The Age Dynamics of Para Rubber Plantations using Landsat Time Series (1991 – 2018) using Machine Learning Algorithms

Natthaphon Somching

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Environmental Management Technology
Prince of Songkla University
2019

The Age Dynamics of Para Rubber Plantations using Landsat Time Series (1991 – 2018) using Machine Learning Algorithms

Natthaphon Somching

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Environmental Management Technology
Prince of Songkla University
2019

Thesis Title	The Age Dynamics of Para Rubber Plantations using Landsat Time Series (1991 – 2018) using Machine Learning Algorithms
Author	Mr. Natthaphon Somching
Major Program	Environmental Management Technology

Major Advisor	**Examining Committee**
....	.... Chairperson
(Asst. Prof. Dr. Werapong Koedsin)	(Assoc. Prof. Dr. Raymond James Ritchie)
	.... Committee
Co-advisor	(Asst. Prof. Dr. Chanida Suwanprasit)
....	.... Committee
(Asst. Prof. Dr. Sangdao Wongsai)	(Asst. Prof. Dr. Werapong Koedsin)
	.... Committee
	(Asst. Prof. Dr. Sangdao Wongsai)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Master of Science Degree in Environmental Management Technology.

....

(Prof. Dr. Damrongsak Faroongsarng)

Dean of Graduate School

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

.... Signature
(Asst. Prof. Dr. Werapong Koedsin)
Major Advisor

.... Signature
(Mr. Natthaphon Somching)
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
(Mr. Natthaphon Somching)
Candidate

Thesis Title The Age Dynamics of Para Rubber Plantations using Landsat Time Series (1991 – 2018) using Machine Learning Algorithms
Author Mr. Natthaphon Somching
Major Program Environmental Management Technology
Academic 2018

ABSTRACT

The investigated the potential of using Landsat time series data and secondary land use and land cover (LULC) data to identify the ages of para rubber plantation in the lowland of Thalang district, Phuket province, southern Thailand. The LULC data, including high spatial resolution historical image from Google Earth Pro^{TM}, were used to identify rubber plantation and the event (year) of rubber planting (T_0). The inter-annual vegetation profiles of 2,168 rubber plantations were extracted from the distribution of sample NDVI values, which depend on the particular plot's size, for each summer period of 129 Landsat NDVI images (October 1991 to April 2018). The predictor variables were generated from difference and ratio of NDVI distribution values (minimum, Q1, median, Q3, and maximum) at different seasons (two years before and six years after T_0) for Recursive Partitioning (RP) supervised classification algorithm. Modeling data (outcome and predictors) from 336 plantations were divided into the training and testing datasets. The predicted RP model was learning on training data (30-time repeated) and we used testing data for cross-validation assessment to optimize an appropriated hyperparameter of the RP model. Then, the RP model with a complexity parameter as 0.01 was applied on both modeling data and predicting data (1,832 plots that unknowns T_0). The predicted T_0 for each plantation was selected based on the maximum nominated in 100-time repeated prediction. Finally, the result validation of T_0 prediction was carried out using 131 records of rubber farmers' registration from Rubber Authority of Thailand (RAOT). There are 15 plots (11.5%) that have a correct prediction, and 54 plots (41.2%) have one-year error prediction because many farmers start planting one or two years after approval from RAOT. The average error prediction is 3.62 years. We found that there is a possibility of using a 30-meter spatial resolution Landsat NDVI time series to identify rubber plantation ages with high accuracy, especially in the larger plots. The high precision of para rubber stands ages database will enable accurate yield prediction that, subsequently, resulting in better decision-making, planning, and development in the agricultural sector.

Keywords: para rubber, plantation age, machine learning, Landsat, time series

ชื่อวิทยานิพนธ์ พลวัตอายุยางพาราจากข้อมูลแลนด์แซทหลายช่วงเวลาปี 2534 – 2561 โดยใช้อัลกอริทึมการเรียนรู้ของเครื่อง
ผู้เขียน นายณัฐพล สมจริง
สาขาวิชา เทคโนโลยีการจัดการสิ่งแวดล้อม
ปีการศึกษา 2561

บทคัดย่อ

การศึกษาถึงศักยภาพของการใช้ข้อมูลหลายช่วงเวลาจากดาวเทียม Landsat และข้อมูลการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดิน (LULC) สำหรับใช้ในการระบุอายุของสวนยางพาราในพื้นที่ราบของอำเภอถลาง จังหวัดภูเก็ต ภาคใต้ของประเทศไทย ข้อมูล LULC รวมไปถึงประวัติของภาพถ่ายดาวเทียมที่มีความละเอียดเชิงพื้นที่สูงจาก Google Earth Pro™ ถูกนำมาใช้เพื่อระบุขอบเขตของแปลงและปีที่พบเหตุการณ์ของการปลูกยางพารา (T_0) การสร้างข้อมูลภาวะการณ์พืชพรรณประจำปีของสวนยางพาราจำนวน 2,168 แปลง ได้รับการดึงข้อมูลจากการกระจายของจุดตัวอย่างค่า NDVI ซึ่งขึ้นอยู่กับขนาดของแปลงในแต่ละช่วงฤดูร้อนที่มีข้อมูล NDVI จากภาพถ่ายดาวเทียม Landsat จำนวน 129 ภาพ (เดือนตุลาคม พ.ศ. 2534 ถึงเดือนเมษายน พ.ศ. 2561) ตัวแปรสำหรับการทำนายชุดข้อมูลถูกสร้างขึ้นจากค่าความแตกต่างและอัตราส่วนของค่าการกระจาย NDVI (ค่าต่ำสุด, ค่าควอไทล์ที่ 1, ค่ามัธยฐาน, ค่าควอไทล์ที่ 3, และค่าสูงสุด) ในฤดูกาลที่แตกต่างกัน (สองปีก่อนและหกปีหลังจาก T_0) สำหรับอัลกอริทึมการจำแนกประเภทภายใต้การกำกับดูแลด้วยวิธี Recursive Partitioning (RP) ข้อมูลการสร้างแบบจำลอง (ผลลัพธ์และตัวทำนาย) จากสวนยางจำนวน 336 แปลงถูกแบ่งออกเป็นชุดข้อมูลสำหรับใช้ในการฝึกอบรมและการทดสอบ แบบจำลอง RP ที่ทำนายไว้เป็นการเรียนรู้บนชุดข้อมูลการฝึกอบรม (การทำซ้ำ 30 ครั้ง) และใช้ข้อมูลการทดสอบสำหรับการประเมินความถูกต้องด้วยวิธี Cross-validation เพื่อเพิ่มประสิทธิภาพของไฮเปอร์พารามิเตอร์ให้เหมาะสมกับแบบจำลอง RP จากนั้นนำแบบจำลอง RP ที่มีพารามิเตอร์ที่ซับซ้อนเท่ากับ 0.01 ไปใช้กับข้อมูลการสร้างแบบจำลองและข้อมูลการทำนาย (1,832 แปลงที่ไม่รู้ T_0) การทำนายเหตุการณ์ T_0 ของสวนยางแต่ละแปลงได้รับการคัดเลือกตามการเสนอเหตุการณ์ที่เป็นไปได้สูงสุดในการทำนายซ้ำจำนวน 100 ครั้ง สุดท้ายการตรวจสอบความถูกต้องของผลลัพธ์จากการทำนาย T_0 ด้วยข้อมูลการขึ้นทะเบียนของเกษตรกรชาวสวนยางจากการยางแห่งประเทศไทย (RAOT) 131 แปลง พบว่าจำนวน 15 แปลง (11.5%) มีการทำนายว่าถูกต้อง และจำนวน 54 แปลง (41.2%) มีการทำนายที่ผิดพลาดไปหนึ่งปี เพราะเกษตรกรจำนวนมากเริ่มปลูกยางในช่วงหนึ่งหรือสองปีหลังจากได้รับการอนุมัติจาก RAOT การคาดคะเนข้อผิดพลาดเฉลี่ยคือ 3.62 ปี เราพบความเป็นไปได้ที่จะใช้ชุดข้อมูล NDVI หลายช่วงเวลาจากภาพถ่ายดาวเทียม Landsat ที่มีความละเอียดของภาพ 30 เมตร เพื่อระบุอายุของสวนยางพาราที่มีความแม่นยำสูงโดยเฉพาะอย่างยิ่งในแปลงที่มีขนาดใหญ่ ความแม่นยำสูงของฐานข้อมูลอายุยางพาราจะช่วยให้สามารถคาดการณ์ผลผลิตได้อย่างถูกต้อง ซึ่งส่งผลให้การตัดสินใจ การวางแผน และการพัฒนาในภาคเกษตรกรรมดีขึ้น

คำสำคัญ: ยางพารา, อายุการปลูก, การเรียนรู้ของเครื่อง, แลนด์แซท, หลายช่วงเวลา

ACKNOWLEDGEMENT

Firstly, I would like to express my gratitude, respect and appreciated to my advisor is Asst. Prof. Dr. Werapong Keodsin, the remote sensing lecturer at Faculty of Technology and Environment, Prince of Songkla University, Phuket Campus and my co-advisor are Asst. Prof. Dr. Sangdao Wongsai, the statistics lecturer in the Department of Mathematics and Statistics at the Faculty of Science and Technology, Thammasat University, Rangsit Campus. They are both patrons and benefactors to continuous encourage to my study at the master’s degree and research. I thank them for their patience as a driving force of the various aspects of my thesis has been accomplished with great success.

I would like to especially thank Mr. Noppachai Wongsai, Ph.D. students in the Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus. He was one of the key people who helped in my Thesis. I am grateful for his advice on my thesis such as research methodology, writing code in the R Project for Statistical Computing.

I would like to thank the rest of my thesis committee; Asst. Prof. Dr. Chanida Suwanprasit, the geoinformatics lecturer in the Department of Geography at the Faculty of Social Sciences, Chiang Mai University, and Assoc. Prof. Dr. Raymond James Ritchie, the plant physiology lecturer at Faculty of Technology and Environment at Prince of Songkla University, Phuket Campus for sacrificing time in helping to express opinions, questions, suggestions, assistance in proofreading. This is a great benefit to my research.

I would like to especially thank the director and staffs, the Rubber Authority of Thailand, Phuket Province for support the information of registration of rubber farmers in Thalang District.

I would like to thank scholarship support from the Graduate School and Faculty of Technology and Environment, Prince of Songkla University, Phuket Campus Finally, I would like to thank my parents and younger brother for my patronage and inspiration during the study including my best friend for accepting my actions and always giving encouragement. Thanks for all your encouragement.

Natthaphon Somching

CONTENTS

CONTENTS (cont.)

CONTENTS (cont.)

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
DNs	Digital Numbers
DTs	Decision Trees
ETM+	Enhanced Thematic Mapper Plus
GIS	Geographic Information System
GISTDA	Geo-informatics and Space Technology Development Agency
GPS	Global Positioning System
LDD	Land Development Department
LTS	Landsat Time Series
LULC	Land Use Land Cover
ML	Machine Learning
NDVI	Normalized Difference Vegetation Index
OLI	Operational Land Imager
RAOT	Rubber Authority of Thailand
RP	Recursive Partitioning
RS	Remote Sensing
SR	Landsat Surface Reflectance
TM	Thematic Mapper
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
VI	Vegetation Index

CHAPTER 1

INTRODUCTION

1.1 Statement of the Problem

Para rubber tree (*Hevea brasiliensis* Muell. Arg.) is one of the significant economic crops of Thailand as well as the South East Asian Nation (ASEAN) region both in terms of income and employment (Koedsin and Yasen, 2016). Originally, they were a native plant in the Amazon Basin of South America. They have grown in numerous locations within the equatorial zone between 10°N and 10°S in areas with continuous twelve-month rainfall (Ratnasingam, Ioraş, and Wenming, 2011). Para rubber trees grow mainly in humid tropical lowlands below 400-meter altitude, covered by dense tropical rainforest (Fox and Castella, 2013). It is a quick-growing perennial plant, rarely exceeding 25 – 30 meters in height in plantations. The para rubber trees start yield rubber latex between 5 – 7 years of age and have a productive lifespan of between 25 – 30 years (Verheye, 2010). The natural rubber latexes are mainly used as a raw material for making various rubber products in many industries, such as vehicle tires, surgeons' gloves, shoes, condoms, sports equipment, balloons, pillow, cosmetics, and other relatively high-value products, etc. (Tekasakul and Tekasakul, 2006). The rubberwood can also be used in the timber industry in furniture, children toys and fuel woods, etc. (Teoh, *et al*., 2011). Currently, about 90% of the total world production of natural rubber is obtained from *Hevea brasiliensis* Muell. Arg. (Li and Fox, 2012).

The background, importance, and problems related to para rubber plantations and the rubber industry in Thailand. The beginning of this industry was 1899 – 1901. Phraya Ratsadanupradit Mahison Phakdi (his former name was Khaw Sim Bee Na-Ranong) as governor of Trang Province and Monthon Phuket, who is being honored as "Father of Thai Rubber", along with Phra Sathon Sathanphitak, who was an adopted son of Phraya Ratsadanupradit Mahison Phakdi and the owner of first para rubber plantation of Thailand. They are both famous people brought in the country's first para rubber tree which was planted in Thailand (Jawjit, *et al*., 2010). They both visited Malaysia and Indonesia and brought some para rubber varieties from there income to be planted in the governor's residence at Kantang District, Trang Province and propagated in Phuket Province to be replaced for the sluggish tin mining industry and pepper. They commissioned public servants to go learned how to manage the para rubber plantations and rubber tapping technique from the neighboring countries and then came back to teach their villagers. Later, the villagers received para rubber varieties to propagate species throughout the south of the country (Agricultural Research Development Agency (Public Organization), 2018). In the early stages, para rubber was planted in 14 southern provinces of the country due to the climate of tropical rainforest, which suitable for the growth of para rubber trees (Riwthong, *et al*., 2017). In the year 1991, Thailand developed the country into the world's largest rubber producer and exporter of natural rubber, accounting for about one-third of global supply, which was followed by Indonesia and Malaysia. The rubber price has generally increased due to world demand and the expansion of the world economy (Romyen, *et al*., 2018). In the middle of 2002, para rubber cultivations are more popular because can export rubber yields to foreign countries, especially China and the USA. The rubber prices have risen steadily until the Thai government at the time, issued policies and encouraged the farmers to grow para rubber plantations increased up to 1,600 square kilometers nationwide (Bhumiratana, *et al*., 2013). About 35% of the latex produced worldwide comes from Thailand in the year 2007 (Jawjit, *et al*., 2010). The rubber prices were up to 180 baht per kilogram in February 2001. The popularity of para rubber plantations in Thailand has increased in almost every region of the country (Romyen, *et al*., 2018). China is Thailand’s most important rubber trading partner has demanded more natural rubber latex used in the manufacture of tires. China imported 2,500,000 tons of rubber per year or 32% from abroad in the year 2013 (Gale, *et al*., 2015). Southeast Asia suffers about 97% of the world's natural rubber, mainly from Thailand 31%, Indonesia 30%, and Malaysia 9% (Li and Fox, 2012). Thailand's rubber exports amounted to 170,418.73 million baht, which ranked 6th of all the country's exports to the world market in the year 2015 (Wadeesirisak, *et al*., 2017). Domestic rubber production has increased to about 4,500,000 tons per year, without quantitative control. Until the world economy slowed down in the later year but the world's popularity of natural rubber has remained. As a result, China wants to reduce imports of rubber from abroad and turn into a producer. China has invested para rubber plantations in the Xishuangbanna, Yunnan, southern China including other neighboring countries are Laos, Vietnam, and Cambodia. The results of this have been that rubber prices in Thailand have dropped sharply (Smajgl, *et al*., 2015; Petchseechaung, 2016). Nowadays, the situation is that the price of rubber is down to 30 – 45 baht per kilogram

(Rubber Authority of Thailand, 2018). Thailand's rubber industry is still suffering from the decline in product prices, due mainly to high global oversupply. The market price of raw natural rubber depends on the global market, which in the futures contract trading system in the stock market including the lack of an accurate database of para rubber plantations to predict natural rubber latex contributions, results in the market price of rubber products in Thailand being uncertain (Petchseechaung, 2016; Petchsawang, 2018). Thailand is still an agricultural country due to the population more than half are cultivators. The government policy of the country is focused on the agriculture industry (Singhapreecha, 2014; Boonyanam, 2018). This problem is a very important issue that the Thai government, under the management of the Ministry of Agriculture and Cooperatives is solving due to the push from the rubber planters nationwide reached 1.54 million families (Post Today, 2017). The Ministry of Agriculture and Cooperatives of Thailand, 2017 has proposed a solution to three measures: 1) increase consumption of domestic rubber products 2) help operators to have the ability to buy rubber products and 3) reduce para rubber planting area (Thairath, 2017). Recently, the Rubber Authority of Thailand (RAOT) under the management of the Ministry of Agriculture and Cooperatives of Thailand, oversees and provides support to research, production, and commercialization of rubber across the entire value chain nationally, from the cultivation of para rubber trees to final processing, partnership with Geo-Informatics and Space Technology Development Agency (GISTDA) under the operation of the Ministry of Science and Technology of Thailand, both have interests and cooperated to generate the project of a spatial database and mapping for monitoring the para rubber plantations nationwide to develop effective monitoring of para rubber plantations (Petchsawang, 2018).

In this study, we need to understand the precision agriculture and yield prediction for the management of para rubber re-planting in the future to help solve the problem of rubber prices falling long-term and sustainable (Satir and Berberoglu, 2016; Boonyanam, 2018). Including knowledge on geoinformation and remotely-sensed imagery time series combined with machine learning algorithms was introduced to investigate the potential of data using precision satellite technology to identify para rubber plantation ages systematically and concretely (Brown, 2015). These are the basic information needed for the management of government projects that are stable and beneficial to all sectors of the rubber industry in the future. We selected Thalang District, Phuket Province, located in Southern Thailand as a case study due to mostly agricultural areas are para rubber plantations and the second largest industry income in the province, after the tourism industry (Phuket Development Strategy, 2016).

1.2 Objective

The objective of this study is to investigate the potential of using dense Landsat time series imagery and land use and land cover (LULC) data to precisely identify para rubber plantation ages using a tree-based machine learning algorithm.

1.3 Scope

The study area was the para rubber plantations situated in lowland (below 50 meters above mean sea level) in Thalang District, Phuket Island, Southern Thailand. The data used in this study consisted of spectral indices time series derived from 30-meter spatial resolution Landsat imagery which obtained in during October 1991 to April 2018. The secondary LULC data of the year 2016 and the registration information of rubber farmers in the study area from 2013 to 2018 acquired from the government agencies were used to identify para rubber plantations and validation, respectively. The decision tree learning algorithm, one of the well-known Machine Learning (ML) techniques, was adopted to examine the potential of using different spectral profiles to identify the age of rubber plantation.

1.4 Expected

A new method of identifying para rubber plantation ages using remote sensing data and LULC data will provide the high precision of para rubber plantation ages data that could enable accurate yield prediction and, subsequently, resulting in better decision-making, management planning and development in the agricultural sector. This approach and developed model in this study will also be able to extend in other agricultural areas of research in the future.

1.5 Definition

"Para Rubber (*Hevea brasiliensis* Muell. Arg.)" means a native rubber tree from tropical of South America. That bark is gray, black or brown, and when tapping on the bark. It will give natural white latexes.

"Rubber Plantation" means the agricultural land for planted para rubber trees, which has an area of not less than 2 rai (rai is a measure of the area of Thailand, 1 rai equal to 1,600 square meters) and easily identified by resolution of satellite images.

"Rubber Farmers" means owners, tenants, growers and tappers who are entitled to receive about various yields from para rubber trees in their own para rubber plantations.

"Re-planting" means the planting of new para rubber trees in order to replace all or part of the old para rubber trees in that area.

"Vegetation Index" means a proportion of vegetation-covered surface by calculating from the wavelength associated with vegetation proportional mutually.

"Time Series" means a set of satellite image data that is collected over a period of time continuously

"Geographic Information System" means a computer application used to store, analysis, and presentation of geographical and statistical information.

"Remote Sensing" means a field of science in acquiring information related to various objects placed in the area from remote data recording tools including aircraft or satellites, based on the properties of electromagnetic waves.

"Machine Learning" means a learning part of the machine being used as a brain of Artificial Intelligence (AI) to create intelligence, often using models derived from the learning of AI, not from human writing.

CHAPTER 2

LITERATURE REVIEW

2.1 Para Rubber (*Hevea brasiliensis* Muell. Arg.)

Para Rubber (scientific name is called "*Hevea brasiliensis* (A. Juss) Muell. Arg.") is a plant in the family Euphorbiaceae. It is a tropical tree and native to the Amazon Basin in Brazil and neighboring countries. It was taken from the Amazon region to many other tropical regions of the world, such as the South and Southeast Asia (ASEAN) by the British Colonial Office (Verheye, 2010). Including Thailand by Phraya Ratsadanupradit Mahison Phakdi (Kosimbee Na-Ranong) as the governor of Trang and Monthon Phuket (1899 – 1901) (Agricultural Research Development Agency (Public Organization), 2018). Para rubber trees grow mainly in tropical lowlands below 400-meter altitude, originally covered by a dense tropical rainforest (Fox and Castella, 2013). It grows best at temperatures between 20 – 28°C with an annual rainfall distributed between 1,800 – 2,000 millimeters. The mature trees on para rubber plantations' height are 25 – 30 meters, with girth tree between 2.0 – 3.0 meters, smooth straight trunks, bark grayish, strong roots, alternate trifoliate leaves, petioles long are 7.5 – 10 centimeters, flowers numerous, female flowers apical, fruit the 3-lobed or 3-seeded ellipsoidal capsule, variable in size between 2.5 – 3 centimeters, mottled brown and weighing 2 – 4 grams each. It requires a special fertilizer for growth in the first 6 years when biomass is built up. The para rubber trees start yield rubber latex between 5 – 7 years of age and have a productive lifespan of a tree between 25 – 30 years (Verheye, 2010). After that, the area will be cleared for re-planting of para rubber trees. The natural rubber latexes are used as a raw material for making various rubber products in many industries such as concentrated latex, block rubber and rubber smoke sheet (Tekasakul and Tekasakul, 2006). The rubberwood can also be used in the timber industry in furniture, children's toys and firewood (Teoh, *et al*., 2011). About 90% of the total world production of natural rubber is obtained from *Hevea brasiliensis* Muell. Arg. Rubber species in Thailand, Malaysia, and Indonesia (Killmann and Hong, 2000; Li and Fox, 2012).

2.1.1 Good Agricultural Practice (GAP) for Thailand

Good Agricultural Practice is the manual of principles, regulations and technical recommendations applicable to production, processing, human health care, environmental protection and improvement of worker conditions and their families for applied to agriculture, create nutrition for consumers or further processing safe and beneficial. Although the definition of competition is that the method used in good agricultural practice, there are many broadly accepted models that manufacturers can follow. The GAP farming system must analyze the history of agriculture. This principle has been designated by the United Nations Food and Agriculture Organization (FAO) (Izquierdo, *et al*., 2007).

In Thailand, the agencies responsible for quality assurance certification of GAP is the Department of Agriculture, under the management of the Ministry of Agriculture and Cooperatives of Thailand, have defined the requirements, rules, and methods of auditing, which according to with the GAP principles of international. This is used as a standard for crop production at the farm level of the country and to prepare a manual for the cultivation of GAP about 24 important plant crops in Thailand including 1) Fruits; Durian, Longan, Pineapple, Pomelo, Mango and Tangerine 2) Vegetables; Tomatoes, Asparagus, Kale, Onion, Cabbage, Chili, Long bean, Peanuts, White cabbage, Baby corn, Onion Sprouts, and Shallots 3) Flowers; Orchid and Siam Tulip and 4) Other plants; Robusta coffee, Cassava, and Para rubber (National Bureau of Agricultural Commodity and Food Standards of Thailand, 2013).

Department of Agriculture of Thailand has provided Good Agricultural Practice for para rubber cultivation in Thailand. Thailand is located in the tropical zone, with an environment is suitable for planting para rubber especially in the southern and some of the eastern provinces, where the original para rubber plantations were set up. Later, para rubber plantations were expanded to new para rubber plantations in the northeast and north, where the environment was less favorable for the para rubber planting such as lack moisture, low temperature, strong winds, high elevation, steep slopes, low soil depth, low poor drainage, and low soil chemical properties. The para rubber trees are adaptable to different environments. The para rubber trees are grown in the south can be opened rubber tapping very quickly and given yield more than the para rubber planted in the north and northeast for approximately 6 months of the year. The yields of para rubber trees are latex or wood. The quality depends on three factors: 1) good varieties 2) the suitability of the area and 3) quality of management of the cultivation (Rubber Research Institute, 2010), as shown in Table 2.1.

Table 2.1 GAP for Thailand's para rubber. (Surat Thani Rubber Research Center, 2018)

Para Rubber Varieties for High Yield	
Latex	RRIT 408, RRIT 251, RRIT 226, BPM 24, RRIM 600
Latex and Wood	RRII 118, PB 235, PB 255, PB 260, PR 255, RRIC 110
Wood	RRIT 402, BPM 1, AVROS 2037
Suitability of the Growing Area	
Area Condition	- The plain area or slope of less than 35°. - If slope exceeds 15° to do the ladder and planted cover crops to prevent soil erosion. - The elevation of the sea level should not exceed 600-m and not flooded. Plants are sensitive to flooding.
Land Form	- The soil is sticky to sandy soil with abundant. - The soil depth of not less than 1-m, no solid rock layer. - Good drainage and ventilation. - The groundwater level less than 1-m. - The pH value is 4.5 – 5.5.
Climate and Weather	- Rainfall is not less than 1,250 millimeters per year. - The rainy days are about 120 – 150 days per year.
Water Source	- Rainwater
Managed Plantations	
Preparation Area	- Adjusts area to suitable and vegetation clearance. - Orient the plantation along the east – west. - Interval rows of old plots are 2.5*8.0-m or 3.0*7.0-m. - Interval rows of new plots are 2.5*7.0-m or 3.0*6.0-m.
Cultivation	- Plants should grow in the early rainy season. - Should repair plots before the end of the rainy season. - Planting repair should not be carried out. If age plots exceed 2 years.
Mixed Plants	- Should grow short-lived plants in during the first 3 years including Banana, Beans, Cane, Cassava, Corn, Forage, Papaya, Pineapple, and Watermelon, etc. - Do not grow mixed plants after 3 years such as cassava. - Understory mixed plants that can tolerate shade trees; Acacia, Anthurium, Dahla, Galangal, Ginger, Krachiew, Rattan, Salacca, and Turmeric, etc.

2.1.2 Rubber Authority of Thailand Act B.E. 2558 (2015)

Rubber Authority of Thailand (RAOT) originated from the combination of three agencies including 1) Office of the Rubber Replanting Aid Fund, 2) Rubber Plantation Organization and 3) Rubber Research Institute. According to the Rubber Authority of Thailand Act B.E. 2558 (2015), announced in the Government Gazette on 14 July 2015 and became effective on July 15, 2015 (Thailand Government, 2015). The objectives of the Rubber Authority of Thailand are the central organization responsible for overseeing the management of the country's rubber system completely, fund management, promote and support the country to be the center of the rubber product industry by providing education, analysis, research, development, dissemination of information about rubber and to ensure stable rubber price levels. As well as providing assistance to rubber farmers, rubber farmers institutions, rubber business operators by providing knowledge in the terms about academic, finance, manufacturing, processing, industry, marketing, business, and other related operations to raise income and quality of better life (Rubber Authority of Thailand, 2019).

2.1.3 Precision Agriculture (PA)

Precision Agriculture (also known as “Satellite Agriculture”) is an integrated crop management system or application of geospatial techniques that combine information technologies including geographic information systems, remote sensing, global positioning system with rational agricultural industries to ensure that the crops and soil receive exactly what they need for optimum health and productivity. The PA's goal is to ensure profitability, sustainability, and environmental protection. (Ashwini, 2017). The concept of PA has become an interesting idea for managing natural resources and recognizing the development of modern sustainable agriculture. The PA’s remote sensing technology is used for collecting and analyzing information about crop and soil characteristics using sensors mounted on satellites, aircraft, UAV, or ground equipment (Far and Rezaei-Moghaddam, 2018). The PA’s technology in Thailand is known as “Smart Farming”, which is used to analyze important agricultural products, such as rice, para rubber, sugarcane, vegetables, and fresh fruits, etc. This technology helps in the mapping of the growing area and plot area, with accuracy and speed, leading to the improvement in managing the means of production (Kraipinit, *et al*., 2017).

2.1.4 Yield Prediction (YP)

Yield Prediction (also called as “Crop Prediction”) is an important agricultural problem. This is very popular among farmers these days, which particularly contributes to the proper selection of crops for sowing. In the past, yield prediction was performed by considering the farmer's previous experience on a particular crop such as topography, climate, rainfall and disease in the agricultural area. Therefore, there is a need to attempt better technique for yield prediction in order to overcome the problem (Paul, *et al*., 2015). From the study about the crop yield prediction using satellite-derived vegetation indices in the year 2016, found that the Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) indices are more effective predictors of yield prediction (Satir and Berberoglu, 2016). The cultivation of para rubber requires high investment and long-term returns. If the rubber farmers decide to invest in a para rubber plantation, then they cannot change the para rubber varieties until the time for vegetation clearance area for re-planting. The rubber farmers must ensure that they can to survive in the global market situation, there are risks and uncertainty in the sustainability of their household because of the rubber cultivations are a long-term investment (Kongmanee and Longpichai, 2017).

2.2 Landsat Time Series (LTS)

Landsat is the name of a series of satellites exploring natural resources and record data the Earth's surface for almost four decades. Initially, the phased project was under the management of NOAA's US organization. The Earth Resources Technology Satellite was launched on July 23, 1972. Then take over by the management of Earth Observing Satellite Company (EOSAT) in the year 1984. The U.S. government has set an obligation’s government mandate for exploring satellite resources continuously in the legal name that “The 1992 Land Remote Sensing Policy Act” and Landsat is backed under the management of USGS and NASA in the U.S. Global Change Research Program (U.S. Geological Survey, 2016). Nowadays, Landsat’s mission is to explore and disseminate information for the benefit of civilians and the contained development of satellite-based exploration equipment. Therefore, scientists have remote sensing data from four decades of monitoring to analyze changes in the world’s phenomena (Loveland and Dwyer, 2012). The information of Landsat TM, ETM+ and OLI satellite imagery including sensor specifications, image characteristics, as shown in Tables 2.2, 2.3 and 2.4.

2.2.1 Landsat-5 Thematic Mapper (TM)

Landsat-5 TM satellite was launched by NASA into orbit with a rocket of the McDonald Douglas Delta 3920 from the Vandenberg Air Force Base, California on March 1, 1984. Landsat 5 transmitted its last image on January 6, 2013.

2.2.2 Landsat-7 Enhanced Thematic Mapper Plus (ETM+)

Landsat-7 ETM+ satellite was launched by NASA, NOAA and USGS into orbit with a rocket of the McDonald Douglas Delta II 7920 from the Vandenberg Air Force Base, California on April 15, 1999. Currently, still working in orbits and records.

2.2.3 Landsat-8 Operation Land Imager (OLI)

Landsat-8 OLI satellite was launched by USGS into orbit by a rocket of the Atlas V401 AV-035 from the Vandenberg Air Force Base, California on February 11, 2013. Nowadays, still working in orbits and records (U.S. Geological Survey, 2016).

Table 2.2 Specifications of Landsat satellite. (U.S. Geological Survey, 2018)

Specifications	**Landsat Satellite**		
	TM	**ETM+**	**OLI**
Perigee – Apogee	694 – 701-km	701 – 703-km	701 – 703-km
Regime	Sun-synchronous	Sun-synchronous	Sun-synchronous
Inclination	98.2°	98.2°	98.2°
Period	98.7-minute	98.8-minute	98.8-minute
Repeat Interval	16-day	16-day	16-day
Pixel Size	30-meter	30-meter	30-meter
Width Band	170*185-km	170*185-km	170*185-km
Quantization	8-bit (0-255)	8-bit (0-255)	16-bit (0-65,535)

Table 2.3 Band applications of Landsat TM/ETM+. (U.S. Geological Survey, 2018) *

Wavelength	Applications
Band 1 Blue (0.45 - 0.52 μm)	Distinguishing Soil and Vegetation
Band 2 Green (0.52 - 0.60 μm)	Assessment Plant Vigor
Band 3 Red (0.63 - 0.69 μm)	Discriminates Vegetation Slopes
Band 4 NIR (0.76 - 0.90 μm)	Biomass and Shorelines
Band 5 SWIR-1 (1.55 - 1.75 μm)	Plants and Soil Moisture, Cloud, Snow
Band 6 Thermal (10.40 - 12.50 μm)	Surface Heat, Soil Moisture, Plant stress
Band 7 SWIR-2 (2.08 - 2.35 μm)	Rocks, Mineral, Deposits

Table 2.4 Band applications of Landsat OLI. (U.S. Geological Survey, 2018) *

Wavelength	Applications
Band 1 Coastal Aerosol (0.43 - 0.45 μm)	Coastal and Aerosol Studies
Band 2 Blue (0.45 - 0.52 μm)	Distinguishing Soil and Vegetation
Band 3 Green (0.53 - 0.59 μm)	Assessment Plant Vigor
Band 4 Red (0.64 - 0.67 μm)	Discriminates Vegetation Slopes
Band 5 Near-infrared (0.85 - 0.88 μm)	Biomass and Shorelines
Band 6 SWIR-1 (1.57 - 1.65 μm)	Plants and Soil Moisture, Cloud, Snow
Band 7 SWIR-2 (2.11 - 2.29 μm)	Plants and Soil Moisture, Cloud, Snow
Band 9 Cirrus (1.36 - 1.38 μm)	Cirrus Cloud Contamination

* The band data of Landsat satellite images with 30-meter spatial resolution only.

2.3 Vegetation Index (VI)

Vegetation Index is the proportion of vegetation-covered surface by calculating from the wavelength associated with vegetation proportional mutually. This index is useful in monitoring the increase or decrease of vegetation, phenological and environmental situations (Fang and Liang, 2008), which one of the most popular indexes is called “Normalized Difference Vegetation Index”.

2.3.1 Normalized Difference Vegetation Index (NDVI)

Normalized Difference Vegetation Index is a measure of the ratio between two wavelengths, which are normally distributed by the calculation NIR bands

(0.86 μm) and Red band (0.66 μm). This index can be used to analyze, measure remotely, and evaluate targets or objects that are being observed for the greenness of vegetation or biomass. The NDVI was first used and developed in the year 1973 by Rouse, *et al*. from the Remote Sensing Centre of Texas A&M University, USA (Rouse, *et al*., 1973). The NDVI values always range from -1 to +1, which can explain the NDVI values, as shown in Table 2.5. The NDVI equation is as follows;

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

Definition of the Terms:

NDVI is the Normalized Difference Vegetation Index

NIR is the Near-infrared band

RED is the Red band

Table 2.5 NDVI values in the ecosystem. (Muradyan, *et al*., 2016)

NDVI Values	Biomass and Vegetation Covers
-1.00 – 0.10	Bare Soil, Water Surface
0.11 – 0.20	Minimum Biomass, Very Low-standing Grass Vegetation
0.21 – 0.30	Middle Biomass, Low-standing Grass Vegetation
0.31 – 0.55	High Biomass, High-standing Grass Vegetation
0.56 – 0.70	Very High Biomass, High-standing Grass Vegetation
0.71 – 1.00	Maximum Biomass, Forested Areas

The advantage or benefits of vegetation index as follows;

- To study the distribution and integrity of plants overall.
- To study the conditions of drought and integrity of the area over time.
- To study dynamics and vegetation phenology of plants.
- To classification of plant types including the amount of vegetation.
- To calculate relative biomass values.

The relationship between the dynamics of NDVI values associated with vegetation phenology of various agricultural areas, including para rubber cultivation can be described as follows. If the level of VIs is higher than the minimum value during the uptrend graph of the planting cycle, it means that the time is “Seedling” status. The VIs increases to a higher level will be considered as a status “Growing”. Until at one level the VIs has the highest value, the plant will be considered as "Complete" status. After that, the VIs will be relatively stable according to the lifespan of the plant, but the value of VIs maybe changes somewhat depending on the season, leaves or climate change. Finally, the level of VIs decreases from the highest to the lowest point during the downward graph will be considered as "Harvested" or "Vegetation Clearance" for re-planting of plants in the new generation (Glenn, *et al*., 2008).

2.4 Machine Learning (ML)

Machine Learning is one of the science that arises from the integration of knowledge in various fields such as computer science, and statistics using the brain of Artificial Intelligence (AI) for making computer systems can self-learning in analyze, predict, and process using data from human input (data scientist) to be used in the decision (Shobha and Rangaswamy, 2018). However, the data scientist must design various variables and find other algorithms to compare looking for the most suitable algorithm in actual use. The current application of ML is widely found in various industries such as science, engineering, medicine, and marketing (Willcock, *et al*., 2018). Machine learning has three main forms of learning:

- Supervised learning is learning under supervision.
- Unsupervised learning is learning without supervision.
- Reinforcement learning is learning from the environment.

For the example of algorithms in ML has been used many purposes such as Support Vector Machine, Naive Bayes, Gradient boosting, K-nearest Neighbor, K-mean, Markov Decision Processes, Linear Regression, Logistic Regression, and Q-learning, etc. (Wakefield, 2019). For the methods was used in this study as supervised learning, namely Recurve Partitioning algorithm.

2.4.1 Decision Trees (DTs)

Decision Trees is a commonly used data mining method for establishing classification systems (supervised learning) based on multiple covariates or for developing prediction algorithms for a target variable or regression models in the form of a tree structure, which this is an important type of algorithm for predictive modeling of ML (Kingsford and Salzberg, 2008). A primary advantage of using a DTs is that it is easy to follow and understand (Quinlan, 1986). The capabilities of the method of DTs in remote sensing applications such as classification LULC, biomass estimation, etc. (Friedl and Brodley, 1997; Graves, *et al*., 2018). For describing the model result is a tree with decision nodes and leaf nodes. A decision node has two or more branches. Leaf node represents a classification or decision. For example, the question in the first node requires a “Yes” or “No” answer, there will be one leaf node for a “Yes” response, and another node for “No”. Figure 2.1 shows an example of a concept of DTs method for making a car purchase decision.

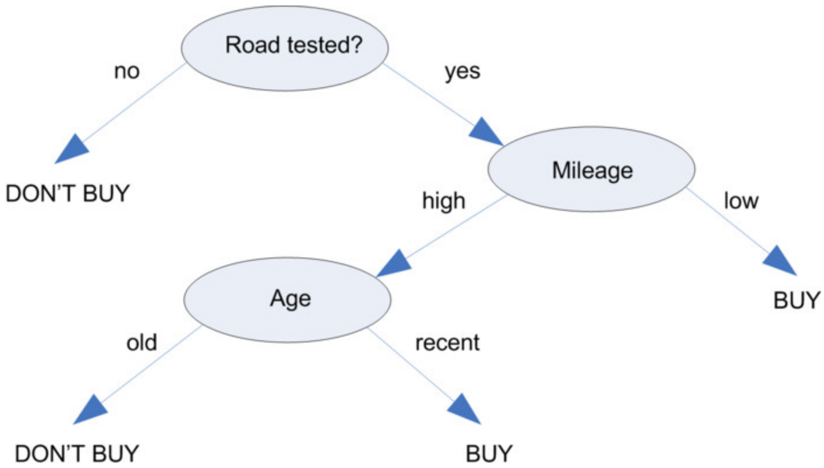


Figure 2.1 A decision tree for making the car purchase. (Gorbachev, Gorbacheva, and Koynov, 2016)

2.4.1.1 Recursive Partitioning (RP)

Recursive Partitioning is a method that predicts the value of a response variable by forming subgroups of a sample data within which the response is relatively homogeneous based on the values of a set of predictor variables (Landau and Barthel, 2010). This algorithm is part of the DTs, which is an important type of algorithm for predictive modeling of ML. It has the same structure as DTs. The RP method was developed since the 1980s. Commonly, this algorithm can be used for establishing classification systems based on multiple covariates or for developing prediction algorithms for a target variable or regression models in the form of a tree structure (Song and Lu, 2015). There are advantages and disadvantages of the RP. This algorithm can easily generate more intuitive models and does not require users to perform calculations, allows for the prioritization of different classifications to create more sensitive or specific decision rules. This algorithm does not work well for situations with continuous variables and may overfit data. (Cook and Goldman, 1984).

2.5 Training, Cross-Validation and Test Datasets

In machine learning, the education and construction of algorithms can learn and make predictions about data. The algorithm can work using data-driven of predictions or decisions, through creating the mathematical modeling from the input of data. Figure 2.2 shows how to divide the dataset into training, cross-validation and testing.

Training Dataset is a set of sample data used for learning, that is to fit the parameters of a classifier. Most methods that search through training data for empirical relationships tend to overfit data, meaning they can identify apparent relationships in training data that are non-conventional.

Cross-validation Dataset (CV) is a set of sample data, that is used to customize the parameters. The datasets can be divided split into training data and validation data. These duplicate partitions can be done in several ways, such as dividing split into two equal datasets and used as training/validation, and validation/training, or select a random subset as the validation dataset repeatedly. In the validation performance of modeling, sometimes additional test data that extended from cross-validation were used (Reunanen, 2003).

Test Dataset is a set of independent data from the training dataset but subsequent probability distributions are the same as training datasets. If that model fitting perfectly with a set of training dataset it will also fit the test dataset as well. The better fit to the training dataset, as opposed to the test dataset, generally indicates overfitting. Therefore, the test dataset is a set of the sample used to estimation the performance of a full-specific classifier (Brownlee, 2017).

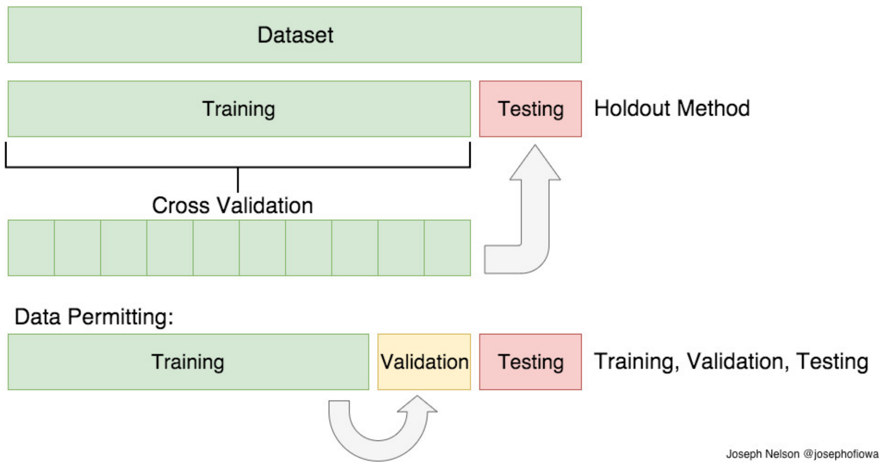


Figure 2.2 Training, cross-validation and testing split. (Bronshtein, 2017)

2.6 Accuracy Assessment

Accuracy Assessment is a method of determining the interpretation or classification of data that has already been analyzed. Accuracy assessment will make the data obtained from the analysis quality and reliability. For accuracy assessment for research related to data from remote sensing, caused by the comparison between data analyzed from satellite imagery and actual data obtained from ground surveys. For example, high-resolution satellite imagery and spatial data. For this method, we chose to use the “Confusion Matrix” (also known as “Error Matrix”) for evaluating the accuracy of the models to find the best algorithm (Story and Congalton, 1986).

2.6.1 Confusion Matrix

Confusion Matrix is a summary table of prediction results on a classification problem (GeeksforGeeks, 2018). The number of correct and incorrect predictions are summarized with count values and broken down by each class or comparison between program results and actual results by human hands (Egghe, 2008; Zhang, *et al*., 2015; Tharwat, 2018) The confusion matrix is the most common way of expressing the accuracy assessment of classification data from remote sensing images, such as type of LULC, type of plants, plant's age, etc. (Comber, *et al*., 2012), as shown in Table 2.6.

Table 2.6 Confusion Matrix Table (Tharwat, 2018)

Confusion Matrix		**Actual Class**		
		0	**1**	**Sum**
Predicted Class	0	TP	FP	TP+FP
	1	FN	TN	FN+TN
	Sum	TP+FN	FP+TN	Accuracy

Definition of the Terms:

True Positives (TP) is where positive observation and predicted to be positive.

True Negative (TN) is where negative observation and predicted to be negative.

False Positives (FP) is where negative observation but predicted to be positive.

False Negatives (FN) is where positive observation but predicted to be negative.

2.6.1.1 Precision

Precision is a measure of the accuracy of the data by considering the class separately. The ratio of correctly predicted true positive (TP) observations to a total predicted positive observation. The high precision is related to a low false positive (FP) rate (Tharwat, 2018). The precision equation is as follows;

$$Precision = \frac{TP}{(TP + FP)}$$

2.6.1.2 Recall

A recall is a measure of the accuracy of a model by considering the class separately. The ratio of correctly predicted true positive (TP) observations to a total observation in the actual class. The high recall is related to a low false negative (FN) rate (Tharwat, 2018). The recall equation is as follows;

$$Recall = \frac{TP}{(TP + FN)}$$

2.6.1.3 Overall Accuracy

Overall Accuracy is a measure of the accuracy of a model by considering all classes. It is the most intuitive performance measure and it is simply a ratio of correctly predicted observation to the total observations (GeeksforGeeks, 2018). The overall accuracy equation is as follows;

$$Overall\ Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN) * 100}$$

2.6.2 F_1-score

F_1-score is a measure of the accuracy of a model by considering both the measured value of Precision and Recall. It is the weighted average of the precision and recall, where an F_1-score reaches is the best value at 1 (perfect precision and recall) and the worst value at 0 (GeeksforGeeks, 2018). The equation is as follows;

$$F_1 - score = \frac{2 * (Recall * Precision)}{(Recall + Precision)}$$

2.7 Related Research

The related research review of mapping the para rubber plantation age using satellite imagery from remote sensing and various approaches to help in this study, can be summarized is as follows;

A study of Zhe Li and Jefferson M. Fox, (2012) about the mapping of para rubber plantations in the mainland of Southeast Asia using MODIS NDVI time series and statistical data. The mainland of Southeast Asia in the study area included China, Thailand, Laos, Vietnam, Cambodia, and Myanmar. This study aims to diagnose the potential of using the Mahalanobis typicality method to manage with mixed pixels of NDVI grids and to explore the potential for MODIS NDVI time series combining with sub-national statistical data on LULC of para rubber to mapping the distribution of para rubber plantations in the mainland of Southeast Asia. The data used were MODIS NDVI time series product (MOD13Q1) from the TERRA satellite between March 2009 and May 2010. The method used Mahalanobis typicality to identify the planting grids in para rubber plantations had the highest probability of occurring and sub-national statistical data about para rubber plantations to quantify the number pixels of para rubber plantations map in each administrative unit. They used Relative Operating Characteristic (ROC) and error matrix analysis respectively to evaluate the potential of Mahalanobis typicality method for validating the accuracy of classification. The high ROC values over 0.8 with the Mahalanobis typicality method showed it was successful in identifying both mature and young of para rubber tree stand ages. The method helped reduce the commission errors for both two types of para rubber plantations to 1.9% and 2.8%, respectively (It is corresponding with user's accuracy to 98.1% and 97.2%, respectively). The results of this study indicated that the integrations used in the Mahalanobis typicality method with MODIS NDVI time series and sub-national statistical data in Southeast Asia can help to successfully overcome the earlier overestimation of plantation survey problems.

A study of Weili Kou, *et al*., (2015) about the mapping of deciduous para rubber plantation ages using PALSAR and Landsat imagery in Xishuangbanna Dai Autonomous Prefecture, Southeast of Yunnan Province, China. This study aimed to present the simple methods for mapping the age of para rubber plantations using an integration of PALSAR 50-meter mosaic imagery and Landsat TM/ETM+ multi-temporal imagery. The mosaic of PALSAR L-band 50-meter images was used for mapping the forest and non-forest areas including both two areas are natural forests and para rubber plantations for PALSAR-based pixels of forested areas. They analyzed the Landsat TM and ETM+ imagery from 2000 to 2009. Firstly, the phenological signatures of deciduous para rubber plantations (seasonal) and natural forest areas by analysis of surface reflectance, NDVI, EVI, and LSWI for generated mapping of para rubber plantations in the year 2009. Then, they analyzed the phenological signatures of para rubber plantations with different of para rubber stand ages and generated mapping in the year 2009 (≤5, 6 – 10, and >10 years) base on Landsat multi-temporal imagery. Finally, they generated the map clearly showing that para rubber plantations expanded into the mountains in the study area over the year. The results of this study demonstrated the potential of integrating microwave using PALSAR imagery and optical remote sensing to describe the characteristics of para rubber plantations and their expansion in that time.

A study of Hayder Dibs, *et al*., (2017) about the mapping of para rubber plantations using hierarchical classification approach for per-pixel and object-oriented classifiers with SPOT-5 imagery in Hulu Selangor, the state of Selangor, Malaysia. This study aimed was to propose a hierarchical classification approach to obtain the accurate mapping of para rubber plantation ages distribution using SPOT-5 satellite imagery for performance evaluation of pixel-based and object-oriented classifiers for classification of para rubber plantations. The general of land use land cover was classified into eight classes using Mahalanobis distance (MD), k-nearest neighbor (k-NN), and Support Vector Machine (SVM) classifiers. After that, they generated the best mapping from the k-NN for classification used to select only pixels that were in the class of para rubber plantations from SPOT-5 imagery. The extracted pixels of imagery served as input the hierarchy of classification were divided into four classifiers: MD, k-NN, SVM and DT classifiers, which were to generated the mapping of para rubber plantations into three intra-class are mature, middle and young of para rubber plantation ages. The results of this study provided produced overall accuracies of 97.48%, 96.90%, 96.25% and 80.80% of k-NN, SVM, MD and DTs classifiers respectively. This indicates that object-oriented classifiers are better than the pixel-based methods for the mapping of para rubber plantations.

A study of Philip Beckschäfer, (2017) about the obtaining of para rubber plantation age information from a dataset of Landsat TM and ETM+ time series and pixel-based image compositing of para rubber plantations in Xishuangbanna, China. He presented a method that facilitates the rapid assessment of para rubber plantations at the regional-scales by using very dense LTS satellite imagery in the second largest para rubber plantations of China. He collected 270 Landsat TM and ETM+ satellite images into annual best-available-pixel composites using the minimum of Normalized Difference Moisture Index (NDMI). The annual composite is classified as pixels of vegetated and non-vegetated applying a global NDMI threshold of 0, which is common in that area to vegetation clearance on the land in last year before the establishment of new-planting in next year. The pixels are located in plantations that classified as non-vegetated, being recorded as the year of re-planting. The result of that study is para rubber plantation age map that has been validated using a stratified random sample of 184 sample points, collected by the visual interpretation of historical Landsat imagery. The Root Mean Square Error (RMSE) of 2.5 years was achieved. He estimated that in the year 2015, the proportion of those plantations in Xishuangbanna was 50% of area that is suitable for natural rubber latexes tapping (8 – 25 years), 27% that is the too young trees to tapping (<8 years) and 24% is the rubber age has reduced latex productivity (>25 years) and will be cutting timber or clear the land in the near future. This proposed technique helps to obtain accurate of para rubber age from NDMI dataset of LST satellite imagery over large areas with the potential to provide in-depth data about spatial dynamics.

A study of Gang Chen, *et al*., (2018) about the estimation of the stand age of para rubber plantations using the integrated pixel and object-based tree growth and annual Landsat time series (LTS) in tri-border region along the junction of China, Myanmar, and Lao. They generated a map of para rubber stand age at a 30-meter spatial resolution to identify para rubber plantations with an accuracy of 87% and an average error of 1.53 years in age estimation. The integration of pixel and object-based image analysis used shows superior performance in generating the annual NDVI time series that reduced spectral noises from ground and vegetation in the canopy of young para rubber trees. The parameters of data prediction models remained relatively stable during sensitivity analysis of the model, resulting in accurate age estimation robust to outliers. They compared to the generally weak statistical relationship between single-date spectral signatures and para rubber tree age. The LTS analysis coupled with tree growth modeling presents a possible alternative for fine-scale age estimation of para rubber plantations.

From the collection or review of relevant research above. This made us decide to use NDVI time series from Landsat satellite imagery because this data set is accessible to the general public and free to download, with a spatial resolution of 30-meter and recording data at the same area repeatedly every 16-day, which this is the advantage of Landsat time series satellites.

CHAPTER 3

RESEARCH METHODOLOGY

The objective of this study is to investigate the potential of using 16-day composite Landsat time series at the 30-meter spatial resolution to identify the abrupt change of vegetation greenness in para rubber plantations. The inter-annual vegetation profile of para rubber planting event that has been extracted and generates predictor variables for supervised classification in Machine Learning. The tree-based recursive partitioning was used to train a model from the training dataset and predict the event of para rubber re-planting and plantation lifespan in Thalang District, Phuket Province of Thailand.

3.1 Study Area

The study area is Thalang District (8° 03' 49.40" N, 98° 21' 03.10" E), the Northern part of Phuket island which is the biggest island located in the South of Thailand (also known as Andaman Sea region). The small-scale study area covers 252 square kilometers in total area, as shown in Figure 3.1. Thalang District has a border with neighboring districts as follow (Phuket Provincial Governor's Office, 2018);

- North is connected to Takua Thung District. (Phang Nga Province).
- East is connected with the Phang Nga Bay, a part of Andaman Sea.
- South is connected to Mueang Phuket District and Kathu District.
- West is connected with the Andaman Sea.

According to the Phuket Provincial Development Plan (2018 – 2021) which based on LULC data in the year 2015, Phuket has a total area of agriculture approximately 169.37 square kilometers. Most of the cultivation area is found in Thalang District where around 87.32 square kilometers is occupied by the para rubber plantations (October 28, 2016) (Phuket Provincial Governor's Office, 2018). Most of the local population is engaged in agricultural occupations and the economic crops are para rubber, palm oil, and pineapple. Rubber farmers often grow pineapple mixture in rubber plantations during the early plantation phase because para rubber trees are young and non-productive (Phuket Development Strategy, 2016).

The Köppen climate classification subtype has a specific code for the climate in Phuket Province is a tropical monsoon climate (Am = Equatorial Monsoonal), under the influence of two seasonal monsoons are southwest monsoon (May – September) and northeast monsoon (October – April) (Köppen, 1918). This local average temperature is about 27.8°C, by the warmest month in April, with an average temperature of about 28.9°C and the coolest month in January, with an average temperature of about 27.2°C. The average amount of precipitation is about 2,230.1 millimeters, by the maximum amount of those in September with about 419.1 millimeters of precipitation and the minimum amount of those in February with an average of about 30.5 millimeters. The average of rainy days during the 1-year period in Phuket Province is about 173 days. The month with the rainiest days are about 23 days in October and the month with the least rainy days are about 4 days in February (Meteorological Development Office, 2017). Therefore, Thalang District is located in weather condition suitable for the para rubber growth (Maggiotto, *et al*., 2014).

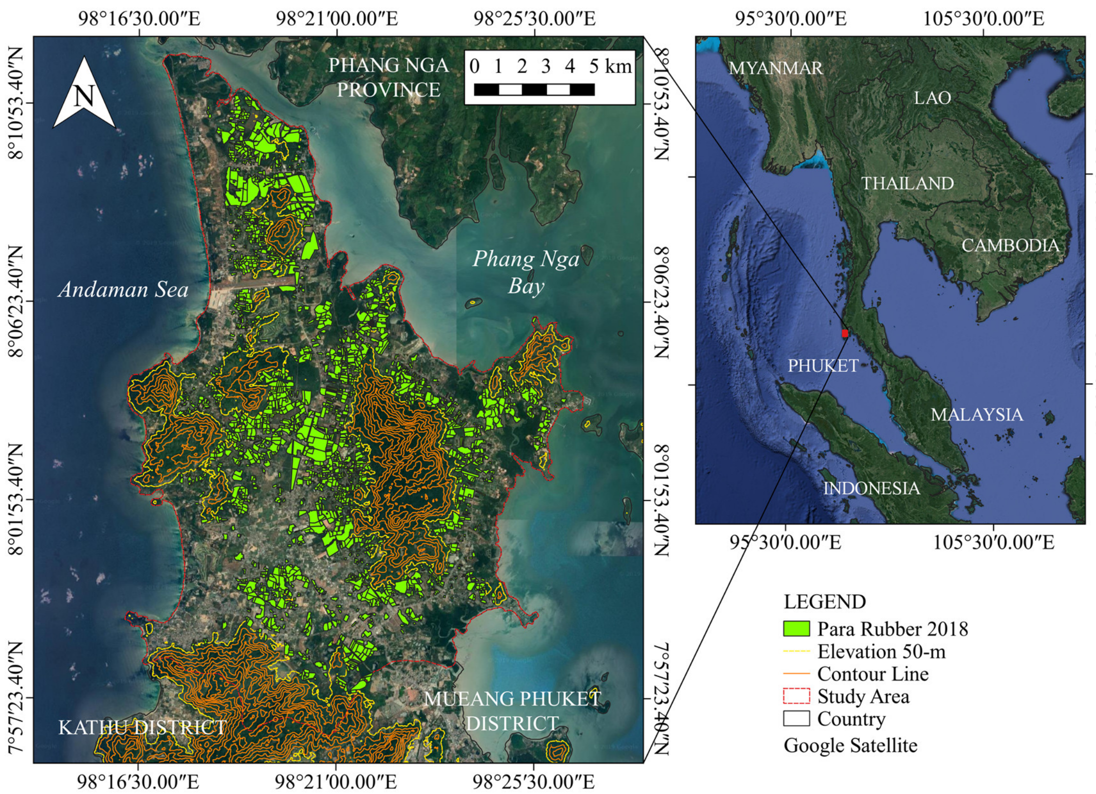


Figure 3.1 Map of para rubber plantations in Thalang District, Phuket Province, Southern Thailand.

3.2 Data and Materials

3.2.1 Landsat Surface Reflectance Data Collections (1991 – 2018)

Landsat imagery was available for download from the Earth Explorer website (http://earthexplorer.usgs.gov/). We downloaded the multispectral bands contain a 30-meter spatial resolution Landsat Surface Reflectance product, Collection 1 Level-2 (on-demand). The obtained satellite images at path 130 and row 54 were covered over the study area. However, we intend to select images available during the dry season (October – April) and contain cloud cover less than 15% over the study area. There were 129 Landsat images from three Landsat platforms, detailed in Table 3.1 and Figure 3.2, were used in this study. The technical detail of Landsat satellite images is also providing in Appendix A.

Table 3.1 Landsat Data Collection

Satellite Imagery	Period of acquisition	Number of Images
Landsat-5 TM	March 16, 1992 to October 15, 2011	55
Landsat-7 ETM+	December 27, 2000 to February 28, 2018	56
Landsat-8 OLI	December 23, 2013 to February 4, 2018	18

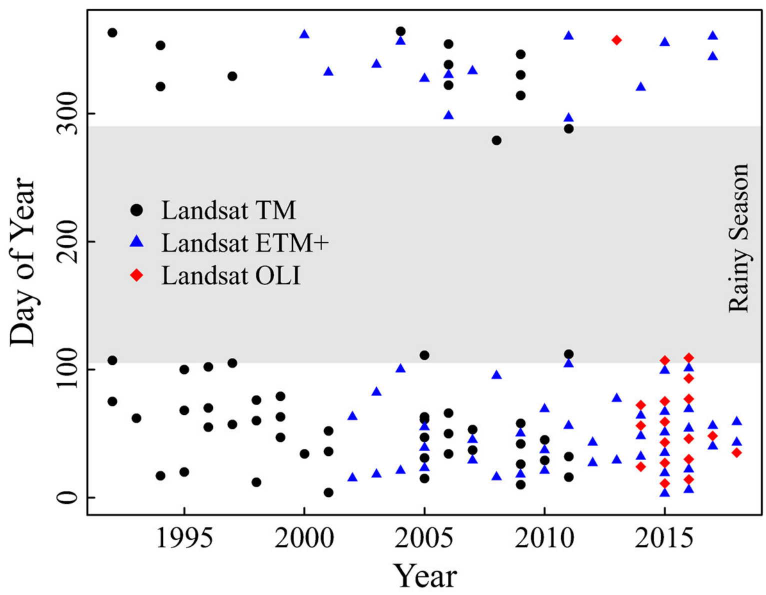


Figure 3.2 The collection of Landsat time series.

3.2.2 Land Use Land Cover Phuket (2016)

LULC data of the year 2016 for Phuket were obtained from the Land Development Department (LDD), under the management of the Ministry of Agriculture and Cooperative of Thailand. The secondary data were in a shapefile vector (.shp) format. LDD is responsible for land use mapping throughout Thailand and distributes the mapping data, based on the interpretation of aerial photography (orthophotos), remote sensing images, and field survey. A 3-level hierarchic classification system is used, which differentiates at the finest level more than 100-classes. The LULC code for para rubber area is A302 (Food and Agriculture Organization of the United Nations, 2017). In this study, we select only LULC data that classified as A302, including mixed A302 with other agriculture plants.

3.2.3 Registration of Rubber Farmers in Thalang District (2013 – 2018)

Registration of Rubber Farmers (2013 – 2018) of Thalang District, Phuket Province, were collected by the Rubber Authority of Thailand (RAOT), Phuket Province, the state enterprise under the management of the Ministry of Agriculture and Cooperative of Thailand. This ground truth data of the study area received and supported by the director and officers of the Rubber Authority of Thailand, Phuket Province (www.rubber.co.th/phuket/). Figure 3.3 shows the website and the example of obtained data. This data has arisen from the registration of Thai's rubber farmers in accordance with the Rubber Authority of Thailand Act B.E. 2558 (2015), which has established the “Rubber Development Fund”. The objective is to support and promote rubber development. The beneficiaries are Thai rubber farmers and Rubber Farmers Institute, which must be registered the para rubber plantation area with the RAOT (Thailand Government, 2015).

The rubber planting registration data used in this study consisted of GPS coordinates (UTM Zone 47), location and size of plots from a survey by ROAT, the year at register and approval for cultivation from RAOT. The summary of the number of plots obtained from 2013 to 2018, with a total of 131 plots spread across the area of Thalang District, which is shown in Appendix B of the Registration of Rubber Farmers in Thalang District.

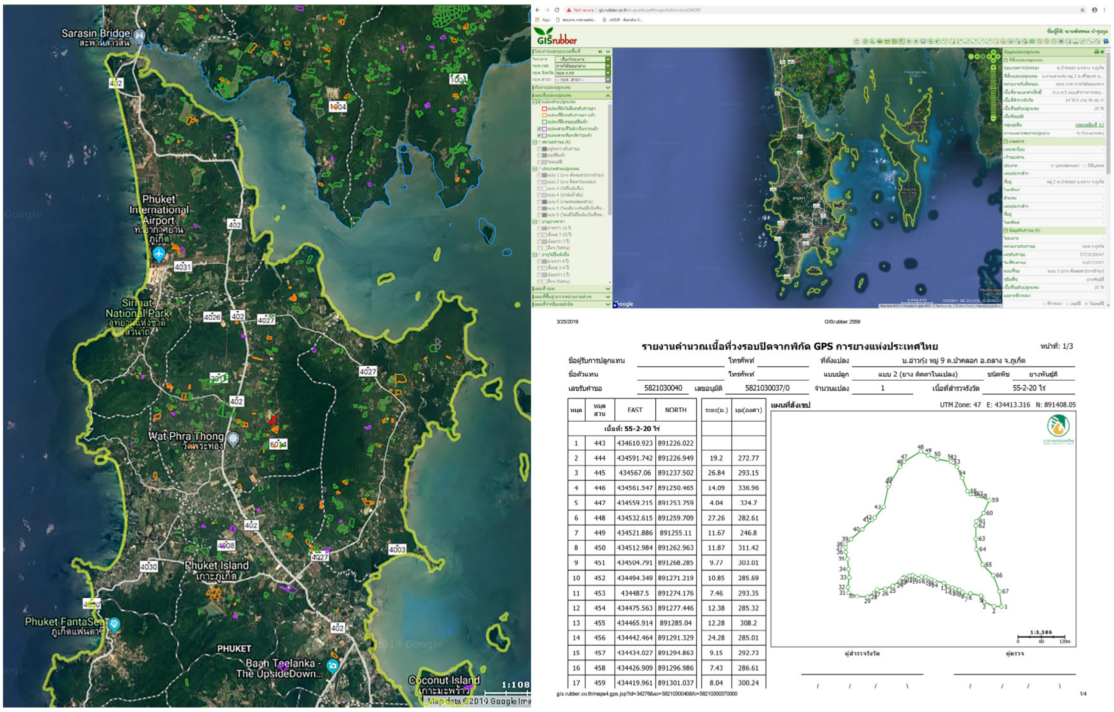


Figure 3.3 The registration of rubber farmers in Thalang District (2013 – 2018).

3.3 Programs and Equipment

1) Geographic Information Processing Software
2) R Project for Statistical Computing
3) RStudio Desktop
4) Google Earth Pro
5) Microsoft Office
6) Computer with Accessories
7) Digital Camera

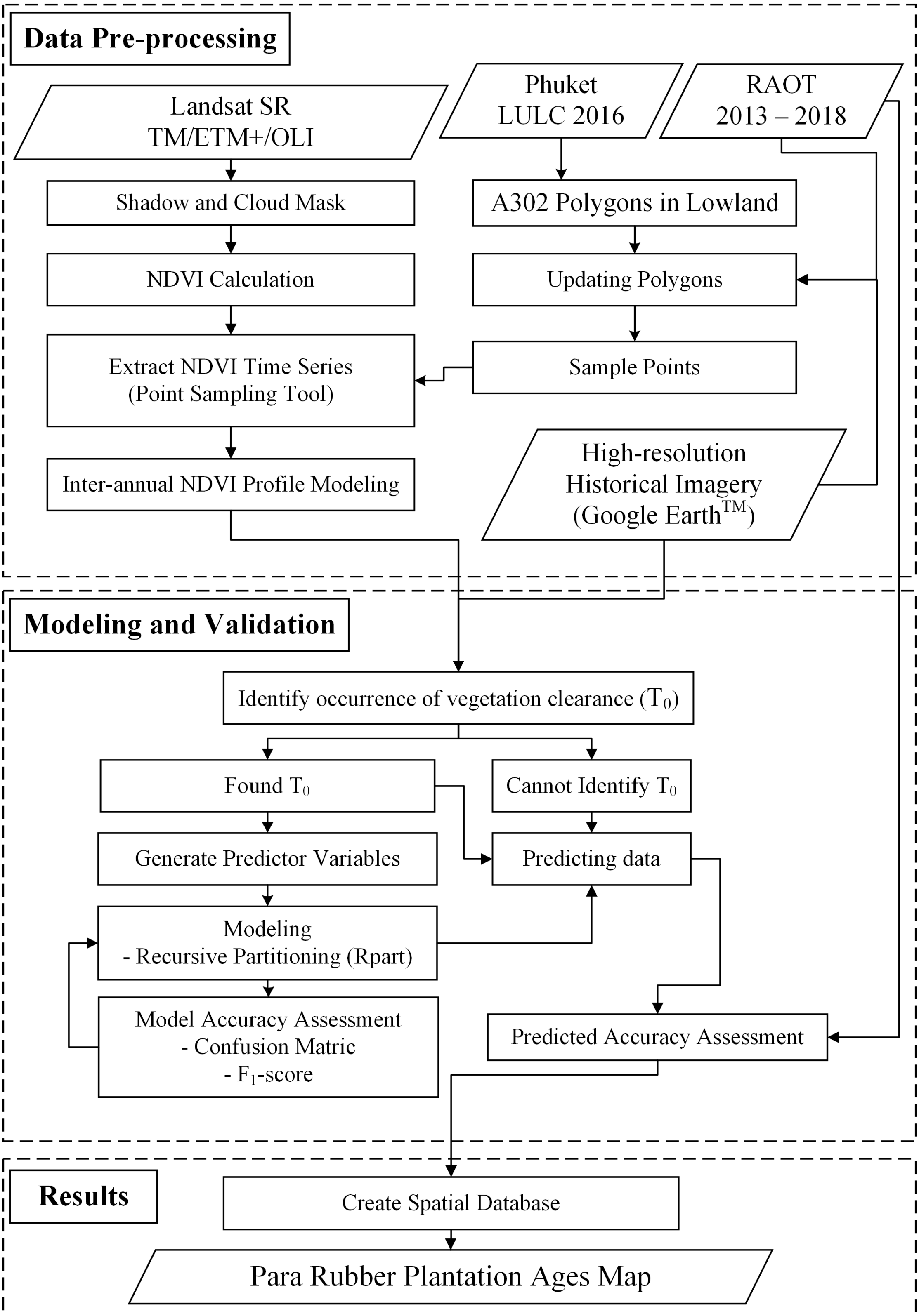


Figure 3.4 Research Framework

3.4 Image Pre-processing

Landsat Surface Reflectance (SR) improves the comparison between multiple images over the same region by accounting for atmospheric effects such as aerosol scattering and thin clouds, which can help in the detection and characterization of Earth Surface change (U.S. Geological Survey, 2019a). Since the obtained SR products were error and radiometrically corrected, we only processed the cloud mask and band combination. Then, the spectral index was calculated before spectral profile extraction and data modeling. The value of DNs to SR to be standardized in the range of 0 – 1 (López-serrano, Corral-rivas, and Díaz-varela, 2016).

3.4.1 Shadow and Cloud Mask

Figure 3.5 shows the step of shadow and cloud mask. We used the information from “Pixel Quality Assessment” or “Pixel QA” for eliminating unwanted pixel data, such as cloud and shadows. Based on the differences in Surface Reflectance processing algorithms (U.S. Geological Survey, 2019a; U.S. Geological Survey, 2019b), the Pixel QA value of the ground surface is 66 and the water surface is 68 for images from Landsat TM and ETM+. But the pixel value of the ground and water surface are 322 and 324, respectively, for Landsat OLI sensor.

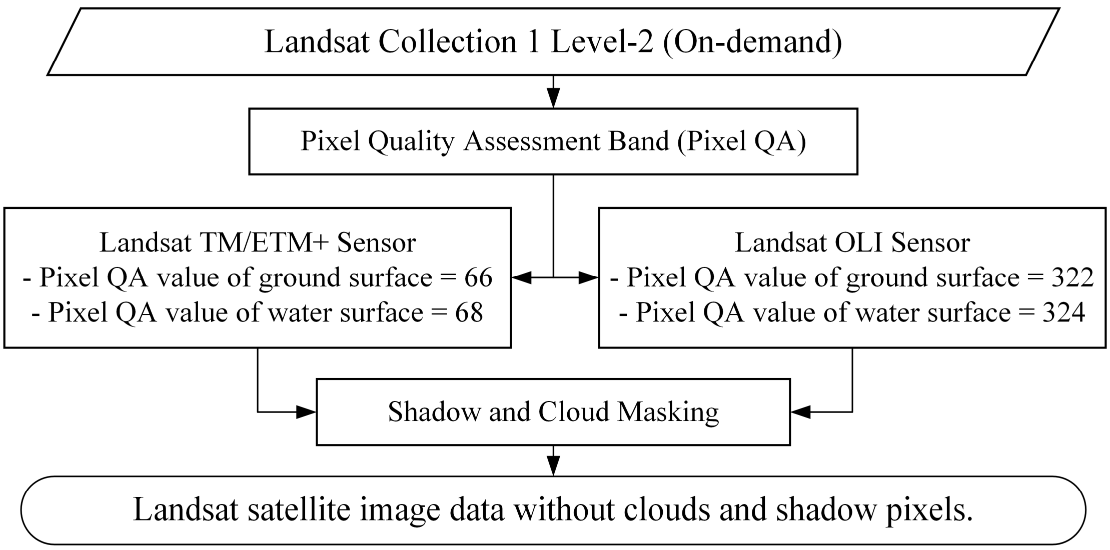


Figure 3.5 The processing of shadow and cloud masking.

Shadow and cloud masking were carried out in QGIS program. We set the pixel value of Pixel QA band to be less than or equal to the ground and water surface in the equation for raster calculating, due to Landsat TM, ETM+ and OLI sensors provide different calculation protocols (Zhu, *et al*., 2018). Therefore, the Pixel QA value of Landsat TM and ETM+ has the value ≤70 and Landsat OLI with the value ≤325 multiplied by the normal value of Pixel QA band, as shown in Table 3.2

Table 3.2 Equation of pixel QA of Landsat TM/ETM+/OLI sensor.

Sensors	Equations
Landsat TM/ETM+ *	(Pixel QA≤70)*Pixel QA
Landsat OLI	(Pixel QA≤325)*Pixel QA

* Landsat TM and ETM+ sensors have the same wavelength.

Later, we brought the Pixel QA bands from the above process to eliminate the pixels of cloud and shadow in other bands of Landsat collection 1 level-SR in the same set by doing one by one. The results are that the pixels of the cloud and shadows are made to zero value. In this calculation, using the raster calculator tool in the geographic information processing software by following this equation:

$$(Pixel\ QA\ \ != 0) * Landsat\ SR$$

Definition of the Terms:

Pixel QA is Pixel Quality Assessment Band

SR are Landsat Surface Reflectance

3.4.2 Band Combination

In this process, we used the band combination method by sort the all bands from the first band to the last band of each Landsat TM, ETM+ and OLI satellite images with a true color combination (including red, green, and blue band) using the remote sensing processing software (Pahlevan, *et al*., 2017), as shown in Table 3.3. The results are the Landsat satellite images that have been processed with shadow-cloud masking, and band combination successfully modified.

Table 3.3 Band combination of Landsat TM/ETM+/OLI imagery.

Landsat Bands	Sensors	
	TM/ETM+ *	OLI
Coastal Aerosol	-	Band 1
Blue	Band 1	Band 2
Green	Band 2	Band 3
Red	Band 3	Band 4
Near-infrared	Band 4	Band 5
SWIR-1	Band 5	Band 6
SWIR-2	Band 7	Band 7
Thermal	Band 6	-
Cirrus	-	Band 9

* Landsat TM and ETM+ sensors have the same wavelength.

3.4.3 NDVI Calculation

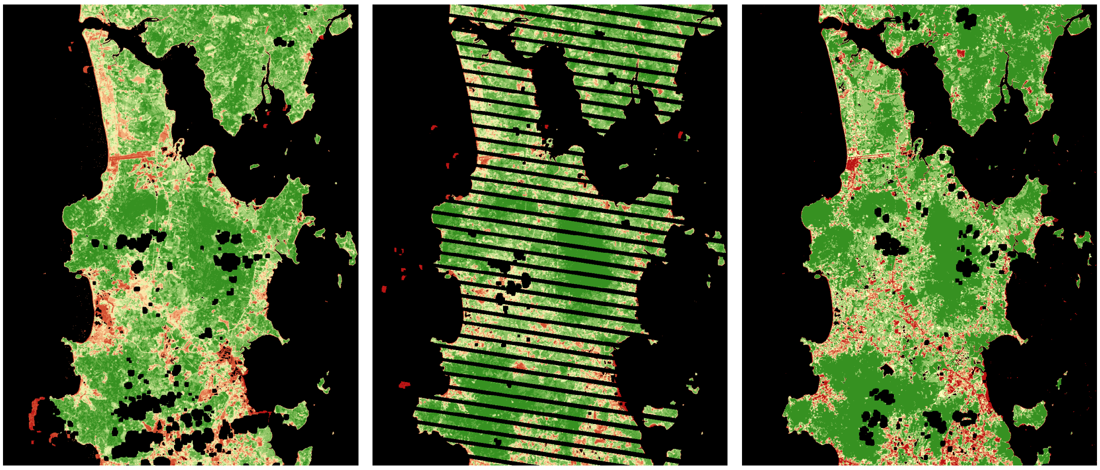
In this study, we used the NDVI equation to calculate the vegetation index. The wavelength of related bands is NIR and red band. The NDVI equation of Landsat TM, ETM+, and OLI sensor to an analysis using the raster calculator tool in the geographic information processing software, as shown in Table 3.4. The NDVI calculation was performed for all 129 cloud-free SR images. The results of this step are the NDVI data in the form of raster data obtained from the Landsat TM, ETM+, and OLI satellite images that have been successfully modified, as shown in Figure 3.6. The NDVI color combination makes the pixel color display in satellite images based on the amount of vegetation found in that area. The NDVI color combination makes the pixel color display in satellite images based on the amount of vegetation found in that area. The forestry areas are dark green pixels. The agricultural areas are light green to yellow colors. The residential space or esplanades are red or orange colors. The black pixels are areas that are water, clouds, and shadows.

Table 3.4 NDVI equation of TM/ETM+/OLI imagery.

Equation Models	Instead Equations	
	Landsat TM/ETM+ *	Landsat OLI
NDVI = (NIR-RED)/(NIR+RED)	(Band4-Band3)/ (Band4+Band3)	(Band5-Band4)/ (Band5+Band4)

* Landsat TM and ETM+ sensors have the same wavelength.

(A) Landsat TM	(B) Landsat ETM+	(C) Landsat OLI



16 March 1992	29 January 2007	4 February 2018

Figure 3.6 Example of Landsat satellite images with NDVI color-level.

3.5 LULC Data Process

3.5.1 Updating of A302 LULC data

Figure 3.7 shows the step of updating LULC 2016 to A302 LULC 2018. After selecting A302 LULC data (polygons) located in lowlands (less than 50 meters above mean sea level) from the original data (LULC 2016), we update and validate the A302 polygons from the historical high-resolution aerial images available in Google Earth Pro™ program. The footprints of rubber plantation and the image interpretation were used to update and reshape the A302 polygon, as well as to re-identify and correct the type of land use. Also, the registration data of rubber farmers in Thalang District from 2013 to 2018 were used to ensure the polygons match with the shape of the approved plantation. There are 131 records of plantation registration which also used in final validation.

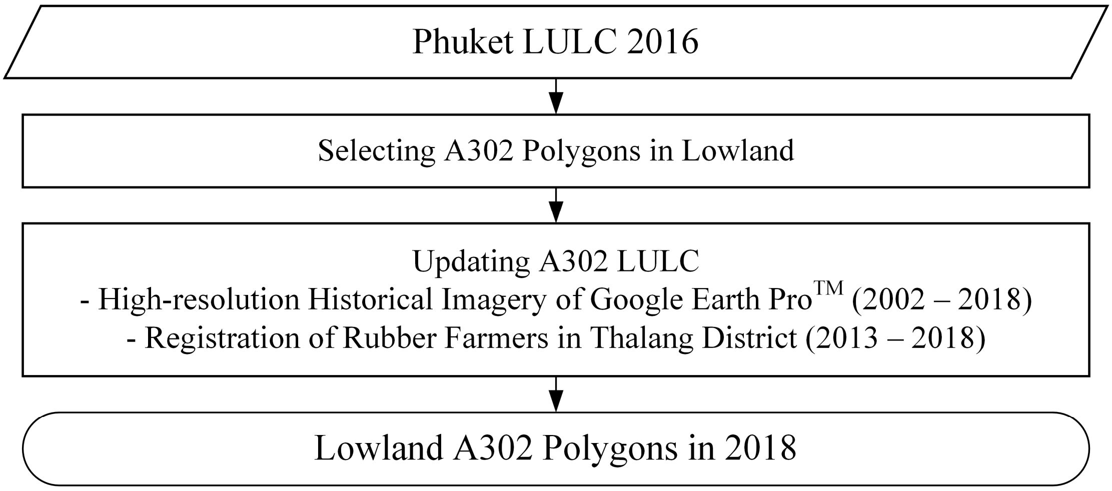


Figure 3.7 Updating para rubber LULC data.

As for the reasons that we are interested in researching the para rubber plantations in the lowlands has details are as follows: Interpretation visually of satellite imagery in order to explore the boundaries of para rubber plots in the lowlands more clearly than high areas due to differences in the terrain and characteristics of cultivation on the steep areas, unlike in the lowlands that look simpler. Para rubber plots in the highlands to explore hard-to-access accuracy. It can be dangerous if we do not specialize in sloping areas. Para rubber cultivation in the lowlands has changed more than because can manage the planting area more easily than in the sloping areas. The expansion of para rubber plots in sloping lands is of more environmental concern (Fox, *et al*., 2014), as shown in Figure 3.8.

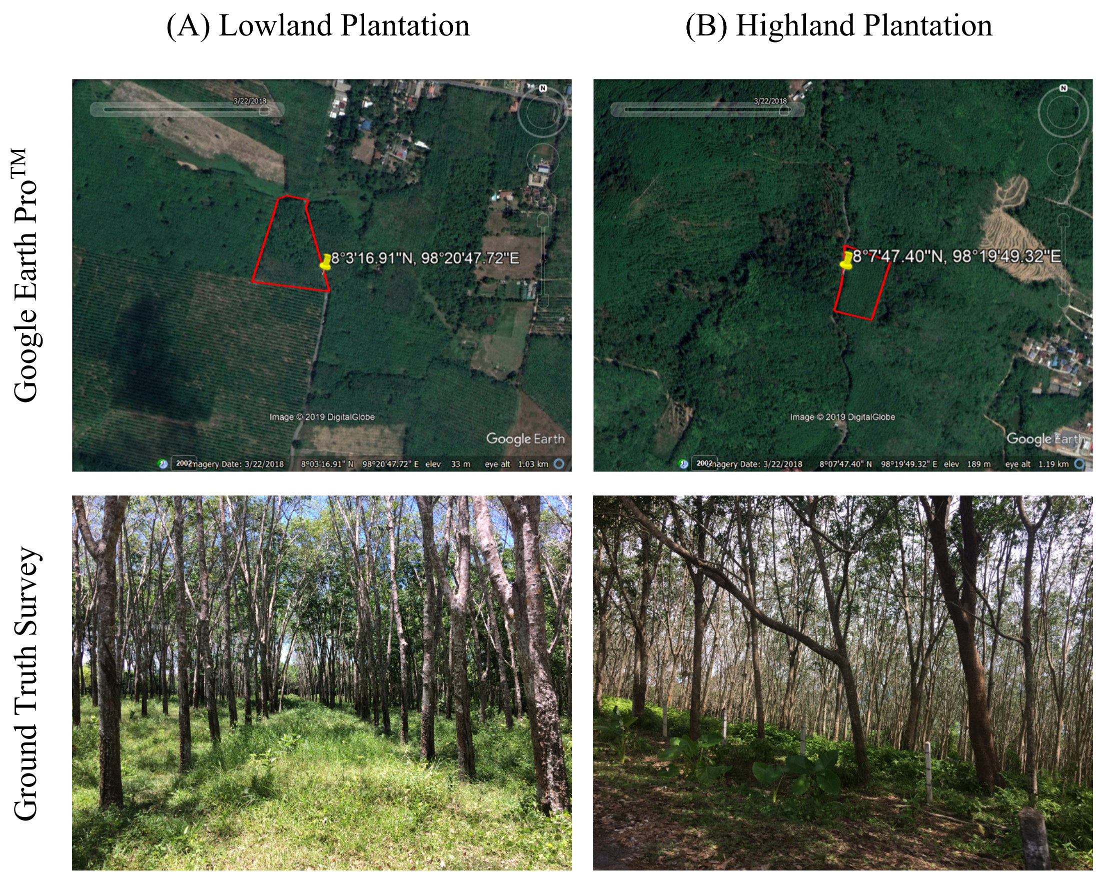


Figure 3.8 The difference between A302 lowland and highland plantations.

3.5.2 Extraction NDVI Time Series

Figure 3.9 displays the process of NDVI time series extraction. The main objective of this process is to extract NDVI values (in pixels) inside each A302 polygon. However, there are many different sizes of a plantation that cover the different number of pixels. For efficiency of data processing, we designed to randomly select a set of pixels for each plantation according to its size. In the first step, the A302 polygons were inward buffered with 15 meters to omit the pixels located at the edge of plantations. This means to avoid the effect of a mixed spectral signal. Then, the extracting points were generated using “Random Point Inside Polygon” command using QGIS software. The maximum number of random points in each polygon was defined by the distribution of the logarithm of the plantation area. The conditions to define the number of random points is detailed in the next section. We also set the offset when randomly generated extracting points to ensure that there are no extracting points reside in the same 30-meter pixel and that the creating points are spread over the polygon area. Once the sample points were generated for each polygon, the IDs of A302 polygons were added to the points using “Add Polygon Attributes to Points” command. The 129 NDVI images were imported into QGIS software as the layers, as well as stacked in order of image acquisition time. Finally, the generated points were used to extracted NDVI values from stacked images using “Point Sampling Tool” plugin. Then, the point layer was saved to the CSV file format for further data modeling.

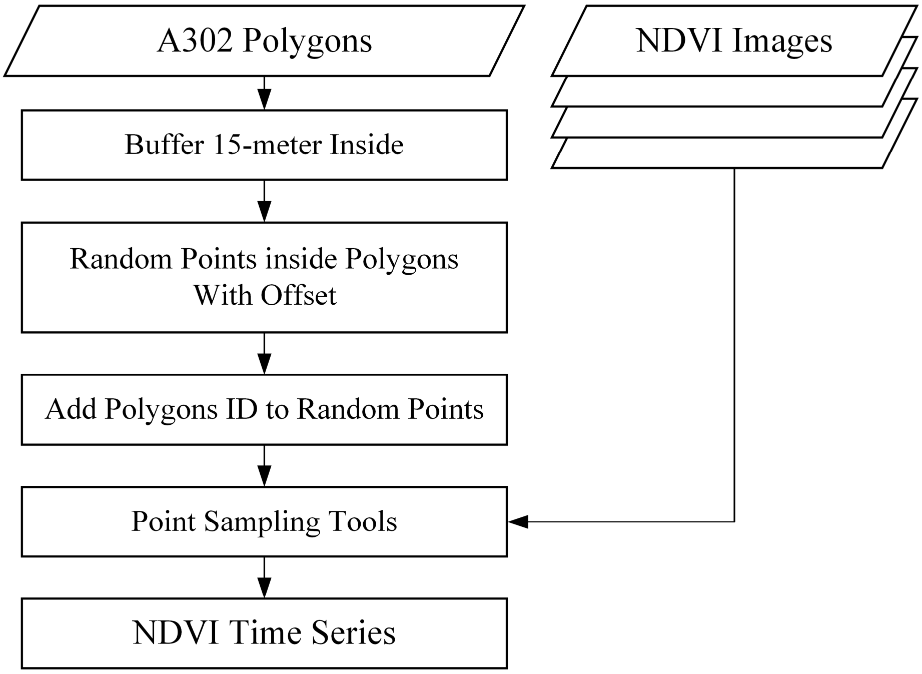


Figure 3.9 Extraction NDVI time series.

3.5.3 Number of Random Points

The number of sample points in each rubber plot depends on the size of the original plot (A302 polygons). We generating a boxplot of the logarithm of the plot’s area to study the proportions of para rubber’s plot size and to determine the optimal random points, as shown in Figure 3.10.

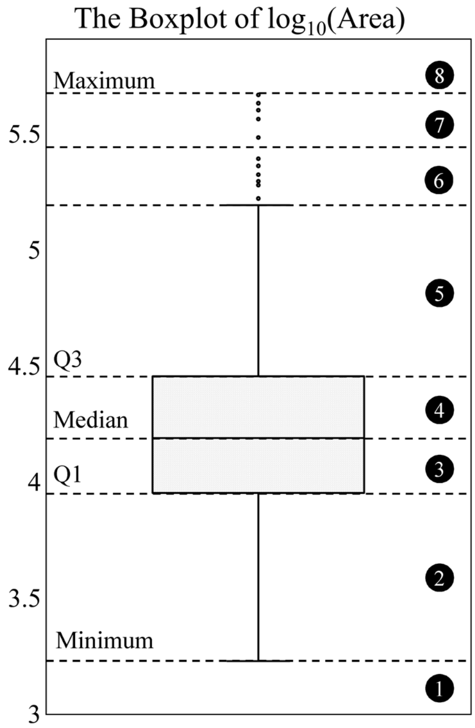


Figure 3.10 The boxplot of $\log_{10}$(area).

From the above figure, we defined the number of random points according to the range of the distribution of $\log_{10}$(area), including the value of the highest area (maximum value), the lowest area (minimum value), median value, first quartile (Q1), third quartile (Q3), and the interquartile range (IRQ) between Q1 and Q3. We used the value of each range from boxplot to calculate the number of random points, as shown in Table 3.5.

Table 3.5 Equation for the number of random points.

The Size of $\log_{10}$(Area)	Number of Random Points
$\log_{10}$(Area)<Minimum	rounddown(Area/($900x2^0$))
$\log_{10}$(Area)≥Minimum&<Q1	rounddown(Area/($900x2^1$))
$\log_{10}$(Area)≥Q1&<Median	rounddown(Area/($900x2^2$))
$\log_{10}$(Area)≥Median&<Q3	rounddown(Area/($900x2^3$))
$\log_{10}$(Area)≥Q3&<Maximum	rounddown(Area/($900x2^4$))
$\log_{10}$(Area)≥Maximum&<Maximum+(IRQ*1)	rounddown(Area/($900x2^5$))
$\log_{10}$(Area)≥Maximum+(IRQ*1)&<Maximum+(IRQ*2)	rounddown(Area/($900x2^6$))
$\log_{10}$(Area)≥Maximum+(IRQ*2)	rounddown(Area/($900x2^7$))

Later, we define the offset of the random sample point with the distance between each point to be at least 30-meters apart, which according to the resolution of NDVI pixels. The equation for calculating offset denoted as the following:

$$Offset = (Number\ of\ Random\ Points - 1) * 30m$$

3.6 Identify the Occurrence of Vegetation Clearance (T_0)

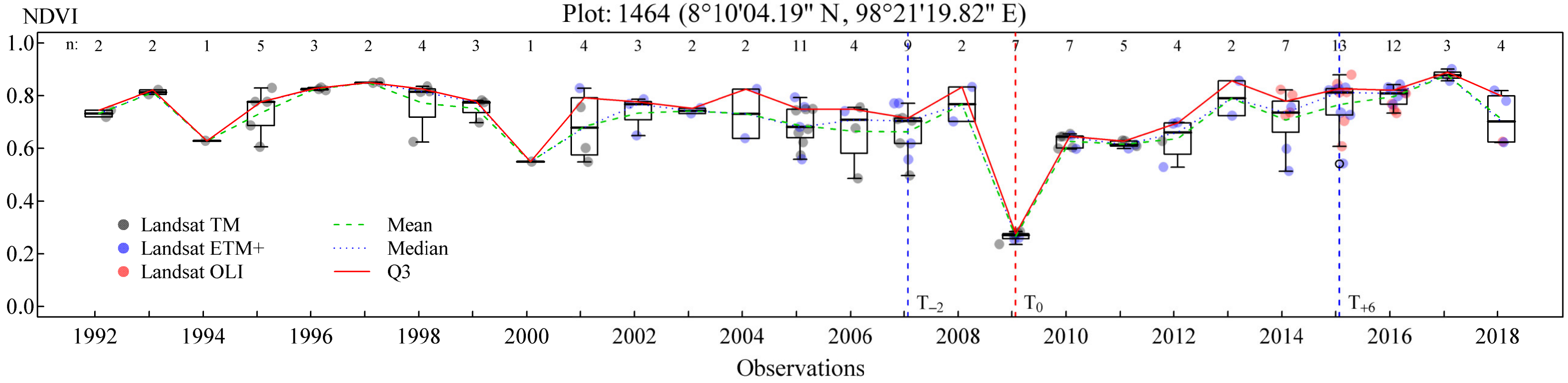
3.6.1 Inter-annual NDVI Profile Modeling

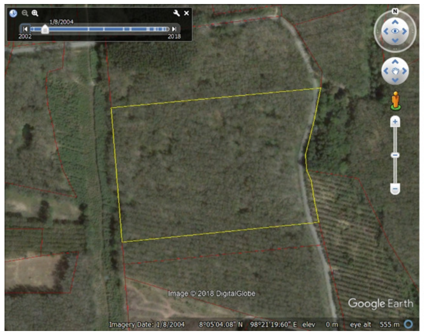
We used the value of NDVI time series obtained from a process of extraction NDVI time series to generating the graph of inter-annual NDVI profile modeling for display the characteristics of the indicative vegetation greenness of a para rubber plantation from October 1991 to April 2018 (27 seasons). We must be done this process to every plot for identifying an occurrence of the vegetation clearance event in the next step (Verbesselt, *et al*., 2010).

In this step, we write command for making the inter-annual NDVI profile graph using the statistical data analysis program. The results of this process are NDVI profile modeling and related the truth event with high-resolution historical satellite imagery from Google Earth Pro™ program in four time periods, including (1) before vegetation clearance event, (2) during the event of re-planted as para rubber trees, (3) the period of plant growth, and (4) the condition of para rubber plantations at the present period. Figure 3.11 displays an example of inter-annual NDVI profile modeling of para rubber plot ID 1464 was the plot that found the NDVI value to be the lowest value at T_0 and has historical satellite imagery at T_0 to confirm that is the true event. We can explain define various symbols in the graph of inter-annual NDVI profile modeling, as follows;

- The y-axis is the level of NDVI values from 0 to 1.
- The x-axis is a duration from 1991 to 2018 (27 years).
- The color points are the type of satellite sensors.
- The graph lines are mean, median and third quartile (Q3) of NDVI values.
- The boxplots are the NDVI value range in each year.
- The red dash-line is a year of vegetation clearance event at time T_0.
- The 1st blue dash-line is a previous event of vegetation clearance 2 years (T_{-2}).
- The 2nd blue dash-line is after the event of para rubber cultivation 6 years (T_{+6}).

Inter-annual NDVI profile modeling at plot ID 1464

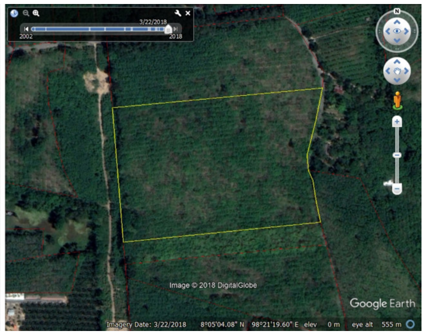




(1) 8 January 2004

(2) 5 February 2009

(3) 22 November 2010



(4) 22 March 2018 (present)

Figure 3.11 Inter-annual NDVI profile modeling and high-resolution historical imagery from Google Earth Pro^{TM} at plot ID 1464.

3.6.2 Identify Event of Re-planting at T_0

From the example of inter-annual NDVI profile above. We examined the events of the vegetation clearance for para rubber re-planting at T_0 using high-resolution historical satellite imagery from Google Earth Pro™ program. It can provide historical satellite data in the study area from 2002 to 2018 (present). However, we must use the registration of rubber farmers in the year 2013 to 2018 for assembling together with this data. This verification process allows us to identify the year for the para rubber re-planting. After that, we organized the para rubber plots divided into two groups, namely the group that can identify the year of para rubber re-planting and group cannot be identified event is clear, this process shown in Figure 3.12.

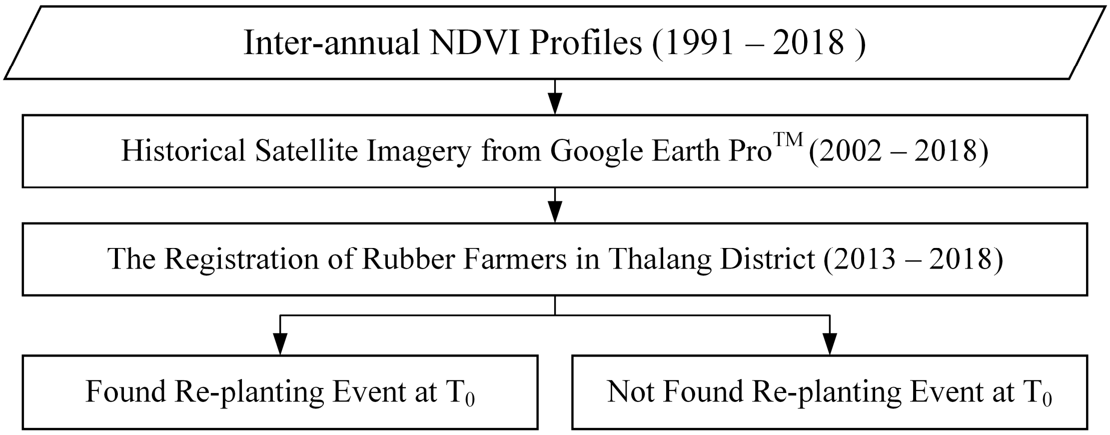


Figure 3.12 Identify event of re-planting at T_0.

3.6.3 Phenological Profile

In this step, we used the NDVI profiles (30 samples), which found para rubber re-planting events at T_0, for creating a phenological profile (T_{-2} – T_{+6}). Then, we generated predictor variables from different of the value at T_0 to other time for machine learning algorithms in data prediction modeling. This profile will display a graph of the distribution of NDVI values before cleared area events of old para rubber plantations for 2 years (T_{-2} – T_{-1}), at the event of para rubber re-planting (T_0), and after the planting of new para rubber for 6 years (T_{+1} – T_{+6}). To train the data prediction model, the ML algorithm learns the values of NDVI time series in the various period of planting para rubber event. Figure 3.13 shows the phenological profile of 30 sample lowland plantations. The information about the distribution of NDVI value from the training dataset (found T_0 in inter-annual NDVI profile) will be used in further modeling.

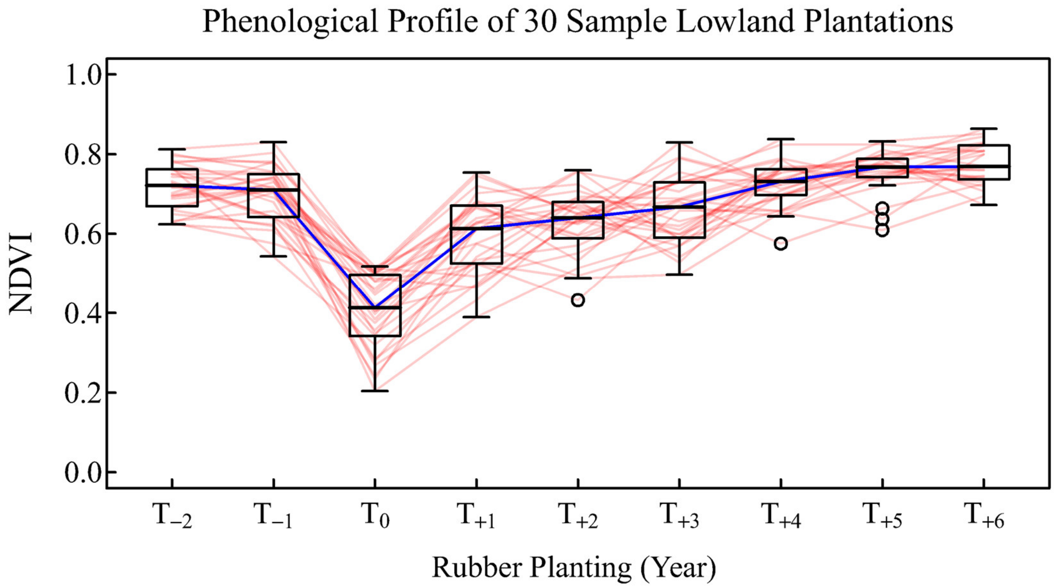


Figure 3.13 Phenological profile of 30 sample lowland para rubber plantations.

3.7 Data Prediction Modeling

3.7.1 Generate Predictor Variables

In this step, we generate predictor variables including the value of the maximum (MAX), minimum (MIN), median (MED), first quartile (Q1) and third quartile (Q3) and the interquartile range (IRQ) from the inter-annual NDVI profiles during the new planting event (T_{-2} – T_{+6}) For the predictor variables from four equations for being used in this study, as shown in Table 3.6. We have set the variables into two groups: variables at T_0 and variables at T_{-2}, T_{-1}, T_{+1}, T_{+2}, T_{+3}, T_{+4}, T_{+5} and T_{+6}.

Table 3.6 Equation for predictor variables.

Time	Predictor Variables					
	MAX	**Q_3**	**MED**	**Q_1**	**MIN**	**IRQ**
T_0	ndvi.MAX.t0	ndvi.Q3.t0	ndvi.MED.t0	ndvi.Q1.t0	ndvi.MIN.t0	ndvi.IRQ.t0

	NDVI.diff.MINt0.MAX	**NDVI.diff.Q1t0.Q3**	**NDVI.ratio.MEDt0.Q3**
T_{-2}	ndvi.diff.MINt0.MAXtn2	ndvi.diff.Q1t0.Q3tn2	ndvi.ratio.MEDt0.Q3tn2
T_{-1}	ndvi.diff.MINt0.MAXtn1	ndvi.diff.Q1t0.Q3tn1	ndvi.ratio.MEDt0.Q3tn1
T_{+1}	ndvi.diff.MINt0.MAXtp1	ndvi.diff.Q1t0.Q3tp1	ndvi.ratio.MEDt0.Q3tp1
T_{+2}	ndvi.diff.MINt0.MAXtp2	ndvi.diff.Q1t0.Q3tp2	ndvi.ratio.MEDt0.Q3tp2
T_{+3}	ndvi.diff.MINt0.MAXtp3	ndvi.diff.Q1t0.Q3tp3	ndvi.ratio.MEDt0.Q3tp3
T_{+4}	ndvi.diff.MINt0.MAXtp4	ndvi.diff.Q1t0.Q3tp4	ndvi.ratio.MEDt0.Q3tp4
T_{+5}	ndvi.diff.MINt0.MAXtp5	ndvi.diff.Q1t0.Q3tp5	ndvi.ratio.MEDt0.Q3tp5
T_{+6}	ndvi.diff.MINt0.MAXtp6	ndvi.diff.Q1t0.Q3tp6	ndvi.ratio.MEDt0.Q3tp6

Definition of the various variables:

- NDVI is the Normalized Difference Vegetation Index.
- MAX is the value of the maximum.
- MIN is the value of the minimum.
- MED is the value of the median.
- Q1 is the value of the first quartile.
- Q3 is the value of the third quartile.
- IRQ is the interquartile range (Q3 – Q1)
- t0 is the NDVI value of para rubber re-planting event.
- tn2 is the NDVI value of before re-planting event 2 years.
- tn1 is the NDVI value of before re-planting event 1 year.
- tp1 is the NDVI value of after re-planting event 1 year.
- tp2 is the NDVI value of after re-planting event 2 years.
- tp3 is the NDVI value of after re-planting event 3 years.
- tp4 is the NDVI value of after re-planting event 4 years.
- tp5 is the NDVI value of after re-planting event 5 years.
- tp6 is the NDVI value of after re-planting event 6 years.

3.7.2 Training and Testing Dataset

For the process of generating the model and accuracy assessment of data prediction models in this study. The methods of data prediction modeling using recursive partitioning (RP). The main reason for adopting the RP algorithm is the ability to handle missing data in the training dataset. Since we generated predictor variables from the different of NDVI value at T_0 and other times (T_{-2}, T_{-1}, T_{+1}, T_{+2}, T_{+3}, T_{+4}, T_{+5} and T_{+6}) there were missing data at the first two and last six variables in some records. To train the model, we divided the NDVI profile data from plots that can identify T_0 with the proportion of 60:40. The 60% of NDVI profile data were used as a training dataset and another 40% of data were used as the testing dataset for assessing the performance of the training model to ensure there are no under-fit or over-fit of the model. The principles of training and testing dataset. We can be explained this process as follows.

- We used the training dataset to learn the appropriate variables and parameters for classification data of modeling.
- We used the testing dataset to determine and customize the parameters need to put in the data prediction modeling such as the complexity parameter in the RP algorithm.

The accuracy assessment index of the training model must be accurate, not very different. If the value is very different, then the model is not good (Reunanen, 2003; Brownlee, 2017).

3.8 Model Accuracy Assessment

In this study, we used the confusion matrix for accurate assessment of the classification age of para rubber plantations in the lowland of Thalang District, Phuket Province, Southern Thailand, which we designed this table for the accuracy assessment of the classification age of para rubber plantations with a training, cross-validation, and testing dataset, the actual class is the truth data from ground survey, and predicted class is the data obtained from the data prediction model, as shown in Table 3.7 to determine the best method for identifying the para rubber plantation ages. In this step, we accuracy assessment using the statistical data analysis program.

Table 3.7 Confusion matrix for accuracy assessment.

Confusion Matrix		Actual Class			Error Matrix
		0	**1**	**Sum**	
Predicted Class	0	TP	FP	TP+FP	OA
	1	FN	TN	FN+TN	F_1-score
	Sum	TP+FN	FP+TN	Accuracy	

The confusion matrix equation is as follows;

$$Precision = \frac{TP}{(TP + FP)}$$

$$Recall = \frac{TP}{(TP + FN)}$$

$$Overall\ Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN) * 100}$$

$$F_1 - score = \frac{2 * (Recall * Precision)}{(Recall + Precision)}$$

CHAPTER 4

RESULTS

Finally, we assessed the accuracy of the predicted model to create a spatial database and mapping the age of para rubber plantations classified at the 1-year level. We found that there is a possibility of using Landsat time series at the 30-meter spatial resolution to identify para rubber plantation ages, especially in this study area, which we can summarize detail of the various results as follows;

4.1 Para Rubber Plots

4.1.1 Size of Plots and Sample Points

After the process of updating LULC data using high-resolution historical images from Google Earth Pro^{TM} program and registration data of rubber farmers (2013 – 2018) of Thalang District, Phuket Province, were collected by the Rubber Authority of Thailand (RAOT). Currently, we found 2,196 plots of lowland para rubber plantation in Thalang District. The total area is 51,490,442.02 square meters. The largest and smallest of plot size were 464,102.28 and 1,037.11 square meters, respectively. The area size, mostly at 14,099.10 square meters. Figure 4.1 shows the lowland rubber plantations for both training and predicting datasets.

Then, we generated a boxplot of the logarithm of the plot's size for determining the number of random points, as shown in Figure 4.2. We found that the maximum value was 5.668, the minimum value was 3.017, the median value was 4.149, the first quartile (Q1) was 3.914, the third quartile (Q3) was 4.410. Thus, the difference between Q1 and Q3 (Interquartile Range) was 0.495. We used the value of each range from boxplot to calculated the number of random points, as shown in Table 4.1.

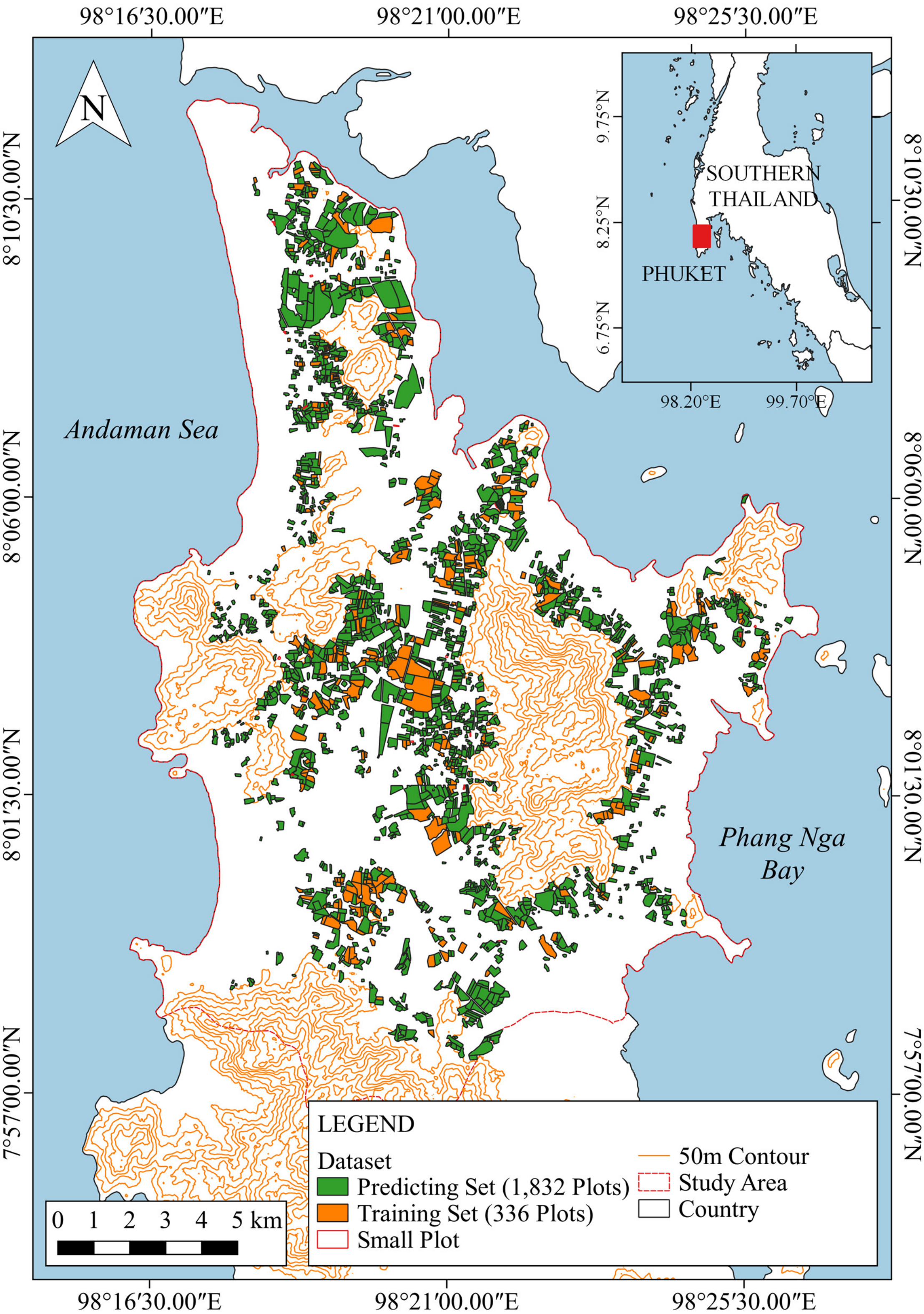


Figure 4.1 Lowland para rubber plantations in Thalang District.

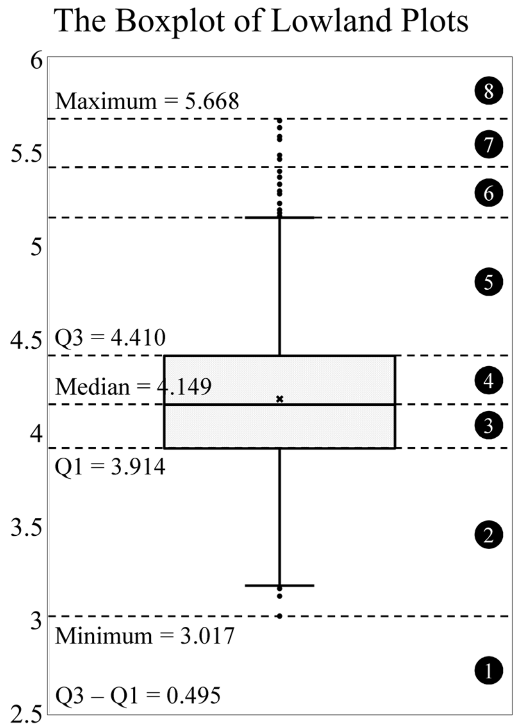


Figure 4.2 The $\log_{10}$(area)'s boxplot of para rubber plantations.

Table 4.1 Equation for the number of random points.

The $\log_{10}$ of Area Size	Number of Random Points
log10(Area)<3.017	rounddown(Area/(900x2^0))
log10(Area)≥3.017&<3.914	rounddown(Area/(900x2^1))
log10(Area)≥3.914&<4.149	rounddown(Area/(900x2^2))
log10(Area)≥4.149&<4.410	rounddown(Area/(900x2^3))
log10(Area)≥4.410&<5.668	rounddown(Area/(900x2^4))
log10(Area)≥5.668&<5.668+(0.495*1)	rounddown(Area/(900x2^5))
log10(Area)≥5.668+(0.495*1)&<5.668+(0.495*2)	rounddown(Area/(900x2^6))
log10(Area)≥5.668+(0.495*2)	rounddown(Area/(900x2^7))

4.2 Data Prediction Model

4.2.1 Identify Event of Re-planting at T_0

There were 2,196 A302 polygons after updating the original LULC data. However, there were 28 small plots which their areas were less than 900 when buffer 15 meters inward. So, the number of random points is zero when we round-down the number. From the interpretation of historical images and the verification of inter-annual NDVI profile, identified the occurrence of vegetation clearance event at T_0 in each plot. We found that para rubber plots can identify para rubber re-planting event was 336 plots, as shown in Table 4.3.

4.2.2 RP Modeling

Recursive partitioning algorithm was used to train the prediction model. The NDVI time series data from 336 plantations were randomly separated into training and testing dataset. We have para rubber 202 plots used for training data, which each plot has 27 values (season), so the number of sample size is 5,454. Figure 4.3 displays the performance of 30 repeated RP algorithm in different complexity parameters modeling over the random testing dataset. We found that using 0.01 as complexity parameters the RP model performed the best prediction. Figure 4.4 shows the structure of a decision tree using 0.01 as a complexity parameter in the RP algorithm.

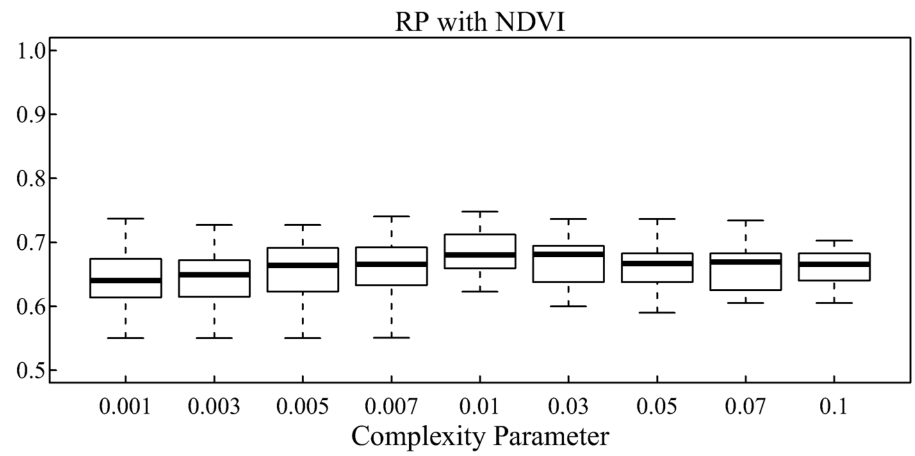


Figure 4.3 Distribution of F_1-score from 30 repeated RP modeling with different complexity parameters.

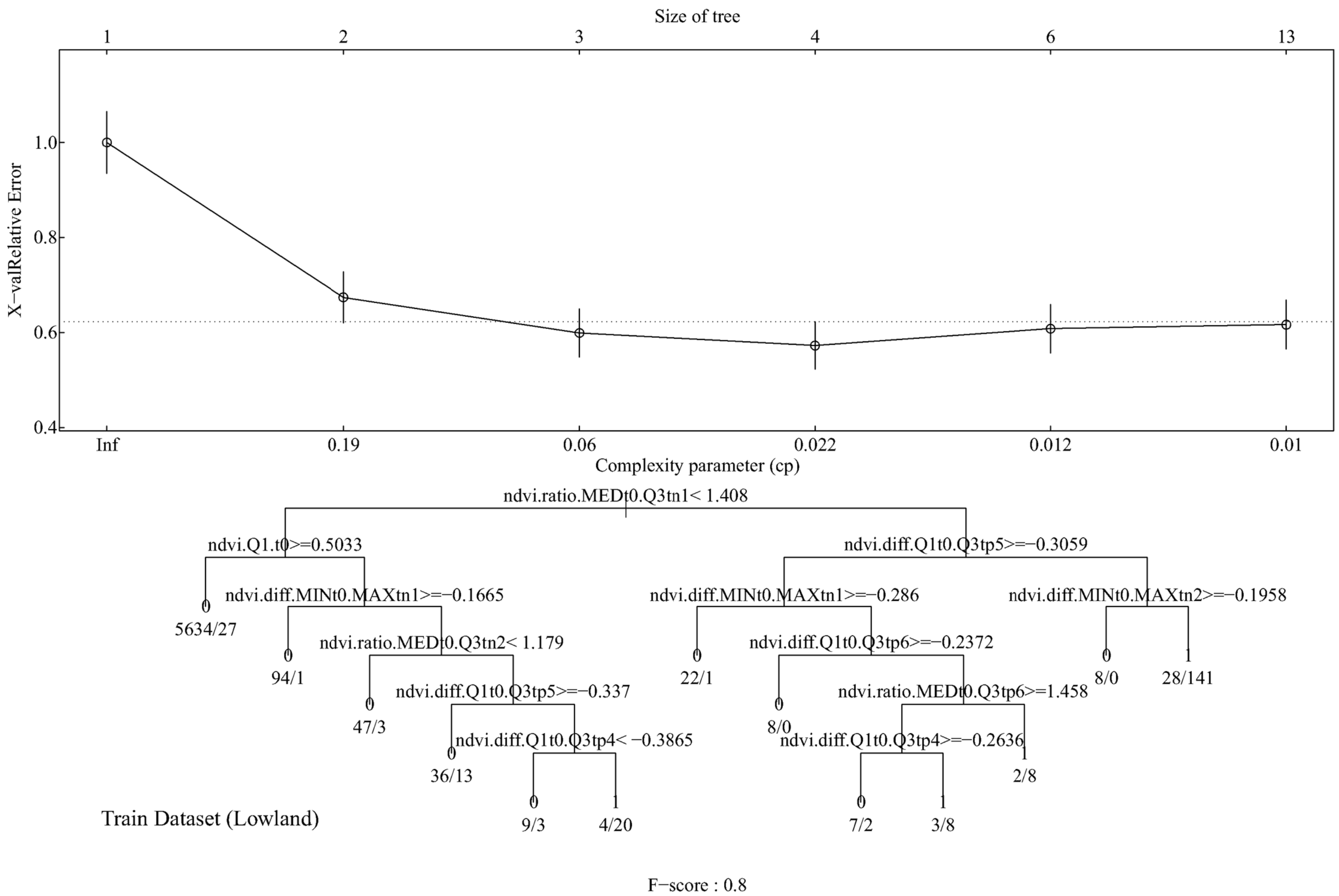


Figure 4.4 Prediction model of Recursive Partitioning.

4.2.3 Model Accuracy Assessment

For the accuracy assessment of data prediction models in this study. We used the confusion matrix to evaluate the potential of modeling from the recursive partitioning algorithm with complexity parameters as 0.01. We can describe the accuracy of the models, as shown in Table 4.2. The sample data was divided into the training and test dataset are 5427 and 3645, respectively. The overall accuracy of training and test datasets were 98.55% and 97.84%, respectively. The accuracy assessment of F_1-score of a training and test dataset were 0.80 and 0.69, respectively. There was insignificant different of F_1-score from training and testing dataset. Thus, the RP model with complexity parameters as 0.01 was acceptable and used to predict the NDVI time series from unknown T_0 from 1,832 plots.

Table 4.2 Model accuracy assessment of Recursive Partitioning.

Training Data		**Actual Class**			**Error Matrix**	
		0	**1**	**Sum**		
Predicted Class	0	5210	37	5247	OA	98.55
	1	42	165	207	F_1-score	0.80
	Sum	5252	202	5454		
Test Data		**Actual Class**			**Error Matrix**	
		0	**1**	**Sum**		
Predicted Class	0	3453	47	3500	OA	97.84
	1	31	87	118	F_1-score	0.69
	Sum	3484	134	3648		

4.2.4 Unknown T_0 Prediction

After we know that using 0.01 as the complexity parameters produced the best prediction result, we performed model training from training (known T_0) dataset and then apply the model to the predicting (unknown T_0) dataset. The process was repeated 100 times and the result of T_0 prediction (year of rubber planting) was recorded. The maximum the predicted T_0 was chosen and the percentage of nominated T_0 was calculated. Table 4.3 shows the prediction result from both training and prediction datasets using the training RP model.

Table 4.3 Results of para rubber plantation ages prediction in Thalang District.

	Lowland (plots)	
Num. of Plots	2,196	
Small Plots	28	
Modeling	2,168	
	Training (known T_0)	Predicting (unknown T_0)
	336 (15.5%)	1,832 (84.5%)
Corrected	285 (84.8%)	
Incorrected	46 (13.7%)	
NA	5 (1.5%)	397 (21.7%)

In training dataset, there are 84.8% of corrected prediction. The miss-prediction account 13.7% which the average of the incorrect year of prediction is 0.721 year. The NA (short for "Not Available") indicates under prediction. The detail of miss-predicted year is displayed in Figure 4.5. The most miss-predicted year is one year (17 out of 46). The results of rubber plantation ages prediction are summarized in Table 4.4 and Figure 4.6 shows the map of predicted rubber plantation ages. The NA indicates the under-prediction in case of predicting training dataset. On another hand, it indicates both under prediction and year of rubber planting is greater than 27 years old in case of predicting dataset.

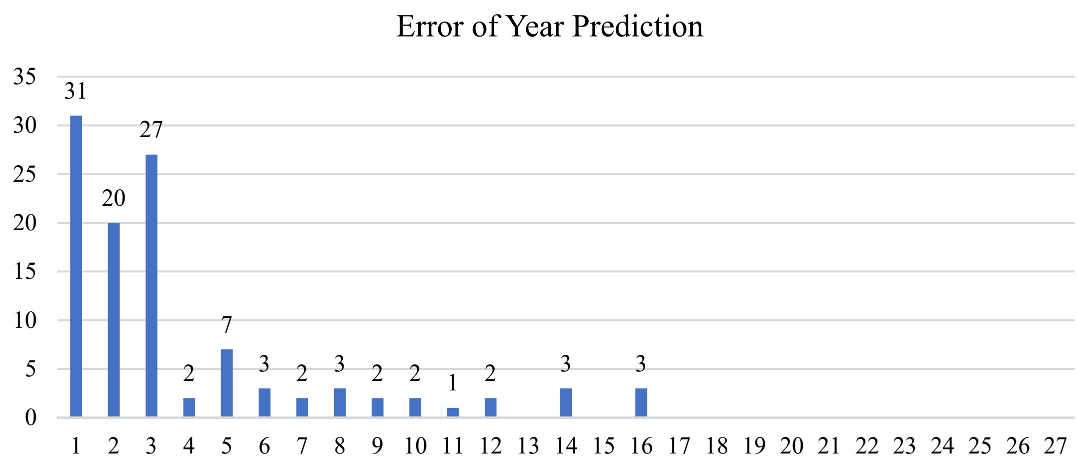


Figure 4.5 Number of the miss-predicted year.

Table 4.4 Results of para rubber plantation ages prediction in Thalang District.

Year of Re-planting	Plot Age (years old)	Number of Plots
1992	26	82
1993	25	18
1994	24	24
1995	23	18
1996	22	54
1997	21	39
1998	20	77
1999	19	63
2000	18	70
2001	17	27
2002	16	92
2003	15	86
2004	14	92
2005	13	188
2006	12	84
2007	11	48
2008	10	67
2009	9	89
2010	8	47
2011	7	79
2012	6	74
2013	5	80
2014	4	20
2015	3	98
2016	2	58
2017	1	36
2018	< 1	48
NA		402
Total		2,168

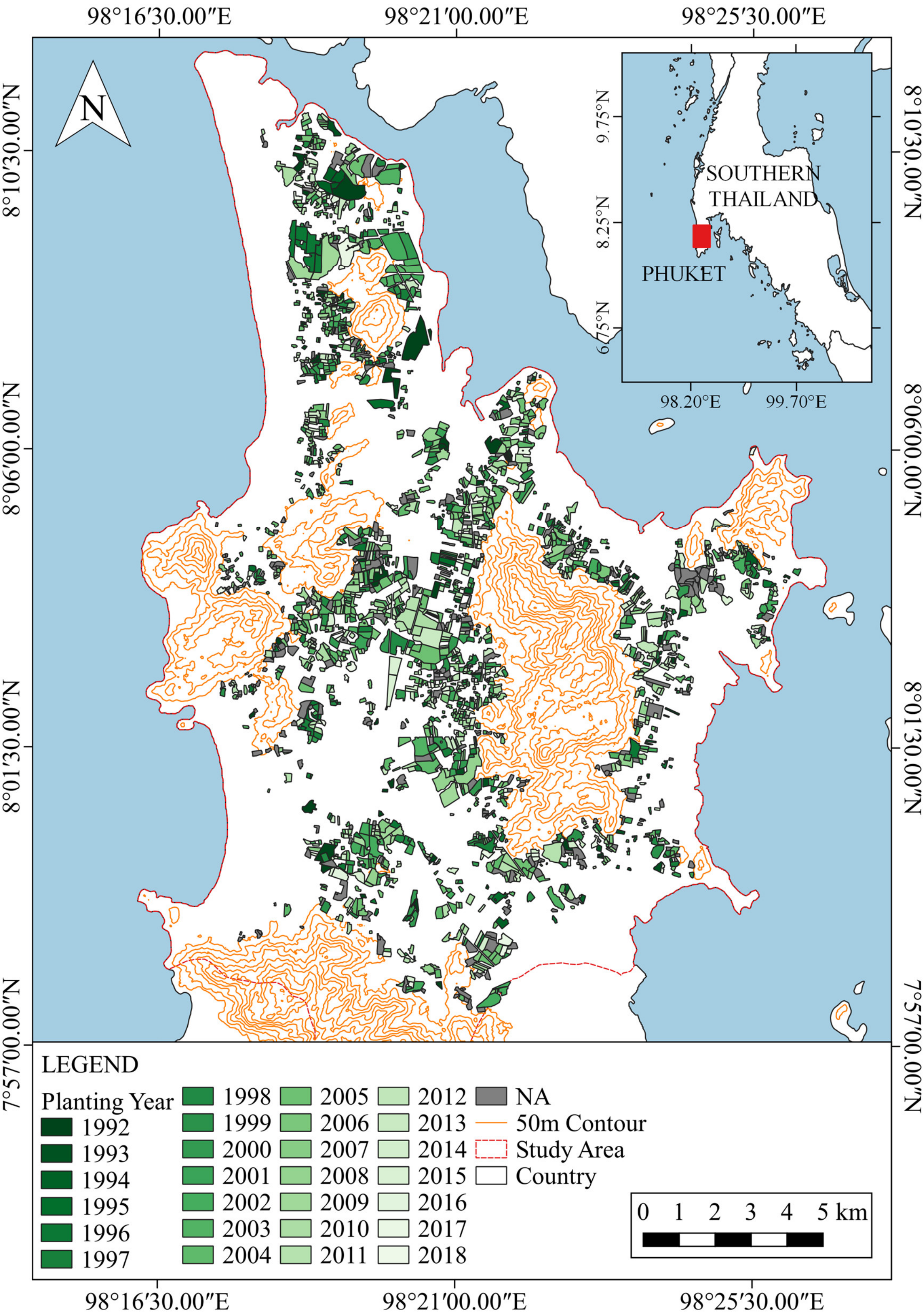


Figure 4.6 Map of lowland para rubber plantation ages in Thalang District, Phuket Province, Southern Thailand in the year 2018.

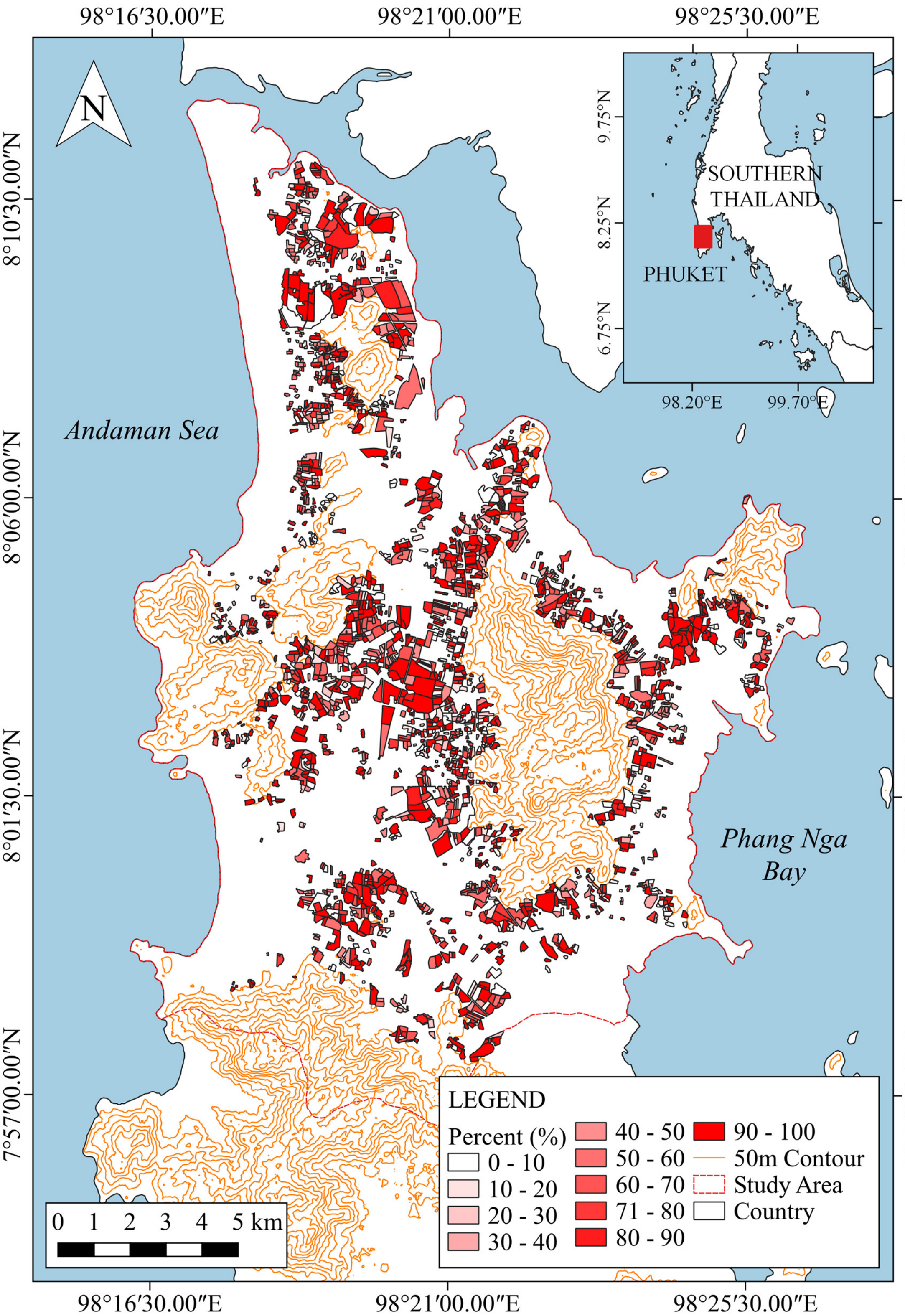


Figure 4.7 Map of percent of age prediction for lowland para rubber plantation in Thalang District, Phuket Province, Southern Thailand in the year 2018.

4.2.5 Prediction Validation

The registration data of rubber plantations from the RAOT then were used to validate the prediction result. There were obtained 131 records from RAOT. However, those are 28 registrations of rubber planting that have been approved by RAOT in the year 2018 which we cannot use as the validated data because the farmers normally start planting rubber in their lands one year after approval. Thus, there were 103 records used in prediction validation. Figure 4.8 summarizes the validation of rubber plantation' age prediction using data from RAOT. The negative number on the x-axis indicates the predicted year is less than the actual planting year. The y-axis is the number of plots that have error prediction. There were 15 plots that have correct prediction and there were 54 plots that have 1-year miss-prediction. The average of the in-corrected year of prediction is 3.62 years.

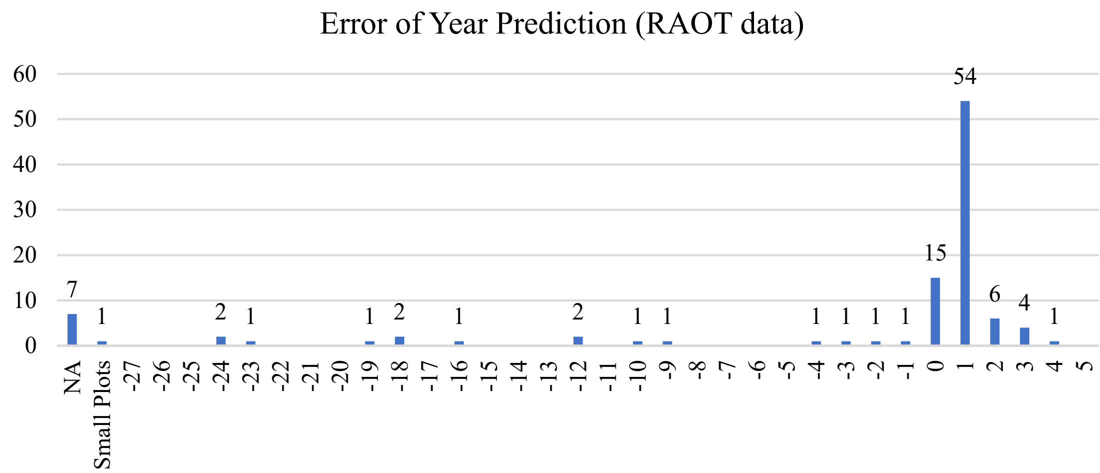


Figure 4.8 Number of the miss-predicted year using data from RAOT.

CHAPTER 5

DISCUSSION AND CONCLUSION

From this study of investigating the potential of using time series of using 16-day composite Landsat time series at the 30-meter spatial resolution to identify the abrupt change of vegetation greenness in para rubber plantations. The inter-annual vegetation profile of para rubber planting event that has been extracted and generates predictor variables for supervised classification in Machine Learning. The tree-based recursive partitioning was used to train a model from the training dataset and predict the event of para rubber re-planting and plantation lifespan. Last, we assessed the accuracy of the predicted model to create a spatial database and mapping the age of para rubber plantations classified at the 1-year level in Thalang District, Phuket Province, Southern Thailand. We can discuss the advantages and disadvantages and summarize all the content from this research as follows.

5.1 Discussion

There are good reasons for using Landsat satellite images for this study. We chose data used NDVI time series every 16-day with a 30-meter spatial resolution from Landsat TM, ETM+, and OLI sensors because the finished product and can download the data for free. It has the potential to identify the age of para rubber plantations that are larger than 30-meter. It can also be used to study other LULC, such as oil palm, forestry, or other agricultural areas. All of these offers advantages for using Landsat NDVI time series. However, we can come across obstacles or factors that affect the accuracy of results from this study are based on seven issues as follows;

1) The data of NDVI time series extracted from the Landsat TM, ETM+, and OLI satellite images affect the accuracy of the results. From this study, we found that the range of NDVI values obtained from Landsat OLI is higher than Landsat TM and ETM+, which it affects to display the graph of inter-annual NDVI profiles.

2) Annual seasons or atmospheric phenomena affects the ability to download satellite images of each year that can be used, due to the number of satellite images affects the accuracy of the results in this study. Also, some satellite images must eliminate the pixels of shadow and clouds, which causes the DNs of satellite images to be changed to NA.

3) The number of sample points will be more or less affect the accuracy of results. If there are many sample points, the results are more accurate, which the number of sample points depends on the size of para rubber plot. Therefore, it indicates that the prediction of large plots is more accurate than small para rubber plots.

4) The deciduous season of the para rubber trees is between December and March. It causes the NDVI values to be lower than usual period, which affects the inter-annual NDVI Profile graphs resulting in the high variation of predictor variables.

5) In this study, we found that the other plants or human-made that are mixed in the para rubber plots. It affects the NDVI time series as well. For example, mixed young rubber tree with pineapples cause NDVI values to be higher than usual. Buildings cause the NDVI value to be lower than average. This negative effect can vary NDVI distribution in the second and third year after year of planting.

6) During the cleared area for the cultivation of plants (T_0). It was found that the NDVI value did not reach zero due to farmers gradually cleared the plot by partition the area into sections, causing the plants to remain in the plot. In addition, there may be objects that are mixed in the plantation such as wood carcasses, plant debris and weeds, etc. It affects the reflection values send to the satellite receiver (Adsavakulchai, 2015).

7) These are 28 registrations of para rubber planting that have been approved by RAOT in the year 2018, which we cannot use as the validated data because the farmers typically start planting rubber in their lands one year after approval. Thus, there were 103 records used in prediction validation.

This presented approach can be applied to other dense satellite time series with a higher spatial resolution effectively. However, higher-resolution satellite images are also some disadvantages, such as enormous file size, longer download time, larger storage space, a longer time for data analysis.

5.2 Conclusion

This research is a study of an abrupt change of vegetation greenness and investigates the potential and methods used in a decision tree algorithm to identify para rubber plantations from 0 to 27 years using 16-day Landsat time series data of the lowland para rubber plantations in Thalang District, Phuket Province, Southern Thailand. First, we extracted data from NDVI time series from Landsat imagery to create the inter-annual NDVI profile modeling were 2,196 plots para rubber plantation to identify the para rubber re-planting event (vegetation clearance occurred) event at T_0 with the high-resolution historical satellite imagery from Google Earth Pro^{TM} program. We divided the data into two groups: modeling dataset (known T_0) and predicting dataset (unknown T_0) which were 336 and 1,832 plots, respectively. Then, we generated the predictors for decision tree classification from inter-annual NDVI profile. The sample data was divided into the training, and test dataset are 5427 and 3645, respectively. The overall accuracy of training and test dataset was 98.55% and 97.84%, respectively. The accuracy assessment of F_1-score of a training and test dataset were 0.80 and 0.69, respectively. There was insignificant different of F_1-score from training and testing dataset. Thus, the recursive partitioning model with complexity parameters as 0.01 was acceptable and used to predict the NDVI time series. Therefore, we used the results obtained prediction to identify para rubber plantation ages to generating the mapping of para rubber plantation ages in Thalang District, Phuket Province, Thailand. The final result is a map of para rubber plantation ages classified at 1-year level was generated using a recursive partitioning algorithm. We found that the 30-meter spatial resolution Landsat NDVI time series successfully identifies para rubber plantation ages, especially for large plots than the image resolution (pixel size) of the Landsat satellite imagery from the remote sensing technology. In addition, this research can be applied research methodology for various research applications, which can be used for inspecting other agricultural areas including monitoring for forestry encroachment, natural disaster risks, etc. The high precision database will enable accurate yield prediction, which help makes the decision-making and planning for management in the agricultural sector better. The success of this research results is an important fundamental step in the further development of Thailand's agricultural technology.

REFERENCES

Adsavakulchai, S. (2015). "Impact of Landuse Practices on Ecosystem Diversity using Remotely Sensed Data." *International Journal of Geoinformatics*, 11(2), 49–56.

Agricultural Research Development Agency (Public Organization). (2018). "History of para rubber plantation in Thailand." (Online) Available on http://www.arda.or.th/kasetinfo/south/para/history/01-05.php (3 October 2018)

Ashwini, R. (2017). "A study on precision agriculture and its technologies." *International Journal of Scientific Research in Science*, Engineering and Technology (IJSRSET), 3(1), 378-382.

Beckschäfer, P. (2017). "Obtaining Rubber Plantation Age Information from Very Dense Landsat TM&ETM+ Time Series data and Pixel-based Image Compositing." *Remote Sensing of Environment*, 196, 89–100.

Bhumiratana, A., Sorosjinda-Nunthawarasilp, P., Kaewwaen, W., Maneekan, P., and Pimnon, S. (2013). "Malaria-associated rubber plantations in Thailand." *Travel Medicine and Infectious Disease*, 11(1), 37–50.

Boonyanam, N. (2018). "Agricultural economic zones in Thailand." *Land Use Policy*, (May), 1–9.

Bronshtein, A. (2017). "Train/Test split and cross-validation in Python." (Online) Available on https://towardsdatascience.com/train-test-split-and-cross-validation-in-python-80b61beca4b6 (7 November 2018).

Brown, M. (2015). "Satellite remote sensing in agriculture and food security assessment.", *Procedia Environmental Sciences*, 29, 307.

Brownlee, J. (2017). "What is the difference between test and validation datasets?" (Online) Available on https://machinelearningmastery.com/difference-test-validation-datasets/ (7 November 2018).

Chen, G., Thill, J., Anantsuksomsri, S., Tontisirin, N., and Tao, R. (2018). "Standage estimation of rubber (Hevea brasiliensis) plantations using an integrated pixel and object-based tree growth model and annual Landsat time series." *ISPRS Journal of Photogrammetry and Remote Sensing*, 144, 94–104.

Comber, A., Fisher, P., Brunsdon, C., and Khmag, A. (2012). "Spatial analysis of remote sensing image classification accuracy." *Remote Sensing of Environment*, 127, 237–246.

REFERENCES (cont.)

Cook, E., and Goldman, L. (1984). "Empiric comparison of multivariate analytic techniques: advantages and disadvantages of recursive partitioning analysis." *Journal of Chronic Diseases*, 37(9), 721–731.

Dibs, H., Idrees, M., Bedawi, G., and Alsalhin, A. (2017). "Hierarchical classification approach for mapping rubber tree growth using per-pixel and object-oriented classifiers with SPOT-5 imagery." *The Egyptian Journal of Remote Sensing and Space Sciences*, 20(1), 21–30.

Egghe, L. (2008). "The measures precision, recall, fallout and miss as a function of the number of retrieved documents and their mutual interrelations." *Information Processing & Management*, 44(2), 856–876.

Fang, H., and Liang, S. (2008). "Leaf area index models." *Encyclopedia of Ecology*, 2014, 2139–2148.

Far, S., and Rezaei-Moghaddam, K. (2018). "Impacts of the precision agricultural technologies in Iran: an analysis experts' perception and their determinants." *Information Processing in Agriculture*, 5(1), 173–184.

Food and Agriculture Organization of the United Nations. (2017). *National Agro-Economic Zoning for major crops in Thailand (NAEZ)*, Food and Agriculture Organization of the United Nations, Rome, Italy.

Fox, J., and Castella, J. (2013). "Expansion of rubber (*Hevea brasiliensis*) in mainland Southeast Asia." *Journal of Peasant Studies*, 40(1), 155–170.

Fox, J., Castella, J., Ziegler, A., and Westley S. (2014). "Rubber Plantations Expand in Mountainous Southeast Asia: What Are the Consequences for the Environment? For centuries, farmers in the mountainous region of the mainland." *Analysis from the East-West Center*, 114, 1–8.

Friedl, M., and Brodley, C. (1997). "Decision tree classification of land cover from remotely sensed data." *Remote Sensing of Environment*, 61(3), 399–409.

Gale, F., Hansen, J., and Jewison, M. (2015). "China's growing demand for agricultural imports." *Economic Information Bulletin Number*, 136, 1–39.

GeeksforGeeks. (2018). "Confusion matrix in machine learning." (Online) Available on https://www.geeksforgeeks.org/confusion-matrix-machine-learning/ (21 October 2018).

REFERENCES (cont.)

Glenn, E., Huete, A., Nagler, P., and Nelson, S. (2008). "Relationship Between Remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape." *Sensors (Basel)*, 8(4), 2136 – 2160.

Gorbachev, S., Gorbacheva, N., and Koynov, S. (2016). "A Synergistic Effect in the Measurement of Neuro-Fuzzy System", *In Proceedings of the MATEC Web of Conferences 76*. Corfu Island, Greece: 14–17 July 2016.

Graves, S., Caughlin, T., Asner, G., and Bohlman, S. (2018). "A tree-based approach to biomass estimation from remote sensing data in a tropical agricultural landscape." *Remote Sensing of Environment*, 218, 32–43.

Izquierdo, J., Fazzone, M., and Duran, M. (2007). *Guidelines of good agricultural practices for family agriculture*, Departmental Program on Food and Nutritional Security. Antioquia, Colombia.

Jawjit, W., Kroeze, C., and Rattanapan, S. (2010). "Greenhouse gas emissions from the rubber industry in Thailand." *Journal of Cleaner Production*, 18(5), 403–411.

Killmann, W., and Hong, L. (2000). "Rubberwood: the success of an agricultural by-product." *Forest Plantations Thematic Papers*, 51, 66–72.

Kingsford, C., and Salzberg, S. (2008). "What are decision trees?" *Nature Biotechnology*, 26(9), 1011–1012.

Koedsin, W., and Yasen, K. (2016). "Estimating leaf area index of rubber plantation using Worldview-2 imagery." *Journal of Life Sciences and Technologies*, 4(1), 1–6.

Kongmanee, C., and Longpichai, O. (2017). *Risk recognition and risk management strategy of rubber farmers in Southern Thailand*, Prince of Songkla University, Songkla, Thailand.

Köppen, W. (1918). "Klassifikation der Klimate nach Temperatur, Niederschlag und Jahresablauf (Classification of Climates According to Temperature, Precipitation and Seasonal Cycle)." *Petermanns Geogr. Mitt*, 64, 193–203.

Kou, W., Xiao, X., Dong, J., Gan, S., Zhai, D., Zhang, G., and Li, L. (2015). "Mapping deciduous rubber plantation areas and stand ages with PALSAR and Landsat images." *Remote Sensing*, 7(1), 1048–1073.

Kraipinit, Y., Chantuk, T., and Siriwong, P. (2017). "New agricultural management of Thailand." *Journal of Research and Development Valaya Alongkorn*, 12(2), 115–127.

REFERENCES (cont.)

Landau, S., and Barthel, S. (2010). "Recursive partitioning." *International Encyclopedia of Education (Third Edition)*, 383–389.

Li, Z., and Fox, J. (2012). "Mapping rubber tree growth in mainland Southeast Asia using time series MODIS 250-m NDVI and statistical data." *Applied Geography*, 32(2), 420–432.

López-serrano, P., Corral-rivas, J., and Díaz-varela, R. (2016). "Evaluation of radiometric and atmospheric correction algorithms for aboveground forest biomass estimation using Landsat 5 TM data." *Remote Sensing*, 8(369), 1–19.

Loveland, T., and Dwyer, J. (2012). "Landsat: building a strong future." *Remote Sensing of Environment*, 122(October 2000), 22–29.

Maggiotto, S., Oliveira, D., Marur, C., Stivari, S., Leclerc, M., and Wagner-Riddle, C. (2014). "Potential carbon sequestration in rubber tree plantations in the northwestern region of the Paraná State, Brazil." *Acta Scientiarum-Agronomy*, 36(1).

Meteorological Development Office. (2017). *Phuket Weather*, Meteorological Department, Phuket, Thailand.

Muradyan, V., Asmaryan, S., and Saghatelyan, A. (2016). "Assessment of space and time changes of NDVI (biomass) in Armenia's mountain ecosystems using remote sensing data." *Current Problems in Remote Sensing of the Earth from Space*, 13(1), 49–60.

National Bureau of Agricultural Commodity and Food Standards of Thailand. (2013). *Thai Agricultural Standard TAS 9001 – 2013: Good Agricultural Practices for Food Crop. Government Gazette*, Ministry of Agriculture and Cooperatives, Bangkok, Thailand. www.acfs.go.th/standard/download/GAP_food%20crop.pdf.

Pahlevan, N., Schott, J., Franz, B., Zibordi, G., Markham, B., Bailey, S., and Strait, C. (2017). "Remote sensing of environment Landsat 8 remote sensing reflectance products: evaluations, intercomparisons, and enhancements." *Remote Sensing of Environment*, 190, 289–301.

Paul, M., Vishwakarma, S., and Verma, A. (2015). "Analysis of soil behavior and prediction of crop yield using data mining approach.", *In 2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, Jabalpur, India: 12–14 December 2015.

REFERENCES (cont.)

Petchsawang, A. (2018). “RAOT and GISTDA bring satellite exploration of para rubber plantations across the country to create database and map to unity.” (Online) Available on http://www.gistda.or.th/main/th/node/2370 (2 October 2018).

Petchseechaung, W. (2016). “Thailand Rubber Industry 2016 – 2018.” *Thailand industry outlook 2016 – 2018*, (May), 1–4.

Phuket Development Strategy. (2016). *General conditions and information of Phuket province 2016*, Phuket Development Strategy, Phuket, Thailand.

Phuket Provincial Governor’s Office. (2018). *General and development situation in Phuket provincial development plan (2018 – 2021)*, Phuket Provincial Governor’s Office, Phuket, Thailand.

Post Today. (2017). “Thailand rubber prices decrease by 3 reasons.” (Online) Available on https://www.posttoday.com/economy/524937 (2 October 2018).

Quinlan, J. R. (1986). “Induction of decision trees.” *In Machine Learning*, 1, 81–106.

Ratnasingam, J., Ioraş, F., and Wenming, L. (2011). “Sustainability of the rubberwood sector in Malaysia.” *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 39(2), 305–311.

Reunanen, J. (2003). “Overfitting in making comparisons between variable selection methods.” *Journal of Machine Learning Research*, 3(2003), 1371–1382.

Riwthong, S., Schreinemachers, P., Grovermann, C., and Berger, T. (2017). “Agricultural commercialization: risk perceptions, risk management and the role of pesticides in Thailand.” *Kasetsart Journal of Social Sciences*, 38(3), 264–272.

Romyen, A., Sausue, P., and Charenjiratragul, S. (2018). “Investigation of the rubber-based intercropping system in southern Thailand. *Kasetsart Journal of Social Sciences*, 39, 135–142.

Rouse, W., Haas, H., and Deering, W. (1973). “Monitoring vegetation systems in the great plains with ERTS.”, *In Proceedings of the Third ERTS symposium*. Washington DC, USA: 10–14 December 1973.

Rubber Authority of Thailand. (2018). “Thailand rubber price today.” (Online) Available on http://www.rubber.co.th/rubber2012/menu5.php (2 October 2018).

Rubber Authority of Thailand. (2019). “Rubber Authority of Thailand.” (Online) Available on http://www.raot.co.th/ewt_news.php?nid=3739 (21 May 2019).

Rubber Research Institute. (2010). *Suitable area for growing rubber in academic information of para rubber*, Rubber Research Institute, Bangkok, Thailand.

REFERENCES (cont.)

Satir, O., and Berberoglu, S. (2016). "Crop yield prediction under soil salinity using satellite-derived vegetation indices." *Field Crops Research*, *192*, 134–143.

Shobha, G., and Rangaswamy, S. (2018). "Machine learning." *Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications*, 38, 197–228.

Singhapreecha, C. (2014). "Economy and agriculture in Thailand." (Online) Available on http://ap.fftc.agnet.org/member.php (2 October 2018).

Smajgl, A., Xu, J., Egan, S., Yi, Z. F., Ward, J., and Su, Y. (2015). "Assessing the effectiveness of payments for ecosystem services for diversifying rubber in Yunnan, China." *Environmental Modelling and Software*, *69*, 187–195.

Song, Y., and Lu, Y. (2015). "Decision tree methods: applications for classification and prediction." *Shanghai Archives of Psychiatry*, *27*(2), 130–135.

Story, M., and Congalton, R. (1986). "Accuracy assessment: a user's perspective." *Society*, 52(3), 397–399.

Surat Thani Rubber Research Center. (2018). "GAP for para rubber in Thailand." (Online) Available on https://www.fio.co.th/south/Powerpoint/GAPRubber.ppt (3 October 2018).

Tekasakul, P., and Tekasakul, S. (2006). "Environmental problems related to natural rubber production in Thailand." *J. Aerosol Res*, 21(2), 122–129.

Teoh, Y., Don, M., and Ujang, S. (2011). "Assessment of the properties, utilization, and preservation of rubberwood (*Hevea brasiliensis*): a case study in Malaysia." *Journal of Wood Science*, 57(4), 255–266.

Thailand Government. (2015). *The Rubber Authority of Thailand Act B.E. 2558 (2015). Government Gazette*, Ministry of Agriculture and Cooperatives, Bangkok, Thailand. https://www.raot.co.th/article_attach/article_20150714182813.pdf.

Thairath. (2017). "Measures to solve the problem of rubber prices." (Online) Available on https://www.thairath.co.th/content/1151493 (October 2, 2018).

Tharwat, A. (2018). "Classification assessment methods." *Applied Computing and Informatics*. (In Press)

U.S. Geological Survey. (2016). "Landsat collections science for a changing world." (Online) Available on https://landsat.usgs.gov/landsat-collections (20 October 2018).

REFERENCES (cont.)

U.S. Geological Survey. (2018). “What are the band designations for the Landsat satellites?” (Online) Available on https://landsat.usgs.gov/what-are-band-designations-landsat-satellites (3 October 2018).

U.S. Geological Survey. (2019). “Landsat 4-7 Surface Reflectance (LEDAPS) Product Guide.” (Online) Available on https://www.usgs.gov/media/files/landsat-4-7-surface-reflectance-code-ledaps-product-guide (22 June 2019).

U.S. Geological Survey. (2019). “Landsat 8 Surface Reflectance Code (LASRC) Product Guide.” (Online) Available on https://www.usgs.gov/media/files/landsat-8-surface-reflectance-code-lasrc-product-guide (22 June 2019).

Verbesselt, J., Hyndman, R., Newnham, G., and Culvenor, D. (2010). “Remote sensing of environment detecting trend and seasonal changes in satellite image time series.” *Remote Sensing of Environment*, 114(1), 106–115.

Verheye, W. (2010). “Growth and production of rubber.” *In Land Use, Land Cover and Soil Sciences*, 295–300.

Wadeesirisak, K., Castano, S., Berthelot, K., Vaysse, L., Bonfils, F., Peruch, F., and Bottier, C. (2017). “Rubber particle proteins REF1 and SRPP1 interact differently with native lipids extracted from *Hevea Brasiliensis* latex.” *Biochimica et Biophysica Acta - Biomembranes*, 1859(2), 201–210.

Wakefield, K. (2019). “A guide to machine learning algorithms and their applications” (Online) Available on https://www.sas.com/en_gb/insights/articles/analytics/ machine-learning-algorithms.html (23 May 2019).

Willcock, S., Martínez-López, J., Hooftman, D., Bagstad, K., Balbi, S., Marzo, A., and Athanasiadis, I. (2018). “Machine learning for ecosystem services.” *Ecosystem Services*, 33, 165–174.

Zhang, X., Feng, X., Xiao, P., He, G., and Zhu, L. (2015). “Segmentation quality evaluation using region-based precision and recall measures for remote sensing images.” *ISPRS Journal of Photogrammetry and Remote Sensing*, 102, 73–84.

Zhu, Z., Qiu, S., and He, B. (2018). “Cloud and cloud shadow detection for Landsat images: the fundamental basis for analyzing Landsat time series.” *Remote Sensing Time Series Image Processing*, (May), 1–23.

APPENDICES

Appendix A
Landsat Time Series

Table I – Landsat Collections (Landsat-5 TM)

ID	Date Taken	DOY	ID	Date Taken	DOY
1	1992-03-16	75	35	2006-02-19	50
2	1992-04-17	107	36	2006-03-07	66
3	1992-12-29	363	37	2006-11-18	322
4	1993-03-03	62	38	2006-12-04	338
5	1994-01-17	17	39	2006-12-20	354
6	1994-11-17	321	40	2007-02-06	37
7	1994-12-19	353	41	2007-02-22	53
8	1995-01-20	20	42	2008-10-06	279
9	1995-03-09	68	43	2009-01-10	10
10	1995-04-10	100	44	2009-01-26	26
11	1996-02-24	55	45	2009-02-11	42
12	1996-03-11	70	46	2009-02-27	58
13	1996-04-12	102	47	2009-11-10	314
14	1997-02-26	57	48	2009-11-26	330
15	1997-04-15	105	49	2009-12-12	346
16	1997-11-25	329	50	2010-01-29	29
17	1998-01-12	12	51	2010-02-14	45
18	1998-03-01	60	52	2011-01-16	16
19	1998-03-17	76	53	2011-02-01	32
20	1999-02-16	47	54	2011-04-22	112
21	1999-03-04	63	55	2011-10-15	288
22	1999-03-20	79	**Total Images**		**55**
23	2000-02-03	34			
24	2001-01-04	4			
25	2001-02-05	36			
26	2001-02-21	52			
27	2004-12-30	364			
28	2005-01-15	15			
29	2005-01-31	31			
30	2005-02-16	47			
31	2005-03-04	63			
32	2005-03-20	61			
33	2005-04-21	111			
34	2006-02-03	34			

Table II – Landsat Collections (Landsat-7 ETM+)

ID	Date Taken	DOY	ID	Date Taken	DOY
1	2000-12-27	361	35	2014-02-01	32
2	2001-11-28	332	36	2014-02-17	48
3	2002-01-15	15	37	2014-03-05	64
4	2002-03-04	63	38	2014-11-16	320
5	2003-01-18	18	39	2015-01-03	3
6	2003-03-23	82	40	2015-01-19	19
7	2003-12-04	338	41	2015-02-04	35
8	2004-01-21	21	42	2015-02-20	51
9	2004-04-10	100	43	2015-03-08	67
10	2004-12-22	356	44	2015-04-09	99
11	2005-01-23	23	45	2015-12-21	355
12	2005-02-08	39	46	2016-01-06	6
13	2005-02-24	55	47	2016-01-22	22
14	2005-11-23	327	48	2016-02-23	54
15	2006-10-25	298	49	2016-03-10	69
16	2006-11-26	330	50	2016-04-11	101
17	2007-01-29	29	51	2017-02-09	40
18	2007-02-14	45	52	2017-02-25	56
19	2007-11-29	333	53	2017-12-10	344
20	2008-01-16	16	54	2017-12-26	360
21	2008-04-05	95	55	2018-02-12	43
22	2009-01-18	18	56	2018-02-28	59
23	2009-02-19	50	**Total Images**		56
24	2010-01-21	21			
25	2010-02-06	37			
26	2010-03-10	69			
27	2011-02-25	56			
28	2011-04-14	104			
29	2011-10-23	296			
30	2011-12-26	360			
31	2012-01-27	27			
32	2012-02-12	43			
33	2013-01-29	29			
34	2013-03-18	77			

Table III – Landsat Collections (Landsat-8 OLI)

ID	Date Taken	DOY
1	2013-12-23	357
2	2014-01-24	24
3	2014-02-25	56
4	2014-03-13	72
5	2015-01-11	11
6	2015-01-27	27
7	2015-02-12	43
8	2015-02-28	59
9	2015-03-16	75
10	2015-04-17	107
11	2016-01-14	14
12	2016-01-30	30
13	2016-02-15	46
14	2016-03-18	77
15	2016-04-03	93
16	2016-04-19	109
17	2017-02-17	48
18	2018-02-04	35
Total Images		**18**

Appendix B
Registration of Rubber Farmers in Thalang District

Table IV – Registration of rubber farmers in Thalang District 2018

ID	PlotID	Latitude	Longitude	Approved
1	24	8° 03' 05.88" N	98° 18' 19.24" E	2015
2	52	8° 01' 41.11" N	98° 21' 36.75" E	2015
3	65	7° 58' 29.77" N	98° 19' 52.53" E	2018
4	68	8° 00' 46.89" N	98° 23' 38.94" E	2018
5	69	8° 01' 53.34" N	98° 21' 15.51" E	2014
6	83	8° 06' 46.99" N	98° 21' 55.81" E	2014
7	100	7° 59' 49.02" N	98° 22' 13.11" E	2015
8	154	8° 10' 53.27" N	98° 18' 59.65" E	2015
9	157	8° 09' 54.63" N	98° 18' 52.44" E	2015
10	312	8° 04' 12.92" N	98° 18' 18.19" E	2015
11	325	8° 07' 03.69" N	98° 18' 52.79" E	2018
12	390	8° 03' 50.45" N	98° 23' 32.89" E	2015
13	403	8° 03' 31.84" N	98° 20' 39.78" E	2015
14	406	8° 02' 57.31" N	98° 20' 16.86" E	2014
15	408	8° 02' 55.43" N	98° 20' 57.56" E	2015
16	473	8° 02' 14.67" N	98° 21' 03.91" E	2017
17	475	8° 02' 12.82" N	98° 21' 05.09" E	2017
18	481	8° 02' 32.75" N	98° 20' 56.78" E	2015
19	492	8° 02' 40.30" N	98° 21' 34.10" E	2017
20	636	8° 04' 32.42" N	98° 21' 13.13" E	2015
21	693	8° 02' 32.74" N	98° 19' 21.02" E	2017
22	748	8° 03' 08.42" N	98° 19' 14.23" E	2015
23	753	8° 04' 53.55" N	98° 21' 10.50" E	2018
24	785	8° 01' 07.29" N	98° 20' 57.67" E	2017
25	789	8° 01' 43.17" N	98° 21' 00.72" E	2014
26	807	8° 00' 50.95" N	98° 21' 45.65" E	2015
27	839	8° 02' 07.81" N	98° 21' 03.42" E	2015
28	886	8° 00' 05.21" N	98° 20' 15.69" E	2016
29	894	7° 59' 40.20" N	98° 19' 21.26" E	2014
30	895	7° 59' 32.81" N	98° 19' 32.99" E	2015
31	897	7° 59' 50.41" N	98° 19' 18.82" E	2015
32	927	7° 59' 45.80" N	98° 21' 39.13" E	2017
33	953	7° 58' 23.53" N	98° 20' 10.95" E	2015
34	960	8° 00' 09.92" N	98° 19' 31.25" E	2016

Table IV – Registration of rubber farmers in Thalang District 2018 (cont.)

ID	Plot ID	Latitude	Longitude	Approved
35	1046	7° 59' 37.19" N	98° 19' 07.87" E	2014
36	1050	8° 05' 58.84" N	98° 22' 02.29" E	2015
37	1128	8° 09' 46.55" N	98° 19' 09.77" E	2015
38	1163	7° 59' 34.54" N	98° 19' 15.55" E	2014
39	1179	8° 02' 22.88" N	98° 20' 51.54" E	2017
40	1184	8° 09' 59.54" N	98° 18' 50.28" E	2015
41	1240	8° 07' 05.45" N	98° 18' 59.53" E	2015
42	1261	8° 07' 21.43" N	98° 18' 34.36" E	2017
43	1272	8° 05' 48.19" N	98° 18' 52.80" E	2015
44	1284	8° 02' 11.04" N	98° 19' 12.04" E	2018
45	1288	8° 02' 45.56" N	98° 19' 13.72" E	2017
46	1292	8° 02' 42.51" N	98° 19' 25.70" E	2017
47	1301	8° 03' 33.28" N	98° 19' 59.34" E	2015
48	1311	8° 04' 33.99" N	98° 21' 08.41" E	2015
49	1325	8° 04' 34.96" N	98° 22' 53.58" E	2017
50	1331	8° 03' 36.21" N	98° 23' 40.73" E	2017
51	1348	8° 04' 22.67" N	98° 25' 07.23" E	2018
52	1376	8° 00' 03.68" N	98° 23' 07.07" E	2017
53	1494	8° 00' 14.65" N	98° 19' 45.62" E	2018
54	1495	8° 00' 17.04" N	98° 19' 49.09" E	2014
55	1551	8° 02' 11.05" N	98° 21' 07.41" E	2015
56	1566	8° 03' 32.96" N	98° 18' 52.72" E	2016
57	1651	8° 03' 30.60" N	98° 21' 08.01" E	2017
58	1656	8° 02' 58.18" N	98° 23' 36.38" E	2015
59	1658	8° 04' 35.49" N	98° 21' 21.53" E	2016
60	1673	8° 10' 51.04" N	98° 19' 00.48" E	2015
61	1674	8° 10' 00.10" N	98° 18' 45.48" E	2017
62	1675	8° 09' 48.54" N	98° 18' 35.47" E	2017
63	1677	8° 05' 10.32" N	98° 21' 26.33" E	2015
64	1678	8° 03' 45.31" N	98° 21' 01.98" E	2016
65	1679	8° 04' 13.18" N	98° 21' 17.65" E	2014
66	1680	8° 04' 12.25" N	98° 18' 21.33" E	2015
67	1682	8° 04' 54.37" N	98° 22' 33.63" E	2015
68	1683	8° 04' 35.57" N	98° 22' 55.20" E	2017

Table IV – Registration of rubber farmers in Thalang District 2018 (cont.)

ID	Plot ID	Latitude	Longitude	Approved
69	1684	8° 04' 31.84" N	98° 23' 00.62" E	2017
70	1685	8° 04' 36.94" N	98° 23' 01.28" E	2017
71	1686	8° 04' 35.84" N	98° 23' 02.10" E	2017
72	1688	8° 02' 46.07" N	98° 20' 28.14" E	2015
73	1689	8° 02' 43.84" N	98° 20' 27.86" E	2015
74	1690	8° 02' 56.57" N	98° 20' 59.07" E	2015
75	1691	8° 07' 49.80" N	98° 19' 00.33" E	2015
76	1694	8° 02' 26.59" N	98° 20' 36.23" E	2017
77	1696	8° 02' 27.44" N	98° 21' 04.36" E	2015
78	1697	8° 03' 49.33" N	98° 19' 26.76" E	2014
79	1700	8° 02' 59.43" N	98° 18' 49.87" E	2015
80	1701	8° 03' 00.14" N	98° 18' 50.07" E	2015
81	1702	8° 03' 01.55" N	98° 18' 50.66" E	2015
82	1703	8° 02' 57.14" N	98° 19' 35.95" E	2016
83	1704	8° 01' 52.39" N	98° 19' 00.01" E	2017
84	1706	8° 00' 01.26" N	98° 18' 53.46" E	2014
85	1707	8° 03' 08.93" N	98° 23' 32.93" E	2017
86	1708	8° 02' 13.17" N	98° 23' 54.46" E	2017
87	1709	8° 01' 03.10" N	98° 23' 47.77" E	2016
88	1710	8° 01' 01.37" N	98° 23' 47.12" E	2016
89	1711	7° 59' 52.49" N	98° 22' 12.60" E	2015
90	1712	7° 59' 53.72" N	98° 21' 28.88" E	2016
91	1713	8° 02' 33.70" N	98° 21' 19.96" E	2018
92	1714	8° 03' 39.96" N	98° 19' 09.54" E	2018
93	1715	8° 02' 09.76" N	98° 19' 14.37" E	2018
94	1716	8° 02' 49.26" N	98° 24' 17.65" E	2018
95	1717	8° 02' 48.50" N	98° 24' 04.20" E	2018
96	1718	8° 02' 12.04" N	98° 23' 57.74" E	2018
97	1719	8° 00' 45.61" N	98° 23' 46.18" E	2018
98	1720	7° 59' 58.40" N	98° 21' 39.80" E	2018
99	1721	8° 00' 00.76" N	98° 21' 29.10" E	2017
100	1722	7° 59' 54.50" N	98° 21' 44.00" E	2018
101	1723	7° 57' 55.32" N	98° 20' 20.26" E	2017
102	1756	8° 07' 23.21" N	98° 19' 11.05" E	2017

Table IV – Registration of rubber farmers in Thalang District 2018 (cont.)

ID	Plot ID	Latitude	Longitude	Approved
103	1765	8° 08' 59.00" N	98° 19' 21.10" E	2016
104	1817	8° 02' 28.79" N	98° 20' 41.07" E	2016
105	1818	8° 02' 26.33" N	98° 20' 39.13" E	2017
106	1847	8° 03' 36.84" N	98° 23' 37.44" E	2015
107	1852	8° 04' 15.27" N	98° 25' 31.52" E	2017
108	1860	8° 02' 21.56" N	98° 17' 49.21" E	2013
109	1894	8° 03' 25.06" N	98° 19' 06.69" E	2014
110	1905	8° 03' 08.15" N	98° 18' 41.23" E	2018
111	1918	7° 58' 25.82" N	98° 21' 44.05" E	2016
112	1932	8° 03' 50.11" N	98° 24' 16.84" E	2015
113	1950	8° 05' 48.47" N	98° 18' 57.32" E	2015
114	1981	8° 02' 36.63" N	98° 21' 18.90" E	2018
115	1997	8° 03' 08.00" N	98° 24' 05.79" E	2016
116	2002	8° 04' 07.12" N	98° 17' 53.36" E	2015
117	2003	8° 04' 36.16" N	98° 21' 17.98" E	2016
118	2004	8° 04' 32.97" N	98° 22' 50.16" E	2017
119	2005	8° 03' 04.70" N	98° 19' 22.92" E	2015
120	2006	8° 05' 53.03" N	98° 21' 43.92" E	2018
121	2007	8° 05' 52.76" N	98° 21' 45.24" E	2018
122	2008	8° 05' 56.88" N	98° 21' 49.75" E	2018
123	2009	8° 05' 53.13" N	98° 21' 47.58" E	2018
124	2010	8° 05' 52.62" N	98° 21' 46.51" E	2018
125	2011	8° 05' 53.82" N	98° 21' 48.59" E	2018
126	2012	8° 05' 55.57" N	98° 21' 49.10" E	2018
127	2013	8° 02' 39.14" N	98° 21' 20.98" E	2018
128	2014	8° 03' 34.80" N	98° 24' 15.21" E	2018
129	2078	7° 59' 46.65" N	98° 22' 09.29" E	2018
130	2175	8° 01' 41.50" N	98° 21' 39.53" E	2015
131	2204	8° 02' 16.32" N	98° 17' 43.75" E	2016

VITAE

Name Natthaphon Somching

Student ID 5930222004

Educational Attainment

Degree	Name of Institution	Year of Graduation
Bachelor of Science (Environmental Geoinformatics)	Prince of Songkla University, Phuket Campus	2014

Scholarship Awards

Graduate School Research Fund Scholarship 2016 – 2018, Prince of Songkla University, Phuket Campus, Thailand