover changes in the coastal zone of Ban Don Bay, Thailand using Landsat 5 TM data. International Journal of Remote Sensing. 26(11), 2311-2323.

Nong, Y. and Du, Q. 2011. Urban growth pattern modeling using logistic regression. Geo-Spatial Information Science. 14(1), 62-67.

Patarasuk, R. and Binford, M.W. 2012. Longitudinal analysis of the road network development and land-cover change in Lop Buri province, Thailand, 1989-2006. Applied Geography. 32(2), 228-239.

Pensuk, A. and Shrestha, R.P. 2008. Effect of paddy area conversion to rubber plantation on rural livelihoods: a case study of Phatthalung Watershed, Southern Thailand. GMSARN International Journal. 2(4), 185-190.

Prabnarong, P. and Thongkao, S. 2006. Land Use Changes in Pak Panang Bay using Arial Photographs and Geographic Information System. Walailak Journal of Science and Technology. 3(1), 93-104.

R Development Core Team 2011. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for statistical Computing. http://www.r-project.org [September 6, 2012].

Rao, J.N.L. and Scott, A.J. 1992. A sample method for the analysis of clustered binary data. Biometrics. 48(2), 577-585.

Stephenne, N., and Lambin, E.F. 2001. A dynamic simulation model of land-use changes in Sudano-sahelian countries of Africa (SALU). Agriculture, Ecosystems and Environment. 85(1), 145-161.

Sudhira, H.S., Ramachandra, T.V. and Jagadish, K.S. 2004. Urban sprawl: metrics, dynamics and modelling using GIS. International Journal of Applied Earth Observation and Geoinformation. 5(1), 29-39.

Tayyebi, A., Perry, P.C. and Tayyebi, A.H. 2014. Predicting the expansion of an urban boundary using spatial logistic regression and hybrid raster–vector routines with remote sensing and GIS. International Journal of Geographical Information Science. 28(4), 639-659.

Thinnukool, O., Chuangchang, P. and Kongkraphan, K. 2014. Analyzing land-use change using grid-digitized method. Songklanakarin Journal of Science and Technology. 36 (2), 235-240.

Underhill, K.K. 2004. Land cover dynamics and urban sprawl indicators: a case study in the Kansas City Metropolitan Area. Unpublished Ph.D. Dissertation, University of Missouri, Kansas City, U.S.A.

Veldkamp, A. and Lambin, E.F. 2001. Predicting land-use change. Agriculture, Ecosystems and Environment. 85(1), 1-6.

Veldkamp, A. and Verburg, P.H. 2004. Modelling land use change and environmental impact. Journal of Environmental Management, 72(1), 1-3.

Verburg, P.H. and Veldkamp, A. 2001. The role of spatially explicit models in land-use change research: a case study for cropping patterns in China. Agriculture, Ecosystems and Environment. 85(1), 177-190.

Weng, Q. (2002). Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. Journal of Environmental Management. 64(3), 273-284.

Zhang, Q., Wang, J., Peng, X., Gong, P. and Shi, P. 2002. Urban built-up land change detection with road density and spectral information from multi-temporal Landsat TM data. International Journal of Remote Sensing. 23(15), 3057-3078.

Appendix III

Modeling Developed Land in Phuket Province of Thailand: 2000-2009

Potjamas Chuangchang, Thanin Sangkhaduang and Phattrawan Tongkumchum*

Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Mueang, 94000 Thailand

Abstract

The objective of this study is to model land development in the 17 sub-districts in Phuket province of Thailand from 2000 to 2009. Logistic regression was used to monitor changes in land-use over this period and predict future changes. The ROC curve was used to measure the performance of the model. Land-use from a previous survey in 2000 and sub-district identity were included as determinants. The area of developed land increased by 4557 ha over the study period. Agricultural land in 2000 was more likely to become developed in 2009 and developed land was more likely to remain developed land in 2009. Land development occurred mostly in Chalong and Talat Nua sub-districts. The area under the ROC curve was 0.83 indicating a reasonably good fit of the model. Prediction of such changes may be used to provide useful information for decision makers and planners.

Keywords: Developed land, Urban growth, Logistic regression model, Land-use data

1. Introduction

Urban growth and land development leads to land-use change in many areas around the world, especially in developing countries where urbanization rates are high, impacting on the environment, the social structure and the economy of the region. Recent studies have shown that urbanized land tends to replace either agricultural land (Lopez *et al.*, 2001; Helmer, 2004; Huang *et al*., 2009; Alsaaideh *et al.*, 2012; Belal, 2011; Forkuor and Cofie, 2011; Mohan *et al.*, 2011; Alsharif and Pradhan, 2014) or forest land (Thomlinson and Rivera, 2000; Schneider and Pontius, 2001).

Statistical models of land-use change and urban growth have become important tools for city planners, economists, ecologists and resource managers, facilitating timely and effective action for sustainable development of urban regions (Herold *et al.*, 2001). These models can provide insight into the dynamics of the urbanization system and can be used to forecast future development trends. Different modeling approaches have been adopted in studies of land-use change from the perspective of their utility for predicting changes in land-use intensification (Lambin *et al.*, 2000). Logistic regression has been widely used to model urban growth (Allen and Lu, 2003; Cheng and Masser, 2003; Hu and Lo, 2007; Nong and Du, 2011; Eyoh *et al*., 2012; Alsharif and Pradhan, 2014; Tayyebi *et al*., 2014). However, these studies do not have predictors comprising land-use type in a previous year, and spatial correlation, which causes violation of the assumption of independent errors, is difficult to handle (Hu and Lo 2007). The method we have used to correct for such correlation is largely new, and extends further the "variance inflation factor" method recently developed by Thinnukool et al. (2014) and Chuangchang and Tongkumchum (2014)

Although land-use change has been extensively studied in Thailand (see, for example, Muttitanon and Tripathi, 2005; Prabnarong and Thongkao, 2006; Swangjang and Iamaram, 2011; Thinnukool *et al.*, 2013), few empirical studies have focused on urban growth. A basic reason for this lack of research is inconsistencies in data structure which complicate appropriate model creation. Urban growth is one of the most important types of land-use change currently affecting Thailand, particularly in the south where tourism has grown in Phuket, now attracting more than 5 million visitors each year. Tourism is now one of the fastest growing economic sectors in the province. The Phuket Provincial Administration Organization has implemented a tourism development plan to increase tourism and to promote Phuket as a world-class center of marine tourism, in turn generating substantial revenue locally and nationally (Sakolnakorn *et al.*, 2013). Tourism has both positive and negative impacts on the economic and social health of a region. The city area has grown with little consideration to the land-use types that are being transformed. Moreover, poorly managed developmental activity leads to environmental damage. Thus, controlling urbanization and creating sustainable development require accurate information about urban growth patterns (Jiang and Yao 2010). Few studies of land-use in Phuket are published (Ratanasermpong et al., 1995; Boupun and Wongsai, 2012). Such historical data also contain valuable information about Phuket history and culture development that is not available elsewhere, and is also valuable to planners and developers.

The objectives of this study are to detect and evaluate the land-use change and to identify the pattern of change in developed land that occurred in Phuket province from 2000 to 2009.

2. Material and Methods

2.1 Study area

Phuket is the largest island in Thailand, located around latitude 7º 53'N and longitude 98º 24'E. Phuket Province is divided into three districts which are further subdivided into 17 sub-districts (Figure 1). Mueang Phuket is the capital district of Phuket Province. This district encompasses the southern part of the island of Phuket. It is subdivided into eight sub-districts (Koh Keao, Rasada, Wichit, Chalong, Karon, Rawai,Talad Yai and Talad Nuea). Kathu district is located in the west of Phuket Island. Nighbouring Thalang to the north, Mueang Phuket to the east and south, and the Andaman Sea to the west, Kathu also covers the famous tourist beach of Phuket, Patong. It is subdivided into three sub-districts (Kamala, Kathu and Patong). Thalang is the district in the north of Phuket Province and is subdivided into six sub-districts (Thep Krasatti, Si Sunthon, Choeng Thale, Pa Khlok, Mai Khao and Sakhu).

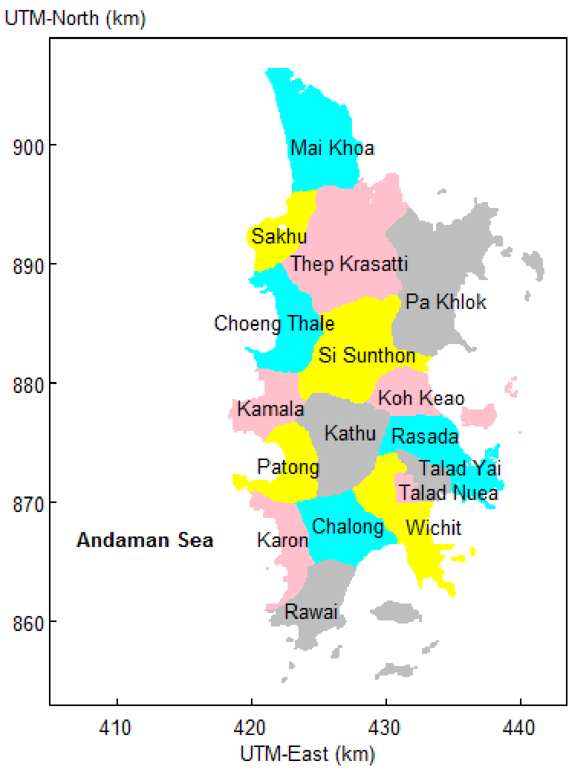


Figure 1 Sub-districts of Phuket province.

The major types of land-use in 2000 included developed land, undeveloped land (forest, grassland, water bodies, marsh and swamp), rubber plantation, and agricultural land (Figure 2).

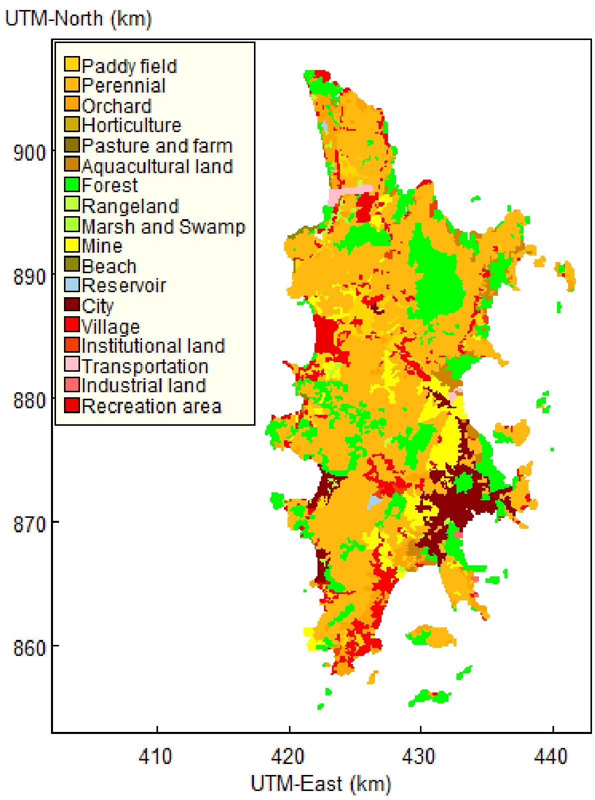


Figure 2 A 2000 land-use thematic map of the study area in Phuket province.

2.2 Data Management

This study presents an analysis of land-use change for time periods based on a grid-digitized method. Land-use data stored in analog (vector or polygonal) form were obtained from the Thailand Department of Land Development for the years 2000 and 2009. Since polygonal land-use data are difficult to analyze because they change in shape and size when land-use changes over time, the data structure can be improved by conversion to data on a fixed grid. This analog-to-digital conversion method is explained by Thinnukool *et al*. (2014), and uses the function *point.in.polygon* in the spatial (sp) library of the R program (Pebesma *et al*., 2014). Duplicated polygons arise when larger polygonal land-use plots contain smaller plots such as farm ponds. The data management steps used in this study are summarized in Figure 3.

Land-use data were thus transformed to Universal Transverse Mercator (UTM) coordinates and all land-use coordinates were converted to be consistent with corresponding current Google Earth coordinates. Each grid point corresponds to one-hectare area because each grid represents 100×100 meters.

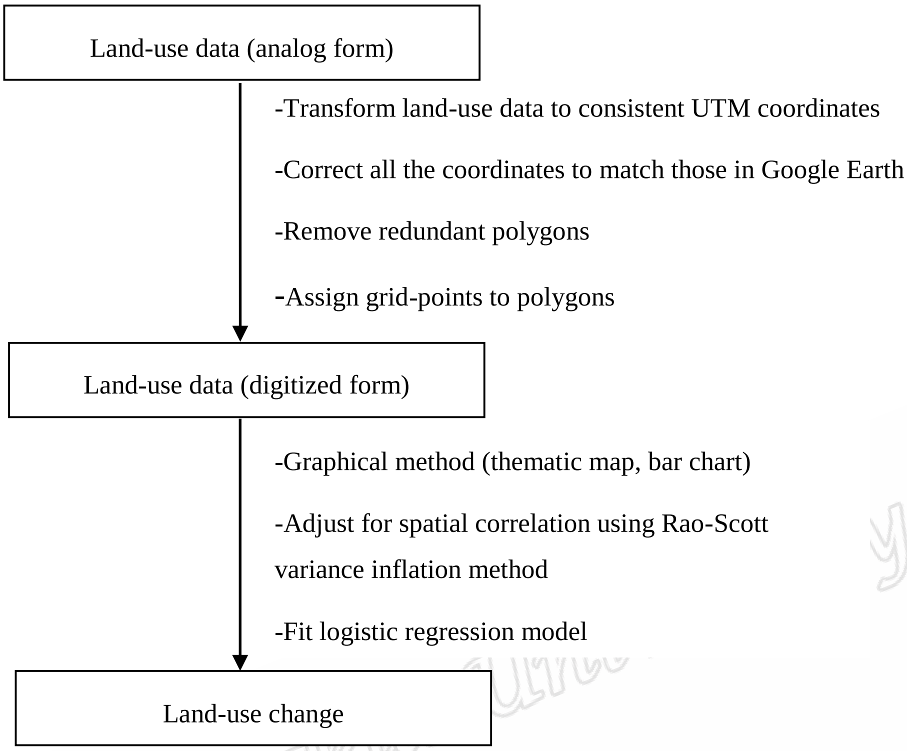


Figure 3 Steps of data management and analysis in this study.

2.3 Land-use categories

The land-use categories were classified into four main groups comprising undeveloped land (UD), rubber plantation (RF), other agriculture (OA), and developed land (Dev). Descriptions of these categories are listed in Table 1.

Table 1 Land-use classification.

Land-use categories	**Descriptions**
Undeveloped land (UD)	Forest, grassland, water bodies, marsh and swamp and miscellaneous land
Rubber plantation (RF)	Rubber plantation
Other agriculture (OA)	Paddy field, field crop perennial, orchard, horticulture, pasture and aquatic plant
Developed land (Dev)	City, town, commercial, village, institutional land, transportation, communication and industrial land

2.4 Logistic regression model

Logistic regression is a powerful empirical method appropriate to model data where the outcome is binary (Hosmer and Lamshow, 2004). In this study, the land-use data were analyzed using this method because the outcome at each grid-point is binary. The outcome is developed land coded as 1 (Dev) and

0 (OA). Two factors were considered as determinants of developed of land, namely sub-district identity and land-use in 2000.

Two logistic probability models were developed to predict the probability of change in land-use to developed land from 2000 to 2009. The simplest model included only one factor, while the full model included two factors. The model formulates the logit of the probability p_{ij} of developed land (Dev) in terms of the two determinant factors as follows:

$$\log\left[\frac{p_{ij}}{1-p_{ij}}\right] = \mu + \alpha_i + \beta_j$$

In this model μ is a constant and the terms α_i and β_j refer to land-use group in 2000 and sub-district identity, respectively. Land-use group in 2000 has four levels (see Table 1) and the sub-district factor has 17 levels, one for each sub-district (see Figure 1).

Variance inflation factors (VIF) were computed to account for spatial correlation between land-use outcomes within sub-districts and thus provide valid confidence intervals, as described by Rao and Scott (1992). The adjusted percentages of developed land for each determinant can thus be presented graphically, showing confidence intervals. The conventional treatment contrasts method gives different estimates depending on which level in the factor is selected as the reference group, and gives wider confidence intervals when this reference group has smaller sample size. To compare sub-district and land-use group effects with their overall mean, rather than with an arbitrary sub-district, the standard errors for the estimated parameters in the model were based on weighted sum contrasts, as described by Tongkumchum and McNeil (2009). An advantage of this method is that it provides a simple criterion for classifying levels of a factor (Chutinantakul *et al.*, 2014; Kongchouy and Sampantarak, 2010). Using this method, the confidence intervals for proportions derived from this method may be classified into three levels; (1) above the mean, (2) crossing the mean, and (3) totally below the mean.

A receiver operating characteristic (ROC) curve describes how well a model predicts a binary outcome. This is a plot of the true positive rate (sensitivity) against the false positive rate (1-specificity) for different possible thresholds of the model. This graph displays the predictive accuracy of the logistic model, which can be evaluated using the area under the ROC curve (AUC). The AUC is particularly important for evaluating how well the method can discriminate between the outcome values. An ideal model would have an AUC of 1. Denoting the predicted outcome as 1 (developed land) if P $\geq$ *c* or 0 (other) if P < *c*, the ROC curve plots the proportion of positive outcomes correctly predicted by the model against the false positive rate (proportion of all outcomes incorrectly predicted), as *c* varies. Choosing *c* to match numbers of predicted and observed outcomes ensures that equal weights are assigned to false positive and false negative prediction errors. The AUC measures the performance of a model and can be used as a measurement of model accuracy (Sakar and Midi, 2010; Takahashi *et al.*, 2006). It shows how well a model predicts a binary outcome (Fan *et al.*, 2006), and varies from 0.5 to

1. An AUC close to 1 signifies that the model has almost perfect discrimination while an AUC close to 0.5 indicates poor discrimination (Hanley and McNeil, 1982).

3. Results

3.1 Land-use change in Phuket province

The bar chart in Figure 4 shows the area (ha) of land-use categories in Phuket province from 2000 and 2009. Most land was used for rubber plantation, with an area of 25250 ha in 2000 and 19791 ha in 2009. Both developed and undeveloped land areas have been increasing, with areas of 7834 ha in 2000, 12391 ha in 2009, 13280 ha in 2000 and 16152 ha in 2009, respectively. Other agriculture land area has been decreasing with and area of 5307 ha in 2000 and 3337 ha in 2009.

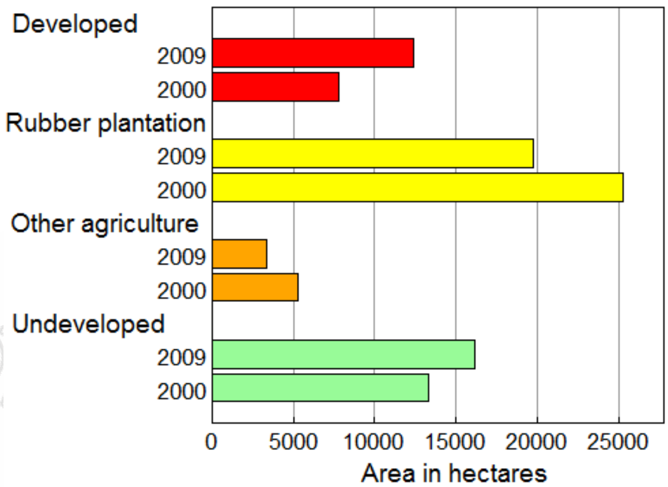


Figure 4 Bar chart of land-use in 2000 and 2009.

The map in Figure 5 illustrates land-use change from 2000 to 2009. The left panel demonstrates land-use in 2000 whereas the right panel shows land-use in 2009 and the middle panel shows loss and gain of land. Undeveloped and developed land gained 4,557 ha and 2,872 ha, respectively. Other agriculture and rubber plantation lost 1,970 ha and 5,459 ha to other categories.

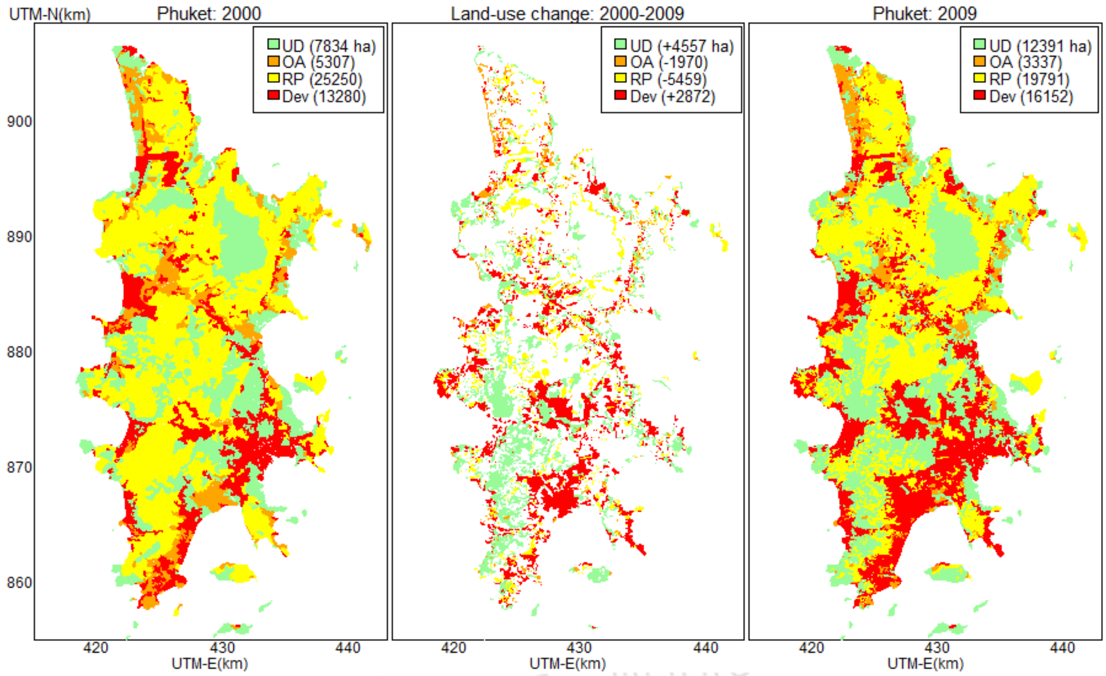


Figure 5 Land-use maps of Phuket province left 2000, for right 2009, middle panel highlights areas where change in land-use.

3.2 Modeling result

Figure 6 shows crude percentages of developed land for the two determinant factors (land-use in 2000 and sub-district identity) as red circles, together with corresponding 95% confidence intervals for differences between these percentages and the overall mean percentage (shown as the horizontal red line). The boxes show 95% confidence intervals coloured according to their location above (cyan), across (yellow) or below (pink) the mean. For each factor, the confidence intervals are adjusted for the effect of the other factor, showing a result that adjusts for any correlation between determinants.

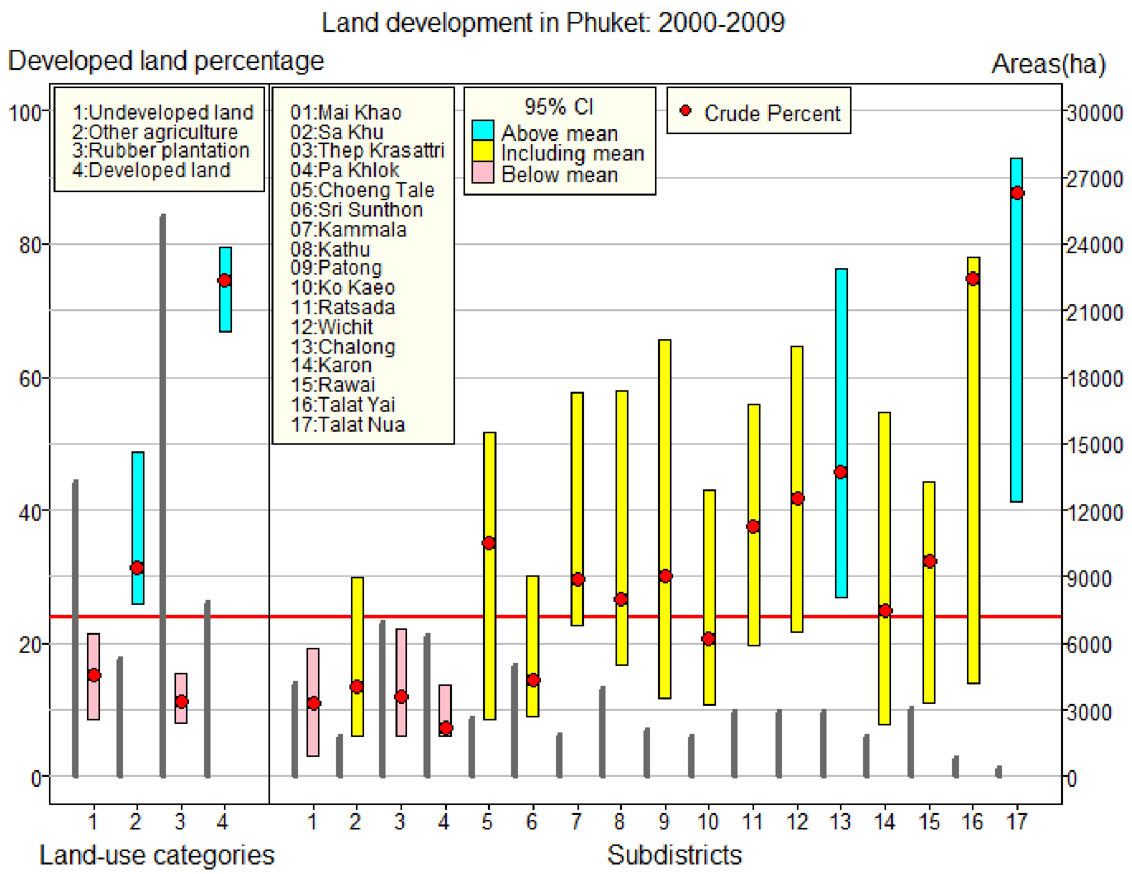


Figure 6 Adjusted percentages of land that changed to developed land in Phuket province from 2000-2009 with coloured rectangles denoting 95% confidence intervals of percentage change by land-use categories and sub-district identity. Vertical bars denote sub-district in hectares.

The 95% confidence interval for both other agriculture and developed land in 2000 are higher than the mean whereas that for rubber plantation is lower. Thus, other agriculture and developed land in 2000 were significantly more likely to change or remain as developed land. The percentages of developed land in Chalong and Talat Nua were all substantially above the mean. Thus, these two sub-districts were more likely to have significantly greater areas of developed land than other sub-districts in 2009.

The full model with two factors was assessed using the ROC curve and compared with the simple model. Figure 7 shows the ROC curve for the simple model with only land-use group in 2000 as determinant factor, together with the ROC curve for the full model with both factors included.
The cut-off point in the curve, where the observed and the predicted number of developed land in 2009 are equal as the criteria, was used to report sensitivity and specificity of the model. The simple model gives AUC 0.751 with 60.6% sensitivity and 85.7% specificity whereas the full model gives AUC 0.833 with 63.1% sensitivity and 88.3% specificity.

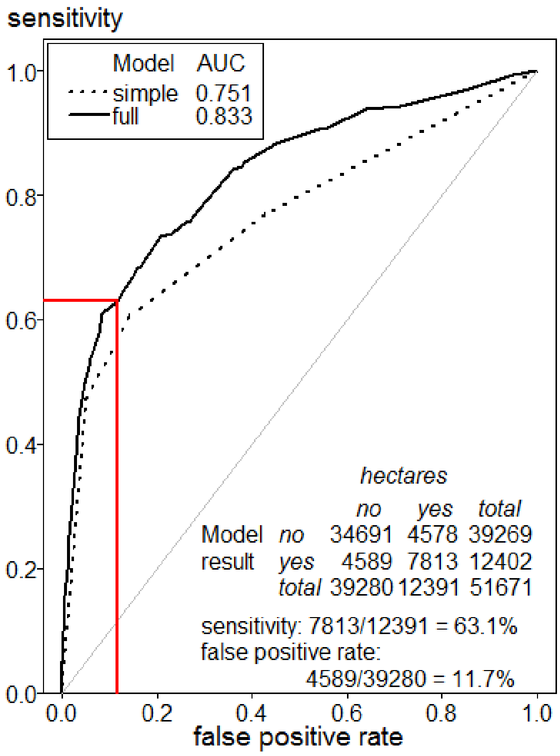


Figure 7 ROC curve for logistic regression models with results from full model.

The map in Figure 8 illustrates the results from the logistic regression model. A false positive result indicates that the grid showed developed land in 2009 when in fact it was not developed (the red area), while a false negative indicates that land not developed in 2009 (the result is *negative*), when in fact it was (the region shaded pink on the map). In both cases, the result is false. The pink and yellow areas show the incidence of developed land in 2009. The yellow areas on the map depicts the area which was correct predicted as developed land, while the pink area also depicts the area of developed land but was incorrectly predicted by the model.

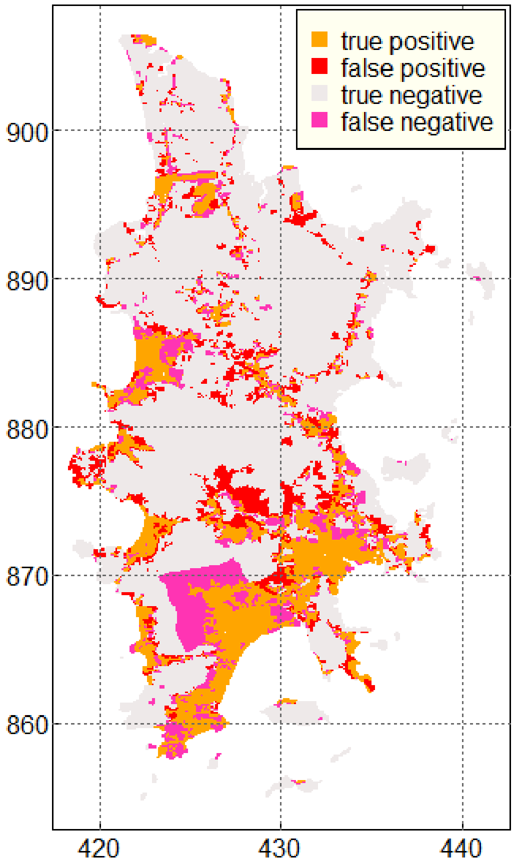


Figure 8 Results from the logistic regression model.

4. Discussion and conclusion

In this study, logistic regression is used to model land development in Phuket from 2000 to 2009. It shows where land development would have probably occurred in 2009 compared to 2000. The results indicate that developed land was more likely to remain as developed land in 2009 and agricultural land was more likely to change to developed land in 2009. Our findings show a similar pattern of land-use change, especially from agricultural land to developed land, as those found elsewhere in the literature. A recent study on Southern Thailand by Chuangchang and Tongkumchum (2014) found that percentage of paddy field and other agriculture (PF+) in the north location along highway 4 became developed land was higher in 2000-2009 than in 1991-2000. In Banir Dar, Ethiopia by Haregeweyn *et al*. (2012) reported that built-up areas increased from 80 ha in 1957 to 155 ha in 1994, primarily converted from agricultural land. For Delhi, during the period 1997-2008, Mohan *et al.* (2011) found that developed land increased by17% mainly due to conversion from agricultural and waste land. For central Jordan over the period 1987-2005, Alsaaideh *et al.* (2011) found that a high percentage of agricultural land was converted to urban areas. A similar study in Freetown, Sierra Leone by Forkuor and Cofie (2011) showed that 27% of agricultural land in 1986 was converted for residential purposes in 2000. In the Al Gharbiya governorate of Egypt from 1972 to 2005 by Belal and Moghanm (2011) reported that urban areas increased by 7.2 and 5.8%, causing loss of productive agricultural lands. Another study of Puerto Rico city by Lopez *et al.* (2001) showed that rapid losses of agricultural lands occurred as a result of urban expansion since 1950.

In our study of Phuket, land development mostly occurred in Chalong and Talat Nua sub-districts. These sub-districts are located in Mueang Phuket (the capital). Chalong is located in the central southern part of Phuket, where most visitors to the islands south of Phuket depart from Chalong pier. Tourists number up to 3,000 per day, and the economy of the sub-district is growing. Moreover, Chalong is also located at the intersection of four roads (4021, 4022, 4024 and 4028). Road networks can influence the conversion of land to developed land. A recent study in Lop Buri province of Thailand by Patarasuk and Binford (2012) reported developed land occurs closer to road networks.

To confirm the model capability, the ROC technique can be used in land-use change modeling studies (Pontius and Schneider, 2001; Hu and Lo, 2007; Wang and Mountrakis, 2011; Arsanjani *et al*., 2013). ROC curves give the proportions of positive and negative outcomes correctly and incorrectly predicted by the model. AUC is currently considered to be the standard method to assess the accuracy of such models. The AUC is well known method used in public health research. There are several scales for the AUC value interpretation, however in general ROC curves with AUC below 0.75 are not considered as clinically useful (Worster *et al*., 2006). Moreover, the benefit of using logistic regression model to analyse land-use data is that the model can be extend when more than one predictors are needed to be investigated.

Understanding the pattern of land-use changes can be useful for planners and policy makers. It may improve their predictions of the amount of land-use change and the location of future developed land, and enhance existing urban strategies for better sustainable land management. Cities need the implementation of an effective urban plan, strict urban development regulations, and reduction of urban growth ratio to save fertile agriculture lands and to protect the environment. Land-use planners can explore different land-use scenarios with different objectives and constraints.
Thailand’s land-use database is updated every few years, and can provide a rich data resource for historians, property investors, environmental scientists, and planning agencies concerned with the sustainable development of land.

The Thailand Department of Land Development has databases from regular surveys of thousands of plots in every province, and data are not readily available later than 2009. Our paper focuses on methods, which we believe are new and important, particularly because they can be integrated with remote sensing data on land cover from Earth-orbiting satellites. However, it is also important to have a method that makes appropriate use of historical land-use data such as those available from Thailand's Department of Land Development. For example, vegetation indices available from the MODIS polar-orbiting satellites could be incorporated into our model, and further investigation of land-use change for tourism services based on the tourist population growth during the last decade would be desirable, using data available in the Phuket Master Plan.

Tourism is one of the major driving forces behind land-use changes. Phuket is one of the most well-known tourist destinations for decades, and currently, the island has been transformed to a major tourism hub with well-prepared infrastructure and services to accommodate millions of tourists around the world (Sakolnakorn *et al.*, 2013). Demand for tourism increases accommodation establishments, transport infrastructure and leisure activities all over Phuket. Difference in land development at the sub-district level is difficult to explain and it is a limitation in our study. Land-use change relates not only to tourism activities but also to other factors including town planning and structure, road networks and population density (Belal and Moghanm, 2011; Patarasuk and Binford, 2012; Alsharif and Pradhan, 2014). However, our study based on available data does not allow us to draw conclusions about land-use changes for tourism services in Phuket, and we hope to address this issue in further research.

Acknowledgement

The authors acknowledge suggestions from Prof Don McNeil. We also would like to thank Thailand's Land Development Department for providing the land-use data. This study was supported by a Prince of Songkla University scholarship.

References

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

Alsaaideh, B., Al-Hanbali, A., and Tateishi, R. (2012). Assessment of Land Use/Cover Change and Urban Expansion of the Central Part of Jordan Using Remote Sensing and GIS. *Asian Journal of Geoinformatics*, *11(3)*.

Alsharif, A. A., and Pradhan, B. (2014). Urban sprawl analysis of Tripoli Metropolitan city (Libya) using remote sensing data and multivariate logistic regression model. *Journal of the Indian Society of Remote Sensing, 42(1),* 149-163.

Arsanjani, J. J., Helbich, M., Kainz, W. and Boloorani, A. D. (2013).
Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation*, *21*, 265–275.

Belal, A. A. and Moghanm, F. S. (2011). Detecting urban growth using remote sensing and GIS techniques in Al Gharbiya governorate, Egypt. *The Egyptian Journal of Remote Sensing and Space Science*, *14(2)*, 73-79.

Boupun, P. and Wongsai, S. (2012). Land use change and the town planning policy of Phuket. *The 33rd Asian Conference On Remote Sensing.* Ambassador City Jomtien Hotel, Pattaya, Thailand.

Cheng, J. and Masser, I. (2003). Urban growth pattern modeling: a case study of Wuhan city, PR China. *Landscape and urban planning*, *62(4)*, 199-217.

Chutinantakul, A., Mayeng, M. and Tongkumchum, P. (2014). Estimation of mortality with missing data using logistic regression. *Songklanakarin Journal of Science and Technology*, *36(2)*, 249-254.

Eyoh, A., Olayinka, D. N., Nwilo, P., Okwuashi, O., Isong, M. and Udoudo, D. (2012). Modelling and predicting future urban expansion of Lagos, Nigeria from remote sensing data using logistic regression and GIS. *International Journal of Applied Science and Technology*, *2(5)*, 116.

Fan, J., Upadhye, S. and Worster, A. (2006). Understanding receiver operating characteristic (ROC) curves. *Canadian Journal of Emergency Medicine*, *8(1)*, 19-20.

Forkuor, G. and Cofie, O. (2011). Dynamics of land-use/land-cover change in Freetown Sierra Leone and its effects on urban and peri-urban agriculture- A remote sensing approach. *International Journal of Remote Sensing*, *32(4)*, 1017-1037.

Hanley, J. A. and McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*, 29-36.

Haregeweyn, N., Fikadu, G., Tsunekawa, A., Tsubo, M. and Meshesha, D. T. (2012). The dynamics of urban expansion and its impacts on land use/land cover change and small-scale farmers living near the urban fringe: A case study of Bahir Dar, Ethiopia. *Landscape and urban planning*, *106(2)*, 149-157.

Helmer E. H. (2004). Forest Conservation and land development in Puerto Rico. *Landscape Ecology*, *19*, 29-40.

Herold, M., Menz, G. and Clarke, K. C. (2001). Remote sensing and urban growth models-demands and perspectives. In Symposium on remote sensing of urban areas, 35.

Hosmer Jr, D. W., and Lemeshow, S. (2004). *Applied logistic regression*. John Wiley & Sons.
Hu, Z. and Lo, C. P. (2007). Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems*, *31(6)*, 667-688.

Huang, B., Zhang, L. and Wu, B. (2009). Spatiotemporal analysis of rural-urban land conversion. *International Journal of Geographical Information Science*, *23*, 379-398.

Jiang, B. and Yao, X. (2010). Geospatial analysis and modeling of urban structure and dynamics: an overview. *In Geospatial Analysis and Modelling of Urban Structure and Dynamics*, Springer, Netherlands, pp. 3-11.

Kongchouy, N. and Sampantarak, U. (2010). Confidence intervals for adjusted proportions using logistic regression. *Modern Applied Science*, *4(6)*, 1-7.
Lambin, E. F., Rounsevell, M. D. A. and Geist, H. J. (2000). Are agricultural land-use models able to predict changes in land-use intensity? *Agriculture, Ecosystems and Environment*, *82*, 321-331.

Lopez, T. del M., Aide, T. M., and Thomlinson, J. R. (2001). Urban expansion and the loss of prime agricultural lands in Puerto Rico. *Ambio: a Journal of the Human environment*, *30(1)*, 49-54.

Muttitanon, W. and Tripathi, N. K. (2005). Land use/land cover changes in the coastal zone of Ban Don Bay, Thailand using Landsat 5 TM data. *International Journal of Remote Sensing*, *26(11)*, 2311-2323.

Mohan, M., Pathan, S. K., Narendrareddy, K., Kandya, A. and Pandey, S. (2011). Dynamics of urbanization and its impact on land use/land cover: A case study of Megacity Delhi. *Journal of Environmental Protection*, *2*, 1274-1283.

Nong, Y. and Du, Q. (2011). Urban growth pattern modeling using logistic regression. *Geo-Spatial Information Science*, *14(1)*, 62-67.

Patarasuk, R. and Binford, M. W. (2012). Longitudinal analysis of the road network development and land-cover change in Lop Buri province, Thailand, 1989-2006. *Applied Geography*, *32(2)*, 228-239.

Pebesma, E., Bivand, R., Rowlingson, B. and Gomez-Rubio, V. (2014).
R topics documented. Available from: http://rspatial.r-forge.r-project.org [July 9, 2014]

Pontius, R. G. and Schneider, L. C. (2001). Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment*, *85*, 239-248.

Prabnarong, P. and Thongkao, S. (2006). Land Use Changes in Pak Panang Bay using Arial Photographs and Geographic Information System. *Walailak Journal of Science and Technology (WJST)*, *3(1)*, 93-104.

Rao, J. N. L. and Scott, A. J. (1992). A sample method for the analysis of clustered binary data. *Biometrics*, *48(2)*, 577-585.

Ratanasermpong, S., PornprasertchaI, J. and Disbunchong, D. (1995), Natural resources and land use changes of Phuket using remote sensing. The 16 Asian Conference on Remote Sensing. Phuket, Thailand.

Sakar, S. and Midi, H. (2010). Importance of assessing the model adequacy of binary logistic regression. *Journal of Applied Sciences*, *10(6)*, 479-486.

Sakolnakorn, T. P. N., Naipinit, A. and Kroeksakul, P. (2013). Sustainable Tourism Development and Management in the Phuket Province, Thailand. *Asian Social Science*, *9(7)*, 75-84.

Swangjang, K. and Iamaram, V. (2011). Change of Land Use Patterns in the Areas Close to the Airport Development Area and Some Implicating Factors. *Sustainability*, *3*, 1517-1530.

Schneider, L. C. and Pontius, R. G. (2001). Modeling land-use change in the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment*, *85(1)*, 83-94.

Takahashi, K. Uchiyama, H. Yanagisawa, S. Kamae, I. (2006). The logistic regression and ROC analysis of group-based screening for predicting Diabetes incidence in four years. *Kobe Journal of Medical Sciences*, *52(6)*, 171-180.

Tayyebi, A., Perry, P. C. and Tayyebi, A. H. (2014). Predicting the expansion of an urban boundary using spatial logistic regression and hybrid raster-vector routines with remote sensing and GIS. *International Journal of Geographical Information Science*, *28(4)*, 639-659.

Thinnukool, O., Chuangchang, P. and Kongkraphan, K. (2014). Analyzing land-use change using grid-digitized method. *Songklanakarin Journal of Science and Technology*, *36 (2)*, 235-240.

Thomlinson, J. R. and Rivera, L. Y. (2000). Suburban growth in Luquillo, Puerto Rico: some consequences of development on natural and semi-natural systems. *Landscape and Urban Planning*, *49(1)*, 15-23.

Tongkumchum, P. and McNeil, D. (2009). Confidence intervals using contrasts for regression model. *Songklanakarin Journal of Science and Technology*, *31(2)*, 151-156.

Wang, J. and Mountrakis, G. (2011). Developing a multi-network urbanization model: a case study of urban growth in Denver, Colorado. *International Journal of Geographical Information Science*, *25*, 229-253.

Worster, A., Fan, J. and Upadhye, S. (2006). Understanding receiver operating characteristic (ROC) curves. *CJEM*, *8(1)*, 19-20.

Appendix IV

Presenting Conference Paper

SD-10: Assessment of Land-use Change and Urban Growth in Phuket Province, Thailand

Potjamas Chuangchang and Phattrawan Tongkumchum
Faculty of Science and Technology, Department of Mathematics and Computer Science, Prince of Songkla University, Pattani Campus, 94000, Thailand. Phone: 66-7-331-2179; Fax: 66-7-331-2179; E-mail: potjamastan@gmail.com, tphattra@bunga.pn.psu.ac.th)

The present study aims to predict change of urban land in Phuket Province for surveys in 1975, 1985, 2000 and 2009. The data of land-use plots were recorded using a digitized-grid method. Land-use was classified broadly into three main groups comprising forest (forest and grassland), agriculture (agriculture and fish farming) and urban (village, city and other developed land including mines). Logistic regression model of urban growth with a combination of location (north or south) and land-use at a previous survey as a determinant was used to explain the patterns. Most of land was used for agriculture, especially in the north of Phuket. More areas of urban land were in the south. Urban land was increasing and average percentages were 14.49% in 1975, 18.76% in 1985, 21.38% in 2000 and 27.52% in 2009. Agricultural land became urban land more than forest whereas urban land tended to remain the same in the south more than in the north. The exception was in period 1985-2000 more urban land in the north than the south. This occurred because of mining activity.

Greater Mekong Subregion Academic and Research Network

CERTIFICATE OF APPRECIATION

ASIAN INSTITUTE OF TECHNOLOGY 1959

The 9th GMSARN International Conference 2014 on

"Connectivity and Sustainability in GMS: Energy, Environmental and Social Issues"

November 12-14, 2014 at Ho Chi Minh City, Vietnam

This is to certify that **Potjamas Chuangchang** of **Prince of Songkla University, Pattani Campus, Thailand** has participated in the conference and presented the paper reference no. **GMSARN-SD-10** with entitled **"Assessment of Land-use Change and Urban Growth in Phuket Province, Thailand"**.

W. Ongsakul

Weerakorn Ongsakul, Ph.D.

GMSARN Secretary General & Conference Coordinator

Appendix V

Program Commands

Program commands for **Figure 3.10** Adjusted percentages of land that changed to developed land in Phuket province from 2000-2009 with rectangles denoting 95% confidence intervals of percentage change by land-use categories and sub-districts.

```
options(scipen=4)                   # display numbers properly
read.table("luPhuket1.txt", h=T,as.is=T) -> p

for (j in c(3,5,7,9,11)) {
 pj <- substr(p[,j],1,2)
 pj <- toupper(pj)
}

p$land10 <- ifelse(is.na(p[,3]),0,1)
p$land18 <- ifelse(is.na(p[,5]),0,1)
p$land28 <- ifelse(is.na(p[,7]),0,1)
p$land43 <- ifelse(is.na(p[,9]),0,1)
p$land52 <- ifelse(is.na(p[,11]),0,1)

p1 <- subset(p,p$land10==1 & p$land18==1 & p$land28==1 &
       p$land43==1 & p$land52==1)

p1[,c(1,2,9,10,11,12,16,17)]-> p1

read.table("Subdistrictsdig.txt", h=T,as.is=T) -> sub
merge(p1,sub,by.x=c("x","y"),by.y=c("x","y")) -> p1

u <- substr(p1$luCode43,1,2)
p1$lug43 <- ifelse(u %in% c("A1","A2","A5","A4","A7","A9","Z1"),2,
       ifelse(u %in% c("A3"),3,
       ifelse(u %in% c("F1","F2","F3","M1","M2","M3","M4","M8","W1","W2"),1,
       ifelse(u %in% c("U1","U2","U3","U4","U5","U6"),4,0))))
v <- substr(p1$luCode52,1,2)
p1$lug52 <- ifelse(v %in% c("A1","A2","A4","A5","A7","A9","Z1"),2,
       ifelse(v %in% c("A3"),3,
       ifelse(v %in% c("F1","F2","F3","M1","M2","M3","M4","M8","W1","W2"),1,
       ifelse(v %in% c("U1","U2","U3","U4","U5","U6"),4,0))))
#-----------------------------------------------------------
u <- p1$subdistrict
p1$sub <- ifelse(u=="Chalong",13,ifelse(u=="Choeng.Tale",5,ifelse(u=="Kammala",7,
       ifelse(u=="Karon",14,ifelse(u=="Kathu",8,ifelse(u=="Ko.Kaeo",10,
       ifelse(u=="Mai.Khao",1,ifelse(u=="Pa.Khlok",4,ifelse(u=="Patong",9,
       ifelse(u=="Ratsada",11,ifelse(u=="Rawai",15,ifelse(u=="Sa.Khu",2,
       ifelse(u=="Sri.Sunthon",6,ifelse(u=="Talat.Nua",17,ifelse(u=="Talat.Yai",16,
       ifelse(u=="Thep.Krasattri",3,ifelse(u=="Wichit",12,0))))) ))))) ))))) ))
p1$yU52 <- ifelse(p1$lug52==4,1,0)
#--------------------------------------------------
rez <- addmargins(table(p1$plotID43,p1$sub,p1$yU52))
rez[,,2] -> x
rez[,,3] -> n
nPlots <- dim(n)[1]-1
```

```
nGrps <- dim(n)[2]-1
n <- n[c(1:nPlots),c(1:nGrps)]
x <- x[c(1:nPlots),c(1:nGrps)]
ph <- colSums(x)/colSums(n)

r <- n
for (i in c(1:nGrps)) {
 r[,i] <- r[,i]*ph[i]
}
r <- x-r
ssr <- colSums(r^2)

m <- ifelse(n>0,1,0)
m <- colSums(m)

n <- colSums(n)
v <- ssr*m/((m-1)*n^2)

d0 <- n*v/(ph*(1-ph))                      # variance inflation factors
d0 <- ifelse(is.na(d0),1,d0)
d0 <- ifelse(d0<1,1,d0)

#-----------------------------------------------------
nSub <- length(unique(p1$sub))
tabS <- table(p1$sub)
namt <- names(tabS)
tab2S <- addmargins(table(p1$sub,p1$yU52))
tab2S <- cbind(tab2S,100*tab2S[,2]/tab2S[,3])
meanPc <- tab2S[nSub+1,4]
tab2S <- tab2S[1:nSub,]

subSiz <- tab2S[,3]
pcCrude1 <- tab2S[1:nSub,4]                # percent developed land by subdistrict
mod <- glm(data=p1,family=binomial,yU52~factor(sub)+factor(lug43))
drop1(mod,test="Chisq")                    # anova tests

rho <- tab2S[,3]/nrow(p1)                  # compute democratic confidence
D1 <- rbind(rho,cbind(diag(nSub-1),0))     # intervals using weighted
C1 <- solve(D1)                            # sum contrasts
C <- C1[,-1]
p1$xsub <- as.factor(p1$sub)
contrasts(p1$xsub) <- C

mod1 <- glm(data=p1,family=binomial,yU52~xsub+     # get coefs & SEs for all
subdistrict
       factor(lug43))                      # except the last one
rez1 <- summary(mod1)

hz <- max(p1$sub)
hy <- max(subset(p1$sub,p1$sub<hz))        # hsub1 has last and second last
p1$sub1 <- ifelse(p1$sub==hy,hz,           # subdistricts interchanged
       ifelse(p1$sub==hz,hy,p1$sub))

rho1 <- rho[c(1:(nSub-2),nSub,nSub-1)]
D1 <- rbind(rho1,cbind(diag(nSub-1),0))
C1 <- solve(D1)
C <- C1[,-1]
p1$xsub1 <- as.factor(p1$sub1)
contrasts(p1$xsub1) <- C
mod1a <- glm(data=p1,family=binomial,yU52~  # refit the model using these
       +xsub1+factor(lug43))               # contrasts
rez1a <- summary(mod1a)
```

```
SubCoef <- c(rez1$coef[2:nSub,1],rez1a$coef[nSub,1])
k <- -1.308458                              # constant needed
adjPcSub <- 100/(1+exp(-k-SubCoef))         # to make sample-size-weighted sum
                                            # of adjusted developed land proportions
sum(adjPcSub*table(p1$sub)/100)             # equal to sample-size-weighted
sum(p1$yU52)                                # sum of observed developed land
proportions

SubSE <- c(rez1$coef[2:nSub,2],rez1a$coef[nSub,2]) # get CIs for adjusted percentages

SubCILB <- 100/(1+exp(-k-(SubCoef-1.96*sqrt(d0)*SubSE)))
SubCIUB <- 100/(1+exp(-k-(SubCoef+1.96*sqrt(d0)*SubSE)))
#----------------------------------------------------
rez <- addmargins(table(p1$plotID43,p1$lug43,p1$yU52))
rez[,,2] -> x
rez[,,3] -> n
nPlots <- dim(n)[1]-1
nGrps <- dim(n)[2]-1
n <- n[c(1:nPlots),c(1:nGrps)]
x <- x[c(1:nPlots),c(1:nGrps)]
ph <- colSums(x)/colSums(n)

r <- n
for (i in c(1:nGrps)) {
 r[,i] <- r[,i]*ph[i]
}
r <- x-r
ssr <- colSums(r^2)

m <- ifelse(n>0,1,0)
m <- colSums(m)

n <- colSums(n)
v <- ssr*m/((m-1)*n^2)

d0 <- n*v/(ph*(1-ph))                       # variance inflation factors
d0 <- ifelse(is.na(d0),1,d0)
d0 <- ifelse(d0<1,1,d0)
#----------------------------------------------------
nlug43 <- length(unique(p1$lug43))
tab <- table(p1$lug43)
namt <- names(tab)
tab2 <- addmargins(table(p1$lug43,p1$yU52))
tab2 <- cbind(tab2,100*tab2[,2]/tab2[,3])
meanPc <- tab2[nlug43+1,4]
tab2 <- tab2[1:nlug43,]
luSiz <- tab2[,3]
pcCrude2 <- tab2[1:nlug43,4]                # percent developed land by land-use
cater
mod <- glm(data=p1,family=binomial,yU52~factor(lug43)+factor(sub))
drop1(mod,test="Chisq")                            # anova tests

rho <- tab2[,3]/nrow(p1)                    # compute democratic confidence
D1 <- rbind(rho,cbind(diag(nlug43-1),0))    # intervals using weighted
C1 <- solve(D1)                                    # sum contrasts
C <- C1[,-1]
p1$xlu <- as.factor(p1$lug43)
contrasts(p1$xlu) <- C

mod1 <- glm(data=p1,family=binomial,yU52~xlu+ # get coefs & SEs for all subdistrict
       factor(sub))                           # except the last one
```

```
rez1 <- summary(mod1)

hz <- max(p1$lug43)
hy <- max(subset(p1$lug43,p1$lug43<hz))
p1$lug431 <- ifelse(p1$lug43==hy,hz,
       ifelse(p1$lug43==hz,hy,p1$lug43))

rho1 <- rho[c(1:(nlug43-2),nlug43,nlug43-1)]
D1 <- rbind(rho1,cbind(diag(nlug43-1),0))
C1 <- solve(D1)
C <- C1[,-1]
p1$xlu1 <- as.factor(p1$lug431)
contrasts(p1$xlu1) <- C
mod1a <- glm(data=p1,family=binomial,yU52~  # refit the model using these
       +xlu1+factor(sub))                   # contrasts
rez1a <- summary(mod1a)

luCoef <- c(rez1$coef[2:nlug43,1],rez1a$coef[nlug43,1])
k <- -1.378417                              # constant needed
adjPclu <- 100/(1+exp(-k-luCoef))           # to make sample-size-weighted sum
                                            # of adjusted developed land proportions
sum(adjPclu*table(p1$lug43)/100)            # equal to sample-size-weighted
sum(p1$yU52)                                # sum of observed developed land
proportions

luSE <- c(rez1$coef[2:nlug43,2],rez1a$coef[nlug43,2]) #get CIs for adjusted percentages
luCILB <- 100/(1+exp(-k-(luCoef-1.96*sqrt(d0)*luSE)))
luCIUB <- 100/(1+exp(-k-(luCoef+1.96*sqrt(d0)*luSE)))

windows(8,6)
par(mar=c(2.5,2,2.8,2.5),mgp=c(1.1,0.2,0),oma=c(0,0,0,0),las=1,tcl=-0.2)
ymax <- 100
xmin <- 0
xmax <- nSub+nlug43+2
xCoord <- c(1:(nSub+1+nlug43))
plot(pcCrude1,type="n",ylim=c(0,100),xlim=c(xmin,xmax),ylab="",xlab="",xaxt="n",
       cex.axis=0.8,xaxs="i")
abline(h=10*c(0:10),col="grey")
abline(h=meanPc,col="red",lwd=2)
abline(v=nlug43+1)

luSizt <- luSiz*0.003333333
for (i in c(1:nlug43)) {
 points(xCoord[i]-c(0,0)-0.4,c(0,luSizt[i]),type="l",lwd=4,col="dimgrey")
}
subSizt <- subSiz*0.003333333
for (i in c(1:nSub)) {
 points(xCoord[i+nlug43+1]-c(0,0)-0.4,c(0,subSizt[i]),type="l",lwd=4,col="dimgrey")
}
d <- 0.15
dx <- d*c(-1,1,1,-1,-1)
dy <- c(0,0)
clrs <- ifelse(SubCIUB<meanPc,"pink",ifelse(SubCILB>meanPc,"cyan","yellow"))
for (i in c(1:nSub)) {
 polygon(xCoord[i+nlug43+1]+dx,c(SubCILB[i]+dy,SubCIUB[i]+dy,SubCILB[i]),col=clrs[i])
}
clrs <- ifelse(luCIUB<meanPc,"pink",ifelse(luCILB>meanPc,"cyan","yellow"))
for (i in c(1:nlug43)) {
 polygon(xCoord[i]+dx,c(luCILB[i]+dy,luCIUB[i]+dy,luCILB[i]),col=clrs[i])
}
xCoord1 <- xCoord[1:nlug43]
xCoord2 <- xCoord[(nlug43+2):(nlug43+nSub+1)]
```

```
points(xCoord1,pcCrude2,cex=1.2,pch=21,bg=2)
points(xCoord2,pcCrude1,cex=1.2,pch=21,bg=2)
points(xCoord1,adjPclu,cex=0.85,pch=3,bg=1)
points(xCoord2,adjPcSub,cex=0.85,pch=3,bg=1)

yat <- c(10*c(0:10))
ylab <- 3000*c(0:10)
axis(side=4,at=yat,lab=ylab,cex.axis=0.8)
xlab1 <- substr(rownames(tab2)[1:nlug43],1,2)
xlab2 <- substr(rownames(tab2S)[1:nSub],1,2)

axis(side=1,at=xCoord1,lab=xlab1,cex.axis=0.8,padj=-0.3)
axis(side=1,at=xCoord2,lab=xlab2,cex.axis=0.8,padj=-0.3)
axis(side=1,at=(1+nlug43)/2,lab="Land-use categories",padj=1.3,tcl=0)
axis(side=1,at=(nlug43+1.5+xmax)/2,lab="Subdistricts",padj=1.3,tcl=0)

mtext(side=3,adj=1.05,line=0.2,"Areas(ha)")
titl <- "Developed land percentage"
mtext(side=3,adj=-0.06,line=0.3,titl)

mtext(side=3,adj=0.5,line=1.5,"Land development in Phuket: 2000-2009")

legend("topleft",leg=c("Above mean","On mean","Below mean"),pch=22,
       pt.cex=2,pt.bg=c("cyan","yellow","pink"),title="95% CI",
       inset=c(0.42,0.02),y.intersp=0.8,bg="ivory",cex=0.8)

legend("topleft",leg=c("Crude Percent","Adj
Percent"),inset=c(0.6,0.02),pch=c(21,3),
       bg="ivory",cex=0.8,pt.cex=1,pt.bg=c(2,1),y.intersp=0.9)

d <- c("01:Mai Khao","02:Sa Khu","03:Thep Krasattri",
       "04:Pa Khlok","05:Choeng Tale","06:Sri Sunthon",
       "07:Kammala","08:Kathu","09:Patong",
       "10:Ko Kaeo","11:Ratsada","12:Wichit",
       "13:Chalong","14:Karon","15:Rawai",
       "16:Talat Yai","17:Talat Nua")

legend("topleft",leg=d,inset=c(0.225,0.01),x.intersp=0,
       bg="ivory",cex=0.77,pt.bg=2,y.intersp=0.8)

ln <- c("1:Undeveloped land","2:Other agriculture",
       "3:Rubber plantation","4:Developed land")

legend("topleft",leg=ln,inset=c(0.01,0.01),x.intersp=0,
       bg="ivory",cex=0.77,pt.bg=2,y.intersp=0.8)

legend("topleft",inset=c(0.8,0.02),leg="sample size",x.intersp=0.5,y.intersp=0.8,
       col="dimgrey",lwd=4,bg="ivory",cex=0.77)
```

Vitae

Name Miss Potjamas Chuangchang

Student ID 5420330015

Educational Attainment

Degree	Name of Institution	Year of Graduation
B.Sc. in Ed. (Biology)	Prince of Songkla University	2006
M.Ed. (Curriculum and Instruction)	Sukhothai Thammathirat Open University	2010

Scholarship Awards during Enrolment

1. Ph.D Scholarship for study from Prince of Songkla University, Songkhla, Thailand
2. Research Scholarship from the Graduate school, Prince of Songkla University, Songkhla, Thailand
3. Scholarship for Visiting at University of Malaya in Malasia from the Faculty of Science and Technology, Prince of Songkla University, Songkhla, Thailand
4. Scholarship for proceeding from the Graduate school and Faculty of Science and Technology, Prince of Songkla University, Songkhla, Thailand

List of Publication and Proceedings

Publications:

Chuangchang, P. and Tongkumchum, P. 2014. Modeling land development along highway 4 in Southern Thailand. Songklanakarin Journal of Science and Technology. 36(6), 719-725.

Chuangchang, P., Thinnukool, O., T. and Tongkumchum, P. 2016a. Modelling Urban Growth over Time using Grid-digitized Method with Variance Inflation Factors applied to Spatial Correlation. Arabian Journal of Geosciences. (In process). Accepted in February 2016.

Chuangchang, P., Sangkhaduang, T. and Tongkumchum, P. 2016b. Modeling Developed Land in Phuket Province of Thailand: 2000-2009. Journal of Social Sciences & Humanities. 24(2), (In process). Accepted in June 2015.

Proceedings:

Chuangchang, P. and Tongkumchum, P. Assessment of Land-use Change and Urban Growth in Phuket Province, Thailand. The 9th GMSARN International Conference 2014 on "Connectivity and Sustainability in GMS: Energy, Environmental & Social Issues". 12-14 November, 2014. Ho Chi Minh City, Vietnam.