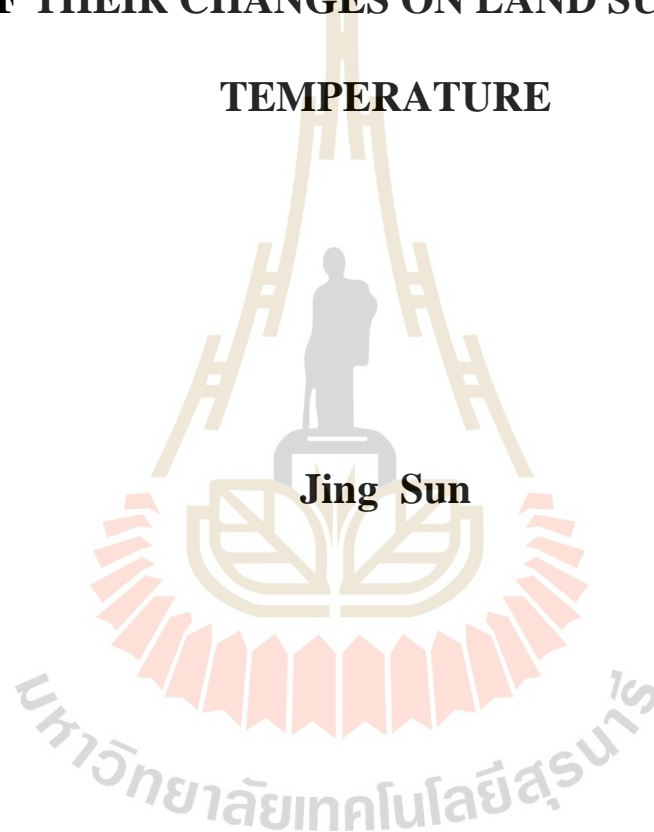


**NEWLY DEVELOPED METHOD OF HARMONIC
ANALYSIS FOR MULTITEMPORAL LAND USE AND
LAND COVER CLASSIFICATION AND IMPACT STUDY
OF THEIR CHANGES ON LAND SURFACE
TEMPERATURE**



**A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Geoinformatics**

Suranaree University of Technology

Academic Year 2019

วิธีการวิเคราะห์แบบฮาร์โมนิกที่พัฒนาขึ้นใหม่สำหรับการจำแนกการใช้ประโยชน์
ที่ดินและสิ่งปกคลุมดินแบบหลายช่วงเวลาและการศึกษาผลกระทบของการ
เปลี่ยนแปลงการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินต่ออุณหภูมิพื้นผิวดิน

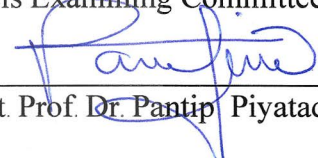


วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาภูมิสารสนเทศ
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2562


**NEWLY DEVELOPED METHOD OF HARMONIC ANALYSIS
FOR MULTITEMPORAL LAND USE AND LAND COVER
CLASSIFICATION AND IMPACT STUDY OF THEIR CHANGES
ON LAND SURFACE TEMPERATURE**

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.


Thesis Examining Committee


(Asst. Prof. Dr. Pantip Piyatadsananon)

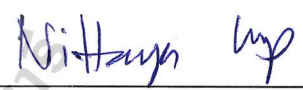
Chairperson


(Assoc. Prof. Dr. Suwit Ongsomwang)

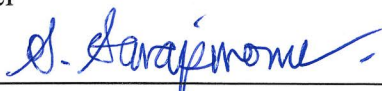
Member (Thesis Advisor)


(Assoc. Prof. Dr. Songkot Dasananda)


Member

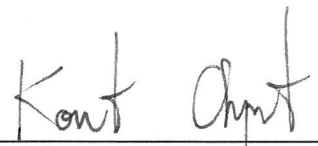

(Assoc. Prof. Dr. Nittaya Kerdprasop)

Member


(Asst. Prof. Dr. Sunya Sarapirome)

Member


(Assoc. Prof. Dr. Worawat Meevasana)


(Assoc. Prof. Flt. Lt. Dr. Kontorn Chamniprasart)

Vice Rector for Academic Affairs
and Internationalization

Dean of Institute of Science

จึง ชวน : วิธีการวิเคราะห์แบบฮาร์โมนิกที่พัฒนาขึ้นมาใหม่สำหรับการจำแนกการใช้
ประโยชน์ที่ดินและสิ่งปกคลุมดินแบบหลายช่วงเวลาและการศึกษาผลกระทบของการ
เปลี่ยนแปลงการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินต่ออุณหภูมิพื้นผิวดิน (NEWLY
DEVELOPED METHOD OF HARMONIC ANALYSIS FOR MULTITEMPORAL
LAND USE AND LAND COVER CLASSIFICATION AND IMPACT STUDY OF
THEIR CHANGES ON LAND SURFACE TEMPERATURE) อาจารย์ที่ปรึกษา :
รองศาสตราจารย์ ดร.สุวิทย์ อ่องสมหวัง, 287 หน้า.

ความเข้าใจในข้อมูลการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินและการเปลี่ยนแปลงที่เกิดขึ้น
ทั้งในอดีตและปัจจุบันเป็นเรื่องสำคัญสำหรับนักวางผังเมือง นักจัดการที่ดิน และนักจัดการ
ทรัพยากร ในขณะที่อุณหภูมิพื้นผิวดินเป็นพารามิเตอร์สำคัญของกระบวนการพื้นผิวโลก เนื่องจาก
อุณหภูมิพื้นผิวดินมีใช้เพียงแต่เป็นตัวชี้วัดของการเปลี่ยนแปลงสภาพภูมิอากาศ แต่ยังเป็นตัวบ่งชี้
ของการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินจากการขยายตัวของเมือง ฉะนั้น เพื่อ
จัดการกับสถานการณ์เหล่านี้ การจำแนกการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินแบบพลวัตและ
การศึกษาผลกระทบการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินต่ออุณหภูมิพื้นผิวดิน
จึงเป็นสิ่งสำคัญ วัตถุประสงค์ของการศึกษาคือ (1) เพื่อประยุกต์ขั้นตอนวิธีการจำแนกแบบไม่กำกับ
ดูแลสำหรับการจำแนกการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินแบบหลายช่วงเวลาจากชุดข้อมูล
แลนด์แซทแบบอนุกรมเวลา (2) เพื่อพัฒนาวิธีการจำแนกและตรวจสอบการเปลี่ยนแปลงการใช้
ประโยชน์ที่ดินและสิ่งปกคลุมดินแบบกำกับดูแลแบบใกล้เคียงเวลาปัจจุบัน โดยอาศัยแบบจำลอง
อนุกรมเวลากับขั้นตอนวิธีการวิเคราะห์ทางเชิงสเปกตรัมน้อยสุด และ (3) เพื่อสร้างชุดข้อมูล
อุณหภูมิพื้นผิวดินและตรวจสอบผลกระทบการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินต่ออุณหภูมิ
พื้นผิวดิน วิธีการศึกษาแบ่งออกเป็น 5 ส่วน ประกอบด้วย (1) การรวบรวมข้อมูลและการ
ประมวลผลชุดข้อมูลแลนด์แซทแบบอนุกรมเวลา (2) การคัดเลือกรูปลักษณะเชิงสเปกตรัมที่
เหมาะสมสำหรับการจำแนกการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดิน (3) การจำแนกการใช้
ประโยชน์ที่ดินและสิ่งปกคลุมดินแบบอนุกรมเวลาด้วยวิธีการแบบไม่กำกับดูแล (4) การจำแนกการ
ใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินแบบอนุกรมเวลาด้วยวิธีการแบบกำกับดูแล และ (5) การ
วิเคราะห์ผลกระทบของการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินต่ออุณหภูมิพื้นผิว
ดิน

ผลการศึกษาของวัตถุประสงค์ที่ 1 พบว่า ความถูกต้องของแผนที่การใช้ประโยชน์ที่ดินและ
สิ่งปกคลุมดินที่ได้รับจากขั้นตอนวิธี ISODATA สูงกว่า ขั้นตอนวิธี K-Means เล็กน้อย และ

รูปลักษณะเชิงสเปกตรัมหลายประเภทสามารถให้ความความถูกต้องมากกว่ารูปลักษณะเชิงสเปกตรัมประเภทเดียว สำหรับผลการศึกษาของวัตถุประสงค์ที่ 2 พบว่า ความถูกต้องโดยรวมเฉลี่ยของการจำแนกการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินในปี พ.ศ. 2543 2549 2554 และ 2560 มีค่าเท่ากับร้อยละ 86.60 81.53 84.55 และ 81.80 ตามลำดับ นอกจากนี้ ความถูกต้องโดยรวมเฉลี่ยที่ใช้รูปลักษณะสเปกตรัมประเภทเดียวและหลายประเภท มีค่าเท่ากับร้อยละ 67.77 84.59 90.80 และ 91.31 ตามลำดับ สำหรับผลการศึกษาของวัตถุประสงค์ที่ 3 ในการพิจารณาข้อมูลตามขอบเขตเชิงเวลา (time domain) พบว่า ค่า intercept และ amplitude ของอนุกรมพื้นที่ผิวดินจะเพิ่มขึ้นทันที เมื่อการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินประเภทอื่น ๆ เปลี่ยนเป็นพื้นที่เมืองและสิ่งปลูกสร้าง และสิ่งนี้บ่งชี้ว่าอนุกรมเฉลี่ยและการเปลี่ยนแปลงของอนุกรมรายปีในพื้นที่เมืองและสิ่งปลูกสร้างสูงกว่าการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินประเภทอื่น ๆ ขณะเดียวกัน ในการพิจารณาข้อมูลตามขอบเขตเชิงพื้นที่ (spatial domain) พบว่า อนุกรมพื้นที่ผิวดินเพิ่มขึ้นเมื่อพื้นที่เกษตรกรรม พื้นที่ป่า และแหล่งน้ำเปลี่ยนเป็นพื้นที่เมืองและสิ่งปลูกสร้าง ปรากฏการณ์นี้บ่งชี้ผลกระทบของการเปลี่ยนแปลงการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินต่ออนุกรมพื้นที่ผิวดิน เมื่อพื้นที่สีเขียว เช่น พื้นที่เกษตรกรรมหรือพื้นที่ป่าเปลี่ยนเป็นพื้นที่เมืองและสิ่งปลูกสร้าง จากผลการศึกษาสามารถสรุปได้ว่า วิธีการจำแนกแบบกำกับดูแลที่พัฒนาขึ้นใหม่โดยใช้การวิเคราะห์ฮาร์โมนิกกับขั้นตอนวิธีการวัดระยะห่างเชิงสเปกตรัมน้อยที่สุดสามารถนำมาใช้จำแนกข้อมูลการใช้ประโยชน์ที่ดินและสิ่งปกคลุมดินได้อย่างมีประสิทธิภาพ

JING SUN : NEWLY DEVELOPED METHOD OF HARMONIC
ANALYSIS FOR MULTITEMPORAL LAND USE AND LAND COVER
CLASSIFICATION AND IMPACT STUDY OF THEIR CHANGES ON
LAND SURFACE TEMPERATURE. THESIS ADVISOR : ASSOC. PROF.
SUWIT ONGSOMWANG, Dr. rer. Nat. 287 PP.

HARMONIC ANALYSIS/ MULTITEMPORAL LAND USE AND LAND COVER
CLASSIFICATION/ LAND SURFACE TEMPERATURE/ NANJING CITY/ CHINA

An understanding of historical and present land use and land cover (LULC) information and its changes is critical for city planners, land managers, and resource managers. Meanwhile, land surface temperature (LST) is also a crucial parameter in land surface processes, not only acting as an indicator of climate change but also a pointer of LULC changes due to urbanization. To deal with these situations, dynamic LULC classification and impact study of land use and land cover change on land surface temperature are necessary. The specific research objectives are (1) to apply unsupervised classification algorithms for multitemporal LULC classification based on time-series Landsat datasets, (2) to develop a nearly real-time supervised LULC classification and change detection method using a time-series model and a minimum spectral distance algorithm, and (3) to reconstruct time-series LST dataset and investigate the impact of LULC on LST. The research methodology consisted of five main components: (1) data collection and preprocessing of time-series Landsat datasets, (2) optimum spectral features selection for LULC classification, (3) time-series LULC

classification using unsupervised method, (4) time-series LULC classification using supervised method, and (5) analyze the impact of LULC change on LST.

Significant results of the first objective revealed that the accuracy of LULC maps derived by the ISODATA algorithm was slightly higher than the K-MEANS algorithm, and the multiple spectral features could provide higher accuracy than a single spectral feature. For the second objective, the average overall accuracy of LULC classification in 2000, 2006, 2011, and 2017 using a newly developed supervised classification method were 86.60%, 81.53%, 84.55%, and 81.80%, respectively. Besides, the average overall accuracy of single and three multiple spectral features were 67.77%, 84.59%, 90.80%, and 91.31%, respectively. For the third objective, by applying the time-domain approach, intercept and amplitude values of LST suddenly increased, when other types of LULC were converted to urban and built land, and it indicates that the average temperature and annual temperature change in urban areas were higher than in other non-urban areas. Meanwhile, by applying the spatial-domain approach, the LST value increased, when agricultural land, forest land, and water bodies were converted into urban and built-up land. This phenomenon indicates the impact of LULC change on LST when the green area, such as agricultural land or forest land, are changed to be urban and built-up land. Consequently, it can be concluded that the newly developed supervised classification method using harmonic analysis with a minimum spectral distance algorithm can be efficiently applied to classify multitemporal LULC data.

School of Geoinformatics

Academic Year 2019

Student's Signature Jing Sun

Advisor's Signature Savit Ang

ACKNOWLEDGEMENTS

I would like to recognize all of those who have given me strength, support, and the opportunity to show that it has never been too late to study.

First and foremost, my special thanks are extended to my supervisor, Assoc. Prof. Dr. Suwit Ongsomwang, for the great support he has given me in completing this study. It has been through his patience, inspiration, continual encouragement, and professional experience that I was able to step on the amazing journey of research and discover my field of interest. He has devoted his time and energy to reading, revising, and polishing my drafts. Moreover, we have spent long hours together discussing the project and were excited about the surprising and promising findings. These precious moments will be cherished in the deepest part of my heart and never fade from memory.

Heartfelt thanks are also expressed to my thesis examination committee: Asst. Prof. Dr. Pantip Piyatadsananon (who served as chairperson), Asst. Prof. Dr. Sunya Sarapirome, Assoc. Prof. Dr. Nittaya Kerdprasop, and Assoc. Prof. Dr. Songkot Dasananda for their valuable suggestions.

I am grateful to the Suranaree University of Technology for awarding me a scholarship to pursue this Ph.D. degree in Geoinformatics at the School of Geoinformatics, Institute of Science, Suranaree University of Technology.

Many thanks are given to my home university, Tongling University, for allowing me this opportunity to engage in professional development. It was a wonderful opportunity for me to upgrade my career in the field of geoinformatics. Thanks also go

to my colleagues in the School of Architectural Engineering for shouldering my job obligations during my absence.

Most of all, I wish to extend my countless thanks to my parents, Mr. Zhaogen Sun and Mrs. Xinxiang Liu. Over the years, their selfless love and wholehearted assistance have always been there and propelled me forward. I will do my utmost to be the pride in their eyes.

To Jing He, my wife, my lifetime companion, without her sincere support, I would not have been able to complete my studies overseas at SUT in Thailand. Words of thanks are also uttered to my son, Yinbo Sun. His merry laughter, childlike innocence, and imaginative ingenuity have always been and will continue to be the inspiration and sunshine of my life.

Finally, I would like to thank my fellow lab mates of School of Geoinformatics for their friendship, guidance, and assistance rendered, and for the times we worked and enjoyed together during my study.



Jing Sun

CONTENTS

	Page
ABSTRACT IN THAI.....	I
ABSTRACT IN ENGLISH	III
ACKNOWLEDGEMENTS	V
CONTENTS.....	VII
LIST OF TABLES	XIII
LIST OF FIGURES	XVIII
LIST OF ABBREVIATIONS.....	XXIX
CHAPTER	
I INTRODUCTION	1
1.1 Background problem and significance of the study	1
1.2 Research objectives.....	5
1.3 Scope of the study	6
1.4 Limitation of the study.....	7
1.5 Study area.....	7
1.6 Benefit of the study	8
II BASIC CONCEPTS AND LITERATURE REVIEWS	10
2.1 Single channel algorithm for LST extraction.....	10
2.2.1 Development background	10

CONTENTS (Continued)

	Page
2.1.2 Single channel algorithm.....	14
2.1.3 Emissivity extraction.....	15
2.2 Time-series of harmonic analysis	17
2.3 Time-series change detection and classification model.....	21
2.3.1 Continuous Change Detection.....	23
2.3.2 Continuous land cover classification	26
2.4 Features selection for LULC change detection and classification.....	27
2.4.1 Band type	27
2.4.2 Index type.....	28
2.4.3 Transformation type.....	30
III RESEARCH METHODOLOGY.....	33
3.1 Data collection and preprocessing of time-series Landsat datasets	35
3.2 Optimum spectral features selection for LULC classification.....	37
3.3 Time-series LULC classification using unsupervised method	38
3.4 Time-series LULC classification using supervised method	38
3.5 Impact study of LULC change on LST.....	39
IV DATA COLLECTION AND PREPROCESSING OF TIME-SERIES	
LANDSAT DATASETS.....	40
4.1 Cloud cover assessment	42

CONTENTS (Continued)

	Page
4.2 Cloud/shadow pixel recognition base on QA band and HA model	44
4.2.1 Cloud/shadow pixel recognition base on QA band.....	44
4.2.2 Cloud/shadow pixel recognition base on HA model.....	45
4.3 Conversion of digital number to spectral reflectance data.....	47
4.4 Time-series spectral reflectance reconstruction	48
V SPECTRAL FEATURE SELECTION FOR LAND USE AND LAND	
COVER CLASSIFICATION	55
5.1 Stable pixel of LULC type extraction.....	56
5.2 Selection of spectral feature.....	62
5.2.1 Spectral feature selection by the fitting effect of harmonic wave	63
5.2.2 Ability of features to distinguish a specific LULC type	73
5.2.3 Optimum spectral features selection	79
VI TIME-SERIES LAND USE AND LAND COVER CLASSIFICATION	
USING AN UNSUPERVISED METHOD	81
6.1 LULC change detection using HA and CCDC	82
6.1.1 Single feature: NIR band.....	84
6.1.2 Multiple features: TOP3 spectral features.....	86
6.1.3 Multiple features: TOP6 and ORI6 spectral features	90

CONTENTS (Continued)

	Page
6.2 Unsupervised LULC classification using K-Means and ISODATA algorithm	99
6.2.1 Suitable coefficient feature for LULC classification	101
6.2.2 LULC classification by K-Means algorithm.....	108
6.2.3 LULC classification by ISODATA algorithm	117
6.3 Accuracy assessment.....	126
6.3.1 Accuracy assessment for change detection	130
6.3.2 Accuracy assessment for LULC classification	141
VII TIME-SERIES LAND USE AND LAND COVER CLASSIFICATION USING A SUPERVISED METHOD	157
7.1 Harmonic function curve transformation and standard harmonic curve construction	158
7.1.1 Spectral harmonic function curve of LULC types	159
7.1.2 Standard harmonic function curve of each LULC type	164
7.2 Spectral distance measurement and probability calculation	169
7.2.1 Spectral distance measurement	171
7.2.2 Probability of an unclassified pixel of being specific LULC type.....	175
7.3 Time-series LULC classification	179

CONTENTS (Continued)

	Page
7.3.1 Average probability of an unclassified pixel being a specific LULC type	180
7.3.2 Supervised LULC classification after post-classification processing.....	183
7.3.3 Result of supervised LULC classification.....	185
7.4 Accuracy assessment.....	193
7.4.1 Accuracy assessment for change detection	193
7.4.2 Accuracy assessment for LULC classification	198
VIII IMPACT STUDY OF LAND USE AND LAND COVER CHANGE ON	
LAND SURFACE TEMPERATURE.....	212
8.1 Time-series LST retrieve using Single Channel algorithm.....	214
8.1.1 Brightness temperature.....	217
8.1.2 Emissivity extraction.....	219
8.1.3 Water vapor content	221
8.2 Time-series LST reconstruction.....	223
8.2.1 Construction of spatiotemporal cube for correct and incorrect LST value	223
8.2.2 Simulated LST recalculation using the HA model.....	229
8.2.3 Time-series LST reconstruction.....	234

CONTENTS (Continued)

	Page
8.2.4 Verification of simulated LST data.....	238
8.3 LST decomposition analysis and impact of LULC on LST.....	240
8.3.1 LST decomposition.....	242
8.3.2 Impact of LULC on LST.....	246
8.4 Spatial analysis and impact of LULC change on LST.....	257
IX CONCLUSION AND RECOMMENDATION.....	262
9.1 Conclusions.....	263
9.1.1 Time-series land use and land cover classification using an unsupervised method.....	263
9.1.2 Time-series land use and land cover classification using a supervised method.....	264
9.1.3 Impact study of land use and land cover change on land surface temperature.....	265
9.2 Recommendations.....	265
REFERENCES	269
CURRICULUM VITAE.....	287

LIST OF TABLES

Table	Page
1.1 Urban population, urban and built-up areas, and gross regional product of Nanjing city from 2000 to 2016.	2
3.1 Characteristics of Landsat 5/7/8 data.	36
4.1 Common pixel values of the QA band of Landsat 5, 7 data.....	44
4.2 Common pixel values of the QA band of Landsat 8 data.....	45
5.1 Area and percentage of stable LULC type and LULC change between 2000 and 2017.	60
5.2 Comparison of mean R^2 value of urban and built-up land from various selected features.....	65
5.3 Comparison of mean R^2 value of agricultural land from various selected features.	66
5.4 Comparison of mean R^2 value of forest land from various selected features.	67
5.5 Comparison of mean R^2 value of water bodies from various selected features.	68
5.6 R^2 significant test (T-test) of three CASEs (significance level of 0.05).	71
5.7 R^2 significant test (Z-test) of three CASEs (significance level of 0.05).	72
5.8 Average R^2 value of four LULC types.	73
5.9 Area of overlapping of any two LULC types.	76

LIST OF TABLES (Continued)

Table	Page
5.10 Distinguishability of any two LULC types.	77
5.11 The ability of the spectral features to distinguish a certain LULC.....	78
5.12 Optimal spectral feature for LULC classification.	80
6.1 Four coefficients of the NIR band at point X location in different periods.	85
6.2 Coefficients of all pixels from the NIR band in the study area.	86
6.3 Four coefficients of TOP3 spectral features in different periods.	88
6.4 Coefficients of all pixels from TOP3 spectral features in the study area.	89
6.5 Four coefficients of TOP6 spectral features in different periods.	92
6.6 Coefficients of all pixels from TOP6 spectral features in the study area.	93
6.7 Four coefficients of ORI6 spectral bands in different periods.	96
6.8 Coefficients of all pixels from ORI6 spectral bands in the study area.	97
6.9 Basic information of the selected spectral feature(s).	100
6.10 Possibility of coefficient feature combination for the most suitable coefficient identification.....	101
6.11 Area and percentage of LULC classification in 2000.	113
6.12 Area and percentage of LULC classification in 2006.	113
6.13 Area and percentage of LULC classification in 2011.	113
6.14 Area and percentage of LULC classification in 2017.	113
6.15 Area and percentage of LULC classification in 2000.	122
6.16 Area and percentage of LULC classification in 2006.	122

LIST OF TABLES (Continued)

Table	Page
6.17 Area and percentage of LULC classification in 2011.	122
6.18 Area and percentage of LULC classification in 2017.	122
6.19 The accuracy assessment of change detection using the NIR feature.	134
6.20 The accuracy assessment of change detection using TOP3 spectral features.	134
6.21 The accuracy assessment of change detection using TOP6 spectral features.	134
6.22 The accuracy assessment of change detection using ORI6 spectral features.	135
6.23 Cloud cover of Landsat in the study area.	139
6.24 Error matrix and accuracy assessment of LULC Classification in 2000.	142
6.25 Error matrix and accuracy assessment of LULC Classification in 2006.	143
6.26 Error matrix and accuracy assessment of LULC Classification in 2011.	144
6.27 Error matrix and accuracy assessment of LULC Classification in 2017.	145
6.28 Average producer's accuracy and user's accuracy using K-Means algorithms.	149
6.29 Average producer's accuracy and average's user accuracy using ISODATA algorithms.	150
6.30 Pairwise Z-test between the K-means and ISODATA algorithm.	153

LIST OF TABLES (Continued)

Table	Page
6.31 Pairwise Z-test between different spectral features using the K-means algorithm.....	153
6.32 Pairwise Z-test between different spectral features using the ISODATA algorithm.....	154
7.1 Median of coefficients for four LULC types.....	165
7.2 Area and percentage of supervised LULC classification in 2000.	190
7.3 Area and percentage of supervised LULC classification in 2006.	190
7.4 Area and percentage of supervised LULC classification in 2011.	190
7.5 Area and percentage of supervised LULC classification in 2017.	190
7.6 The accuracy assessment of change detection using the NIR feature.	195
7.7 The accuracy assessment of change detection using TOP3 spectral features.	195
7.8 The accuracy assessment of change detection using TOP6 spectral features.	195
7.9 The accuracy assessment of change detection using ORI6 spectral features.	196
7.10 Accuracy assessment of LULC classification in 2000.	199
7.11 Accuracy assessment of LULC Classification in 2006.	200
7.12 Accuracy assessment of LULC classification in 2011.	201
7.13 Accuracy assessment of LULC classification in 2017.	202

LIST OF TABLES (Continued)

Table	Page
7.14 Average producer's accuracy and average user's accuracy using supervised classification.....	205
7.15 Pairwise Z-test between different spectral features.....	208
7.16 Comparison of accuracy between unsupervised (ISODATA) and developed supervised classification methods.....	211
8.1 The mean and standard deviation of LST data in different processing steps.....	228
8.2 The mean and SD of LST data.....	237
8.3 Final result of harmonic parameters of some pixels from all periods.....	246
8.4 Scale of the correlation coefficient.....	250
8.5 The time-weighted average of different persistent time of four harmonic terms of four different LULC types.....	251
8.6 Linear regression of persistent years and values of four harmonic terms.....	254
8.7 Thermal signatures and thermal signatures change of the LULC change area.....	256
8.8 LST change of the LULC change area.....	259

LIST OF FIGURES

Figure	Page
1.1 Location map of the study area.	8
2.1 One-dimensional Fourier analysis.....	18
2.2 Using the threshold of three times RMSE for continuous change detection.....	25
3.1 Overview of research methodology framework.....	34
4.1 Workflow of Component 1: Data collection and preprocessing of time-series Landsat datasets.....	41
4.2 The distribution of all Landsat images from 2000 to 2017.	43
4.3 Percent of total images and percent of total clear pixels based on all available Landsat 5/7/8 images between 2000 and 2017.	43
4.4 Examples of clearly observed and contaminated pixels in the study area: (a) Landsat 5, date 20 January 2000, and (b) Landsat 7, date 4 May 2009.	46
4.5 Spatial distribution of spectral reflectance composite images: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.....	48
4.6 Time-series spectral reflectance reconstruction.	49
4.7 Comparison of (a) original time-series spectral reflectance data of six bands from one pixel between 2000 and 2017 and (b) time-series spectral reflectance data of six bands from the same pixel after removal of contaminated values.	51

LIST OF FIGURES (Continued)

Figure	Page
4.8 Examples of the derived spectral indices (TCB): (a) one scene of the whole study area, date 20 January 2000, and (b) one pixel from 2000 to 2017.....	53
5.1 Workflow of Component 2: Spectral features selection for LULC classification.	55
5.2 Example of stable history pixel selected from Jiangning district in different 4 years: (a) location of different stable LULC types and (b) zoom in maps.	57
5.3 Spatial distribution of LULC classification map in: (a) 2000, (b) 2006, (c) 2011, and (d) 2017.....	59
5.4 Stable and change of LULC type between 2000 and 2017.	60
5.5 Sample points of the historical stable pixel of each LULC type.	61
5.6 An example of HA for three cases of urban and built-up land by TCB data.	63
5.7 An example of HA for three cases of agricultural land by TCB data.	64
5.8 An example of HA for three cases of forest land by TCB data.....	64
5.9 An example of HA for three cases of water bodies by TCB data.	64
5.10 Distribution of R^2 of agricultural land by PC1 data for Case 1.....	69
5.11 Distribution of R^2 of agricultural land by PC1 data for Case 2.....	69
5.12 Distribution of R^2 of agricultural land by PC1 data for Case 3.....	70
5.13 Probability of four LULC types on the blue band.....	74
5.14 Probability density function of four LULC types on the blue band.	75

LIST OF FIGURES (Continued)

Figure	Page
6.1 Workflow of Component 3: Time-series LULC classification using unsupervised classification method.....	82
6.2 The geographical location of the Jiangning District and the location of point X (Row 1443, column 1174): (a) Image from Google Earth in 2000 and (b) Image from Google Earth in 2017.....	83
6.3 Time-series model of NIR band for point X.	84
6.4 Time-series model of TOP3 spectral features for point X.	87
6.5 Time-series model of TOP6 spectral features for point X.	91
6.6 Time-series model of ORI6 spectral bands for point X.	94
6.7 LULC classification map by K-Means algorithm with 15 coefficient features of NIR band.	102
6.8 LULC classification map by ISODATA algorithm with 15 coefficient features of the NIR band.....	105
6.9 LULC classification map by K-Means with amplitude coefficient of NIR band: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) 18 May 2017.	109
6.10 LULC classification map by K-Means with amplitude coefficients of TOP3 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) 18 May 2017.....	110

LIST OF FIGURES (Continued)

Figure	Page
6.11 LULC classification map by K-Means with amplitude coefficients of TOP6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) 18 May 2017.....	111
6.12 LULC classification map by K-Means with amplitude coefficients of ORI6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) 18 May 2017.....	112
6.13 Comparison of the proportional LULC area in 2000 using the K-Means algorithm with various spectral features.....	114
6.14 Comparison of the proportional LULC area in 2006 using the K-Means algorithm with various spectral features.....	114
6.15 Comparison of the proportional LULC area in 2011 using the K-Means algorithm with various spectral features.....	115
6.16 Comparison of the proportional LULC area in 2017 using the K-Means algorithm with various spectral features.....	115
6.17 LULC classification map by ISODATA with amplitude coefficient of NIR band: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.....	118
6.18 LULC classification map by ISODATA with amplitude coefficients of TOP3 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.....	119

LIST OF FIGURES (Continued)

Figure	Page
6.19 LULC classification map by ISODATA with amplitude coefficients of TOP6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.....	120
6.20 LULC classification map by ISODATA with amplitude coefficients of ORI6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.....	121
6.21 Comparison of the proportional LULC area in 2000 using the ISODATA algorithm with various spectral features.....	123
6.22 Comparison of the proportional LULC area in 2006 using the ISODATA algorithm with various spectral features.....	123
6.23 Comparison of the proportional LULC area in 2011 using the ISODATA algorithm with various spectral features.....	124
6.24 Comparison of the proportional LULC area in 2017 using the ISODATA algorithm with various spectral features.....	124
6.25 Pan-sharpened image from Landsat data in 4 different years: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.....	128
6.26 Images from Google Earth in 4 different years: (a) Date 2000, (b) Date 2006, (c) Date 2011, and (d) Date 2017.	129

LIST OF FIGURES (Continued)

Figure	Page
6.27 LULC change detection by the K-Means algorithm: (a) Using the NIR band, (b) Using TOP3 spectral features, (c) Using TOP6 spectral features, and (d) Using ORI6 spectral features.	131
6.28 LULC change detection by the ISODATA algorithm: (a) Using the NIR band, (b) Using TOP3 spectral features, (c) Using TOP6 spectral features, and (d) Using ORI6 spectral features.	132
6.29 Comparison of the overall accuracy of the two algorithms.	146
6.30 Comparison of the overall accuracy of four spectral features.	147
6.31 Comparison of the overall accuracy of four years.	148
6.32 Producer's accuracy and user's accuracy based on the K-Means algorithm.	151
6.33 Producer's accuracy and user's accuracy based on the ISODATA.	151
7.1 Workflow of Component 4: Time-series LULC classification using supervised method.	158
7.2 Harmonic function curves of four LULC type.	161
7.3 Standard reference harmonic waves of four LULC type.	166
7.4 Spectral distance between the standard harmonic function curves and polyline of the unclassified pixels between 2000 and 2017.	172
7.5 Probability of an unclassified pixel for each LULC type between 2000 and 2017 in each spectral feature.	176

LIST OF FIGURES (Continued)

Figure	Page
7.6 Average probability of the unclassified pixel for each LULC type between 2000 and 2017 from TOP3, TOP6, and ORI6 features.....	182
7.7 Multitemporal LULC classification of one unclassified pixel between 2000 and 2017: (a) before post-classification and (b) after post-classification.....	184
7.8 LULC classification map by the supervised algorithm with NIR band: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.....	185
7.9 LULC classification map by the supervised algorithm with TOP3 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.....	186
7.10 LULC classification map by the supervised algorithm with TOP6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.....	187
7.11 LULC classification map by the supervised algorithm with ORI6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.....	188
7.12 Comparison of the proportional LULC area in 2000 using supervised classification with various spectral features.....	191
7.13 Comparison of the proportional LULC area in 2006 using supervised classification with various spectral features.....	191

LIST OF FIGURES (Continued)

Figure	Page
7.14 Comparison of the proportional LULC area in 2011 using supervised classification with various spectral features.	192
7.15 Comparison of the proportional LULC area in 2017 using supervised classification with various spectral features.	192
7.16 LULC change detection by the supervised algorithm: (a) Using NIR band, (b) Using TOP3 spectral features, (c) Using TOP6 spectral features, and (d) Using ORI6 spectral features.....	194
7.17 Comparison of the overall accuracy of four spectral features.....	203
7.18 Comparison of the overall accuracy of four years.....	203
7.19 Producer's accuracy and user's accuracy.....	206
8.1 Workflow of Component 5: Time-series LST retrieves and time-series LST reconstruction.	213
8.2 Workflow of time-series LST retrieves using a single channel algorithm.....	215
8.3 Spatial distribution of spectral radiance data: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.....	218
8.4 Spatial distribution of brightness temperature data: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.....	218
8.5 Spatial distribution of NDVI data: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.	220

LIST OF FIGURES (Continued)

Figure	Page
8.6 Spatial distribution of emissivity data: (a) 3 May 2000 and (b) 4 May 2009.	221
8.7 Spatial distribution of water vapor content from the MODIS product: (a) Date 3 May 2000 and (b) Date 4 May 2009.	222
8.8 Spatial distribution of land surface temperature data: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.	222
8.9 Workflow of Section 8.2.1: Identification of correct and incorrect LST values.	226
8.10 Histogram distribution of LST values in different stages of processing: (a) original LST values by SC algorithm, (b) after applying data error cube, (c) after applying algorithm cube, and (d) after applying LST outlier cube.	228
8.11 Workflow of Section 8.2.2: Simulated LST recalculation using the HA model.	230
8.12 Main calculation process and results of one row of data.	232
8.13 Workflow of Section 8.2.3: Time-series LST reconstruction.	234
8.14 Schematic diagram of LST reconstruction based on the LST indicator, the original LST, and the simulated LST.	235
8.15 Reconstructed LST data.	236

LIST OF FIGURES (Continued)

Figure	Page
8.16 Reconstruction of LST data of Landsat 7, date 30 October 2010: (a) LST from original LST, (b) LST from simulated LST, and (c) LST from reconstructed LST.	237
8.17 Schematic diagram of LST reconstruction accuracy verification.	239
8.18 Workflow of Section 8.3: LST decomposition analysis and impact study of LULC on LST.....	241
8.19 LST decomposition procedure using time-series LST at one point: (a) Original and reconstructed LST, (b) Trend component, (c) Seasonality component, and (d) Residue.	244
8.20 Distribution of frequency and probability of intercept, slope, amplitude, phase for four different LULC types: (a) original histogram distribution of harmonic term and (b) normalization of histogram distribution of harmonic term.....	248
8.21 Comparison of four harmonic terms between four different LULC types.....	249
8.22 The relationship between persistent time and the value of four harmonic terms for four different LULC types: (a) intercept, (b) slope, (c) amplitude, and (d) phase.....	252
8.23 LULC map of four years: (a) date 30 July 2000, (b) date 31 July 2006, (c) date 29 July 2011, (d) date 29 July 2017.....	258

LIST OF FIGURES (Continued)

Figure	Page
8.24 LST map of four years: (a) date 30 July 2000, (b) date 31 July 2006, (c) date 29 July 2011, (d) date 29 July 2017.....	259



LIST OF ABBREVIATIONS

CCDC	=	Continuous Change Detection and Classification
EVI	=	Enhanced Vegetation Index
HA	=	Harmonic Analysis
LST	=	Land Surface Temperature
LULC	=	Land Use and Land Cover
MLC	=	Maximum Likelihood Classifier
MNDBI	=	Modified Normalized Difference Built-up Index
MNDWI	=	Modified Normalized Difference Water Index
NDBI	=	Normalized Difference Built-up Index
NDVI	=	Normalized Difference Vegetation Index
NDWI	=	Normalized Difference Water Index
PCA	=	Principal Component Analysis
QA	=	Quality Assessment
SC	=	Single Channel
TCB	=	Tasseled Cap Brightness
TCG	=	Tasseled Cap Greenness
TCW	=	Tasseled Cap Wetness
TIR	=	Thermal Infrared

CHAPTER I

INTRODUCTION

1.1 Background problem and significance of the study

Land use and land cover (LULC) relates to the biophysical cover of the Earth's terrestrial surface, identifying vegetation, inland water, bare soil, or human infrastructure. Distinct LULC types provide specific habitats and determine the exchange of energy and carbon between terrestrial and atmospheric regions (DeVries, 2015; Houghton et al., 2012; Running et al., 1999). Additionally, LULC is an essential element for description and studying of the environment (Herold, Latham, Di Gregorio, and Schmullius, 2006) and is a crucial input to model ecosystem services (Andrew, Wulder, and Nelson, 2014).

In the meantime, urban growth induces the replacement of natural land covers with the impervious urban materials, the modifications of the biophysical environment, and the alterations of the land surface energy processes (Lo and Quattrochi, 2003). As reports of Ministry of Housing and Urban-Rural Development of the People's Republic of China (MOHURD, 2017) and Nanjing Municipal Bureau Statistics (NJMBS, 2017) in 2017, the gross regional product of the city during the last two decades had continuously increased from 107.35 billion Yuan in 2000 to 1,050.30 billion Yuan in 2016. The urban population of the city had increased from 3.10 million persons in 2000 to 5.91 million persons in 2016; these phenomena induce LULC change in the city.

Urban and built-up areas had been increased from 201.40 km² in 2000 to 773.79 km² in 2016 (Table 1.1).

Table 1.1 Urban population, urban and built-up areas, and gross regional product of Nanjing city from 2000 to 2016.

Year	Urban Population (10,000 Persons)	Urban and built-up areas (Km ²)	Gross Regional Product (Billion Yuan)
2000	309.52	201.40	107.35
2001	323.86	211.80	121.85
2002	339.35	438.63	138.51
2003	372.39	446.79	169.08
2004	394.80	484.27	206.72
2005	410.54	512.60	245.19
2006	431.32	574.94	282.28
2007	466.75	577.44	334.01
2008	478.16	592.07	381.46
2009	482.60	598.14	423.03
2010	484.87	618.64	513.07
2011	503.76	637.71	614.55
2012	537.27	653.31	720.16
2013	567.11	713.29	808.02
2014	572.65	734.34	882.08
2015	581.65	755.27	972.08
2016	590.60	773.79	1,050.3

Further, changes in land surface conditions (e.g., albedo, soil moisture, land surface temperature, and surface roughness) and atmospheric composition (e.g., CO₂ and methane) have had significant effects on regional and global climates due to LULC change (Bonan, Pollard, and Thompson, 1992; Foley et al., 2005; Lee, He, Zhou, and Liang, 2015; Mahmood et al., 2014; McPherson, 2007; Pielke, 2005). Thus, many researchers attempt to observe and model the effects of LULC change on climate based

on Earth observation satellites (Douglas et al., 2006; Kaufmann and Stern, 1997; Lawrence and Chase, 2010; Lawrence et al., 2012; Lee et al., 2009; Lee et al., 2015; Lee, Sacks, Chase, and Foley, 2011). Both observation and modeling studies require LULC maps for characterizing the pattern of LULC changing over time.

Remote sensing data have always been an essential source for LULC and change detection because of their macroscopic and real-time nature. In the previous studies, maps that relate change between two or several specific year dates typically lack information regarding underlying processes and are not enable insights on the nature of the transformations present, such as rate or persistence of change (Gillanders, Coops, Wulder, Gergel, and Nelson, 2008). Ideally, both subtle modifications and rapid changes should be accounted for by LULC change studies (Lambin, Geist, and Lepers, 2003). A time-series of LULC maps can capture the complexities of Earth's changing surface (Liu and Cai, 2012; Sexton, Urban, Donohue, and Song, 2013) and can be used to parameterize biogeochemical models (Feddes, 2005; Running, Loveland, and Pierce, 1994). Therefore, obtaining a continuous sequence of LULC maps over an extended time, for at least multiple decades, is critical for quantifying the effects of LULC change on climate.

Besides, land surface temperature (LST) provides a valuable set of observations for characterizing land surface states and land-atmosphere exchange. LST is widely used in a variety of fields including evapotranspiration, climate change, hydrological cycle, vegetation monitoring, urban climate and environmental studies (Arnfield, 2003; Bastiaanssen, Menenti, Feddes, and Holtslag, 1998; Hansen, Ruedy, Sato, and Lo, 2010; Kogan, 2001; Su, 2002; Voogt and Oke, 2003; Weng, 2009; Weng, Lu, and Schubring, 2004). Besides, it has been recognized as one of the high priority parameters of the

International Geosphere and Biosphere Program (IGBP) (Townshend et al., 1994). Therefore, LST is a crucial parameter in land surface processes, not only acting as an indicator of climate change but also a pointer of LULC changes due to urbanization.

Early studies on LULC and LST from Earth observation satellite commonly use of coarse spatial resolution sensors such as the Advanced Very High Resolution Radiometer (AVHRR) and the Moderate Resolution Imaging Spectroradiometer (MODIS) (Keenan et al., 2014). With the opening of the United States Geological Survey (USGS) archive in 2008 (Woodcock et al., 2008), time-series Landsat imageries have become a new potential data source for monitoring LULC and LST change. The Landsat sensor series span more than 40 years of continuous earth observation data at a spatial resolution of 30 m, and a potential eight days repeat the cycle in the periods where different Landsat sensors overlap (Roy et al., 2014). It is promising to utilize the temporal feature better to reconstruct a long-term history of urban expansion and LST change for any city.

Nevertheless, it is challenging to automatically characterize urban LULC changes consistently at an acceptable accuracy (Loveland and Defries, 2004; Sexton et al., 2013; Zhu and Woodcock, 2014b). Furthermore, thermal characteristics over time may change to respond to land cover changes and thus become non-stationary, e.g., the mean and yearly amplitude of LSTs may change over time. A temporal analysis of thermal landscapes, therefore, requires the consideration of time-varying thermal characteristics. One way to avoid non-stationary in modeling, the temporal thermal landscape patterns are divided time-series observations into an individual segment that correspond to different land covers. This individual segment refers to as temporally homogeneous segments in the study. As such, consistent time-series LST datasets

reveal the urban thermal dynamics caused by land cover conversions (Weng, Fu, and Gao, 2014).

In this study, an unsupervised classification algorithm based on the Harmonic Analysis (HA) model by using intercept, slope, amplitude, and phase as coefficient values of a harmonic wave is first applied to classify time-series LULC data using K-Means and ISODATA (Iterative Self-Organizing Data Analysis) algorithms. Then, a nearly-real-time supervised classification using a time-series model with a minimum spectra distance is developed to classify multitemporal LULC maps and to detect LULC change. For the validation of these two classification methods, a case study in Nanjing City, Jiangsu Province, China is here examined using time-series Landsat 5, 7, 8 datasets between 2000 and 2017. Also, the impact of LULC change on LST between 2000 and 2017 is investigated.

1.2 Research objectives

The ultimate goal of the study is to develop a new supervised classification method using harmonic analysis with a minimum spectral distance algorithm for multitemporal LULC mapping. The specific research objectives are as follows:

(1) To apply unsupervised classification algorithms for multitemporal LULC classification based on time-series Landsat datasets.

(2) To develop a nearly real-time supervised LULC classification and change detection method using a time-series model and a minimum spectral distance algorithm.

(3) To reconstruct the time-series LST dataset and investigate the impact of LULC on LST.

1.3 Scope of the study

Scope of the study can be summarized as follows:

(1) The LST data of time-series Landsat datasets between 2000 and 2017 are retrieved using the Single Channel (SC) algorithm. Meanwhile, spectral reflectance data are extracted using a standard conversion method that recommends by USGS from Landsat 5, 7, and 8 Data User Handbook. Also, the MODIS product (MOD05) that provides water vapor content is downloaded for LST extraction.

(2) Two generic unsupervised classification algorithms, namely K-Means and ISODATA, are applied to classify multitemporal LULC maps based on the derived coefficient of HA. Herein, the K-Means and ISODATA algorithms are implemented using the MATLAB software. Main LULC types include urban and built-up land (U), agricultural land (A), forest land (F), and water bodies (W).

(3) A new supervised classification method using harmonic analysis with a minimum spectral distance algorithm is developed to classify multitemporal LULC maps using MATLAB software.

(4) An optimum method for multitemporal LULC classification and change detection from unsupervised and supervised classification methods is identified based on overall accuracy and Kappa hat coefficient that are derived by comparison between classified LULC in 2000, 2006, 2011, and 2017 and high spatial resolution image of Google Earth from corresponding years.

(5) The impact of LULC on LST is assessed based on the derived LULC data between 2000 and 2017 using decomposition analysis.

1.4 Limitation of the study

Due to the limitation of ground reference data in 2000, 2006, 2011, and 2017 for accuracy assessment of LULC classification maps, available high spatial resolution, imageries from the Google Earth, and Landsat data are applied as reference ground information for accuracy assessment.

1.5 Study area

Jiangning, as the largest district from eleven districts of Nanjing city, Jiangsu Province, China, is selected as a study area. It covers an area of 1,587 km² or about 24% of Nanjing City (6,598 km²) and situates between 118°28'-119°7'E and 31°37'-32°7'N. The climate of Jiangning belongs to a humid subtropical climate, and its landform is characterized by hillocks, low maintains, hills, plains, rivers, and lakes (Figure 1.1). The selected study area represents the rapid urbanization and economic growth of Nanjing City, as mentioned earlier in this chapter. Many agricultural lands in the city were converted to urban and built-up land due to population growth and rapid urbanization.

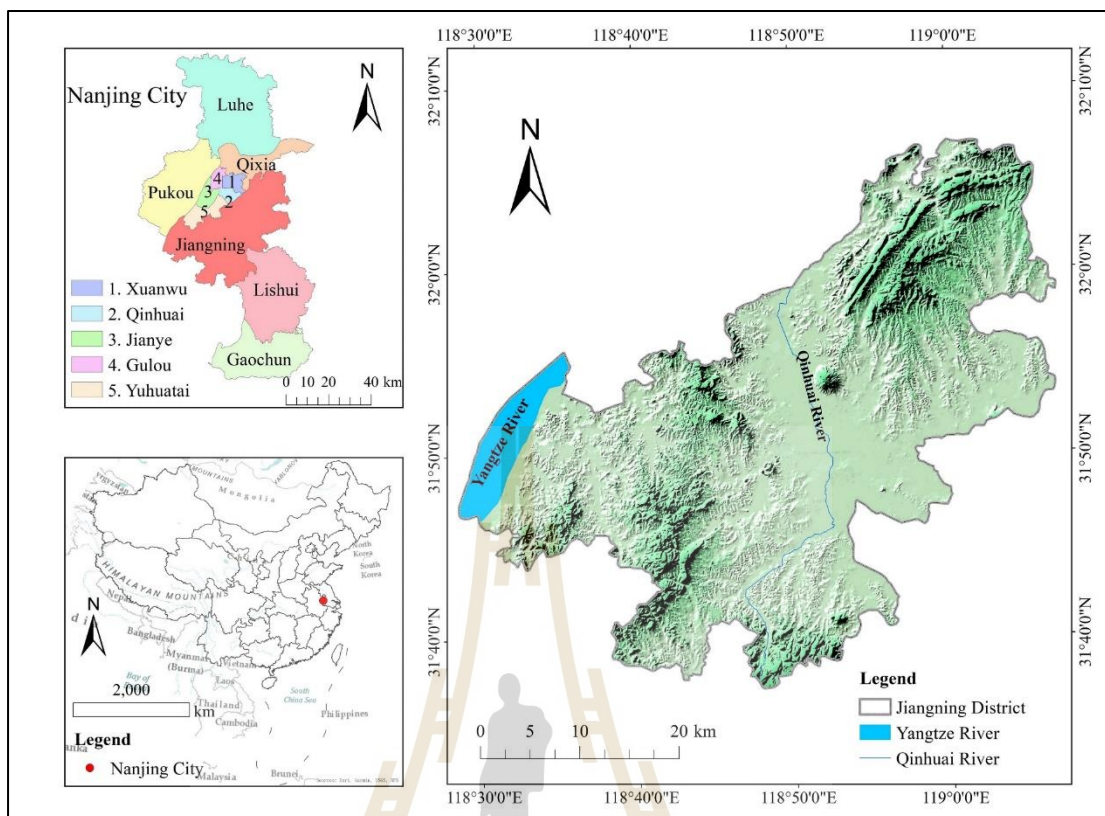


Figure 1.1 Location map of the study area.

1.6 Benefit of the study

The benefits of the study are as follows:

(1) The traditional classification method mostly uses the spectral features from a single date image (single sensor or multiple sensors) for LULC classification. However, if the single date image is affected by unforeseen factors (such as cloud coverage, cloud shadow, snowfall, heavy rainfall, and flood), the LULC classification will have a significant error. This research introduces time-series features for LULC classification to remove the influence of these unforeseen factors so that we can understand the changes of LULC over a long period.

(2) The comprehensive utilization of spectral and temporal features that generate a large number of feature information can be used for multitemporal LULC classification. This research explores the suitable feature information for LULC classification of time-series data through a systematic comparison of the results.

(3) Results about the impact of long-term LULC changes on LST can be further applied to describe the relationship between LULC and LST in the context of global climate change and urbanization.

(4) A new supervised classification method using harmonic analysis with minimum spectral distance algorithm for multitemporal land use and land cover mapping from the time-series Landsat datasets (2000-2017) was developed in this study. The results revealed that the newly developed supervised classification method could provide an average overall accuracy and Kappa hat coefficient of more than 80%. With a newly developed classification method for multitemporal land use and land cover mapping, it will be a benefit to operators, managers, and researchers who are interested in the field of remote sensing, particularly multitemporal LULC classification and mapping.

CHAPTER II

BASIC CONCEPTS AND LITERATURE REVIEWS

Basic concepts and literature reviews related to the research are here summarized, including (1) Single Channel algorithm for LST extraction, (2) time-series of the harmonic analysis, (3) time-series change detection and classification model, and (4) feature selection for LULC change detection and classification.

2.1 Single channel algorithm for LST extraction

2.2.1 Development background

The LST is an essential parameter in the surface process of land and air interaction. Over the past several decades, LST estimation from satellite thermal infrared (TIR) band measurements has been significantly improved and applied in various research aspects e. g. urbanization, urban heat island, etc. Meanwhile, due to high spatial resolution of the thermal infrared band of the Landsat series, such as the band 6 of Landsat 5 Thematic Mapper (TM), band 6 of the Landsat 7 Enhanced Thematic Mapper (ETM), and band 10, band 11 of the Landsat 8, these TIR data have been widely applied for LST extraction. Thus, to reduce the dependence on radiosonde data, several Single Channel (SC) algorithms have been proposed within the past decade to estimate the LST from satellite data assuming that the Land Surface Emissivity (LSE) is known.

Based on the thermal radiance transfer equation, Qin, Karnieli, and Berliner (2001) had developed a mono-window algorithm for retrieving LST from Landsat TM data. Three parameters are required for the algorithm: emissivity, transmittance, and effective mean atmospheric temperature. Sensitivity analysis of the algorithm indicated that the possible error of ground emissivity, which is difficult to estimate, has a relatively insignificant impact on the probable LST estimation error, which is sensitive to the possible error of transmittance and atmospheric temperature. Validation of the simulated data for various situations of seven typical atmospheres indicated that the algorithm could provide an accurate LST retrieval from TM data. The LST difference between the retrieved and the simulated ones is less than 0.4 °C for most situations.

Wang et al. (2015) presented an Improved Mono-Window (IMW) algorithm for LST extraction from the Landsat 8 TIR band 10 data. Sensitivity analysis conducted for the IMW algorithm revealed that the possible error in estimating the required atmospheric water vapor content has the most significant impact on the probable LST estimation error. Under moderate errors in both water vapor content and ground emissivity, the algorithm had an accuracy of 1.4 K for LST estimation. Validation of the IMW algorithm using the simulated datasets for various situations indicated that the LST difference between the estimated and the simulated ones was 0.67 K on average, with a root mean square error (RMSE) of 0.43 K. Comparison of the IMW algorithm with the SC algorithm for three-atmosphere profiles indicated that the average error and RMSE of the IMW algorithm were -0.05 K and 0.84 K, respectively, which were less than the -2.86 K and 1.05 K of the SC algorithm.

Jiménez-Muñoz and Sobrino (2003) proposed a generalized SC algorithm that only uses the total atmospheric water vapor content and the effective channel wavelength, which can apply to thermal sensors characterized with a full width half maximum (FWHM) of around $1 \mu\text{m}$ operative onboard satellites. The main advantage of this algorithm compared with the other SC methods is that in-situ radio soundings or effective mean atmospheric temperature values are not needed. Meanwhile, the main advantage of this algorithm compared with split-window and dual-angle methods is that it can be applied to different thermal sensors using the same equation and coefficients. The validation for different test sites showed a root mean square deviations (RMSD) is lower than 2 K for AVHRR channel 4 and ATSR-2 channel 2, and lower than 1.5 K for Landsat TM band 6.

Jiménez-Muñoz et al. (2009) presented a revision, an update, and an extension of the generalized SC algorithm developed by Jiménez-Muñoz and Sobrino (2003), which was applied to the TIR channel located in the Landsat 5 TM sensor. The SC algorithm relies on the concept of atmospheric functions, which are dependent on atmospheric transmissivity and upwelling and downwelling atmospheric radiances. These atmospheric functions are fitted versus the atmospheric water vapor content for operational purposes. They presented updated fits using MODTRAN 4 radiative transfer code, and they also extended the application of the SC algorithm to the TIR channel of the TM sensor onboard the Landsat 4 platform and the ETM sensor onboard the Landsat 7 platform. Five different atmospheric sounding databases had been examined to create simulated data used for testing the algorithm. The test from independent simulated data provided RMSE values below 1 K in most cases when

atmospheric water vapor content is lower than two $\text{g}\cdot\text{cm}^{-2}$. For values higher than three $\text{g}\cdot\text{cm}^{-2}$, errors are not acceptable, as what occurs with other SC algorithms.

Jiménez-Muñoz, Sobrino, Skoković, Mattar, and Cristóbal (2014) proposed SC algorithm for LST estimation from Landsat 8 TIRs data. The algorithm was tested with simulated data obtained from forwarding simulations using atmospheric profile databases and emissivity spectra extracted from spectral libraries. Results showed mean errors typically were below 1.5 K for both SC and Split Window (SW) algorithms, with slightly better results for the SW algorithm than for the SC algorithm with increasing atmospheric water vapor contents.

Cristóbal, Jiménez-Muñoz, Sobrino, Ninyerola, and Pons (2009) presented an improved methodology to extract LST from Landsat 4 TM, Landsat 5 TM, and Landsat 7 ETM using four atmospheric databases covering different water vapor ranges (from 0 to 8 $\text{g}\cdot\text{cm}^{-2}$) to build the LST retrieval models and using both water vapor and air temperature as input variables. They also compared this method with LST estimation using only water vapor or only air temperature, as well as with an existing LST retrieval algorithm. Herewith, seventy-seven Landsat images taken between 2002 and 2006 and two sources of water vapor (radiosonde data and remote sensing estimations) and air temperature (radiosonde data and air temperature modeling) were selected to validate the results. The best results using radiosonde data are obtained when both air temperature and water vapor are present in the LST retrieval models with a mean RMSE of 0.9 K, followed by only water vapor models with a mean RMSE of 1.5 K and only air temperature models with a mean RMSE of 5.6 K. The results obtained using Terra MODIS Level 2 water vapor product and at-satellite-pass air temperature modeling as input data also showed that this kind of input data offers best results, with

a mean RMSE of 0.9 K, followed by water vapor models with a mean RMSE of 2.1 K and only air temperature models with a mean RMSE of 5.6 K. Similar errors when using radiosonde or modeled water vapor and air temperature as input data suggested the avoidance of radiosonde data to retrieve LST over extensive areas.

In this study, the SC algorithm based on Jiménez-Muñoz and Sobrino (2003) is selected for LST extraction from Landsat series data. Since the USGS (2017) pointed out that, given the more substantial uncertainty in the band 11 values, users should work with band 10 data as a single spectral band (like TM, ETM) and should not attempt a split-window correction using both TIRs bands 10 and 11.

2.1.2 Single channel algorithm

The Single Channel (SC) algorithm developed by Jiménez-Muñoz and Sobrino (2003) retrieves LST (T_s) using the following equation:

$$T_s = \gamma \left[\frac{1}{\varepsilon} (\varphi_1 L_{sen} + \varphi_2) + \varphi_3 \right] + \delta \quad (2.1)$$

where T_s is land surface temperature, L_{sen} is the at-sensor radiance, ε is the surface emissivity, and (γ, δ) are two parameters given by:

$$\gamma \approx \frac{T_{sen}^2}{b_\gamma L_{sen}}, \quad \delta \approx T_{sen} - \frac{T_{sen}^2}{b_\gamma} \quad (2.2)$$

where T_{sen} is the at-sensor Brightness Temperature (BT); $b_\gamma = c_2/\lambda$; and φ_1 , φ_2 , and φ_3 are so-called atmospheric functions, given by:

$$\varphi_1 = 1/\tau; \quad \varphi_2 = -L_d - L_u/\tau; \quad \varphi_3 = L_d \quad (2.3)$$

where τ is the total atmospheric transmissivity, L_d is the down-welling radiance, L_u is the up-welling atmospheric radiance.

The practical approach proposed in the SC algorithm consists of the approximation of the atmospheric functions defined in (Eq. 2.3) versus the atmospheric water vapor content w from a second-order polynomial fit, expressed in matrix notation as follows ($\varphi = c w$):

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} w^2 \\ w \\ 1 \end{bmatrix} \quad (2.4)$$

where coefficients c_{ij} are obtained by simulation.

Alternatively, if atmospheric parameters τ , L_u , and L_d are known, the atmospheric functions can be calculated from Eq. 2.3, thus avoiding the empirical relationship versus w .

The atmospheric functions φ_1 , φ_2 , and φ_3 for Landsat series data can be obtained as a function of the total atmospheric water vapor content (w) using the following equations. Eq. 2.5 (Jiménez-Muñoz et al., 2009; Sobrino, Jiménez-Muñoz, and Paolini, 2004) is used for Landsat 5/7 band 6 data while Eq. 2.6 (Jiménez-Muñoz et al., 2014) is used for Landsat 8 band 10 data.

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix} = \begin{bmatrix} 0.14714 & -0.15583 & 1.1234 \\ -1.1836 & -0.37607 & -0.52894 \\ 0.04554 & 1.8719 & -0.39071 \end{bmatrix} \begin{bmatrix} w^2 \\ w \\ 1 \end{bmatrix} \quad (2.5)$$

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix} = \begin{bmatrix} 0.04019 & 0.02916 & 1.01523 \\ -0.38333 & -1.50294 & 0.20324 \\ 0.00918 & 1.36072 & -0.27514 \end{bmatrix} \begin{bmatrix} w^2 \\ w \\ 1 \end{bmatrix} \quad (2.6)$$

2.1.3 Emissivity extraction

Three required parameters for the SC algorithm include emissivity, atmospheric water vapor content, and BT. Herein emissivity is an important one. In principle, the emissivity of an object is mainly determined by its thermo-physical

characteristics. For the ground surface, the components composing the surface are the main factors determining the ground emissivity. At present, many effective methods have been approved to estimate the emissivity for LST retrieval. Since the emissivity is variable with the wavelength, the Normalized Difference Vegetation Index (NDVI) threshold method (Sobrino et al., 2008) can be used to estimate the emissivity of different land surfaces in the 10-12 μm range. Additionally, the spectral range of band 6 of Landsat 5/7 and band 10 of Landsat 8 is suitable in this range. At this wavelength range, the emissivity could be estimated as follows:

$$\varepsilon_{\lambda} = \begin{cases} \varepsilon_w & NDVI \leq 0 \\ \varepsilon_{s\lambda} & 0 < NDVI < NDVI_s \\ \varepsilon_{v\lambda} \cdot P_v + \varepsilon_{s\lambda} \cdot (1 - P_v) + C_{\lambda} & NDVI_s \leq NDVI \leq NDVI_v \\ \varepsilon_{v\lambda} \cdot P_v + C_{\lambda} & NDVI > NDVI_v \end{cases} \quad (2.7)$$

subject to:

$$C_{\lambda} = (1 - \varepsilon_{s\lambda}) \cdot \varepsilon_{v\lambda} \cdot F' \cdot (1 - P_v) \quad (2.8)$$

where ε_{λ} is the band emissivity, $\varepsilon_{v\lambda}$ and $\varepsilon_{s\lambda}$ are respectively the vegetation and soil emissivity, P_v is the proportion of vegetation, C is a term due to surface roughness ($C = 0$ for a flat surface), $NDVI_v$ and $NDVI_s$ are the $NDVI$ for a fully vegetated pixel and a soil one, respectively, and F' is a geometrical factor ranging between zero and one.

Usually, the vegetation cover fraction at the pixel scale can be computed from its $NDVI$ as follows (Sobrino et al., 2004):

$$P_v = \left(\frac{NDVI - NDVI_s}{NDVI_v - NDVI_s} \right)^2 \quad (2.9)$$

Over particular areas, $NDVI_v$ and $NDVI_s$ values can be extracted from the $NDVI$ histogram. Values of $NDVI_v = 0.5$ and $NDVI_s = 0.2$ were proposed to apply in global conditions (Sobrino et al., 2008).

Meanwhile, the TIR band 10 of Landsat 8 is different from band 6 of Landsat 5/7. It is worth noting that most satellite sensors measure the outgoing radiation with a finite spectral-bandwidth, and the channel-effective quantities of interest, therefore a weighted average are expressed by Li et al. (2013):

$$X = \frac{\int_{\lambda_1}^{\lambda_2} f(\lambda)X(\lambda)d\lambda}{\int_{\lambda_1}^{\lambda_2} f(\lambda)d\lambda} \quad (2.10)$$

where X is a weighted average value of emissivity in a thermal infrared band, $X(\lambda)$ is various spectral quantity considered as emissivity, λ_1 and λ_2 are the lower and upper boundaries of the wavelength in TIR channel, and $f(\lambda)$ is the spectral response function.

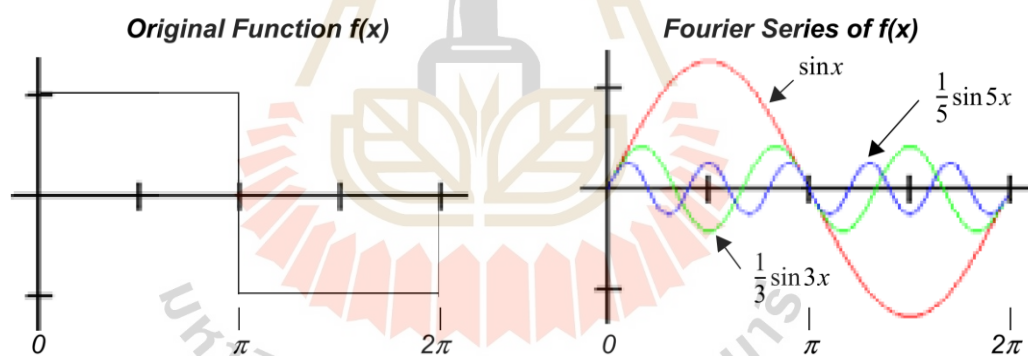
2.2 Time-series of harmonic analysis

Many time-series reconstruction models have been developed to eliminate the effect of the negative deviations which may be contaminated by atmospheric conditions or snow cover, and most of them perform differently in different applications and regions. The harmonic analysis (HA) is one of the most widely used time-series reconstruction models owing to their simplicity of implementation or the capability to extract phenological metrics from the time-series.

A series of algorithms based on discrete Fourier analysis (Immerzeel, Quiroz, and de Jong, 2005; Menenti, Azzali, Verhoef, and van Swol, 1993; Roerink, Menenti, and Verhoef, 2000; Sellers et al., 1996; Verhoef, Menenti, and Azzali, 1996) is also

called HA model. The time curve for each data point is expressed as the sum of a series of residual (positive) sine wave additive terms, each determined by a unique amplitude and phase. These continuous harmonic terms are summed to produce a complex curve. Each harmonic term occupies a certain percentage of the total variance of the original time-series curve. Since some input parameters cannot be determined by specific objective criteria, they are obtained through empirical comparisons, which puts the method at a disadvantage. However, it is one of the favorite algorithms for time-series remotely sensed data reconstruction.

Figure 2.1 shows how a function $f(x)$ can represent a linear combination of sine and cosine (Leica, 2005). In this example, the function is a square wave whose cosine coefficients are zero, leaving only sine terms.



Source: (Leica, 2005)

Figure 2.1 One-dimensional Fourier analysis.

The harmonic form of a time-series y can be expressed as:

$$y = a_0 + b_0 t + \sum_{i=1}^{\infty} \left[A_i \cos \left(\frac{2\pi}{T} i t - \varphi_i \right) \right] \quad (2.11)$$

where i is the number of frequencies, T is period (365 days), A_i and φ_i are the amplitude and phase of i^{th} harmonic, and t is point time.

Many researchers applied HA to reconstruct time-series of NDVI for detecting forest lost (Azzali and Menenti, 2000; Geerken, Zaitchik, and Evans, 2005; Menenti et al., 1993; Menenti et al., 2010; Verhoef et al., 1996). Menenti et al. (2016) applied HA for leaf area index estimation and LST extraction as well as the polarization difference BT by Shang, Jia, and Menenti (2015) for removing random noise or eliminating cloud/snow/ice contamination. Weng et al. (2004) explored the potential for the cloud removal of HA to reconstruct ten days of composited AVHRR data over the mountainous regions of the Tibetan Plateau. The capability of HA for processing NDVI time-series from SPOT-VEGETATION and for defining the start and end of the growing season was investigated by Wit and Su (2005). Additionally, the global applicability of HA had been evaluated by de Jong, de Bruin, de Wit, Schaepman, and Dent (2011) in the study of global greening and browning trends using GIMMS NDVI dataset (1981-2006). Besides cloud removal, the amplitude and phase of harmonic components are quantitative measures of vegetation phenology (Azzali and Menenti, 2000; Menenti et al., 1993), which makes HA immensely appealing for phenological studies. By concurrently processing LST and NDVI time-series derived from the Pathfinder AVHRR dataset, Julien, Sobrino, and Verhoef (2006) estimated changes in vegetation in the European continent between 1982 and 1999, which highlights the cloud removal as well as the extraction of quantitative phenological information by HA. Once more, building upon both these two capabilities, Jia, Shang, Hu, and Menenti (2011) analyzed the phenological response of vegetation to upstream river flow in the Heihe river basin of China based on NDVI time-series from MODIS. Alfieri, Lorenzi, and Menenti (2013) developed an approach to map air temperature based on daily LST products from MODIS. Likewise, Jakubauskas, Legates, and Kastens (2002) and

Geerken (2009) used the frequency domain information (i.e., amplitude and phase of harmonic components) provided by HA without the cloud detection procedure for cropland classification and phenology monitoring.

Zhu, Woodcock, Holden, and Yang (2015) had developed and tested an HA modeling approach that does not depend on any sensor. The approach has evolved from an initial version intended for change and near-real-time change applications (Zhu, Woodcock, and Olofsson, 2012). It had been tested for post-seasonal cloud screening with Landsat (Zhu and Woodcock, 2014b), and land cover monitoring (Zhu and Woodcock, 2014a). In their approaches, time-series are modeled with harmonic sine and cosine basis, linear trend, and offset terms. The fit was carried out for segments of the time-series where land cover change had not been detected. For abrupt changes, the time-series is modeled separately before and after the change event. In their works, the model fit has been determined by least squares, weighted least squares, and regularized least-squares methods to enhance stability and efficiency. As stated in Zhu et al. (2015), the minimum sampling requirements are twelve samples to obtain a reliable result.

Since no systematic study on the accuracy of such reconstruction had been done, Zhou, Jia, and Menenti (2015) took the global MODIS vegetation index as an example to develop a generic method to evaluate the reconstruction performance of HA of time-series. The overall reconstruction errors were divided into gap-related errors and fitting-method related errors. The results suggested that the gap-related errors for most of the high latitude forest area (between 50°N and 70°N) were rather large (average root mean squared deviation (RMSD) reached 0.15), which may be attributed to the fact that large gaps appeared in the NDVI profiles between snow melting and vegetation re-greening season. The gap-related errors were found negligible for the other areas of the globe

except for the North China Plain, North India, and several mountainous areas where the average RMSD was around 0.1. The inadequate capability of low-frequency harmonics to capture the rapid transition during snowmelt in spring at the high latitude region of the North Hemisphere made the fitting-method related error in this region rather large (RMSD can reach 0.1).

2.3 Time-series change detection and classification model

Many algorithms have been developed for detecting land cover change by analyzing time-series of satellite data (Hermosilla, Wulder, White, Coops, and Hobart, 2015; Huang et al., 2010; Kennedy, Cohen, and Schroeder, 2007; Verbesselt, Hyndman, Newnham, and Culvenor, 2010; Yang and Lo, 2002).

A new algorithm, “Continuous Change Detection and Classification (CCDC) model,” was developed by Zhu and Woodcock (2014b). It can detect many kinds of land cover changes continuously as new images are collected and providing land cover maps for any given time. A two-step cloud removal, cloud shadow, and snow masking algorithm is used for eliminating “noisy” observations. A time-series model that has components of seasonality, trend, and break estimates surface reflectance and BT. The time-series model is updated dynamically with newly acquired observations. Due to the differences in spectral response for various kinds of land cover change, the CCDC algorithm uses a threshold derived from all seven Landsat bands. When the difference between observed and predicted images exceeds a threshold three consecutive times, a pixel is identified as land surface change. Land cover classification is done after change detection. Coefficients from the time-series models and the RMSE from model estimation are used as input to the Random Forests Classifier. A random stratified

sample design was used for assessing the change detection accuracy, with 250 pixels selected within areas of persistent land cover and 250 pixels selected within areas of change identified by the CCDC algorithm. The accuracy assessment shows that CCDC results were accurate for detecting the land surface change, with producer's accuracy of 98% and user's accuracies of 86% in the spatial domain and temporal accuracy of 80%. Land cover reference data were used as the basis for assessing the accuracy of the land cover classification. The land cover map with 16 categories resulting from the CCDC algorithm had an overall accuracy of 90%.

Fu and Weng (2016) took the Atlanta metropolitan area as a case study to illustrate LULC change and its impact on LST variations. The Landsat L1T images from TM/ETM from 1984 to 2011 were consistently calibrated to surface reflectance and BT. The cloud, cloud shadow, and snow contaminated pixels were excluded in the analysis according to the metadata, and a further screening procedure based on the Robust Iteratively Reweighted Least Squares (RIRLS) technique was implemented. The LULC classification and change detection were accomplished by using the CCDC algorithm. Results showed that the overall LULC classification and change detection accuracies were 89% and 92%, respectively.

Zhu et al. (2016) comparatively analyzed the potential of the simple linear trend and CCDC algorithm in the detection of the green trend of the urban suburbs. The results show that the simple linear trend method is more accurate than the CCDC algorithm in detecting the general trend of the greenness change; when covering the trend of the greenness of the changing area, the CCDC algorithm can provide more detailed and more accurate information (evaluating gradients and mutations,

respectively). Its mapping accuracy and user accuracy are 67.88% to 85.19% and 68% to 97.30%, respectively.

The CCDC algorithm (Zhu and Woodcock, 2014b) includes two steps of change detection and classification. Herein, continuous change detection was developed based on the HA model, while continuous land cover classification used Random Forest Classifier.

2.3.1 Continuous Change Detection

Using all available TM/ETM observations from Landsat 4, 5, and 7, the CCDC algorithm (Zhu and Woodcock, 2014b) was developed. The use of the term “continuous” here refers to the capability to detect change detection every time a new image is collected. If it is updated as new images are collected, then the approach begins to approach near-real-time change detection. The use of “continuous” in the name of the algorithm also follows from the idea that a land cover map can be produced for any given time within the period covered by images. One additional note is that this new algorithm can detect many kinds of land cover change.

Generally, land surface change can be divided into three categories: (1) intra-annual change, caused by vegetation phenology driven by seasonal patterns of environmental factors like temperature and precipitation; (2) gradual inter-annual change, caused by climate variability, vegetation growth or gradual change in land management or land degradation; and (3) abrupt change, caused by deforestation, floods, fire, insects, and urbanization. Therefore, a time-series model that has components of seasonality, trend, and breaks is used to captures all three categories of surface change.

The model coefficients are estimated by the Ordinary Least Squares (OLS) method based on the remaining clear Landsat observations.

$$\hat{P}(i, d) = a_i + b_{1i} \cos\left(\frac{2\pi d}{T}\right) + b_{2i} \sin\left(\frac{2\pi d}{T}\right) + c_i d \quad (2.12)$$

where $\hat{P}(i, d)$ is the predicted value for reflectance and LST at Julian day d , a_i is overall value, b_{1i} , b_{2i} are the coefficients for the seasonal component (or intra-annual changes), c_i are the coefficients for the trend component (inter-annual change), d is Julian date, and T is 365.

Ideally, the more the coefficients included, the more accurate the model will be. However, when there are too many coefficients, the model may start to fit noise.

The basis of the method is the comparison of model predictions with clear satellite observations to find change. Ideally, a single date comparison would be definitive for detecting change. However, there is sufficient noise in the system due to factors like undetected clouds, cloud shadows, snow, atmospheric haze, smoke, and changes in soil wetness, which lead to numerous false-positive errors in change detection when using a single date for comparison. While noise factors tend to be ephemeral, land cover change is more persistent through time. The CCDC algorithm minimizes ephemeral effects by processing a set of dates together as a group for identifying land cover change (Zhu and Woodcock, 2014b).

Based on previous studies (Zhu et al., 2012), changes identified in three successive dates showed the best results. Therefore, pixels showing change for one or two consecutive times will be flagged as “possible change”, and if a third consecutive change is found, the pixel is assigned to the “change” class.

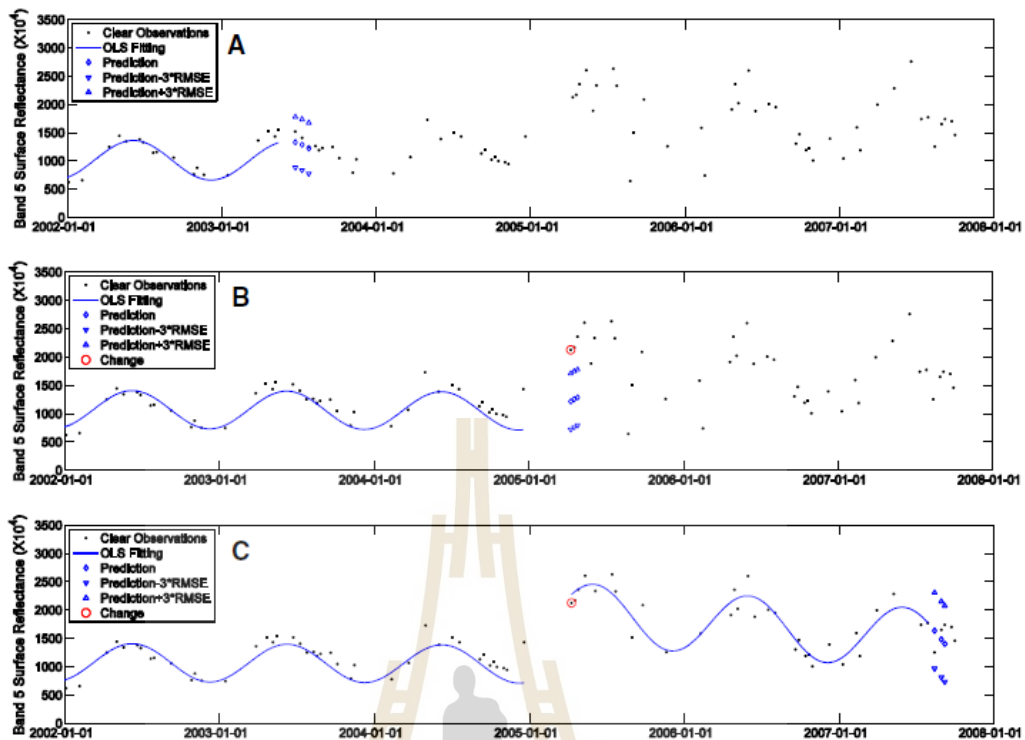


Figure 2.2 Using the threshold of three times RMSE for continuous change detection.

In Figure 2.2, Zhu and Woodcock (2014b) illustrated how the “three times RMSE” criterion was used for detecting land cover change for a deforestation pixel. When there is no land cover change, the next three clear observations are always within the model predicted ranges ($\pm 3 \times \text{RMSE}$) (Figures 2.2A and 2.2C). Figure 2.2B showed how change is initially detected by comparing the next three consecutive clear observations with model predictions.

Therefore, the continuous change detection model of the CCDC algorithm averages the difference between observations and model predictions that have been normalized by three times RMSE for all seven Landsat bands. If the result is larger than 1 for three consecutive clear observations, a change is identified (Eq. 2.13).

Otherwise, if the values for only one or two consecutive observations are larger than 1, it is regarded as an ephemeral change, and the observations are flagged as outliers. The CCDC algorithm updates the time-series model when new clear observations become available, adding a dynamic character to the process that allows the time-series to adjust over time. Eq. 2.13 provides the details of the land cover change algorithm.

$$\frac{1}{k} \sum_{i=1}^k \frac{|\rho(i,d) - \hat{\rho}(i,d)|}{3 \times RMSE_i} > 1 \quad (\text{three times consecutively}) \quad (2.13)$$

while d is Julian date, i is the i^{th} Landsat band, k is the number of Landsat bands, $\rho(i, d)$ is the observed value for the i^{th} Landsat band at Julian date, d , $\hat{\rho}(i, d)$ is predicted value for the i^{th} Landsat band at Julian date, d , and $RMSE$ is a root mean square error that shows the difference between observed and predicted values.

The model began to detect the land cover change when there are 15 clear sky observations (the model initialization phase). The first 12 observations were used to determine outliers. The last three observations allowed the model to determine whether changes in land cover occurred.

2.3.2 Continuous land cover classification

Instead of classifying the original Landsat images as conventional methods would, the CCDC algorithm uses the coefficients of time-series models as the inputs for land cover classification. After the change detection process, each pixel will have its time-series models before and after any changes. By classifying the time-series model coefficients, this algorithm can provide a land cover type for the entire period for each time-series model. Time-series observations from all seven Landsat bands were used for the land cover classification. The main idea of the land cover

classification is that different land cover classes will have different shapes for the estimated time-series models.

The classification was accomplished through a Random Forests Classifier with inputs of the coefficients derived from the time-series model. The principle of the classification supposed that different land covers showed different modeling characteristics (i.e., different coefficients for seasonal and trend components and modeling errors) for reflectance and LST. The variables, \bar{P} , b_{1i} , b_{2i} , c_i , and RMSE derived from the time-series model were used as inputs for the classification with \bar{P} representing the mean value:

$$\bar{P} = a + c \times \frac{t_1 + t_n}{2} \quad (2.14)$$

with t_1 and t_n are the Julian dates for the starting and ending time of the model.

2.4 Features selection for LULC change detection and classification

Many spectral features (also named variables) can be used as input data in change detection and classification scheme. These features can be divided into three categories (Tang, Fan, and Zhang, 2017): band type, index type, and transformation type (linear and non-linear transformation).

2.4.1 Band type

Band type refers to the direct use of the original band for change detection, the most commonly used is SWIR band, which is a famous band that

describes the vegetation structure and the detection of forest changes, or inferred vegetation burns (Brown, Chen, Leblanc, and Cihlar, 2000; Healey, Yang, Cohen, and Pierce, 2006).

However, a single band cannot make full use of all band information, and it is challenging to diagnose different types of changes. Therefore, it is generally recommended to use multiple bands or multiple indicators for comparative analysis (Zhu et al., 2016; Zhu and Woodcock, 2014b; Zhu et al., 2012).

Under this study, top of atmosphere (TOA) spectral reflectance data of Landsat time-series include Blue, Green, Red, NIR, SWIR1, and SWIR2 bands are applied as band type for time-series data analysis.

2.4.2 Index type

The spectral indices had been developed for remotely sensed data to improve the appearance of an image for human visual analysis or occasionally for subsequent machine analysis (Jensen, 2015). Among them, NDVI (Normalized Difference Vegetation Index) developed by Rouse, Haas, Schell, and Deering (1973) is the most widely used vegetation index (Myneni, 1995). NDVI and Enhanced Vegetation Index (EVI) have a high correlation with chlorophyll content, leaf area index, photosynthesis ability, etc. They are widely used to analyze the greenness trend of vegetation (Carlson and Ripley, 1997; Huete et al., 2002; Zhu et al., 2016), to detect forest disturbance (Zhu et al., 2012), to identify vegetation change (vegetation restoration, pests, and diseases) (Vogelmann, Gallant, Shi, and Zhu, 2016). However, EVI is less affected by atmospheric conditions and soil background when it is compared with NDVI (Huete, Justice, and Leeuwen, 1999).

Zha, Gao, and Ni (2003) proposed a method based on NDBI (Normalized Difference Built-up Index) to automate the process for built-up areas mapping. Shingare, Hemane, and Dandekar (2014) modified the NDBI formula by using SWIR2 instead of SWIR1. The final results of the decision tree algorithm using original NDBI and MNDBI (Modified Normalized Difference Built-up Index) were compared, and it was found that the decision tree algorithm using MNDBI provided more accurate results for urban area classification.

The NDWI (Normalized Difference Water Index) has been developed by McFeeters (1996) to delineate open water features and enhance their presence in remotely sensed data. The NDWI makes use of reflected near-infrared radiation and visible green light to enhance the presence of such features while eliminating the presence of soil and terrestrial vegetation features. It was suggested that the NDWI might also provide researchers with turbidity estimations of water bodies using remotely sensed data. Meanwhile, the MNDWI (Modified Normalized Difference Water Index) by Xu (2006) can enhance open water features while efficiently suppressing and even removing built-up land noise as well as vegetation and soil noise. The enhanced water information using the NDWI is often mixed with built-up land noise, and the area of the extracted water is thus overestimated. Accordingly, the MNDWI is more suitable for enhancing and extracting water information for a water region with a background dominated by built-up land because of its advantage in reducing and even removing built-up land noise over the NDWI.

In this research, spectral indices for time-series data analysis consist of NDVI, EVI, NDWI, MNDWI, NDBI, and MNDBI. The equations for these selected spectral indices are listed as below:

$$NDVI = \frac{NIR-Red}{NIR+Red} \quad (2.15)$$

$$EVI = 2.5 * \frac{NIR-Red}{NIR+6*Red-7.5*Blue+1} \quad (2.16)$$

$$NDWI = \frac{Green-NIR}{Green+NIR} \quad (2.17)$$

$$MNDWI = \frac{Green-SWIR1}{Green+SWIR1} \quad (2.18)$$

$$NDBI = \frac{SWIR1-NIR}{SWIR1+NIR} \quad (2.19)$$

$$MNDBI = \frac{SWIR2-NIR}{SWIR2+NIR} \quad (2.20)$$

2.4.3 Transformation type

In general, transformation types are composed of linear and non-linear transformation. The most frequent use of transformation type is linear form, including tasseled cap and principal component transformations.

Tasseled cap transformation that describes characteristics of crops and other types (Crist and Cicone, 1984; J. Kauth and S. Thomas, 1976) is applied only for Landsat images. Huang, Wylie, Yang, Homer, and Zylstra (2002) had developed a new tasseled cap transformation based on Landsat 7 at-satellite reflectance. They found that this transformation is the most appropriate for regional applications where atmospheric correction is not feasible. Baig, Zhang, Shuai, and Tong (2014) applied the new development of tasseled cap transformation for biomass estimation of Landsat 8 images.

The first tasseled cap band corresponds to the Tasseled Cap Brightness (TCB) of the image. The second tasseled-cap band corresponds to Tasseled Cap Greenness (TCG) and is typically used as an index of photosynthetically active vegetation. The third tasseled-cap band is often interpreted as an index of Tasseled Cap

Wetness (TCW) (e.g., soil or surface moisture). The equations for calculation of tasseled cap transformation of Landsat series data are summarized as follows:

For Landsat 5:

$$TCB = 0.2909 \times B + 0.2493 \times G + 0.4806 \times R + 0.5568 \times NIR + 0.4438 \times SWIR1 + 0.1706 \times SWIR2 \quad (2.21)$$

$$TCG = -0.2728 \times B - 0.2174 \times G - 0.5508 \times R + 0.7221 \times NIR + 0.0733 \times SWIR1 - 0.1648 \times SWIR2 \quad (2.22)$$

$$TCW = 0.1446 \times B + 0.1761 \times G + 0.3322 \times R + 0.3396 \times NIR - 0.621 \times SWIR1 - 0.4186 \times SWIR2 \quad (2.23)$$

For Landsat 7 and 8:

$$TCB = 0.3561 \times B + 0.3972 \times G + 0.3904 \times R + 0.6966 \times NIR + 0.2286 \times SWIR1 + 0.1596 \times SWIR2 \quad (2.24)$$

$$TCG = -0.3344 \times B - 0.3544 \times G - 0.4556 \times R + 0.6966 \times NIR - 0.0242 \times SWIR1 - 0.2630 \times SWIR2 \quad (2.25)$$

$$TCW = 0.2626 \times B + 0.2141 \times G + 0.0926 \times R + 0.0656 \times NIR - 0.7629 \times SWIR1 - 0.5388 \times SWIR2 \quad (2.26)$$

In the meantime, the Principal Component Transformation (PCT) or Principal Component Analysis (PCA) is a feature space transformation designed to remove the spectral redundancy (Ready and Wintz, 1973). PCA has been widely used in remote sensing studies to isolate specific components related to vegetation dynamics (Anyamba and Eastman, 1996; Eastman and Fulk, 1993; Gurgel and Ferreira, 2003; Hall-Beyer, 2003; Hirosawa, Marsh, and Kliman, 1996) or to perform classification based on time-series similarity (Benedetti, Rossini, and Taddei, 1994; Lobo and Maisongrande, 2008).

To calculate the PCA, the eigenvectors and eigenvalues of the n principal components are derived from the covariance matrix of original remotely sensed data, as shown in the following equation (Leica, 2005):

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \quad (2.27)$$

$$E \cdot Cov \cdot ET = \Lambda \quad (2.28)$$

where Cov is the covariance matrix, E is the matrix of eigenvectors, T is the transposition function, Λ is a diagonal matrix of eigenvalues, in which all non-diagonal elements are zeros.

Λ is computed so that its nonzero elements are ordered from highest to least, so that $\lambda_1 > \lambda_2 > \lambda_3 \dots > \lambda_n$.

$$R_{kp} = \frac{a_{kp} \times \sqrt{\lambda_p}}{\sqrt{Var_k}} \quad (2.29)$$

where $R_{k,p}$ is a correlation between band k and each principal component, $a_{k,p}$ is an eigenvector for band k and component p , λ_p is p^{th} eigenvalue, Var_k is a variance of band k in the covariance matrix.

$$newBV_{i,j,p} = \sum_{k=1}^n R_{kp} BV_{i,j,k} \quad (2.30)$$

where, $BV_{i,j,k}$ is brightness value in band k for the pixel at row i , column j , n is the number of bands.

CHAPTER III

RESEARCH METHODOLOGY

The overview research methodology framework, according to the research objectives, is presented in Figure 3.1. Brief information with significant tasks of each component, including (1) data collection and preprocessing of time-series Landsat datasets, (2) spectral features selection for LULC classification, (3) time-series LULC classification using unsupervised method, (4) time-series LULC classification using supervised method, and (5) impact study of LULC change on LST, is separately summarized in the following sections. Meanwhile, details of the research methodology of each component and results were described in the following chapters.

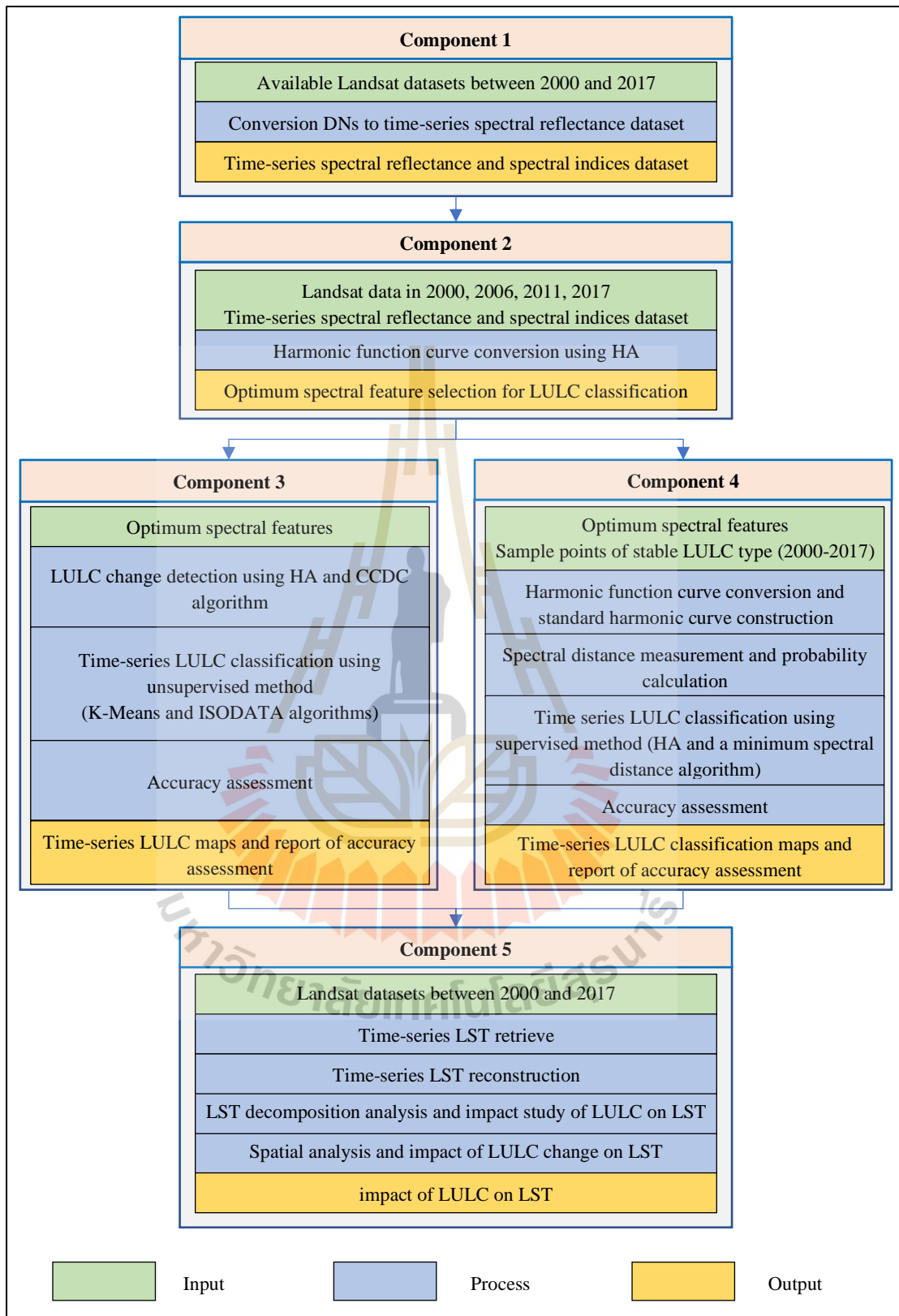


Figure 3.1 Overview of research methodology framework.

3.1 Data collection and preprocessing of time-series Landsat datasets

All available Landsat 5/7/8 images at Level-1 product of path 120 and row 38 (WRS-2) between 2000 and 2017 were firstly downloaded through USGS online portal. Characteristics of Landsat 5/7 and 8 data were summarized in Table 3.1. Then, the images were selected according to the metadata of the Landsat image, if cloud coverage in the study area is less than 90%. This percentage criterion is set to remove images since they are severely contaminated by clouds.



Table 3.1 Characteristics of Landsat 5/7/8 data.

Landsat – 5				Landsat – 7				Landsat – 8			
Sensor	Band	Spectral Resolution (mm)	Spatial Resolution (m)	Sensor	Band	Spectral Resolution (mm)	Spatial Resolution (m)	Sensor	Band	Spectral Resolution (mm)	Spatial Resolution (m)
TM	1	0.45 – 0.52	30	ETM	1	0.450–0.515	30	OLI	1	0.42 – 0.45	30
	2	0.52 – 0.60	30		2	0.525–0.605	30		2	0.45 – 0.51	30
	3	0.63 – 0.69	30		3	0.63 – 0.69	30		3	0.53 – 0.59	30
	4	0.76 – 0.90	30		4	0.75 – 0.90	30		4	0.64 – 0.67	30
	5	1.55 – 1.75	30		5	1.55 – 1.75	30		5	0.85 – 0.88	30
	6	10.4 – 12.5	120		6	10.4 – 12.5	60		6	1.57 – 1.65	30
	7	2.08 – 2.35	30		7	2.08 – 2.35	30		7	2.11 – 2.29	30
			8		0.52 – 0.90	15	8		0.50 – 0.68	15	
						9	1.36 – 1.38		30		
						10	10.6–11.19		100		
						11	11.5–12.51		100		



In this component, four essential tasks of data preprocessing of time-series Landsat datasets include (1) cloud cover assessment, (2) cloud/shadow pixel recognition based on quality assessment (QA) band and HA model, (3) conversion of digital number to spectral reflectance data, and (4) time-series spectral reflectance construction.

The main output of this component is time-series spectral reflectance and spectral indices datasets.

3.2 Optimum spectral features selection for LULC classification

Under this component, LULC data in 2000, 2006, 2011, and 2017 were first classified using the Maximum Likelihood Classifier under ERDAS Imagine software. Then the historical stable pixels of LULC types between 2000 and 2017 were extracted based on the classified LULC maps in 2000, 2006, 2011, and 2017 using overlay analysis under ESRI ArcGIS software. After that, sample points from each stable LULC type were randomly selected by using the create random points function of the ESRI ArcMap software. The sample points of each LULC type were further used to transform into harmonic function curves for six original spectral bands and twelve spectral indices. After that, two criteria based on the fitting effect of the harmonic wave function using R^2 and the ability of features to distinguish a specific LULC type using probability were applied to select optimum spectral featured by multiplication operation.

The main output of this component is the optimum spectral feature selection for LULC classification.

3.3 Time-series LULC classification using unsupervised method

Under this component, three essential tasks of time-series LULC classification using the unsupervised method included (1) LULC change detection using HA and CCDC, (2) unsupervised LULC classification using K-Means and ISODATA algorithm and (3) accuracy assessment. In brief, the harmonic wave coefficient (intercept, slope, amplitude, and phase) of the selected spectral features were here applied for unsupervised LULC classification using K-Means and ISODATA algorithms. The advantage of unsupervised classification is that it does not require the selection of a training area, and the result is objective, and there is no artificial influence.

Two main outputs of this component are time-series LULC maps and reports of accuracy assessment of unsupervised classification methods for time-series LULC classification.

3.4 Time-series LULC classification using supervised method

Under this component, four essential tasks of time-series LULC classification using supervised method included (1) harmonic function curve conversion and standard harmonic curve construction, (2) spectral distance measurement and probability calculation, (3) time-series LULC classification and (4) accuracy assessment. In brief, time-series LULC maps were classified using the developed supervised classification method using HA with a minimum spectral distance algorithm. Herein, the constructed harmonic function curve of each LULC type was applied to calculate spectral distance measurement and probability calculation for time-series LULC classification.

Two main outputs of this component are time-series LULC maps and reports of accuracy assessment of supervised classification method for time-series LULC classification.

3.5 Impact study of LULC change on LST

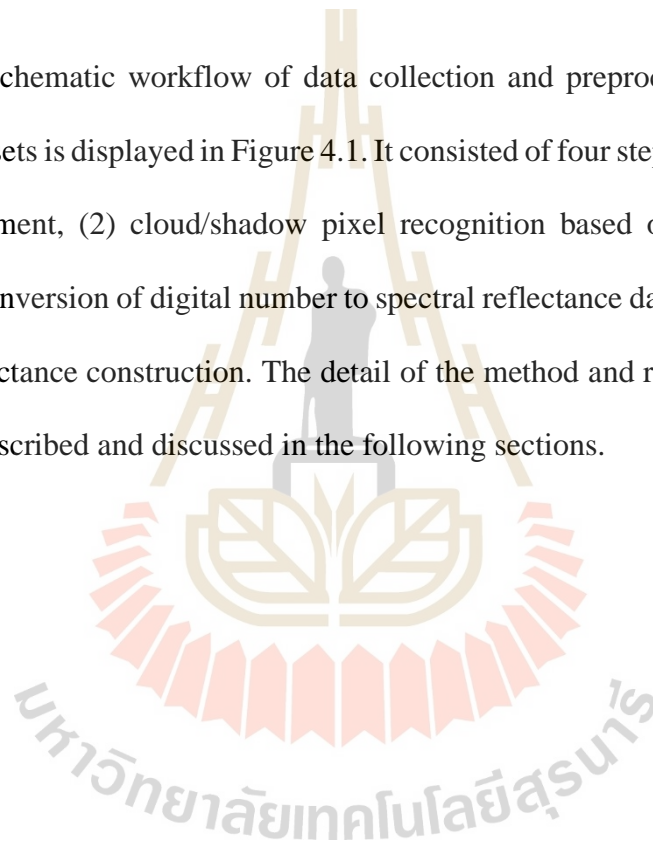
Under this component, four tasks included (1) time-series LST retrieve, (2) time-series LST reconstruction, (3) LST decomposition analysis and impact study of LULC on LST, and (4) spatial analysis and impact of LULC change on LST were implemented. In brief, the time-series LST dataset was extracted for each independent scene using the SC algorithm. Then, the time-series LST dataset was reconstructed using the HA model due to some incorrect LST values. Later, the reconstructed time-series LST dataset was further used in decomposition analysis and spatial analysis for an impact study of LULC change on LST. In this study, two approaches based on time and spatial domains were applied to study the impact of LULC change on LST.

The main output of this component is reconstructed time-series LST between 2000 and 2017 and report on the impact of LULC change on LST.

CHAPTER IV

DATA COLLECTION AND PREPROCESSING OF TIME-SERIES LANDSAT DATASETS

The schematic workflow of data collection and preprocessing of time-series Landsat datasets is displayed in Figure 4.1. It consisted of four steps, including (1) cloud cover assessment, (2) cloud/shadow pixel recognition based on QA band and HA model, (3) conversion of digital number to spectral reflectance data, and (4) time-series spectral reflectance construction. The detail of the method and result of each step was separately described and discussed in the following sections.



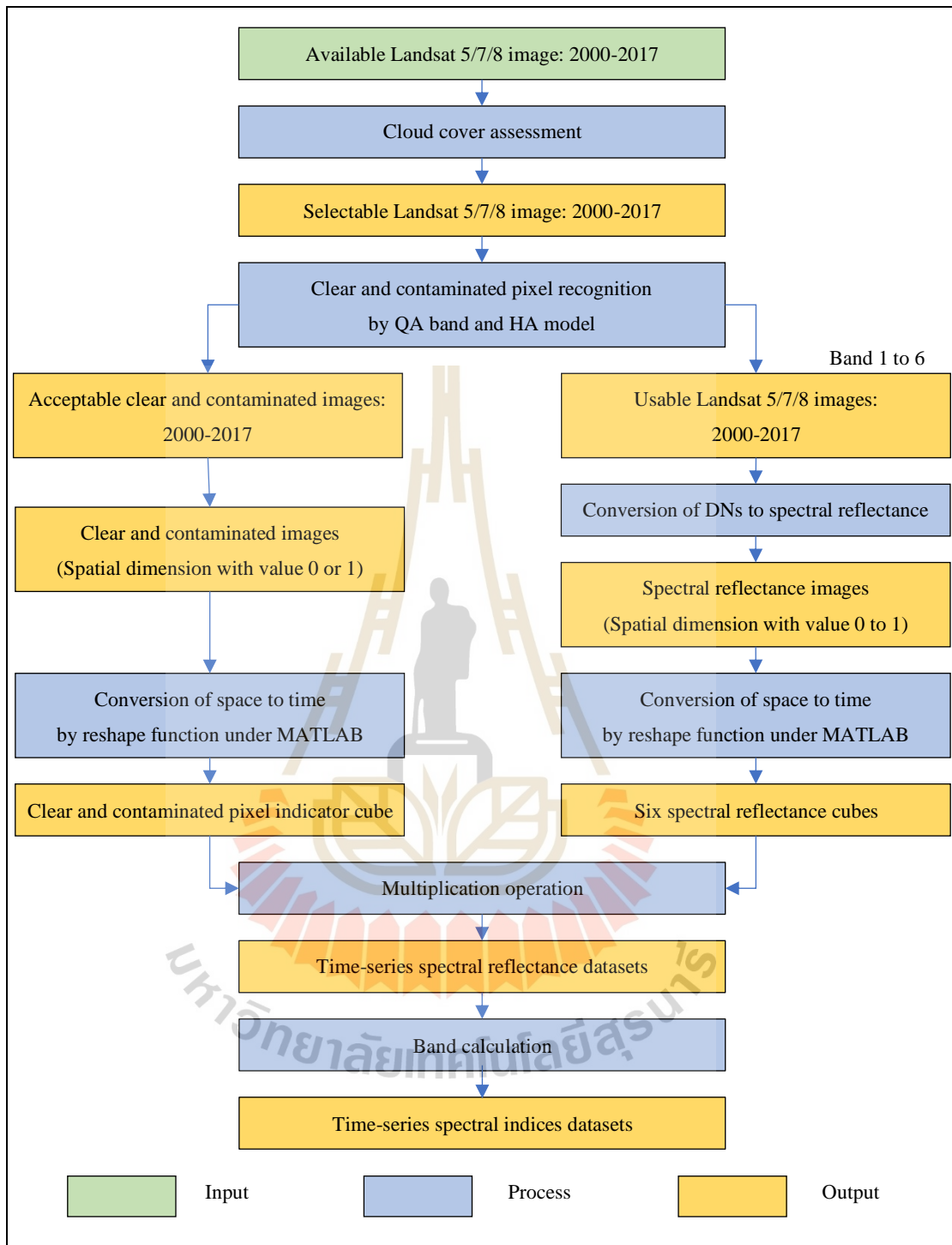


Figure 4.1 Workflow of Component 1: Data collection and preprocessing of time-series Landsat datasets.

4.1 Cloud cover assessment

Under this step, the percent of total images and percent of total clear observed pixels by percent of cloud cover were calculated for reducing the number of scenes and costs and keeping the highest percent of total clear observed pixels for data analysis.

In practice, the metadata of cloud cover percent from the USGS web portal of available Landsat 5/7/8 images between 2000 and 2017 was observed, recorded, and calculated. The percent of total image and percent of total clear observed pixels according to 10 interval classes (≤ 10 , ≤ 20 , ≤ 30 , ≤ 40 , ≤ 50 , ≤ 60 , ≤ 70 , ≤ 80 , ≤ 90 , and ≤ 100) of cloud cover (%) was calculated using Eq. 4.1 and 4.2.

$$P_{image,i} = \frac{N_{image,i}}{N_{image}} * 100\% \quad (4.1)$$

$$P_{pixel,j} = \frac{\sum_0^j N_{image,j} * (1-j) * M}{N_{image} * M} * 100\% \quad (4.2)$$

where, i is the cloud cover (%) less than or equal to interval class (10, 20, ..., 100), $P_{image,i}$ is the percent of total images, $N_{image,i}$ is the number of images, N_{image} is the number of total images. j is the cloud cover (%), $P_{pixel,j}$ is the percent of total clear observed pixels, and M is the number of pixels in the image.

All downloaded Landsat 5/7/8 images at Level-1 product of path 120 and row 38 (WRS-2) between 2000 and 2017 with a total of 673 images (Figure 4.2) were assessed cloud coverage from metadata of Landsat image. According to metadata, all images with cloud coverage was less than 90% are here selected for preprocessing. Figure 4.3 shows a cumulative histogram of the percentage of images and clear pixels over the Nanjing City scene between 2000 and 2017. Based on this information, if we choose an image with cloud cover less than or equal 10%, about 26% of total images can be used. These images have about 50% of total clear pixels; therefore, about 50%

of clear pixels are ignored. Likewise, we choose an image with cloud cover less than or equal 60%, about 56% of total images can be used. These images have about 90% of total clear pixels; therefore, about 10% of clear pixels are ignored.

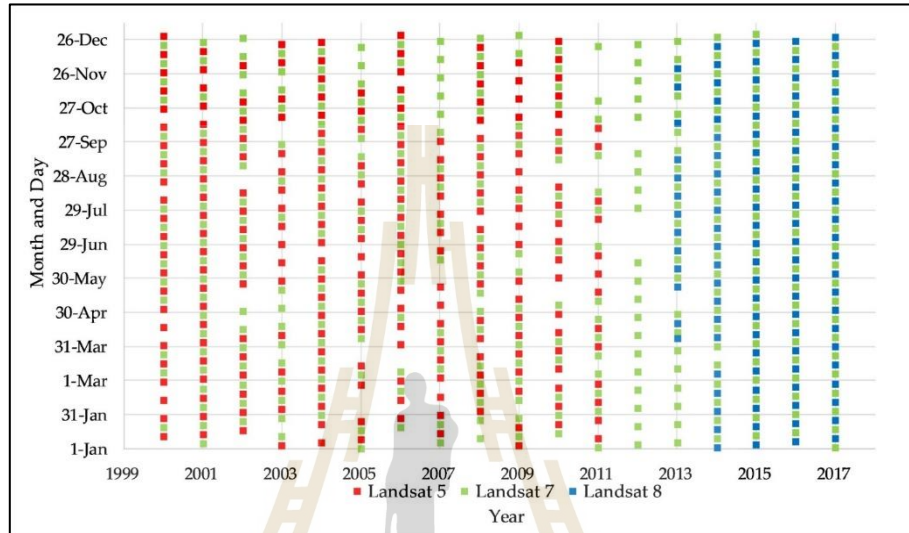


Figure 4.2 The distribution of all Landsat images from 2000 to 2017.

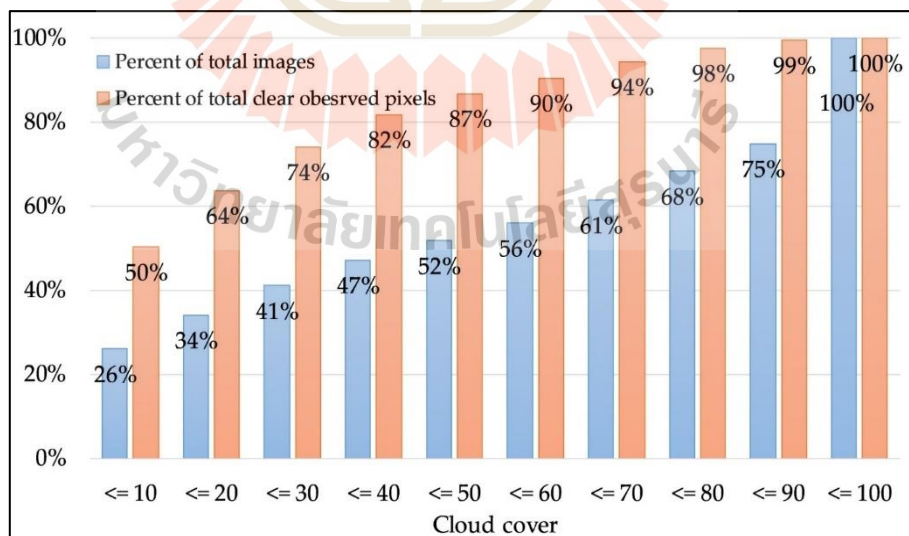


Figure 4.3 Percent of total images and percent of total clear pixels based on all available Landsat 5/7/8 images between 2000 and 2017.

After cloud coverage assessment, 168 images of Landsat 5, 251 images of Landsat7, and 76 images of Landsat 8 with a total of 495 images were finally selected to recognize cloud/shadow using the QA band and HA model.

4.2 Cloud/shadow pixel recognition base on QA band and HA model

4.2.1 Cloud/shadow pixel recognition base on QA band

The Landsat collection Level-1 quality assessment (QA) band allows users to apply per pixel filters to the Landsat 5 TM, Landsat 7 ETM, and Landsat 8 OLI/TIRS data products. Each pixel in the QA band contains unsigned integers that represent bit-packed combinations of surface, atmospheric, and sensor conditions that can affect the overall usefulness of a given pixel. Some common pixel values and their meanings of the QA band are summarized in Tables 4.1 and 4.2.

Table 4.1 Common pixel values of the QA band of Landsat 5, 7 data.

Pixel Value	Snow/Ice Confidence	Cloud Shadow Confidence	Cloud Confidence	Cloud	Radiometric Saturation	Dropped Pixel	Fill
672	Low	Low	Low	No	No	No	No
676	Low	Low	Low	No	1-2 bands	No	No
680	Low	Low	Low	No	3-4 bands	No	No
684	Low	Low	Low	No	5+ bands	No	No
704	Low	Low	Medium	No	No	No	No
752	Low	Low	High	Yes	No	No	No
756	Low	Low	High	Yes	1-2 bands	No	No
760	Low	Low	High	Yes	3-4 bands	No	No
928	Low	High	Low	No	No	No	No
1696	High	Low	Low	No	No	No	No
1700	High	Low	Low	No	1-2 bands	No	No
1704	High	Low	Low	No	3-4 bands	No	No

Source: <https://landsat.usgs.gov/collectionqualityband>

Table 4.2 Common pixel values of the QA band of Landsat 8 data.

Pixel Value	Cirrus Confidence	Snow/Ice Confidence	Cloud Shadow Confidence	Cloud Confidence	Cloud	Radiometric Saturation	Terrain Occlusion	Fill
2720	Low	Low	Low	Low	No	No	No	No
2724	Low	Low	Low	Low	No	1-2 bands	No	No
2728	Low	Low	Low	Low	No	3-4 bands	No	No
2752	Low	Low	Low	Medium	No	No	No	No
2800	Low	Low	Low	High	Yes	No	No	No
2804	Low	Low	Low	High	Yes	1-2 bands	No	No
2808	Low	Low	Low	High	Yes	3-4 bands	No	No
2976	Low	Low	High	Low	No	No	No	No
3744	Low	High	Low	Low	No	No	No	No
3748	Low	High	Low	Low	No	1-2 bands	No	No
3752	Low	High	Low	Low	No	3-4 bands	No	No
6816	High	Low	Low	Low	No	No	No	No
7072	High	Low	High	Low	No	No	No	No

Source: <https://landsat.usgs.gov/collectionqualityband>

4.2.2 Cloud/shadow pixel recognition base on HA model

Because cloud and snow make band 2 (Green band) brighter and cloud shadow and snow makes band 5 (SWIR1 band) darker, Zhu and Woodcock (2014b) estimated time-series models for band 2 and band 5 by comparing the actual Landsat observations with the corresponding model predictions. It was comparatively easy to identify any remaining clouds, cloud shadow, snow, and other ephemeral changes. If any of the conditions in Eq. 4.3 or 4.4 are met; this observation suggests that these pixels are clouds.

$$\rho(2, x) - \hat{\rho}(2, x)_{RIRLS} > 0.04 \quad (4.3)$$

$$\rho(5, x) - \hat{\rho}(5, x)_{RIRLS} < -0.04 \quad (4.4)$$

where x is Julian date, $\rho(i, x)$ is the observed value for the i^{th} Landsat band at Julian date x , $\hat{\rho}(i, x)_{RIRLS}$ is the predicted value for the i^{th} Landsat band at Julian date x .

After cloud coverage assessment, images with less than or equal to 90% of contaminated pixels in the study area were reselected using the QA band and HA model. As a result, 121 images of Landsat 5, 201 images of Landsat 7, 66 images of Landsat 8 with a total of 388 images are finally selected and used in this study. Additionally, the digital values of clear observed and contaminated pixels of all chosen images are 0 and 1, respectively. Examples of clear observed and contaminated pixels distribution from Landsat 5 and 7 are displayed in Figure 4.4. In Figure 4.4(a), most of the contaminated pixels are cloud and cloud shadow, whereas contaminated pixels in Figure 4.4(b) are cloud, cloud shadow, and gaps.

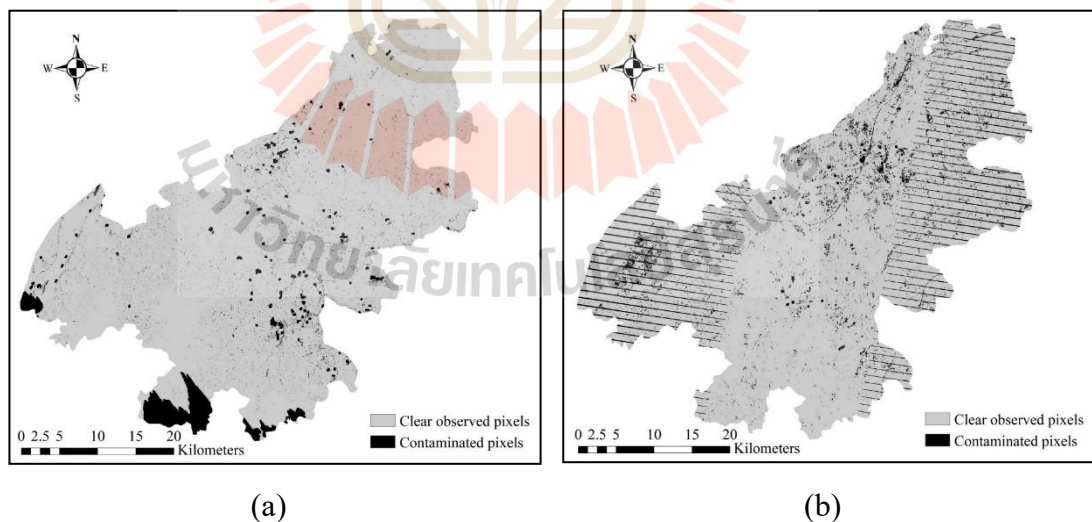


Figure 4.4 Examples of clearly observed and contaminated pixels in the study area: (a) Landsat 5, date 20 January 2000, and (b) Landsat 7, date 4 May 2009.

4.3 Conversion of digital number to spectral reflectance data

Radiative correction (including sensor calibration, atmospheric correction, terrain correction, relative radiation normalization, etc.) is a crucial step to ensure the homogeneity of the change detection from time-series data, and ignoring the results of this step is often ineffective (Vicente-Serrano, Pérez-Cabello, and Lasanta, 2008). The USGS uses the Landsat Ecosystem Interference Adaptive Processing System (LEDAPS) and the Landsat 8 Surface Reflection Code (LaSRC) to convert Landsat TM, ETM, and OLI level-1 data to surface reflectance Landsat Surface Reflectance Higher-Level Data Products (USGS, 2019a, 2019b). To make consistent Landsat Level 1 data products, known-quality archive data for timing analysis is required, USGS re-archives all Landsat data to Tier 1, Tier 2, and Real-Time. Among them, Level 1 data have been calibrated between sensors, and the accuracy of image geometry registration is within 0.5 pixels, so this high-quality data product is suitable for Landsat timing change detection.

Herein, the 16-bit integer values in the Level 1 product were converted to spectral reflectance using Eq. 4.5.

$$\rho\lambda' = M_{\rho} * Q_{cal} + A_{\rho} \quad (4.5)$$

where $\rho\lambda'$ is spectral reflectance, without correction for solar angle (Unitless), M_{ρ} is reflectance multiplicative scaling factor for the band from the metadata, A_{ρ} is reflectance additive scaling factor for the band from the metadata, Q_{cal} is the L1 pixel value in DN. The derived outputs are further applied for Component 2 to Component 4.

In this study, the final selection of all Landsat data with a total of 388 scenes was converted to spectral reflectance using Eq. 4.5 for data analysis. The digital values

of spectral reflectance of six bands of all selected scenes vary between 0 and 1. Examples of the spectral reflectance data distribution of NIR band from Landsat 5 and 7 are displayed in Figure 4.5. As a result, the existing gaps with a value less than 0.0 in Landsat 7 data, as shown in Figure 4.5(b), are here considered as contaminated pixels, and they are ignored in data analysis.

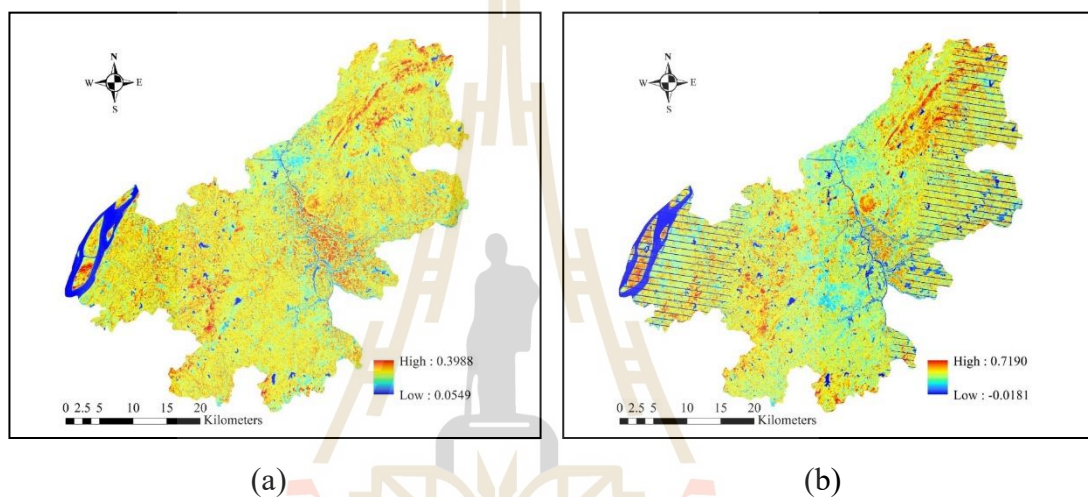


Figure 4.5 Spatial distribution of spectral reflectance composite images: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.

4.4 Time-series spectral reflectance reconstruction

Under this step, the clearly observed and contaminated pixels of Landsat imageries using the QA band and HA model (with value 0 or 1) from the step 2, Section 4.2 were used to create the spatiotemporal cube using reshape function under MATLAB software. Likewise, the derived six spectral reflectance bands (with values 0 to 1) from step 3, Section 4.3 were also used to create spectral reflectance spatiotemporal cube in the same manner. Then the information of clear observed and contaminated pixels and spectral reflectance of six bands were converted from spatial dimension to time

dimension using reshape function (read the data pixel by pixel from spatiotemporal cubes) of MATLAB software. After that, two-time dimension datasets were combined to remove contaminated pixels for time-series spectral reflectance reconstruction. (See Figure 4.6).

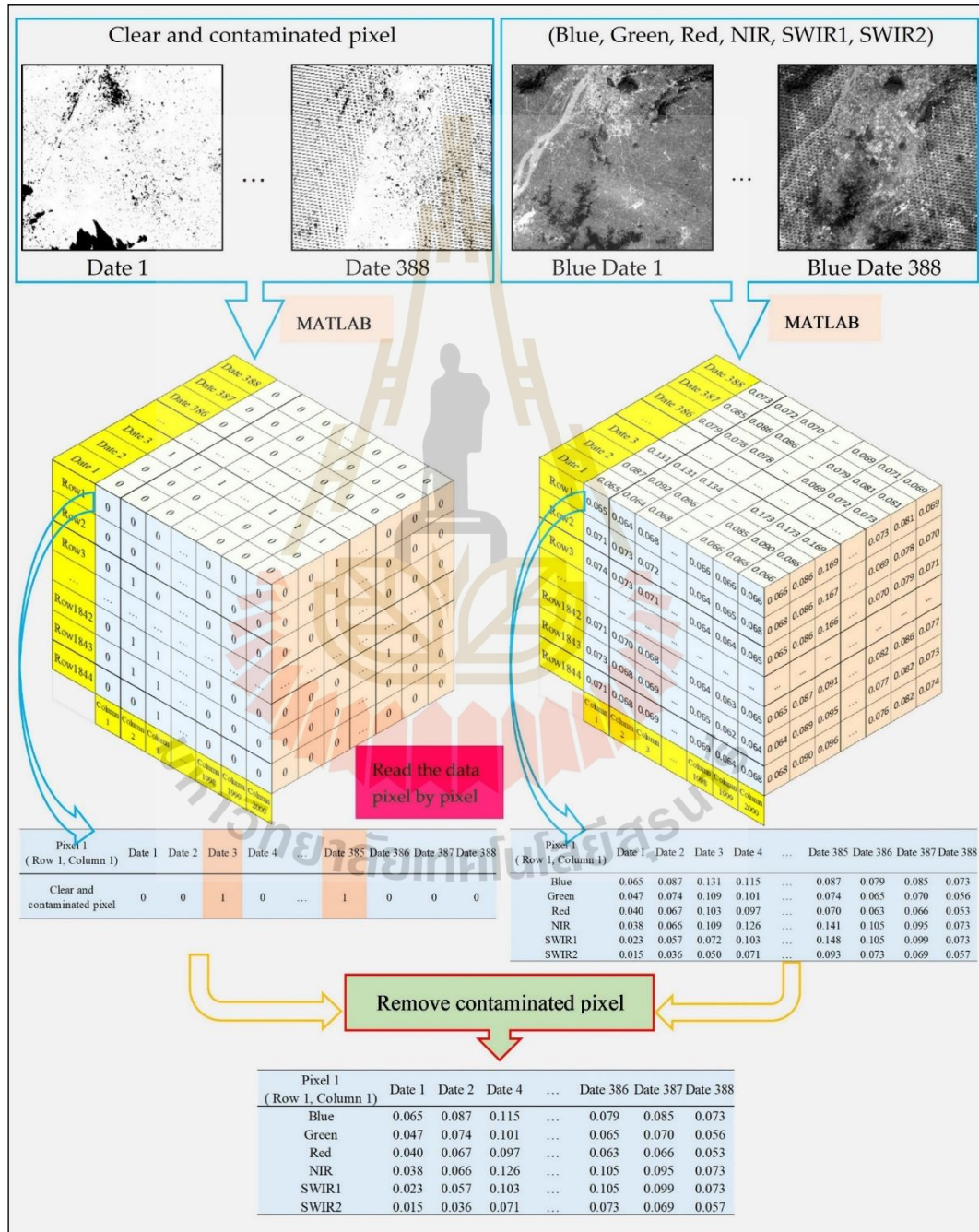


Figure 4.6 Time-series spectral reflectance reconstruction.

The derived time-series spectral reflectance data are further applied to calculate spectral indices (NDVI, NDBI, NDWI, EVI, MNDBI, MNDWI, TCB, TCG, TCW, PC1, PC2, and PC3) for time-series LULC classification.

By combination between the spatiotemporal cube of the recognized clear observed and contaminated pixels (with value 0 or 1) and the spatiotemporal cube of the derived spectral reflectance data of six bands (with value 0 to 1) using the reshape function of MATLAB software, the contaminated pixels were removed from original time-series spectral reflectance data of six bands.

Figure 4.7(a) on the left shows the original time-series spectral reflectance data of six bands (with value 0 to 1) from one pixel of 388 dates. It reveals that some pixel values deviate from the normal reflectance range due to contamination, such as cloud/cloud shadows. Meanwhile, Figure 4.7(b) on the right shows the time-series spectral reflectance data of six bands from the same pixel after the removal of contaminated values on some dates. It reveals that the reflectance values oscillate up and down with temporal variations. The reconstructed time-series spectral reflectance data of 388 images are the primary input data for multitemporal LULC mapping using HA with a minimum spectral distance algorithm in this study.

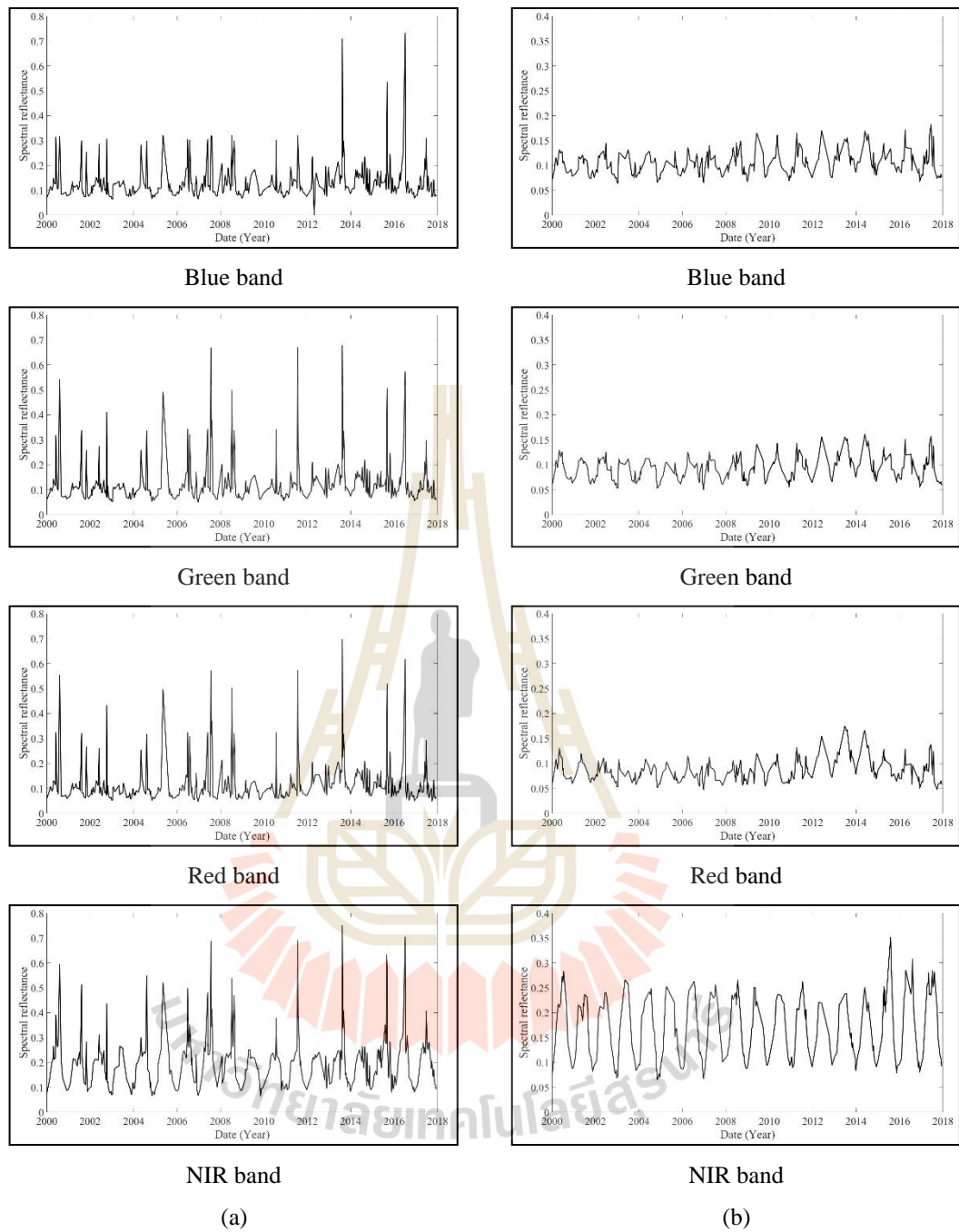


Figure 4.7 Comparison of (a) original time-series spectral reflectance data of six bands from one pixel between 2000 and 2017 and (b) time-series spectral reflectance data of six bands from the same pixel after removal of contaminated values.

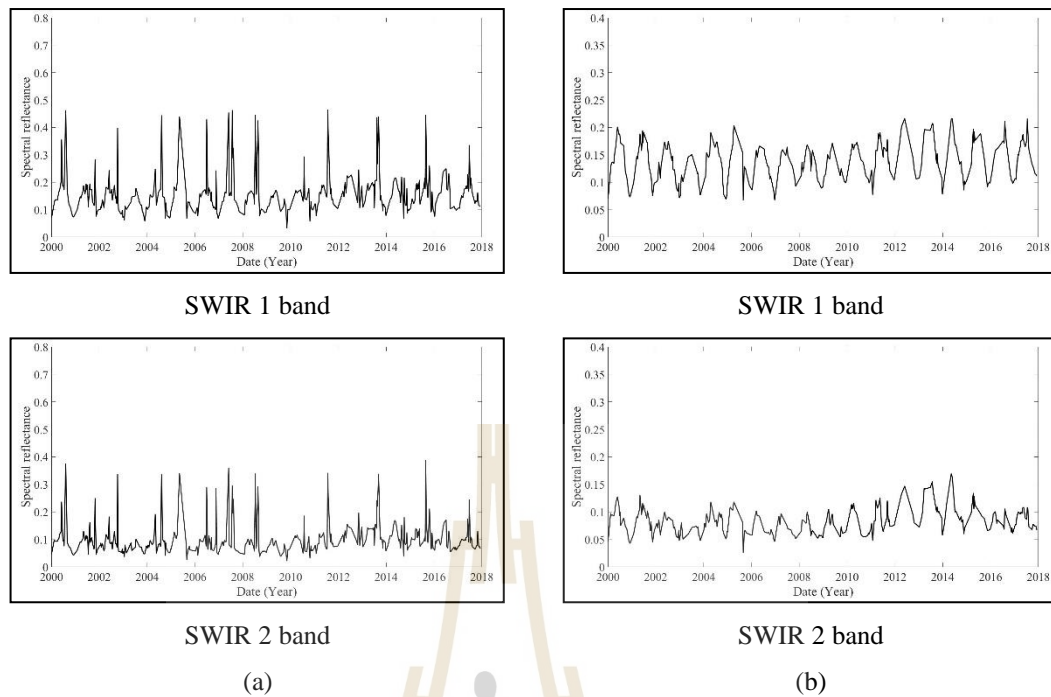
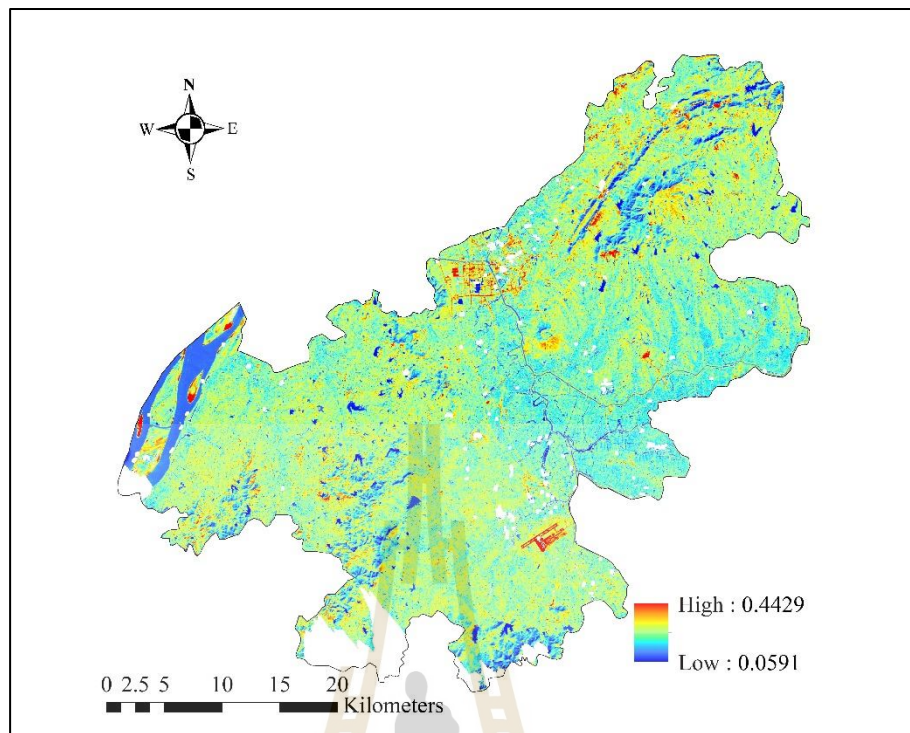
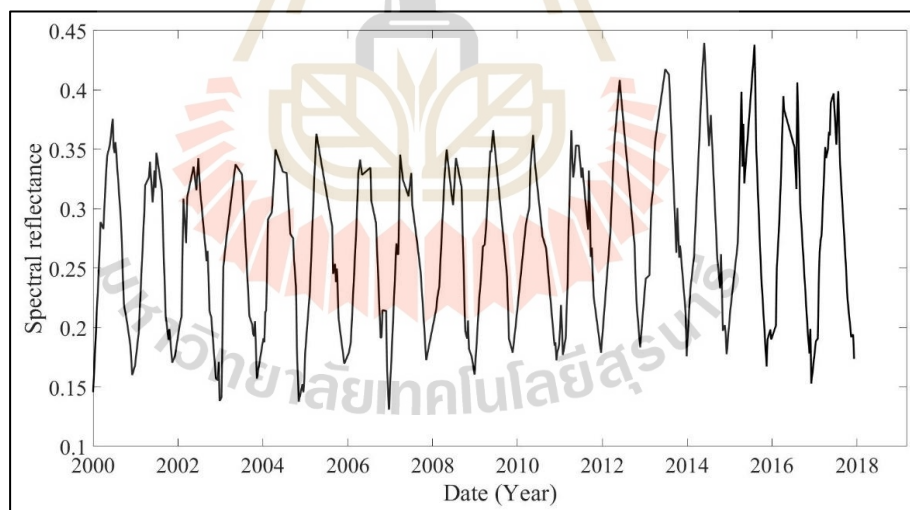


Figure 4.7 (Continued).

Figure 4.8 displays examples of the derived spectral indices (TCB), which were selected as a candidate image for spectral feature selection. Figure 4.8(a) on the top shows the one scene of TCB values on 20 January 2000 from a total of 388 scenes, the white areas in the study area show the contaminated pixels, and they are ignored in data analysis. Figure 4.8(b) on the bottom shows the TCB values of one selected pixel of the whole period from 2000 to 2017.



(a)



(b)

Figure 4.8 Examples of the derived spectral indices (TCB): (a) one scene of the whole study area, date 20 January 2000, and (b) one pixel from 2000 to 2017.

SUMMARY

The preprocessing of time-series Landsat data was here systematically implemented to produce corrected images due to internal and external errors that generally exist in remotely sensed data. In general, internal errors are introduced by the remote sensing system (e.g., uncalibrated detectors), while external errors are introduced by natural phenomena (e.g., atmosphere, terrain elevation, slope, aspect) that vary in nature through space and time. As a result, 121 images of Landsat 5, 201 images of Landsat 7, 66 images of Landsat 8 with a total of 388 images were finally selected for data analysis in this study.

Besides, time-series spectral reflectance dataset was constructed from a clearly observed and contaminated pixel indicator cube and six spectral reflectance cubes. These cubes were converted from 388 images by reshape function under MATLAB. Finally, the derived time-series spectral reflectance dataset was further applied to calculate time-series spectral index dataset (NDVI, NDBI, NDWI, EVI, MNDBI, MNDWI, TCB, TCG, TCW, PC1, PC2, and PC3). Time-series spectral reflectance and spectral indices datasets are further selected as optimum spectral features for time-series LULC classification using unsupervised and supervised methods.

CHAPTER V

SPECTRAL FEATURE SELECTION FOR LAND USE AND LAND COVER CLASSIFICATION

The workflow of spectral features selection for LULC classification is schematically displayed in Figure 5.1. The detail of two major tasks of spectral features selection and their result included (1) stable pixel of LULC type extraction and (2) selection of spectral features are separately described and discussed in the following section.

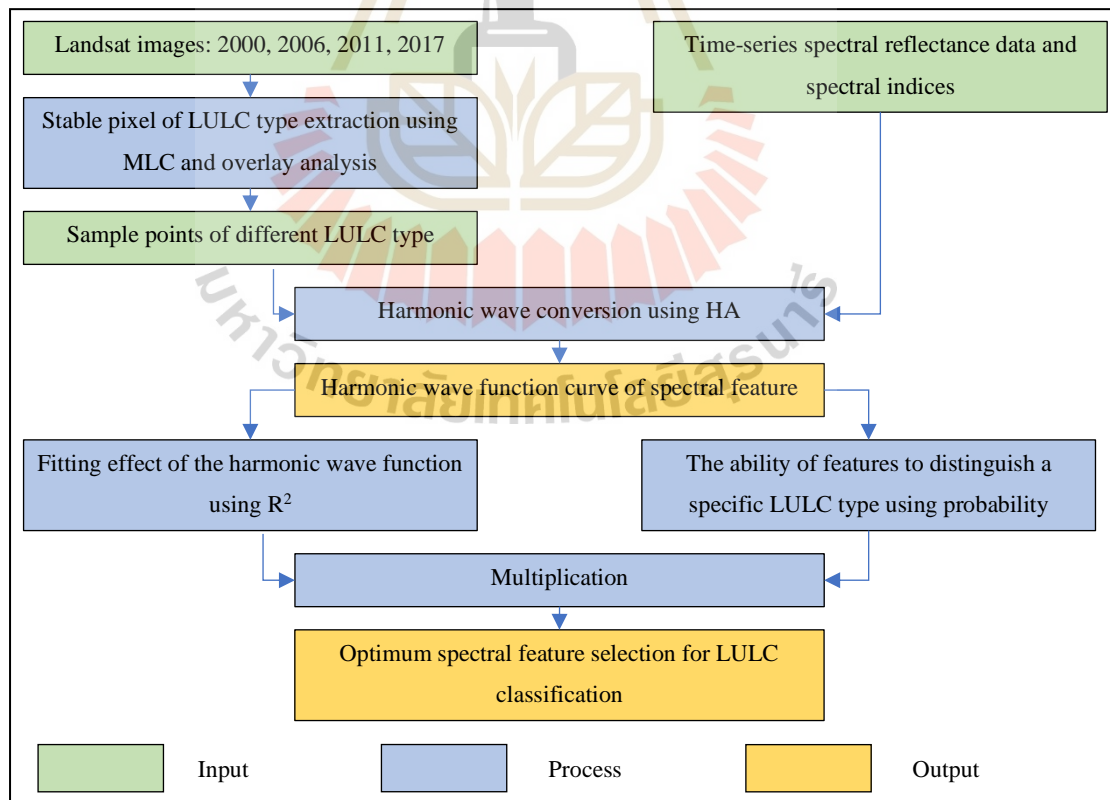
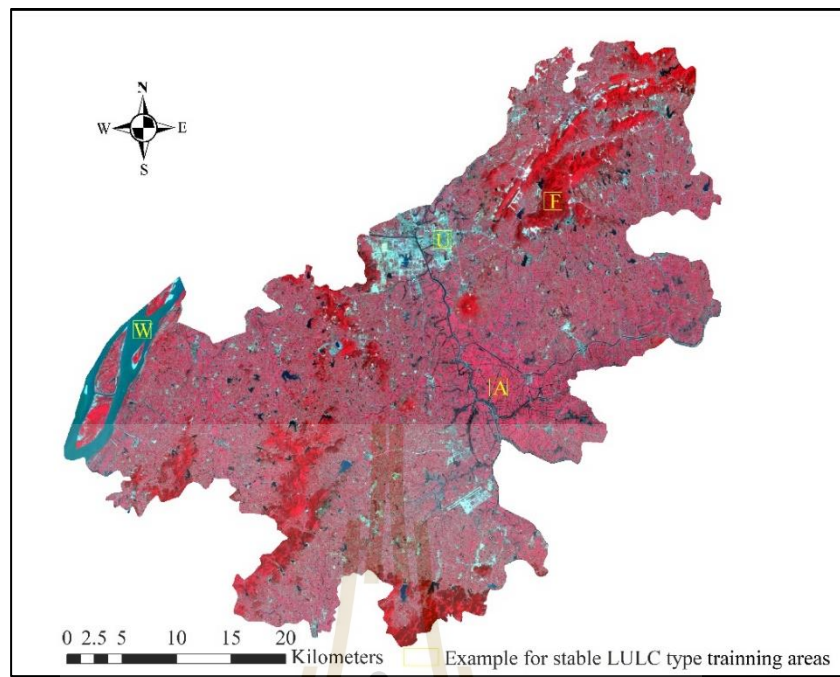


Figure 5.1 Workflow of Component 2: Spectral features selection for LULC classification.

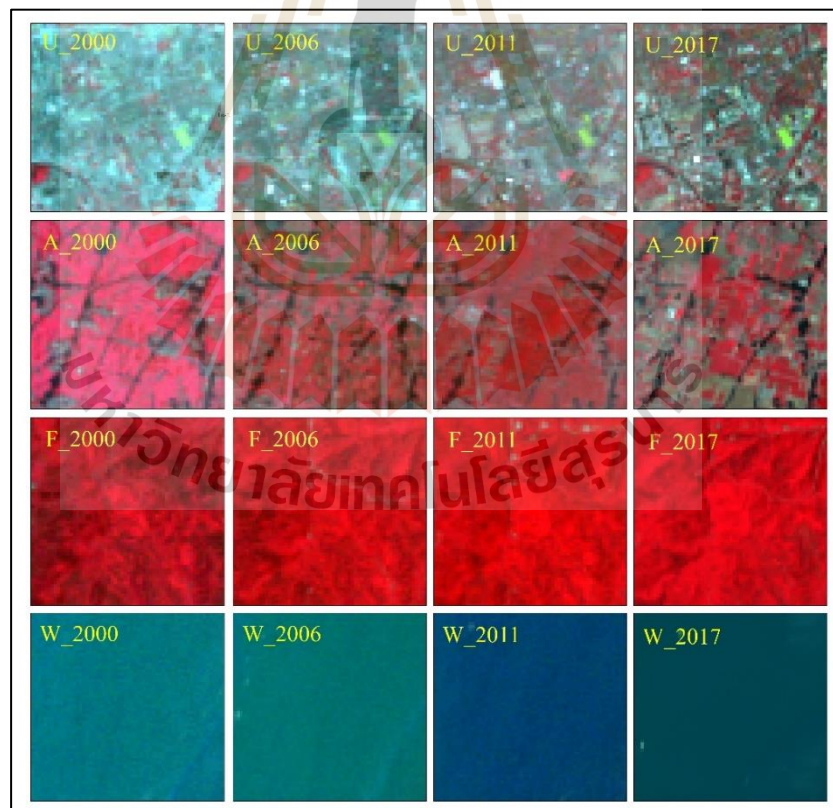
5.1 Stable pixel of LULC type extraction

In the study area, there are hundreds of Landsat images between 2000 and 2017. The traditional method of LULC classification using spectral features is time-consuming and labor-intensive. This research attempts to establish a method for LULC classification using temporal features, but it requires to understand the spectral characteristics of Brightness Value (BV) of each LULC type throughout times, particularly historical stable pixels of four main LULC types (urban and built-up land, U; agricultural land, A; forest land, F; and water bodies, W between 2000 and 2017 (18 years).

The different LULC types (U, A, F and W) in 2000, 2006, 2011, and 2017 were first classified using maximum likelihood classifier (MLC) under the ERDAS Imagine software (See example for stable LULC type selection between 2000 and 2017 in Figure 5.2). Then, four classified LULC maps in 2000, 2006, 2011, and 2017 were simultaneously superimposed to identify common areas of each LULC type from four different years by overlay analysis under the ESRI ArcMap software. The derived output presents the stable pixels of each LULC category between 2000 and 2017 and LULC change in this period.



(a)



(b)

Figure 5.2 Example of stable history pixel selected from Jiangning district in different four years: (a) location of different stable LULC types and (b) zoom-in maps.

After that, sample points from each stable LULC type were randomly selected by using the create random points function of the ESRI ArcMap software, and they then were converted to grid format points (30 * 30m).

Finally, the location of sample points of each LULC type was further used to transform standard harmonic function curves of 18 spectral features for four different LULC types using three categories of time-series models in the following equations.

CASE 1: HA models include the 12-month harmonic components.

$$y = a_0 + b_0 t + A_1 \cos\left(\frac{2\pi}{T} t - \varphi_1\right) \quad (5.1)$$

CASE2: HA models include the 12-month, 6-month harmonic components.

$$y = a_0 + b_0 t + A_1 \cos\left(\frac{2\pi}{T} t - \varphi_1\right) + A_2 \cos\left(\frac{4\pi}{T} t - \varphi_2\right) \quad (5.2)$$

CASE3: HA models include the 12-month, 6-month, and 3-month harmonic components.

$$y = a_0 + b_0 t + A_1 \cos\left(\frac{2\pi}{T} t - \varphi_1\right) + A_2 \cos\left(\frac{4\pi}{T} t - \varphi_2\right) + A_3 \cos\left(\frac{8\pi}{T} t - \varphi_3\right) \quad (5.3)$$

Where, t is the Julian date, T is the number of days per year ($T= 365$), a_0 is the coefficient of intercept value, b_0 is the coefficient of slope value, A_i is the coefficient of amplitude value, φ_i is the coefficient of phase value, y is the reconstructed reflectance value at Julian date t .

The result of LULC classification by using MLC for stable LULC extraction is presented in Figure 5.3.

As a result, it can be observed that urban and built-up land has been continuously increased while agriculture land has been continuously decreased. In the meantime, some forest land and most of the water bodies are stable.

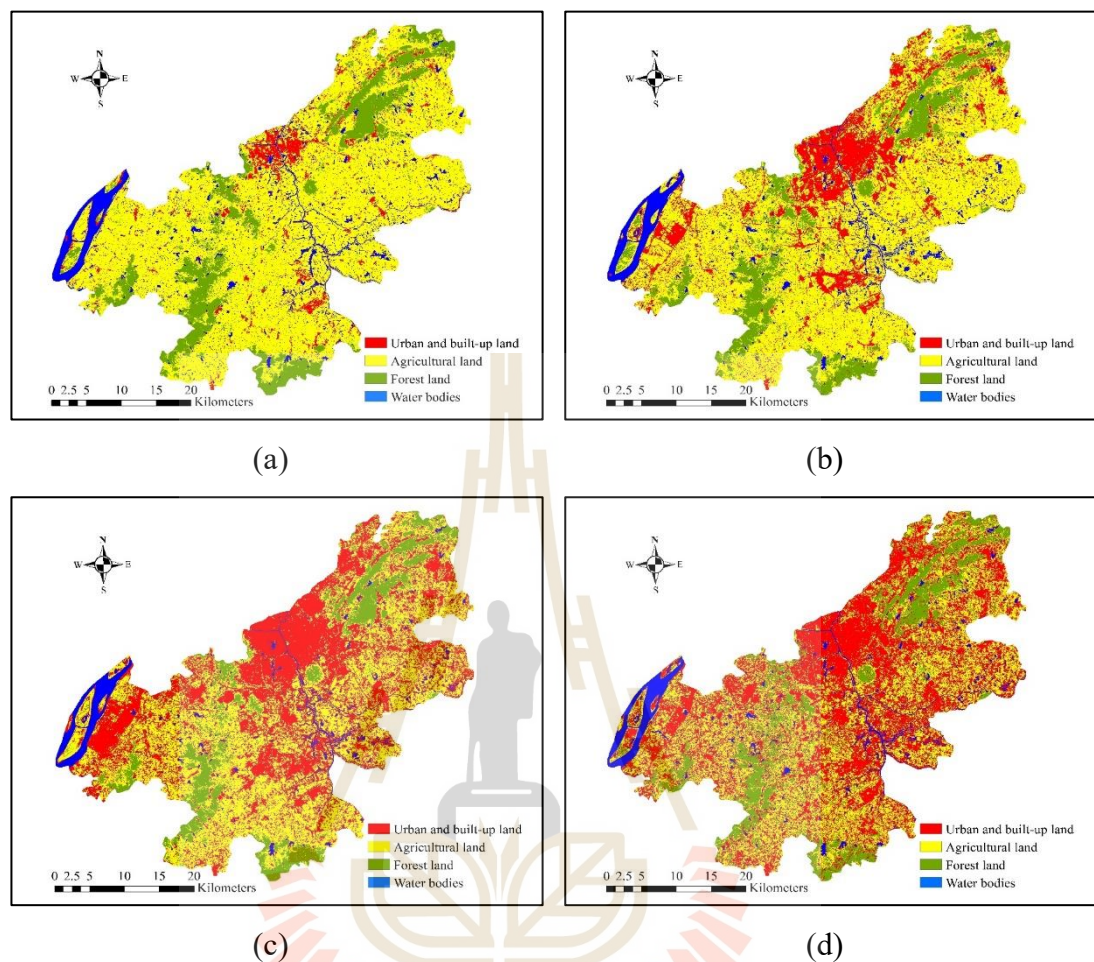


Figure 5.3 Spatial distribution of LULC classification map in: (a) 2000, (b) 2006, (c) 2011, and (d) 2017.

These classified LULC maps were applied to identify common areas of each LULC type in four different years (2000, 2006, 2011, and 2017), as results show in Figure 5.4 and Table 5.1. The results indicate change and unchanged among LULC types in four different years. It was found that stable areas of urban and built-up land, agriculture land, forest land, and water bodies are 62.05 km² (3.91%), 335.40 km² (21.14%), 129.59 km² (8.17%) and 49.82 km² (3.14%), respectively while area of changes in four different years was 1,009.87 km² (63.64%). This result indicated

dramatic LULC change due to urbanization occurs from 2000 to 2017, particularly increasing urban and built-up land. This finding is consistent with the statistical report of Nanjing Municipal Bureau Statistics (NJMBS, 2017).

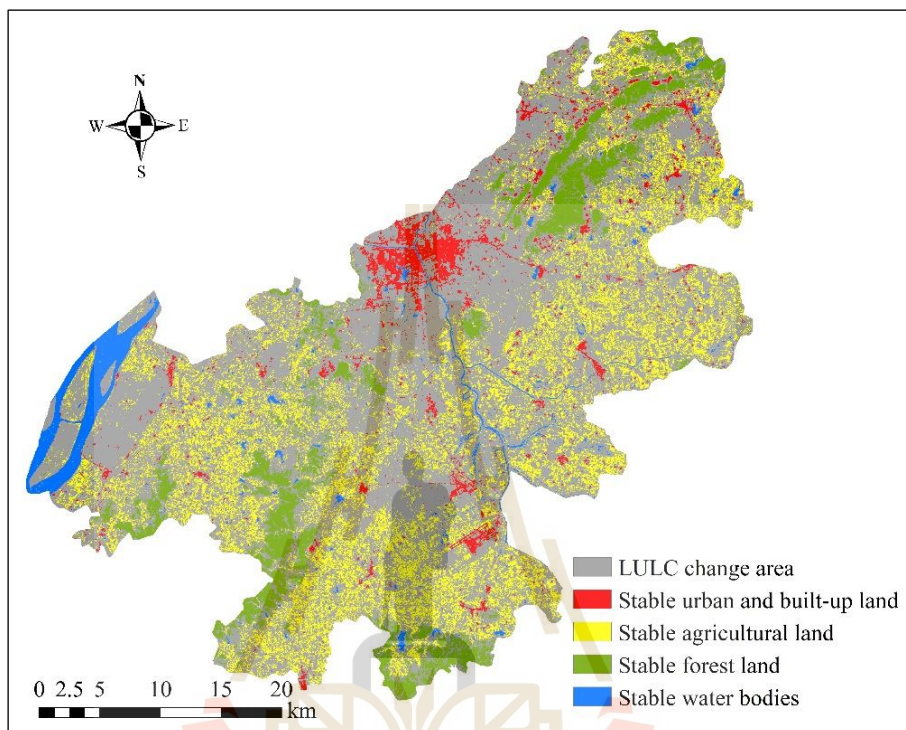


Figure 5.4 Stable and change of LULC type between 2000 and 2017.

Table 5.1 Area and percentage of stable LULC type and LULC change between 2000 and 2017.

No	LULC type	Area in km ²	Percent (%)
1	Stable urban and built-up land	62.05	3.91
2	Stable agriculture land	335.40	21.14
3	Stable forest land	129.59	8.17
4	Stable water bodies	49.82	3.14
5	LULC change during 2000 to 2017	1,009.87	63.64
Total		1,586.74	100

Besides, the random sample points of each stable LULC type spatially display in Figure 5.5. As a result, random sample points of U, A, F, and W are 9,395 pixels, 9,893 pixels, 9,711 pixels, and 9,181 pixels, respectively. These sample points are further used to construct the harmonic function curve of each LULC type by time-series model for optimum spectral feature selection.

Then, these vector points were converted to grid format points (30x30m). As a result, 9,395 pixels of urban and built-up land, 9,893 pixels of agricultural land, 9,711 pixels of forest land, and 9,181 pixels of water bodies are further applied to select spectral features for LULC classification.

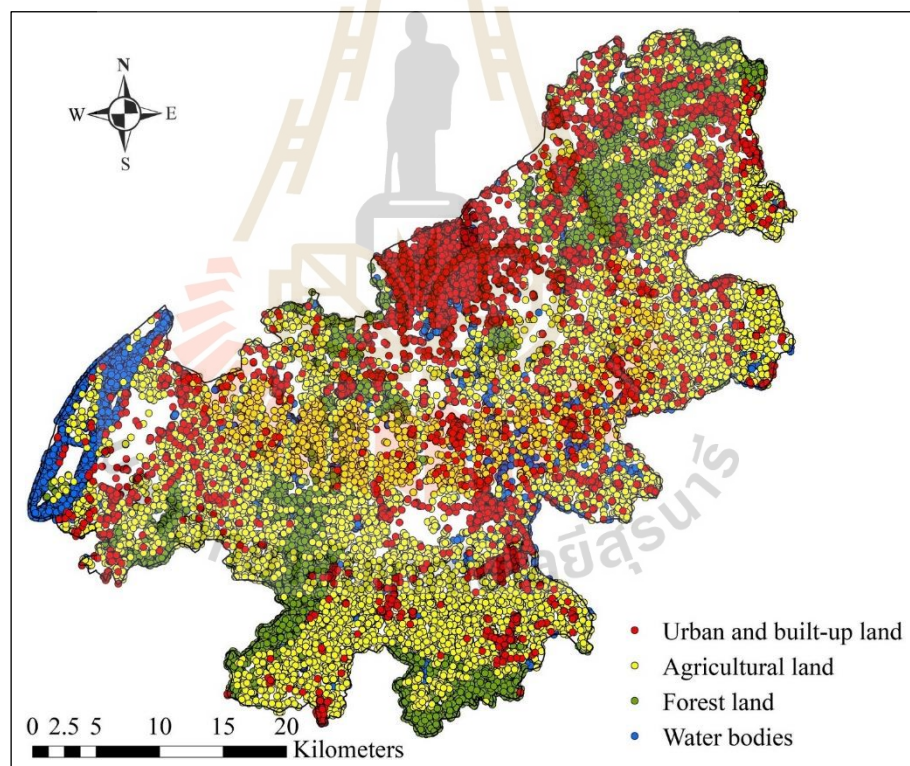


Figure 5.5 Sample points of the historical stable pixel of each LULC type.

5.2 Selection of spectral feature

Standard harmonic curves of four different LULC types from 18 spectral features included Blue, Green, Red, NIR, SWIR1, SWIR2, NDVI, NDBI, NDWI, EVI, MNDBI, MNDWI, TCB, TCG, TCW, PC1, PC2 and PC3 were here applied for spectral feature selection. Selection of spectral feature was operated in two steps:

Step 1. Fitting effect of the harmonic wave function

The fitting effect of the harmonic wave function of each LULC type from three cases was calculated using the coefficient of determination (R^2) for spectral feature selection. Besides, the derived R^2 values from different cases were compared and examined the significant mean difference using the T-test. The spectral features obtained in this step indicates that these features work best when fitted with the HA model.

Step 2. The ability of features to distinguish a specific LULC type

The result of step 1 indicates features that can work best when they fitted with the HA model, but it does not mean that these features can effectively distinguish four different LULC types. Therefore, the ability of different spectral features to distinguish four LULC types is here examined by calculating the overlapping area between two LULC types (pair by pair) based on probability density function under the MATLAB software.

Finally, average R^2 which represents fitting of the harmonic wave function and the average value of ability of features to distinguish a specific LULC type, were combined using multiplication operation for LULC classification with optimum spectral feature (single and multiple features).

The main result of spectral feature selection is firstly described and discussed according to two main steps, and optimum spectral feature selection is then reported.

5.2.1 Spectral feature selection by the fitting effect of harmonic wave

The selected historical stable pixels for each LULC type were firstly applied to construct harmonic wave and calculated coefficients for three cases using Eq. 5.1, 5.2, and 5.3 and then calculated R^2 for measuring the fitting effect of the harmonic wave that reflects changes of spectral features over time.

Figures 5.6 to 5.9 show an example of HA for three cases from one selected stable pixel by using Tasseled Cap Brightness (TCB) data for each LULC type, namely, urban and built-up land, agricultural land, forest land, and water bodies.

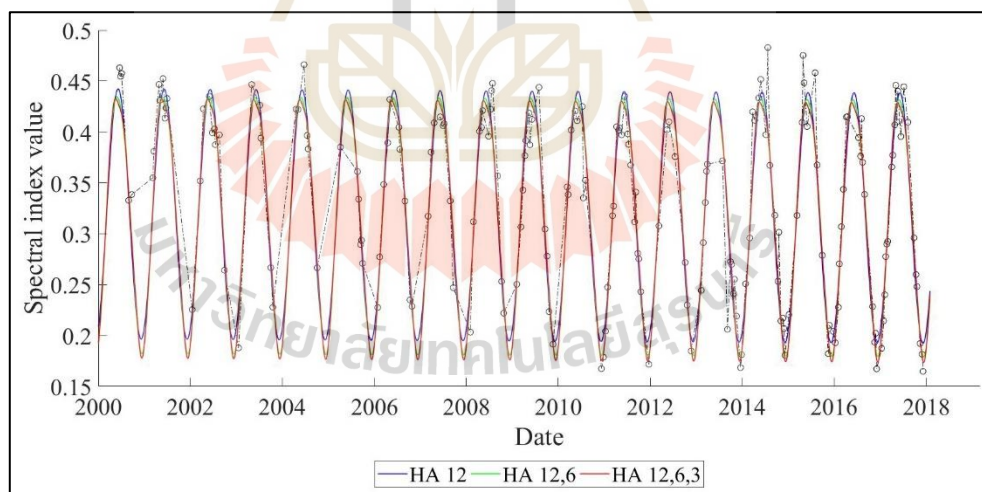


Figure 5.6 An example of HA for three cases of urban and built-up land by TCB data.

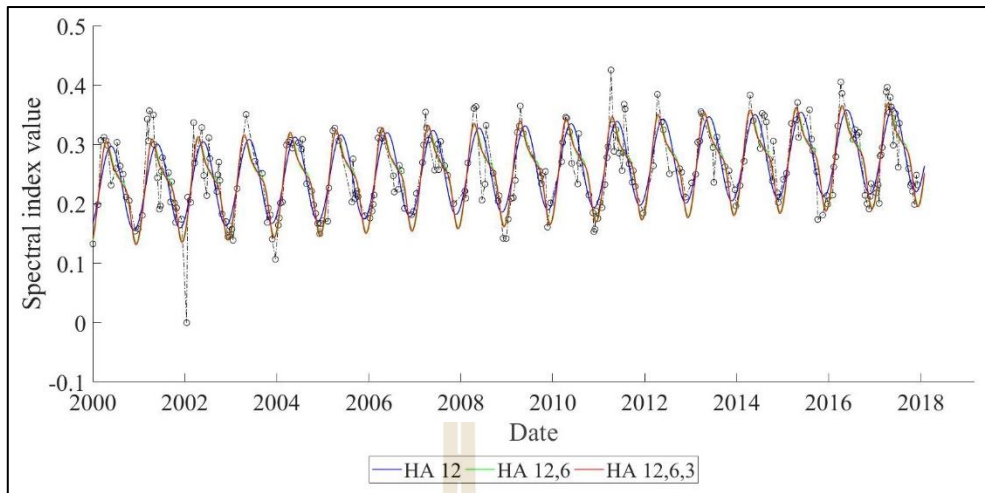


Figure 5.7 An example of HA for three cases of agricultural land by TCB data.

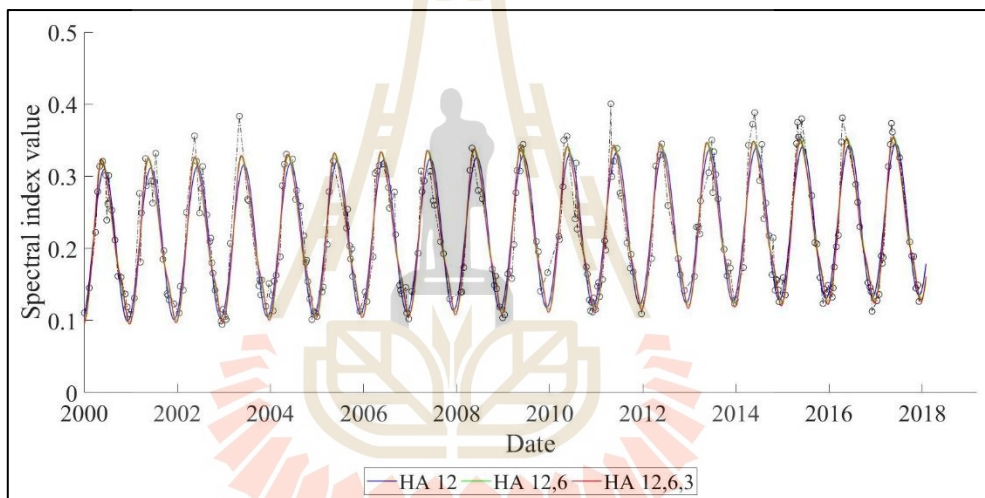


Figure 5.8 An example of HA for three cases of forest land by TCB data.

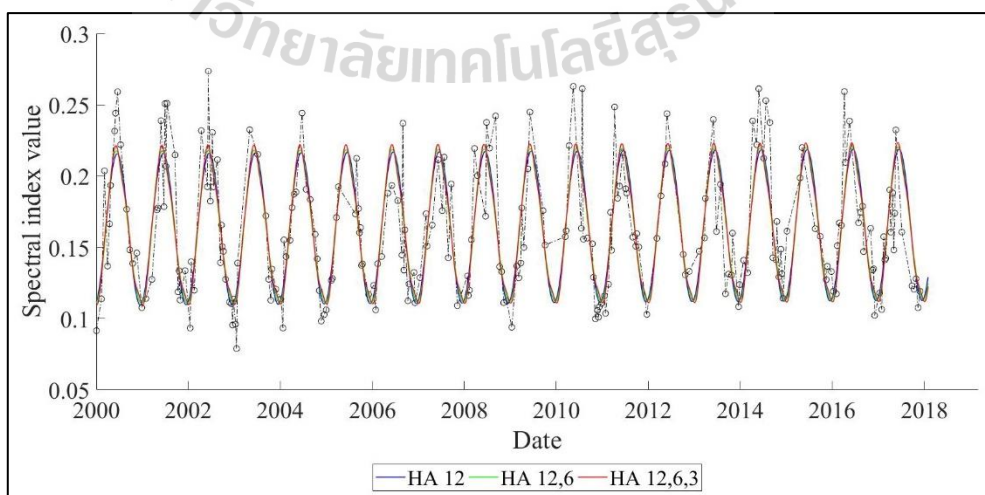


Figure 5.9 An example of HA for three cases of water bodies by TCB data.

As a result, it can be observed that the same case of the HA model has different fitting effects for different LULC types. A different case of the HA model has different fitting effects on the same LULC type. In order to quantitatively understand the difference in the fitting effect, coefficient of determination (R^2) is introduced here. So, each sample pixel of each LULC type, the fitting effect of the harmonic wave function were calculated using mean R^2 as a summary in Tables 5.2 to 5.5.

Table 5.2 Comparison of mean R^2 value of urban and built-up land from various selected features.

Feature	Urban and built-up land		
	HA12	HA12,6	HA12,6,3
BLUE	0.722	0.728	0.731
GREEN	0.767	0.778	0.780
RED	0.739	0.757	0.759
NIR	0.853	0.865	0.867
SWIR1	0.800	0.818	0.819
SWIR2	0.703	0.741	0.743
NDVI	0.476	0.501	0.509
EVI	0.681	0.691	0.696
NDBI	0.225	0.237	0.244
MNDBI	0.302	0.317	0.323
NDWI	0.458	0.490	0.498
MNDWI	0.241	0.264	0.270
TCB	0.852	0.866	0.867
TCG	0.194	0.204	0.217
TCW	0.268	0.277	0.279
PC1	0.821	0.836	0.837
PC2	0.314	0.318	0.344
PC3	0.024	0.066	0.080

Table 5.3 Comparison of mean R^2 value of agricultural land from various selected features.

Feature	Agricultural land		
	HA12	HA12,6	HA12,6,3
BLUE	0.528	0.536	0.544
GREEN	0.599	0.627	0.633
RED	0.420	0.484	0.497
NIR	0.781	0.822	0.827
SWIR1	0.632	0.672	0.678
SWIR2	0.401	0.480	0.491
NDVI	0.526	0.614	0.626
EVI	0.706	0.761	0.770
NDBI	0.448	0.503	0.516
MNDBI	0.470	0.521	0.531
NDWI	0.488	0.579	0.591
MNDWI	0.152	0.176	0.189
TCB	0.802	0.834	0.836
TCG	0.439	0.502	0.521
TCW	0.220	0.234	0.239
PC1	0.784	0.811	0.813
PC2	0.055	0.078	0.096
PC3	0.048	0.099	0.119

Table 5.4 Comparison of mean R^2 value of forest land from various selected features.

Feature	Forest land		
	HA12	HA12,6	HA12,6,3
BLUE	0.495	0.500	0.507
GREEN	0.566	0.579	0.582
RED	0.351	0.365	0.376
NIR	0.877	0.902	0.908
SWIR1	0.860	0.875	0.877
SWIR2	0.656	0.707	0.721
NDVI	0.651	0.680	0.694
EVI	0.869	0.888	0.895
NDBI	0.559	0.614	0.635
MNDBI	0.622	0.661	0.676
NDWI	0.651	0.680	0.693
MNDWI	0.260	0.280	0.285
TCB	0.873	0.891	0.893
TCG	0.664	0.695	0.714
TCW	0.216	0.222	0.227
PC1	0.877	0.891	0.893
PC2	0.259	0.341	0.351
PC3	0.049	0.098	0.108

Table 5.5 Comparison of mean R^2 value of water bodies from various selected features.

Feature	Water bodies		
	HA12	HA12,6	HA12,6,3
BLUE	0.638	0.641	0.649
GREEN	0.702	0.710	0.716
RED	0.650	0.658	0.665
NIR	0.556	0.564	0.567
SWIR1	0.358	0.396	0.403
SWIR2	0.316	0.366	0.374
NDVI	0.312	0.332	0.337
EVI	0.382	0.411	0.418
NDBI	0.220	0.253	0.265
MNDBI	0.205	0.235	0.246
NDWI	0.326	0.347	0.351
MNDWI	0.158	0.196	0.205
TCB	0.616	0.624	0.629
TCG	0.483	0.499	0.509
TCW	0.269	0.286	0.291
PC1	0.393	0.405	0.412
PC2	0.247	0.263	0.267
PC3	0.067	0.112	0.160

Results in Tables 5.2 to 5.5 show that mean R^2 values of CASE 3 are better than CASE2, and CASE 2 is better than CASE 1 for different spectral bands and index of four LULC types. As a consequence, the mean R^2 values for different three CASEs are similar.

Meanwhile, Figures 5.10 to 5.12 depict the R^2 distribution of PC1 data in different three cases of agricultural land sample pixels.

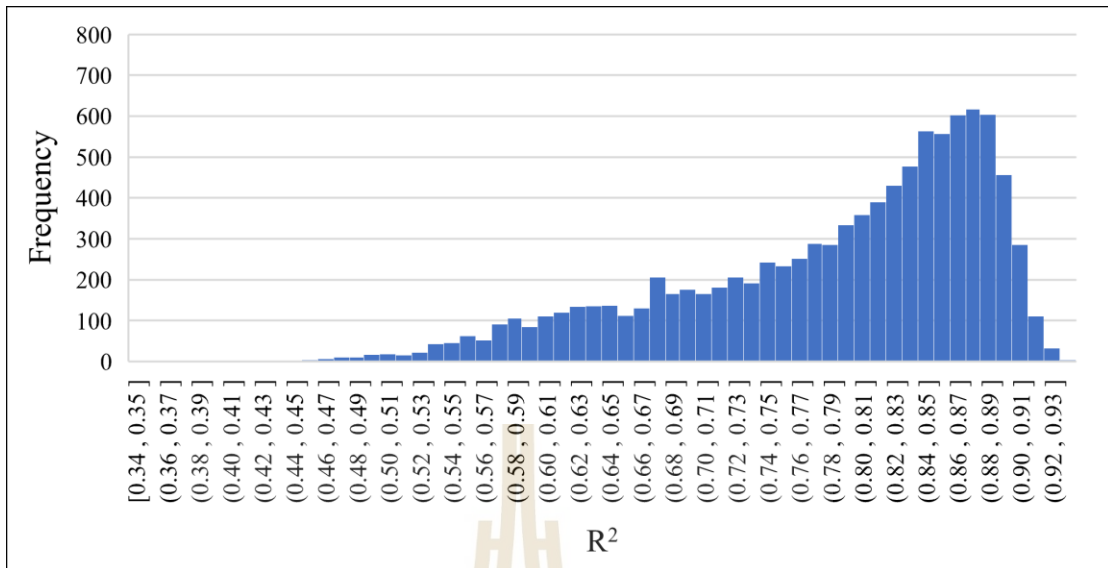


Figure 5.10 Distribution of R^2 of agricultural land by PC1 data for Case 1.

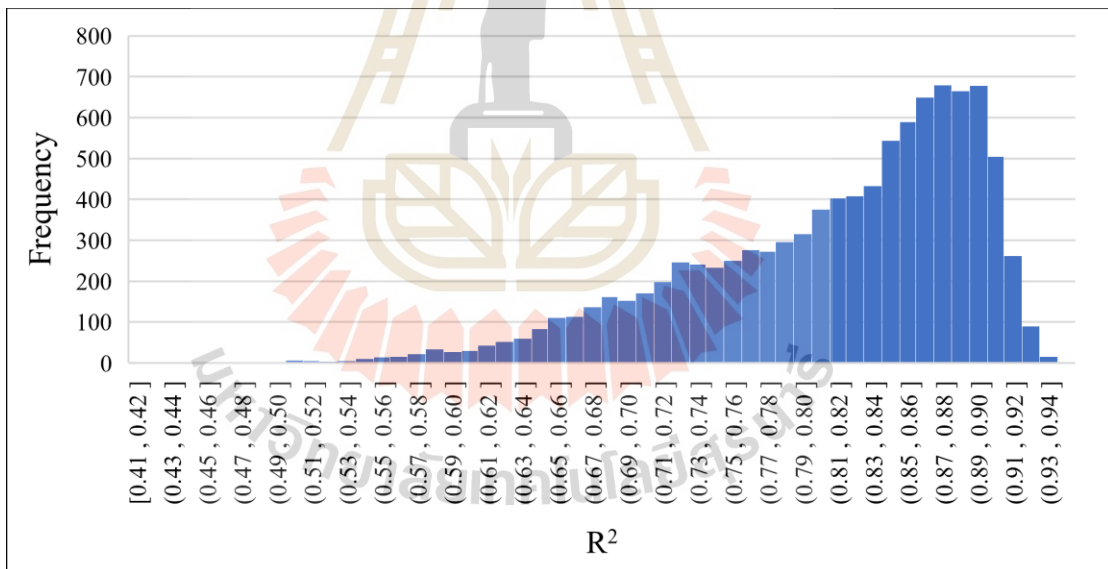


Figure 5.11 Distribution of R^2 of agricultural land by PC1 data for Case 2.

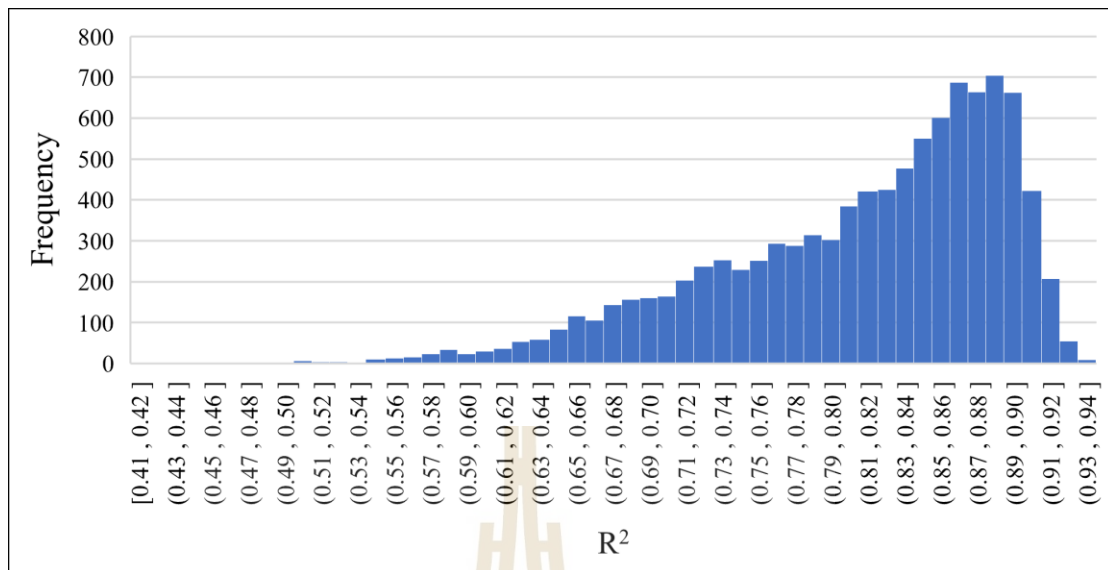


Figure 5.12 Distribution of R^2 of agricultural land by PC1 data for Case 3.

As results in Figures 5.10 to 5.12, R^2 distribution of PC1 data in different three cases of agricultural land sample pixels indicates that R^2 is different in different cases.

Therefore, in order to understand whether the difference in R^2 is significant in different cases, the R^2 from all sample pixels were here used for significant different test using T-test and Z-test methods.

Table 5.6 and Table 5.7 shows a result of the significant test (T-test and Z-test) for four LULC types of 18 features at a significance level of 0.05. All values are close to 0 and less than 1.96, which means that CASE1 and CASE2 are not significantly different. Similarly, CASE1 and CASE3 are also not significantly different.

Table 5.7 R^2 significant test (Z-test) of three CASEs (significance level of 0.05).

Feature	Urban and built-up land		Agricultural land		Forest land		Water bodies	
	CASE1,	CASE1,	CASE1,	CASE1,	CASE1,	CASE1,	CASE1,	CASE1,
	CASE 2	CASE 3	CASE 2	CASE 3	CASE 2	CASE 3	CASE 2	CASE 3
BLUE	1.61E-13	0	0	0	0	0	2.81E-06	0
GREEN	0	0	0	0	0	0	7.07E-12	0
RED	0	0	0	0	0	0	0.000109	2.11E-13
NIR	0	0	0	0	0	0	1.55E-15	0
SWIR1	0	0	0	0	0	0	0	0
SWIR2	0	0	0	0	0	0	0	0
NDVI	0	0	0	0	0	0	0	0
EVI	2.91E-09	0	0	0	0	0	0	0
NDBI	2.16E-07	4.44E-16	0	0	0	0	0	0
MNDBI	1.45E-08	6.66E-15	0	0	0	0	0	0
NDWI	0	0	0	0	0	0	0	0
MNDWI	0	0	0	0	2.29E-10	2.66E-15	0	0
TCB	0	0	0	0	0	0	7.22E-14	0
TCG	1.34E-09	0	0	0	0	0	2.54E-10	0
TCW	4.51E-09	5.71E-14	0	0	1.64E-05	0	0	0
PC1	0	0	0	0	0	0	0	0
PC2	0.000334	0	0	0	0	0	0	0
PC3	0	0	0	0	0	0	0	0

Therefore, CASE 1: HA model with the 12-month harmonic components is used in this study due to simple operation. However, the previously calculated R^2 is the fitting effect of the HA model on a certain LULC for different spectral features. In practice, the LULC type of each pixel is unknown before the LULC classification. Consequently, the average R^2 of four LULC types were applied for ranking spectral features for various LULC types classification as a summary in Table 5.8.

In summary, the first six spectral features with larger average R^2 values include TCB, NIR, PC1, GREEN, SWIR1, and EVI.

Table 5.8 Average R² value of four LULC types.

No	Feature	Urban and built-up land	Agricultural land	Forest land	Water bodies	Average	Ranking
1	BLUE	0.722	0.528	0.495	0.638	0.596	7
2	GREEN	0.767	0.599	0.566	0.702	0.658	6
3	RED	0.739	0.420	0.351	0.650	0.540	8
4	NIR	0.853	0.781	0.877	0.556	0.767	2
5	SWIR1	0.800	0.632	0.860	0.358	0.662	4
6	SWIR2	0.703	0.401	0.656	0.316	0.519	9
7	NDVI	0.476	0.526	0.651	0.312	0.491	10
8	EVI	0.681	0.706	0.869	0.382	0.660	5
9	NDBI	0.225	0.448	0.559	0.220	0.363	14
10	MNDBI	0.302	0.470	0.622	0.205	0.400	13
11	NDWI	0.458	0.488	0.651	0.326	0.481	11
12	MNDWI	0.241	0.152	0.260	0.158	0.203	17
13	TCB	0.852	0.802	0.873	0.616	0.786	1
14	TCG	0.194	0.439	0.664	0.483	0.445	12
15	TCW	0.268	0.220	0.216	0.269	0.243	15
16	PC1	0.821	0.784	0.877	0.393	0.719	3
17	PC2	0.314	0.055	0.259	0.247	0.219	16
18	PC3	0.024	0.048	0.049	0.067	0.047	18

5.2.2 Ability of features to distinguish a specific LULC type

The spectral feature selection obtained in the previous section only indicates that those features work best when fitted with the HA model, but it does not mean that these features can effectively distinguish four different LULC types. So, the ability of spectral features to distinguish four LULC types was here examined by calculating the overlapping area between two LULC types (pair by pair) by using probability density function.

In order to understand the histogram distribution of the four LULC types in different features, one LULC map that has been classified at a specific time is

required. Herein, the LULC classification map dated 18 May 2017 was used to identify spectral features that can distinguish four LULC types.

Figure 5.13 shows an example of the probability distribution of four LULC types from a blue band with specific symbols: red for urban and built-up land, yellow for agricultural land, green for forest land, and blue for water bodies. This histogram shows overlapping areas of spectral values among LULC types when it only compares its probability value. In fact, the sum of the probability (height) of each LULC type equals 1, but the distributed ranges (width) of the bin from each LULC type are different. Therefore, it is difficult to directly understand the size of the overlap of different LULC types in the histogram. A more elegant solution is to approximate the distribution of the probability density function as a curve.

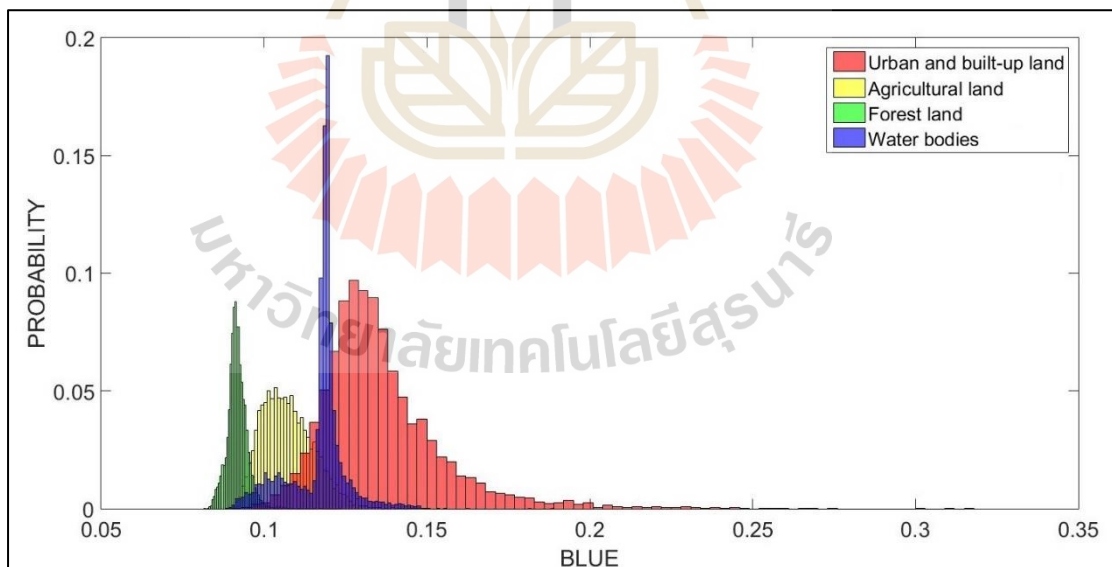


Figure 5.13 Probability of four LULC types on the blue band.

Figure 5.14 shows an example of the probability density function curve of four LULC types on the blue band. In this figure, the area enclosed by any curve and the abscissa is equal to 1. Therefore, it is easy to calculate the overlap area between any two curves. The area of the overlap area equals to zero or small means that the blue band can distinguish between the two LULC types. Conversely, a large area of overlap means that it is difficult to distinguish between the two LULC types by using a blue band. Therefore, the method of calculating the area of the overlapping area under MATLAB was used to explore the optimal spectral features for distinguishing four LULC types.

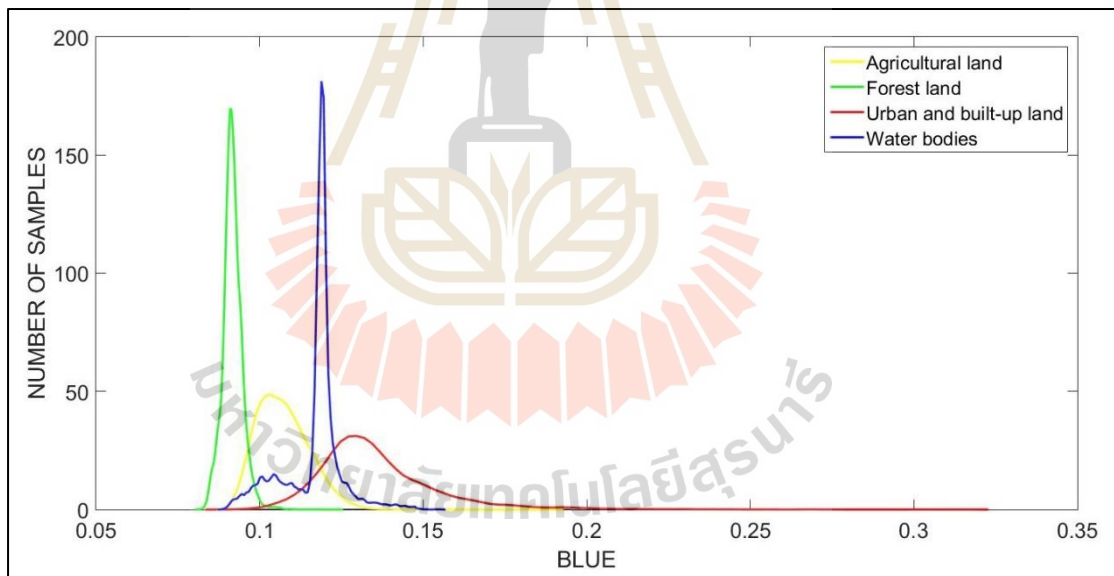


Figure 5.14 Probability density function of four LULC types on the blue band.

Table 5.9 Area of overlapping of any two LULC types.

Feature	U_A	U_F	U_W	A_F	A_W	F_W
BLUE	0.216	0.012	0.335	0.125	0.394	0.072
GREEN	0.418	0.029	0.423	0.143	0.444	0.160
RED	0.392	0.020	0.301	0.094	0.517	0.111
NIR	0.386	0.102	0.037	0.427	0.022	0.003
SWIR1	0.778	0.608	0.028	0.623	0.021	0.012
SWIR2	0.528	0.087	0.028	0.259	0.032	0.030
NDVI	0.265	0.010	0.062	0.118	0.035	0.003
EVI	0.304	0.021	0.057	0.223	0.031	0.002
NDBI	0.400	0.030	0.109	0.217	0.305	0.268
MNDBI	0.326	0.019	0.162	0.150	0.583	0.541
NDWI	0.252	0.010	0.041	0.156	0.023	0.002
MNDWI	0.457	0.158	0.028	0.506	0.017	0.002
TCB	0.657	0.655	0.039	0.851	0.023	0.012
TCG	0.253	0.012	0.346	0.230	0.052	0.002
TCW	0.802	0.402	0.027	0.485	0.022	0.011
PC1	0.638	0.595	0.031	0.831	0.021	0.009
PC2	0.284	0.011	0.388	0.168	0.200	0.003
PC3	0.852	0.478	0.043	0.507	0.025	0.008

Table 5.9 shows the area of overlapping of any two LULC types of 18 features. Herein, U denotes urban and built-up land, A denotes agricultural land, F denotes forest land, and W denotes water bodies. As a result, it is easy to see in this table that the area of the overlap of any two curves when using different spectral features. This area illustrates the proportion of misclassification when using a feature to distinguish between two LULC types. For example, if NIR feature is applied to distinguish between water bodies and other LULC types, pair of forest land and water bodies show the lowest overlapped area than other pairs, This finding indicates that NIR feature can be used to separate water bodies from forest land better than water bodies from urban and built-land or agricultural land.

However, the goal of this operation is to find spectral features that can distinguish between different LULC types. Therefore, the values in Table 5.9 is subtracted from the constant 1, and the results obtained, as shown in Table 5.10. This Table shows the ability of spectral features to distinguish between any two different LULC types.

Table 5.10 Distinguishability of any two LULC types.

Feature	U_A	U_F	U_W	A_F	A_W	F_W
BLUE	0.784	0.988	0.665	0.875	0.606	0.928
GREEN	0.582	0.971	0.577	0.857	0.556	0.840
RED	0.608	0.980	0.699	0.906	0.483	0.889
NIR	0.614	0.898	0.963	0.573	0.978	0.997
SWIR1	0.222	0.392	0.972	0.377	0.979	0.988
SWIR2	0.472	0.913	0.972	0.741	0.968	0.970
NDVI	0.735	0.990	0.938	0.882	0.965	0.997
EVI	0.696	0.979	0.943	0.777	0.969	0.998
NDBI	0.600	0.970	0.891	0.783	0.695	0.732
MNDBI	0.674	0.981	0.838	0.850	0.417	0.459
NDWI	0.748	0.990	0.959	0.844	0.977	0.998
MNDWI	0.543	0.842	0.972	0.494	0.983	0.998
TCB	0.343	0.345	0.961	0.149	0.977	0.988
TCG	0.747	0.988	0.654	0.770	0.948	0.998
TCW	0.198	0.598	0.973	0.515	0.978	0.989
PC1	0.362	0.405	0.969	0.169	0.979	0.991
PC2	0.716	0.989	0.612	0.832	0.800	0.997
PC3	0.148	0.522	0.957	0.493	0.975	0.992

Table 5.10 only shows the ability of features to distinguish between two different LULC types. Therefore, it further requires calculating the ability of spectral features to distinguish each specific type of LULC, as shown in Table 5.11.

Table 5.11 The ability of the spectral features to distinguish a certain LULC.

No	Feature	U	A	F	W	Average	Ranking
1	BLUE	0.813	0.755	0.930	0.733	0.808	8
2	GREEN	0.710	0.665	0.889	0.658	0.730	12
3	RED	0.763	0.666	0.925	0.690	0.761	11
4	NIR	0.825	0.722	0.823	0.979	0.837	6
5	SWIR1	0.529	0.526	0.586	0.980	0.655	16
6	SWIR2	0.786	0.727	0.875	0.970	0.839	5
7	NDVI	0.888	0.861	0.956	0.966	0.918	2
8	EVI	0.873	0.814	0.918	0.970	0.893	3
9	NDBI	0.820	0.693	0.828	0.773	0.778	10
10	MNDBI	0.831	0.647	0.763	0.572	0.703	14
11	NDWI	0.899	0.856	0.944	0.978	0.919	1
12	MNDWI	0.786	0.673	0.778	0.984	0.805	9
13	TCB	0.550	0.490	0.494	0.975	0.627	18
14	TCG	0.796	0.822	0.919	0.866	0.851	4
15	TCW	0.590	0.564	0.701	0.980	0.708	13
16	PC1	0.579	0.504	0.522	0.980	0.646	17
17	PC2	0.773	0.783	0.939	0.803	0.824	7
18	PC3	0.542	0.538	0.669	0.974	0.681	15

In consequence, it indicates that almost all spectral features make it easy to distinguish between water bodies and other types of LULC. However, some spectral features are not suitable for distinguishing between agricultural land, forest land, and urban and built-up land. Table 5.11 shows an average distinguishing ability of four LULC types. Larger values indicate that those spectral features are highly capable of distinguishing four LULC types.

In summary, the first six spectral features with larger values to distinguish four LULC types are NDWI, NDVI, EVI, TCG, SWIR2, and NIR.

5.2.3 Optimum spectral features selection

By combining the spectral features that can better fit into the harmonic wave function and can better distinguish four LULC types, multiplication operation between average R^2 of four LULC types from Table 5.8 and average value of the ability of spectral features to distinguish a certain LULC from Table 5.11 was applied to identify optimum spectral features selection for LULC classification as summarized in Table 5.12.

As a consequence, the suitable spectral features for all LULC type classification should include NIR, EVI, TCB, BLUE, GREEN, and PC1.

These derived optimum spectral features under this chapter are further applied for LULC classification with K-Means and ISODATA clustering algorithm in Chapter 6: Time-series LULC extraction using unsupervised classification algorithm and Chapter 7: Time-series LULC extraction using supervised classification.

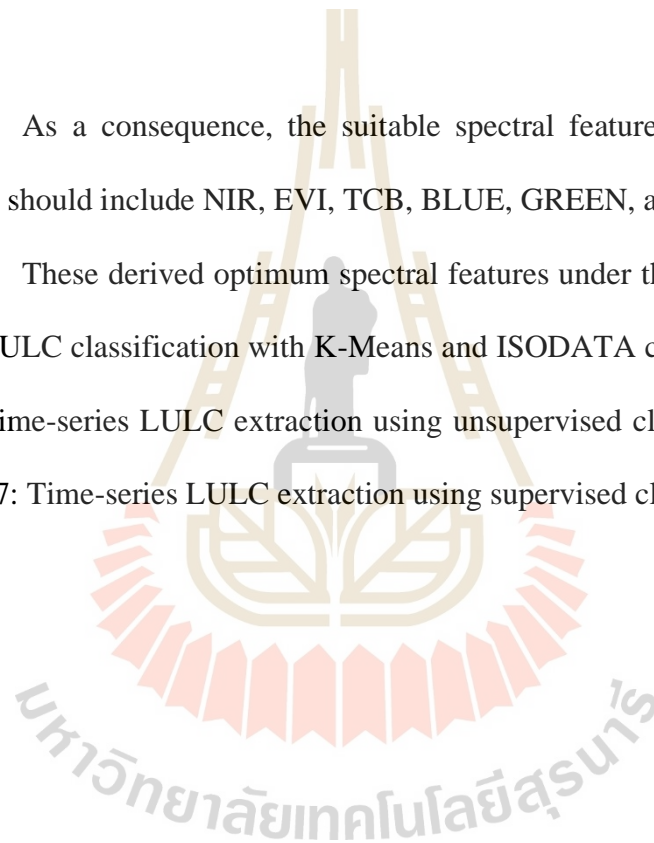


Table 5.12 Optimal spectral feature for LULC classification.

No	Feature	Average R ²	Average ability value	Multiplication	Ranking
1	BLUE	0.596	0.808	0.482	4
2	GREEN	0.658	0.73	0.480	5
3	RED	0.54	0.761	0.411	11
4	NIR	0.767	0.837	0.642	1
5	SWIR1	0.662	0.655	0.434	10
6	SWIR2	0.519	0.839	0.435	9
7	NDVI	0.491	0.918	0.451	7
8	EVI	0.66	0.893	0.589	2
9	NDBI	0.363	0.778	0.282	13
10	MNDBI	0.4	0.703	0.281	14
11	NDWI	0.481	0.919	0.442	8
12	MNDWI	0.203	0.805	0.163	17
13	TCB	0.786	0.627	0.493	3
14	TCG	0.445	0.851	0.379	12
15	TCW	0.243	0.708	0.172	16
16	PC1	0.719	0.646	0.464	6
17	PC2	0.219	0.824	0.180	15
18	PC3	0.047	0.681	0.032	18

SUMMARY

The spectral features include TCB, NIR, PC1, GREEN, SWIR1, and EVI have a high fitting effect of the harmonic wave function. Meanwhile, the spectral features include NDWI, NDVI, EVI, TCG, SWIR2, and NIR have a high ability to distinguish a specific LULC type. By considering both characteristics from spectral features, the optimum spectral features for LULC classification included NIR, EVI, TCB, BLUE, GREEN, and PC1 since these spectral features can better fit into the harmonic wave function and can better distinguish four LULC types.

CHAPTER VI

TIME-SERIES LAND USE AND LAND COVER

CLASSIFICATION USING AN UNSUPERVISED

METHOD

In principle, unsupervised classification is the process whereby numerical operations have performed that search for natural groupings of the spectral properties of pixels, as examined in multispectral feature space. The clustering process results in a classification map consisting of spectral classes (Lo and Yeung, 2007).

The workflow of time-series LULC classification using an unsupervised classification algorithm is schematically displayed in Figure 6.1. The detail of three major tasks and their results of this component included (1) LULC change detection using HA and CCDC, (2) unsupervised LULC classification using K-Means and ISODATA algorithm and (3) accuracy assessment were separately described and discussed in the following section.

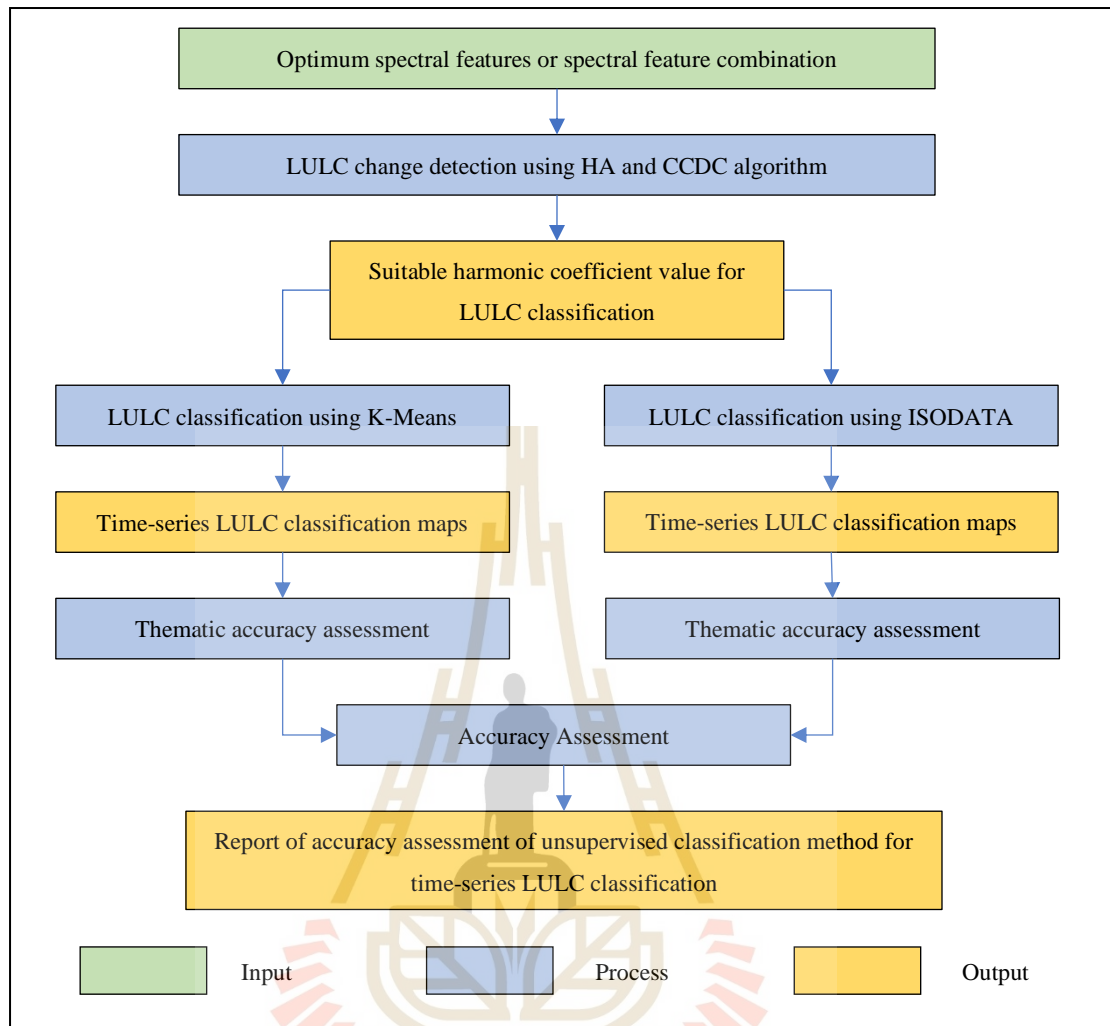


Figure 6.1 Workflow of Component 3: Time-series LULC classification using unsupervised classification method.

6.1 LULC change detection using HA and CCDC

The HA and CCDC were here applied to characterize LULC change detection using the single spectral feature and multiple spectral features (Table 5.12), which were derived from Chapter 5, Optimum spectral features selection for LULC classification.

To illustrate the change detection procedure of CCDC algorithm, one pixel (Point X in Figure 6.2) that experienced a conversion from one LULC type to another

type was here characterized using the optimum spectral feature: single and multiple spectral features. In this study, the single feature is NIR band while multiple spectral features are TOP3 spectral features (NIR, EVI, and TCB), TOP6 spectral features (NIR, EVI, TCB, BLUE, GREEN, and PC1) and ORI6 spectral features of Landsat data (BLUE, GREEN, RED, NIR, SWIR1, and SWIR2).

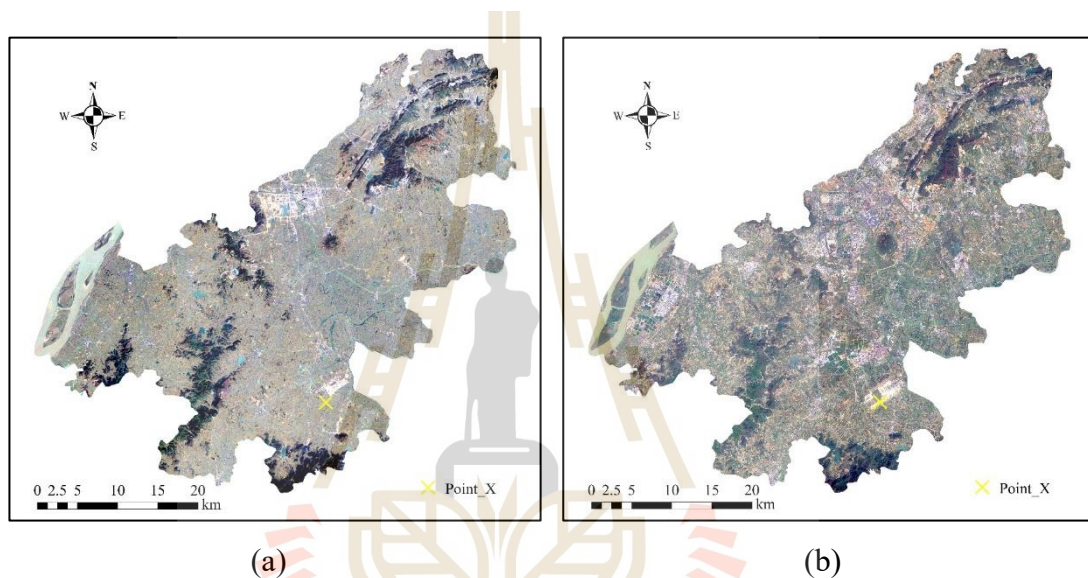


Figure 6.2 The geographical location of the Jiangning District and the location of point X (Row 1443, column 1174): (a) Image from Google Earth in 2000 and (b) Image from Google Earth in 2017.

Figure 6.2 demonstrates two images of Jiangning district obtained from Google Earth. The acquisition time of the image in Figure 6.2(a) was 2000, and the acquisition time of Figure 6.2(b) was 2017. Point X was taken from the new airport runway area of Nanjing Lukou International Airport. The LULC type was an undisturbed area in 2000, and the LULC type had been converted into urban and built-up land in 2017.

Details of characteristics of LULC change detection based on the HA and CCDC from optimum single and multiple spectral features are described and discussed in the following sections.

6.1.1 Single feature: NIR band

The NIR band is the most suitable spectral feature that can fit the HA model and distinguish four LULC types (see detail in Chapter V). Characteristics of LULC change detection on point X location based on the NIR band over the study period (2000-2017) display in Figure 6.3.

In principle, if the harmonic of a pixel is a complete wave, it means that the LULC type of the pixel does not change. If the harmonic of a pixel is not a complete wave and it is divided into multiple sections, the LULC type of the pixel may change. On the contrary, if the harmonic function of the pixel has only one group of coefficients, the LULC type of the pixel does not change. If the LULC type of the pixel has changed, the harmonic function of the pixel shows multiple groups of coefficients.

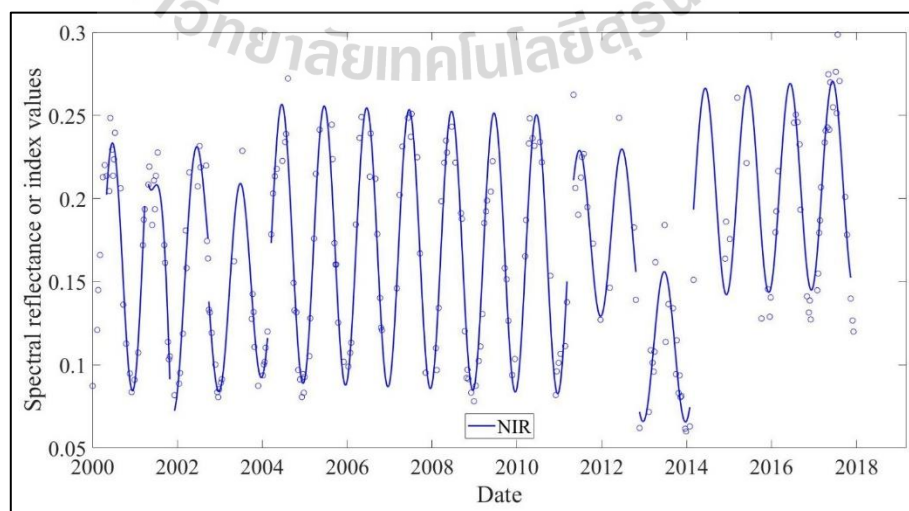


Figure 6.3 Time-series model of NIR band for point X.

As shown in Figure 6.3, there are 8 harmonic waves for 18 years; it indicates that the LULC type at location X may have changed 7 times, and the NIR feature can be applied to detect LULC change. Mainly, it can be observed that the harmonic curve of the LULC type on 27 September 2013 at X location shows LULC change. In the meantime, Table 6.1 shows the detailed information of the harmonic wave coefficient of NIR from point X location.

As shown in Table 6.1, the LULC type at Point X may be changed in seven periods with Julian day (481-665, 705-993, 1001-1505, 1537-4081, 4137-4673, 4705-5137, and 5169-6521).

Furthermore, the NIR band can be used to detect all possible LULC changes over the eighteen years of the whole study area. Table 6.2 shows the results of LULC change detection over specific locations in the study area. As a result, LULC changes are taken place at locations (1, 1), (1, 2), (1, 2000), and (2, 1).

Table 6.1 Four coefficients of the NIR band at point X location in different periods.

Period (Julian day)	Intercept	Slope (E-04)	Amplitude	Phase
121-449	0.126	1.257	-0.086	0.273
481-665	0.866	-12.487	-0.078	-1.238
705-993	-0.017	2.112	-0.060	-5.813
1001-1505	0.120	0.228	-0.060	-6.163
1537-4081	0.178	-0.029	-0.084	-6.038
4137-4673	0.170	0.022	-0.050	-6.113
4705-5137	0.114	-0.006	-0.045	-6.108
5169-6521	0.183	0.039	-0.063	-5.890

Table 6.2 Coefficients of all pixels from the NIR band in the study area.

Row and Column	Period (Julian day)	Intercept	Slope (E-04)	Amplitude	Phase
(1,1)	89-777	0.098	0.781	-0.157	0.139
	833-1113	0.290	-1.769	-0.056	0.264
	1217-1665	0.106	0.040	-0.069	0.371
	1681-1953	-0.248	2.016	-0.068	-0.023
	2057-2473	-0.265	1.771	-0.074	0.149
	2481-6521	0.129	0.105	-0.105	0.115
(1,2)	121-833	0.141	0.252	-0.118	0.192
	881-1113	0.464	-3.304	-0.058	0.372
	1217-4137	0.106	0.161	-0.087	0.227
	4153-4785	0.246	-0.200	-0.064	0.303
	4827-6521	0.225	-0.051	-0.111	0.140
...
(1,2000)	1-745	0.219	-0.716	-0.102	0.263
	753-1025	0.002	2.178	-0.086	0.737
	1073-1465	0.199	-0.189	-0.086	0.159
	1481-6521	0.213	-0.065	-0.093	0.322
(2,1)	1-833	0.162	0.116	-0.102	0.175
	881-1145	0.075	0.743	-0.095	0.137
	1217-4137	0.130	0.112	-0.087	0.289
	4153-6521	0.100	0.177	-0.105	0.155
...

6.1.2 Multiple features: TOP3 spectral features

It is relatively simple to use only one feature for LULC change detection. However, if this feature is not sensitive to change in a specific type of LULC, it is difficult to find a change in LULC type through this feature, or if it is susceptible to change in LULC type, it will result in a large number of the unwanted subcategories such as turbid and clear water bodies. So, it can lead to an inaccurate result of LULC change detection. Three spectral features were, therefore, applied to detect LULC change in this study.

Similar to the previous section, NIR, EVI, and TCB are the first three suitable spectral features that can be applied to distinguish four different LULC types (see detail in Chapter V). Figure 6.4 shows the TOP3 spectral features (NIR, EVI, and TCB) that are used for change detection using point X as an example.

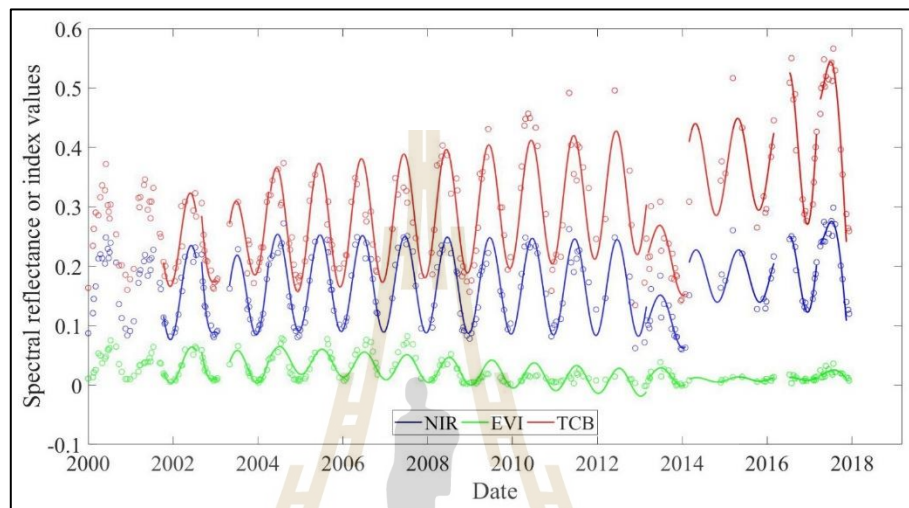


Figure 6.4 Time-series model of TOP3 spectral features for point X.

Similar to one feature for LULC change detection, if the LULC type of a pixel does not change, the harmonic function of each spectral feature of the pixel will have only one group of coefficients. If the LULC type of pixel changes, the harmonic function of each feature of the pixel will have multiple groups of coefficients.

As shown in Figure 6.4, there are also eight harmonic waves for 18 years; it indicates that LULC type at location X may have changed in seven periods, and TOP3 spectral features can be applied to detect LULC change. Table 6.3 shows the detail information of harmonic wave coefficients of TOP3 spectral features at point X location.

As it can be seen from Table 6.3, the LULC type at point X may have changed in seven periods with the dates of Julian day (977-1113, 1217-1553, 1569-4801, 4817-5137, 5169-5897, 6033-6245 and 6297-6521).

Furthermore, the TOP3 spectral features were used to detect all possible LULC changes over the eighteen years of the whole study area. Table 6.4 shows the results of LULC change detection over specific locations in the study area. As a result, LULC changes are taken place at locations (1, 1), (1, 2), and (2, 1), whereas LULC type at the location (1, 2000) does not change.

Table 6.3 Four coefficients of TOP3 spectral features in different periods.

Period (Julian day)	Feature	Intercept	Slope (E-04)	Amplitude	Phase
649-929	NIR	0.194	-0.485	-0.084	0.414
	EVI	0.044	-0.123	-0.031	0.319
	TCB	0.344	-1.260	-0.090	0.499
977-1113	NIR	2.683	-23.938	-0.142	1.761
	EVI	1.099	-10.073	-0.060	1.816
	TCB	2.194	-18.501	-0.111	1.625
1217-1553	NIR	-0.168	2.341	0.087	3.225
	EVI	-0.044	0.546	-0.032	6.277
	TCB	-0.117	2.688	-0.085	0.260
1569-4801	NIR	0.178	-0.031	0.081	3.361
	EVI	0.063	-0.125	-0.022	6.188
	TCB	0.225	0.213	-0.106	0.385
4817-5137	NIR	0.442	-0.667	0.038	3.157
	EVI	0.016	-0.005	-0.015	6.322
	TCB	0.854	-1.284	-0.047	0.045
5169-5897	NIR	0.183	0.000	0.044	4.240
	EVI	-0.010	0.034	-0.004	7.321
	TCB	0.235	0.241	-0.080	1.233
6033-6256	NIR	-2.277	4.031	0.098	3.189
	EVI	-0.159	0.277	-0.005	6.281
	TCB	-3.344	6.129	-0.181	0.122
6297-6521	NIR	5.601	-8.428	0.080	2.375
	EVI	0.499	-0.752	-0.010	5.344
	TCB	8.679	-12.908	-0.131	-0.570

Table 6.4 Coefficients of all pixels from TOP3 spectral features in the study area.

Row and Column	Period (Julian day)	Feature	Intercept	Slope (E-04)	Amplitude	Phase
(1, 1)	89-833	NIR	0.106	0.593	-0.154	0.118
		EVI	0.016	0.342	-0.060	0.032
		TCB	0.169	0.619	-0.161	0.196
	881-1113	NIR	0.363	-2.491	-0.049	0.300
		EVI	-0.038	0.633	-0.032	0.223
		TCB	1.011	-8.287	-0.012	1.355
	1217-2313	NIR	0.098	0.103	-0.067	0.344
		EVI	0.014	0.040	-0.021	0.258
		TCB	0.185	0.062	-0.089	-5.896
2425-6521	NIR	0.130	0.103	-0.105	0.114	
	EVI	0.026	0.034	-0.039	-0.014	
	TCB	0.206	0.139	-0.112	-6.066	
...	
(1, 2)	121-881	NIR	0.148	0.087	-0.114	0.190
		EVI	0.033	0.108	-0.043	0.058
		TCB	0.218	0.138	-0.122	0.332
	905-1113	NIR	0.402	-2.697	-0.062	0.330
		EVI	0.148	-1.146	-0.023	0.491
		TCB	0.359	-1.633	-0.077	0.102
	1217-5257	NIR	0.107	0.153	-0.090	0.219
		EVI	0.021	0.036	-0.030	0.029
		TCB	0.184	0.216	-0.111	0.342
5287-5601	NIR	2.573	-4.356	-0.116	0.613	
	EVI	1.142	-2.000	-0.044	0.518	
	TCB	3.141	-5.221	-0.130	0.818	
5625-5913	NIR	-0.216	0.741	-0.133	0.188	
	EVI	-0.187	0.419	-0.051	0.120	
	TCB	-0.013	0.545	-0.129	0.246	
5945-6265	NIR	-0.306	0.804	-0.095	0.145	
	EVI	-0.182	0.367	-0.032	0.022	
	TCB	0.245	0.060	-0.101	0.349	
6297-6489	NIR	22.828	-35.397	-0.300	-1.310	
	EVI	10.646	-16.571	-0.143	-1.421	
	TCB	16.268	-25.007	-0.219	-1.025	

Table 6.4 (Continued).

Row and Column	Period (Julian day)	Feature	Intercept	Slope (E-04)	Amplitude	Phase
(1, 2)	6521-6545	NIR	22.828	-35.397	-0.300	-1.310
		EVI	10.646	-16.571	-0.143	-1.421
		TCB	16.268	-25.007	-0.219	-1.025
...
(1,2000)	41-6521	NIR	0.199	-0.036	-0.094	0.332
		EVI	0.042	-0.010	-0.031	0.216
		TCB	0.289	-0.001	-0.103	0.394
...
(2,1)	89-833	NIR	0.159	0.161	-0.103	0.169
		EVI	0.034	0.192	-0.043	-0.013
		TCB	0.238	0.037	-0.100	0.376
	881-1113	NIR	0.150	-0.017	-0.089	0.132
		EVI	-0.066	1.042	-0.043	0.108
		TCB	0.767	-5.459	-0.045	0.244
	1145-4137	NIR	0.130	0.111	-0.087	0.290
		EVI	0.030	0.011	-0.029	0.062
		TCB	0.210	0.178	-0.106	0.446
4153-6521	NIR	0.100	0.177	-0.105	0.155	
	EVI	-0.001	0.088	-0.037	0.002	
	TCB	0.245	0.094	-0.114	0.309	
...

6.1.3 Multiple features: TOP6 and ORI6 spectral features

Refer to optimum spectral features' selection based on standard multispectral features of Landsat that consists of six bands (BLUE, GREEN, RED, NIR, SWIR-1, and SWIR-2) and their derived spectral indices (NDVI, EVI, NDBI, MNDWI, NDWI, MNDWI, TCB, TCG, TCW, PC1, PC2, and PC3), TOP6 spectral features (NIR, EVI, TCB, BLUE, GREEN, and PC1) as optimum features and ORI6 spectral features (BLUE, GREEN, RED, NIR, SWIR-1, and SWIR-2) as original multispectral

bands of Landsat data that can better distinguish four LULC types and detect LULC change was here described and discussed.

Figure 6.5 shows TOP6 spectral features (NIR, EVI, TCB, BLUE, GREEN, and PCI) that were used to detect LULC change at point X location. As shown in Figure 6.5, there are four harmonic waves for 18 years, and it indicates that the LULC type at location X may have changed in three periods. Therefore, TOP6 spectral features can be applied to detect LULC change.

Table 6.5 shows the detail information of harmonic wave coefficients of TOP6 spectral features at point X location.

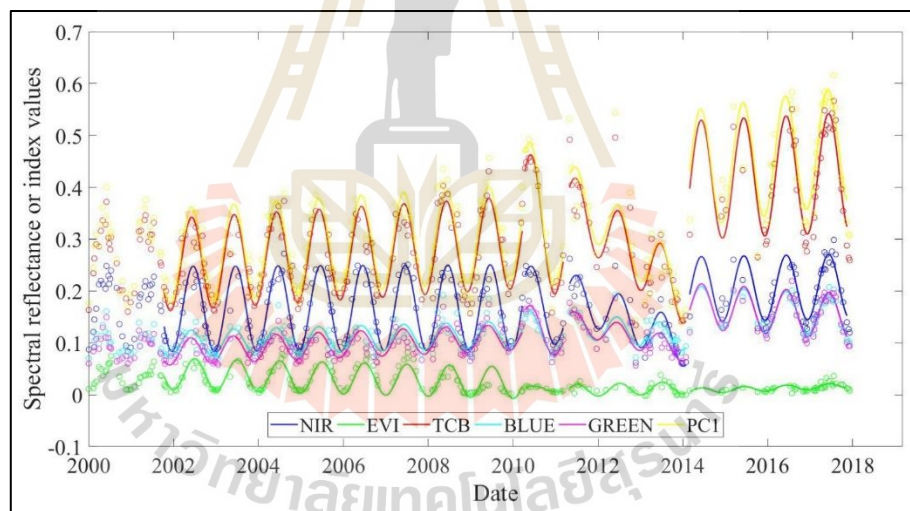


Figure 6.5 Time-series model of TOP6 spectral features for point X.

As shown in Table 6.5, it was found that the LULC type at point X may have changed in three periods with the dates of Julian day (3753-4081, 4137-5137, and 5169-6521). Furthermore, the TOP6 spectral features were used to detect all possible changes over the eighteen years of the whole study area. Table 6.6 shows the results of LULC change detection over specific locations in the study area. As a result, it was

found that LULC changes are taken place at locations (1, 1) and (1, 2000), whereas LULC type at locations (1, 2) and (2, 1) do not change.

Table 6.5 Four coefficients of TOP6 spectral features in different periods.

Period (Julian day)	Feature	Intercept	Slope (E-04)	Amplitude	Phase
649-3729	NIR	0.165	0.010	-0.082	0.246
	EVI	0.045	-0.058	-0.031	0.002
	TCB	0.241	0.142	-0.089	0.461
	BLUE	0.094	0.062	-0.025	0.495
	GREEN	0.076	0.093	-0.025	0.586
	PC1	0.257	0.154	-0.092	0.489
3753-4081	NIR	0.958	-2.009	-0.056	0.285
	EVI	0.068	-0.145	-0.004	-0.034
	TCB	1.698	-3.508	-0.100	0.333
	BLUE	0.507	-0.992	-0.031	0.599
	GREEN	0.630	-1.297	-0.035	0.436
	PC1	2.021	-4.289	-0.095	0.444
4137-5137	NIR	0.597	-0.976	-0.043	-0.015
	EVI	-0.015	0.057	-0.012	-0.086
	TCB	1.080	-1.725	-0.060	0.079
	BLUE	0.244	-0.263	-0.025	0.321
	GREEN	0.374	-0.564	-0.022	0.248
	PC1	1.209	-1.971	-0.056	0.077
5169-6521	NIR	0.183	0.039	-0.063	0.393
	EVI	-0.031	0.074	-0.005	0.355
	TCB	0.355	0.113	-0.115	0.418
	BLUE	0.242	-0.137	-0.040	0.379
	GREEN	0.251	-0.153	-0.044	0.367
	PC1	0.254	0.350	-0.113	0.524

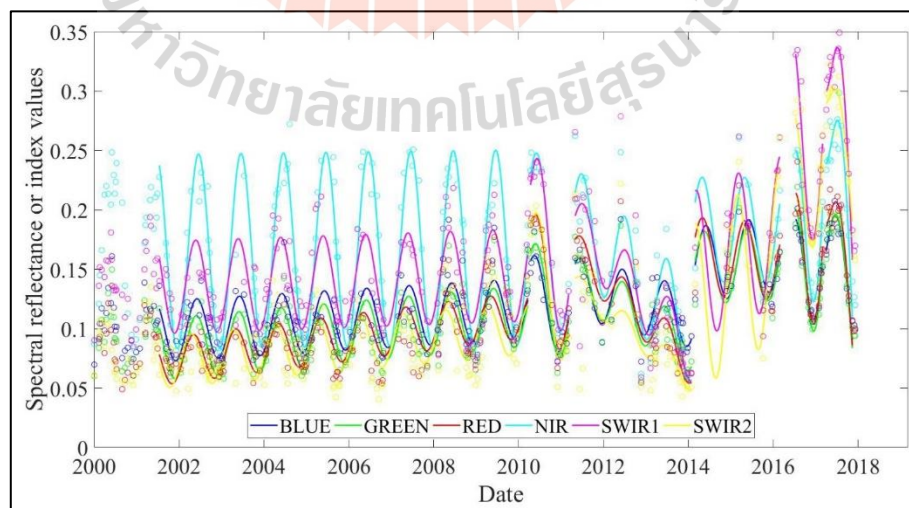
Table 6.6 Coefficients of all pixels from TOP6 spectral features in the study area.

Row and Column	Period (Julian day)	Feature	Intercept	Slope (E-04)	Amplitude	Phase
	89-833	NIR	0.106	0.593	-0.154	0.118
		EVI	0.016	0.342	-0.060	0.032
		TCB	0.169	0.619	-0.161	0.196
		BLUE	0.087	0.107	-0.032	0.411
		GREEN	0.077	0.011	-0.032	0.473
		PC1	0.158	0.815	-0.184	0.184
(1,1)	881-1113	NIR	0.363	-2.491	-0.049	0.300
		EVI	-0.038	0.633	-0.032	0.223
		TCB	1.011	-8.287	-0.012	1.355
		BLUE	0.432	-3.371	0.004	-0.116
		GREEN	0.360	-2.829	0.002	-0.369
		PC1	0.777	-6.048	-0.037	0.390
	1217-6521	NIR	0.098	0.166	-0.097	0.136
		EVI	0.016	0.055	-0.035	0.009
		TCB	0.172	0.205	-0.108	-6.044
		BLUE	0.105	0.005	-0.028	-5.954
		GREEN	0.089	0.009	-0.027	-12.207
		PC1	0.171	0.235	-0.114	0.227
...	
(1,2)	145-6521	NIR	0.128	0.106	-0.099	0.194
		EVI	0.028	0.028	-0.035	0.041
		TCB	0.206	0.150	-0.114	0.316
		BLUE	0.100	0.018	-0.028	0.396
		GREEN	0.084	0.023	-0.028	0.449
		PC1	0.211	0.172	-0.121	0.318
...	
(1,2000)	41-3353	NIR	0.192	0.017	-0.097	0.350
		EVI	0.044	-0.004	-0.033	0.212
		TCB	0.276	0.054	-0.099	0.419
		BLUE	0.096	0.026	-0.022	0.573
		GREEN	0.087	0.027	-0.025	0.653
		PC1	0.292	0.074	-0.105	0.391

Table 6.6 (Continued).

Row and Column	Period (Julian day)	Feature	Intercept	Slope (E-04)	Amplitude	Phase
(1,2000)	3417-6521	NIR	0.196	-0.034	-0.093	0.298
		EVI	-0.008	0.082	-0.030	0.173
		TCB	0.424	-0.254	-0.106	0.396
		BLUE	0.165	-0.102	-0.027	0.513
		GREEN	0.178	-0.149	-0.027	0.527
		PC1	0.461	-0.287	-0.106	0.430
...
(2,1)	145-6521	NIR	0.144	0.090	-0.097	0.222
		EVI	0.033	0.021	-0.035	0.044
		TCB	0.224	0.132	-0.108	0.378
		BLUE	0.102	0.017	-0.026	0.439
		GREEN	0.085	0.024	-0.027	0.505
		PC1	0.234	0.150	-0.115	0.390
...

Likewise, Figure 6.6 shows the ORI6 spectral bands (BLUE, GREEN, RED, NIR, SWIR1, and SWIR2) that were used to detect LULC change at point X location.

**Figure 6.6** Time-series model of ORI6 spectral bands for point X.

As shown in Figure 6.6, there are four harmonic waves during 18 years that indicate LULC change at location X in five periods, and they were applied to detect LULC change. Table 6.7 shows the detail information of harmonic wave coefficients of ORI6 spectral bands at point X location.

As shown in Table 6.7, the LULC type at point X location may have changed in 5 periods with the dates of Julian day (3753-4081, 4137-5137, 5169-5897, 6033-6265, and 6297-6521).

Furthermore, ORI6 spectral bands were used to detect all possible LULC changes over the eighteen years of the whole study area. Table 6.8 shows the results of LULC change detection over specific locations in the study area. As a result, LULC changes are taken place at locations (1, 2000) and (2, 1), whereas LULC type at locations (1, 1) and (1, 2) do not change.

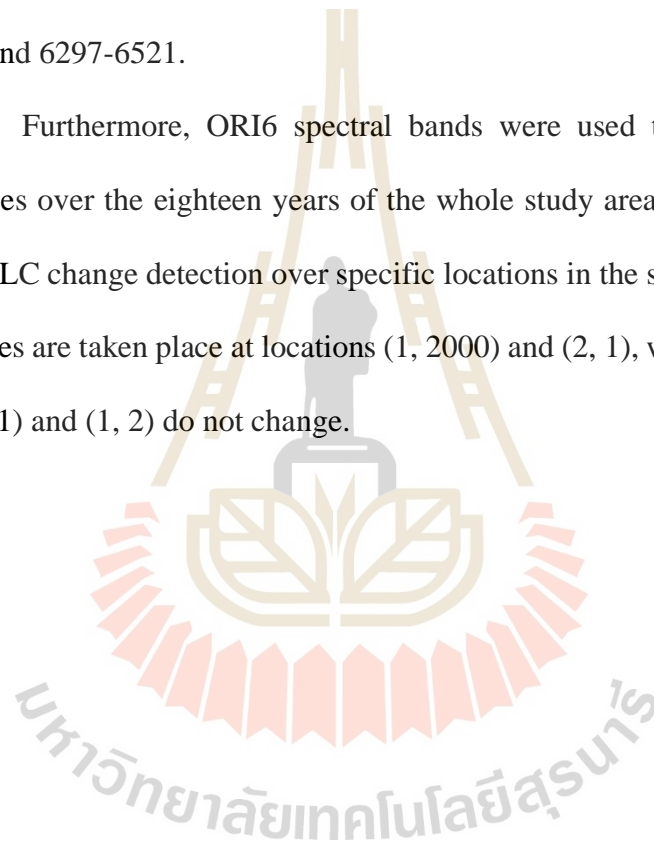


Table 6.7 Four coefficients of ORI6 spectral bands in different periods.

Period (Julian day)	Feature	Intercept	Slope (E-04)	Amplitude	Phase
561-3729	BLUE	0.165	0.061	-0.082	0.246
	GREEN	0.045	0.090	-0.031	0.002
	RED	0.241	0.127	-0.089	0.461
	NIR	0.094	0.013	-0.025	0.495
	SWIR1	0.076	0.033	-0.025	0.586
	SWIR2	0.257	0.091	-0.092	0.489
3753-4081	BLUE	0.958	-0.992	-0.056	0.285
	GREEN	0.068	-1.297	-0.004	-0.034
	RED	1.698	-1.473	-0.100	0.333
	NIR	0.507	-2.009	-0.031	0.599
	SWIR1	0.630	-1.767	-0.035	0.436
	SWIR2	2.021	-1.989	-0.095	0.444
4137-5137	BLUE	0.597	-0.263	-0.043	-0.015
	GREEN	-0.015	-0.564	-0.012	-0.086
	RED	1.080	-0.945	-0.060	0.079
	NIR	0.244	-0.976	-0.025	0.321
	SWIR1	0.374	-1.070	-0.022	0.248
	SWIR2	1.209	-0.789	-0.056	0.077
5169-5897	BLUE	0.183	0.136	-0.063	0.393
	GREEN	-0.031	0.057	-0.005	0.355
	RED	0.355	-0.069	-0.115	0.418
	NIR	0.242	0.000	-0.040	0.379
	SWIR1	0.251	0.401	-0.044	0.367
	SWIR2	0.254	0.654	-0.113	0.524
6033-6265	BLUE	0.597	1.852	-0.043	-0.015
	GREEN	-0.015	2.449	-0.012	-0.086
	RED	1.080	2.999	-0.060	0.079
	NIR	0.244	4.031	-0.025	0.321
	SWIR1	0.374	0.430	-0.022	0.248
	SWIR2	1.209	2.629	-0.056	0.077
6297-6521	BLUE	0.183	-2.189	-0.063	0.393
	GREEN	-0.031	-3.784	-0.005	0.355
	RED	0.355	-4.936	-0.115	0.418
	NIR	0.242	-8.428	-0.040	0.379
	SWIR1	0.251	-9.030	-0.044	0.367
	SWIR2	0.254	-4.783	-0.113	0.524

Table 6.8 Coefficients of all pixels from ORI6 spectral bands in the study area.

Row and Column	Period (Julian day)	Feature	Intercept	Slope (E-04)	Amplitude	Phase
(1,1)	89-6521	BLUE	0.098	0.020	-0.027	0.336
		GREEN	0.082	0.023	-0.028	0.376
		RED	0.066	0.032	-0.020	0.608
		NIR	0.117	0.126	-0.103	0.152
		SWIR1	0.060	0.146	-0.053	0.244
		SWIR2	0.030	0.095	-0.022	0.596
...
(1,2)	153-6521	BLUE	0.100	0.018	-0.028	0.396
		GREEN	0.084	0.023	-0.028	0.449
		RED	0.070	0.028	-0.022	0.685
		NIR	0.128	0.106	-0.099	0.193
		SWIR1	0.084	0.113	-0.058	0.352
		SWIR2	0.048	0.074	-0.028	0.718
...
(1,2000)	625-3337	BLUE	0.096	0.027	-0.023	0.486
		GREEN	0.086	0.032	-0.024	0.534
		RED	0.074	0.056	-0.014	0.672
		NIR	0.188	0.037	-0.096	0.340
		SWIR1	0.133	0.058	-0.041	0.173
		SWIR2	0.072	0.073	-0.013	-6.220
	3353-6521	BLUE	0.163	-0.099	-0.027	0.503
		GREEN	0.179	-0.151	-0.027	0.533
		RED	0.199	-0.203	-0.022	0.673
		NIR	0.197	-0.036	-0.093	0.300
		SWIR1	0.232	-0.155	-0.044	0.387
		SWIR2	0.196	-0.166	-0.022	-5.466
...
(2,1)	193-545	BLUE	0.098	0.003	-0.017	0.743
		GREEN	0.086	-0.055	-0.018	0.922
		RED	0.079	-0.233	-0.018	1.469
		NIR	0.177	-0.290	-0.106	0.300
		SWIR1	0.124	-0.254	-0.050	0.575
		SWIR2	0.078	-0.333	-0.024	1.205

Table 6.8 (Continued).

Row and Column	Period (Julian day)	Feature	Intercept	Slope (E-04)	Amplitude	Phase
(2,1)	561-3321	BLUE	0.098	0.045	-0.031	0.368
		GREEN	0.081	0.046	-0.029	0.425
		RED	0.068	0.047	-0.021	0.608
		NIR	0.151	0.022	-0.085	0.270
		SWIR1	0.104	0.035	-0.046	0.467
		SWIR2	0.059	0.033	-0.021	0.823
	3353-6521	BLUE	0.119	-0.019	-0.025	0.492
		GREEN	0.115	-0.035	-0.027	0.539
		RED	0.120	-0.065	-0.022	0.920
		NIR	0.120	0.141	-0.103	0.170
		SWIR1	0.148	0.012	-0.057	0.462
		SWIR2	0.117	-0.045	-0.032	1.004
...

As a result, spectral feature (i.e., NIR) and the combination of spectral features (i.e., TOP3, TOP6, and ORI6) only indicate the possible changes in the LULC type of a particular pixel, and it might not be the actual LULC type on the ground that represented by the pixel.

In this study, taking location at point X as an example, NIR band indicates that the LULC type of the pixel has changed in seven periods, TOP3 spectral features also indicate that the LULC type of the pixel has changed in seven times, TOP6 spectral features indicate that the LULC type of the pixel has changed in three periods and ORI6 spectral bands from Landsat indicate that the LULC type of the pixel has changed in five periods. Subsequently, spectral features and the combination of spectral features indicate when the possible LULC types change.

However, LULC change detection based on these features may be affected by many factors include:

(1) Actual LULC changes among the classes, e.g., the water body to be the urban and built-up area.

(2) Misjudgment of LULC change that is highly vulnerable due to temporal change, mainly urban and built-up land, and agricultural land. In the different states of urban development and the growing stage of crops, misclassification of LULC maybe happened.

(3) A particular spectral characteristic of a pixel cannot identify the difference between forests and crops.

(4) Even, could detection and remove had been here conducted, some cloudy pixels may exist, and they cannot classify the corrected LULC types.

Therefore, the coefficients of the spectral feature and the combination of spectral features were here applied to classify LULC types using unsupervised classification methods (K-Means and ISODATA) and to assess accuracy as one of the primary research objectives in this study.

6.2 Unsupervised LULC classification using K-Means and ISODATA algorithm

The unsupervised LULC classification (K-Means and ISODATA algorithm) was implemented under the MATLAB software environment. Mathematically, there are possible 60 combinations (4 spectral feature(s) combinations * 15 coefficient feature(s) combinations) that can be applied to classify LULC type. However, the number of coefficient feature(s) combinations (15) is rather high, so the reduction of

this number is required at the first step before they were applied to classify LULC with spectral feature(s) combinations by using K-Means and ISODATA algorithms. Table 6.9 summarizes the necessary information of the selected spectral feature(s) for unsupervised LULC classification by using K-Means and ISODATA algorithm. Meanwhile, Table 6.10 displays all possibility of coefficient feature combination for the most suitable coefficient identification.

In order to explore the most suitable coefficient feature, fifteen coefficient feature(s) combinations (different harmonic function curves) from the NIR band of time-series Landsat data were applied to classify LULC by using K-Means and ISODATA algorithms. So, the LULC types represented by different harmonic function curves are obtained to identify the most suitable coefficient feature(s) combinations.

After that, the most suitable coefficient features were chosen for time-series LULC classification with four optimum spectral features (NIR, TOP3, TOP6, and ORI6) under the unsupervised classification approach in this study.

Table 6.9 Basic information of the selected spectral feature(s).

NO.	Spectral feature and features combination	Abbreviation
1	NIR	NIR
2	NIR, EVI, TCB	TOP3
3	NIR, EVI, TCB, BLUE, GREEN, PC1	TOP6
4	BLUE, GREEN, RED, NIR, SWIR1, SWIR2	ORI6

Table 6.10 Possibility of coefficient feature combination for the most suitable coefficient identification.

No	Number of coefficient feature	Coefficient	Abbreviation
1	1	Intercept	I
2		Slope	S
3		Amplitude	A
4		Phase	P
5	2	Intercept and slope	IS
6		Intercept and amplitude	IA
7		Intercept and phase	IP
8		Slope and amplitude	SA
9		Slope and phase	SP
10		Amplitude and phase	AP
11	3	Intercept, slope, and amplitude	ISA
12		Intercept, slope, and phase	ISP
13		Intercept, amplitude, and phase	IAP
14		Slope, amplitude, and phase	SAP
15	4	Intercept, slope, amplitude, and phase	ISAP

6.2.1 Suitable coefficient feature for LULC classification

LULC classification based on NIR band from Landsat data acquired on 18 May 2017 for identifying suitable coefficient feature using K-Means and ISODATA algorithms is presented in Figures 6.7 and 6.8, respectively. Herein, the initial number of clusters for LULC classification using the K-Means algorithm was ten spectral classes, while the initial number of clusters for LULC classification using the ISODATA algorithm varied from 1 to 15. The initial number of clusters was assigned by observing feature space images. Then, the classified spectral classes of both algorithms were reclassified thematic classes based on ground reference data in 2017 (e. g., central business district, Yangtze river, forest area, paddy field, etc.) and display as thematic maps with four LULC types (U, A, F and W).

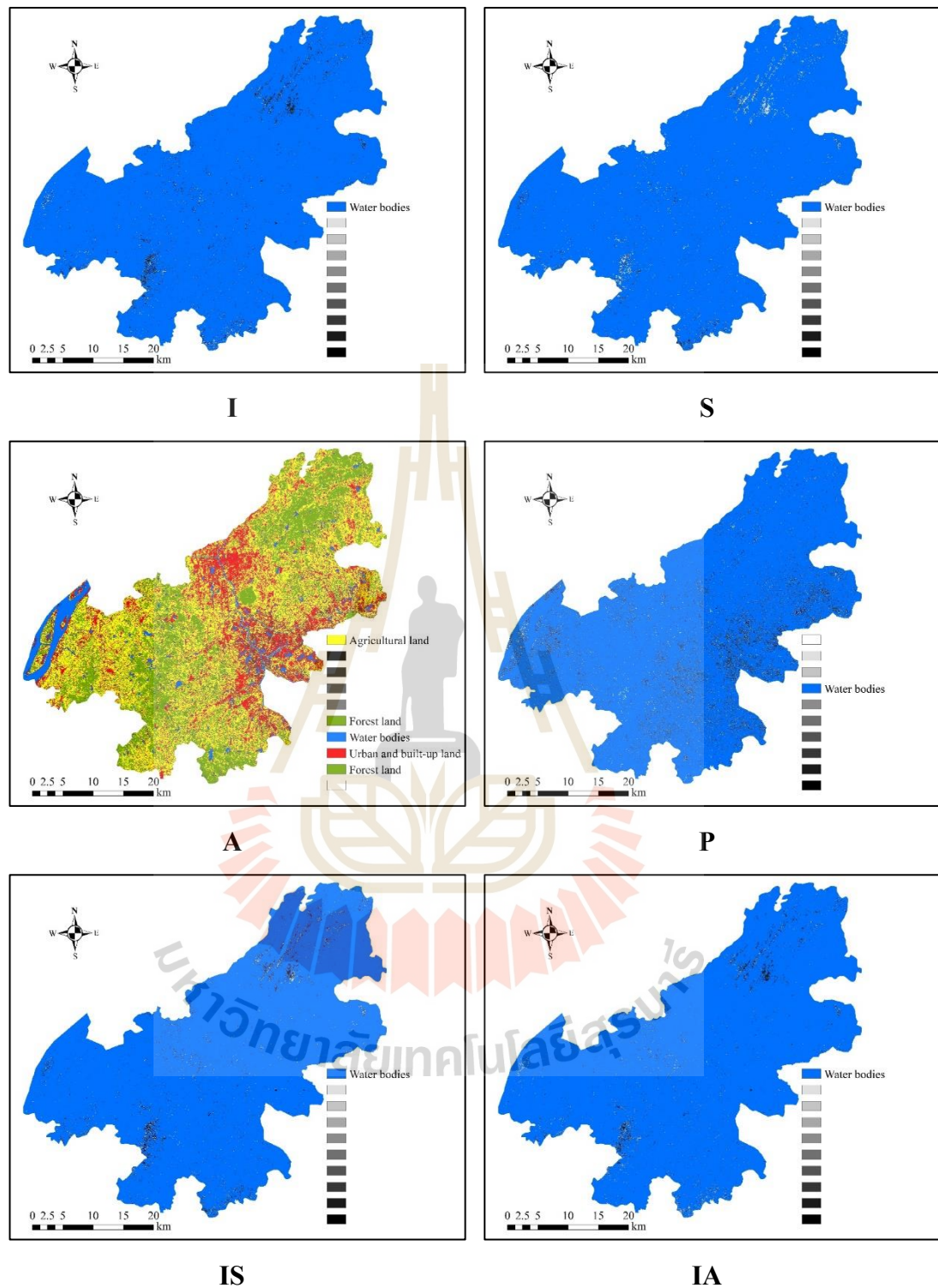


Figure 6.7 LULC classification map by K-Means algorithm with 15 coefficient features of NIR band.

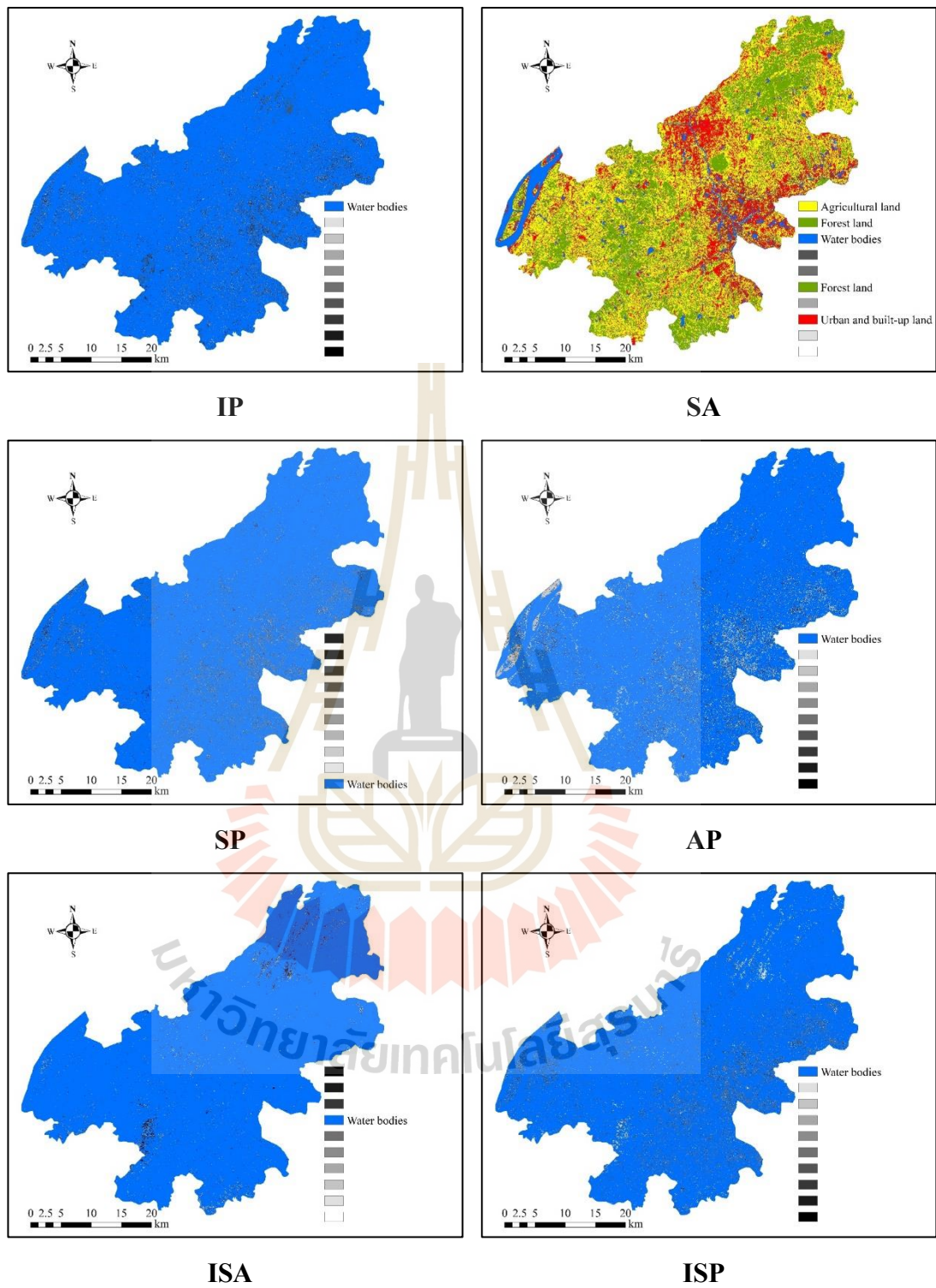


Figure 6.7 (Continued).

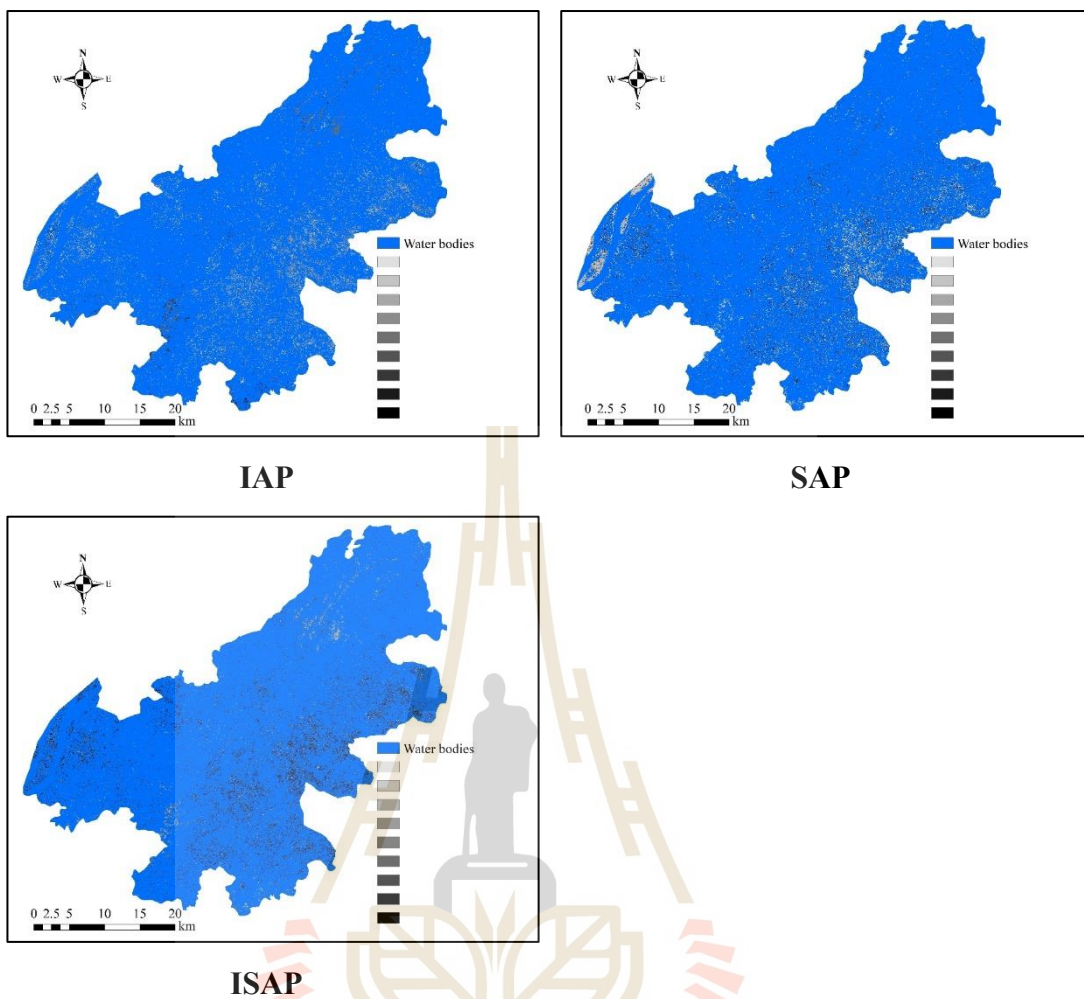


Figure 6.7 (Continued).

มหาวิทยาลัยเทคโนโลยีสุรนารี

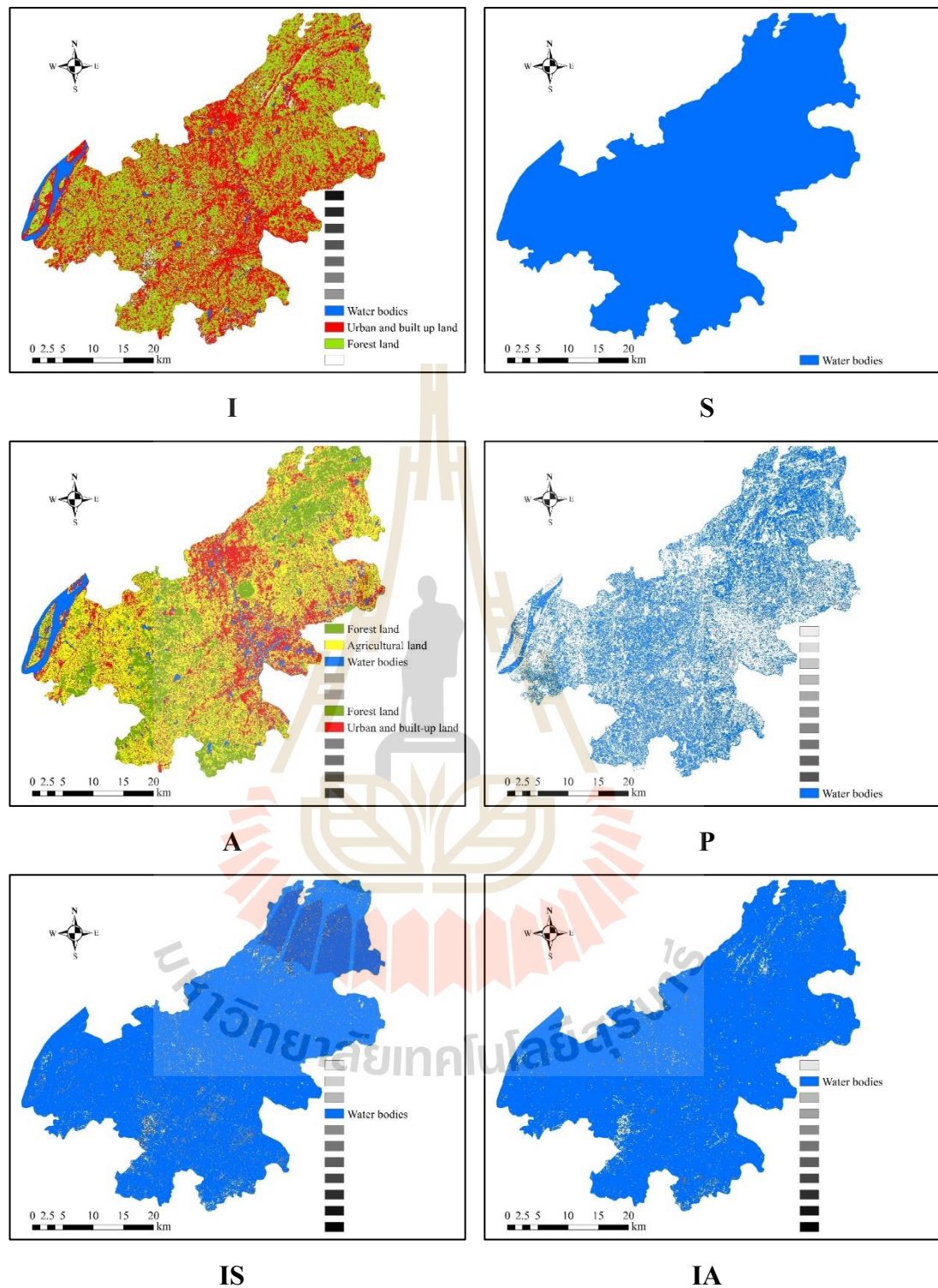


Figure 6.8 LULC classification map by ISODATA algorithm with 15 coefficient features of the NIR band.

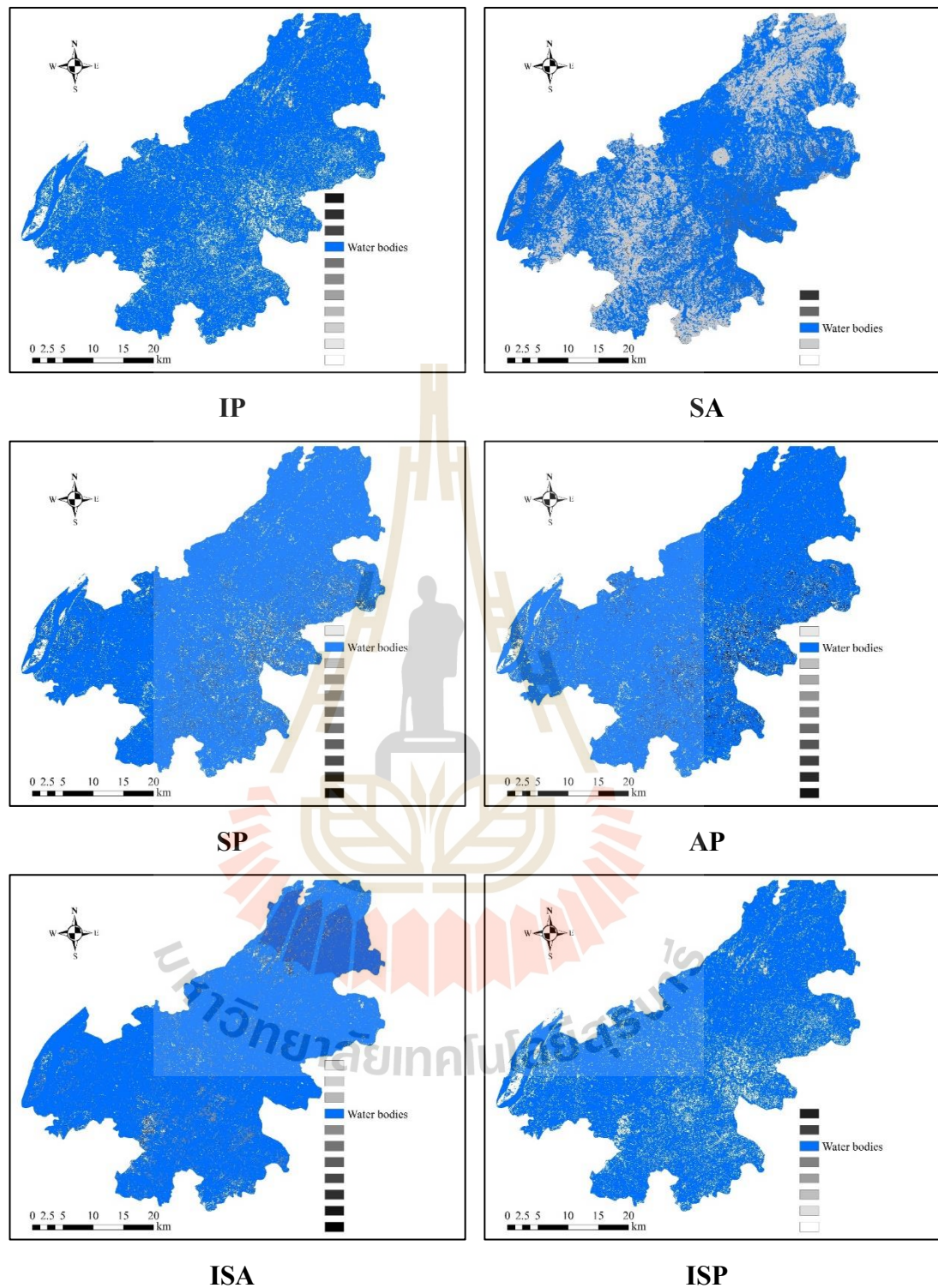


Figure 6.8 (Continued).

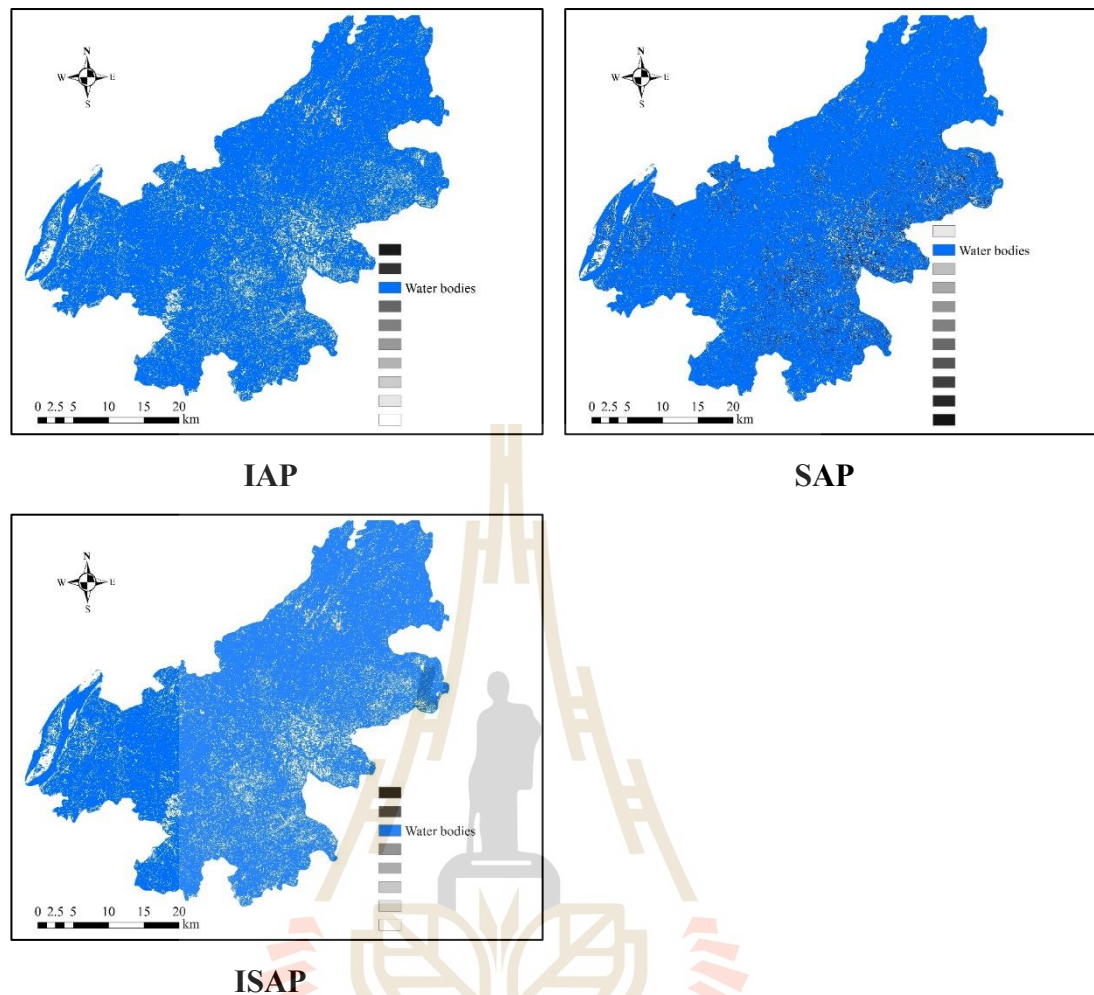


Figure 6.8 (Continued).

As a result, the amplitude (A) and slope and amplitude (SA) coefficient features can classify four different LULC types using the K-Means algorithm, and the spatial distribution of LULC type is almost the same due to the value of slope (S) is very small. In the meantime, only the amplitude (A) coefficient feature can classify all LULC types using the ISODATA algorithm.

On the contrary, most of the LULC classification maps from other coefficients features by using K-Means and ISODATA algorithms are mostly classified as water bodies. Additionally, the intercept (I) or amplitude (A) coefficient features can

easily separate water bodies from other classes, but the intercept (I) coefficient feature cannot separate agricultural land from forest land under the ISODATA algorithm.

All results above reveal that the amplitude (A) coefficient feature can provide the high potential to distinguish four different LULC types by using unsupervised K-Means and ISODATA algorithms. So, the amplitude (A) coefficient feature was here chosen as the most suitable coefficient feature for time-series LULC classification with four optimum spectral features (NIR, TOP3, TOP6, and ORI6) under unsupervised classification approach.

6.2.2 LULC classification by K-Means algorithm

The K-Means algorithm (Macqueen, 1967) is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) without fixing a priori probability.

Figures 6.9 to 6.11 show the LULC classification map in 2000, 2006, 2011 and 2017 using K-Means algorithm with ten initial clusters based on amplitude coefficients of NIR, TOP3 (NIR, EVI, and TCB) and TOP6 (NIR, EVI, TCB, BLUE, GREEN, and PC1) spectral feature(s), respectively.

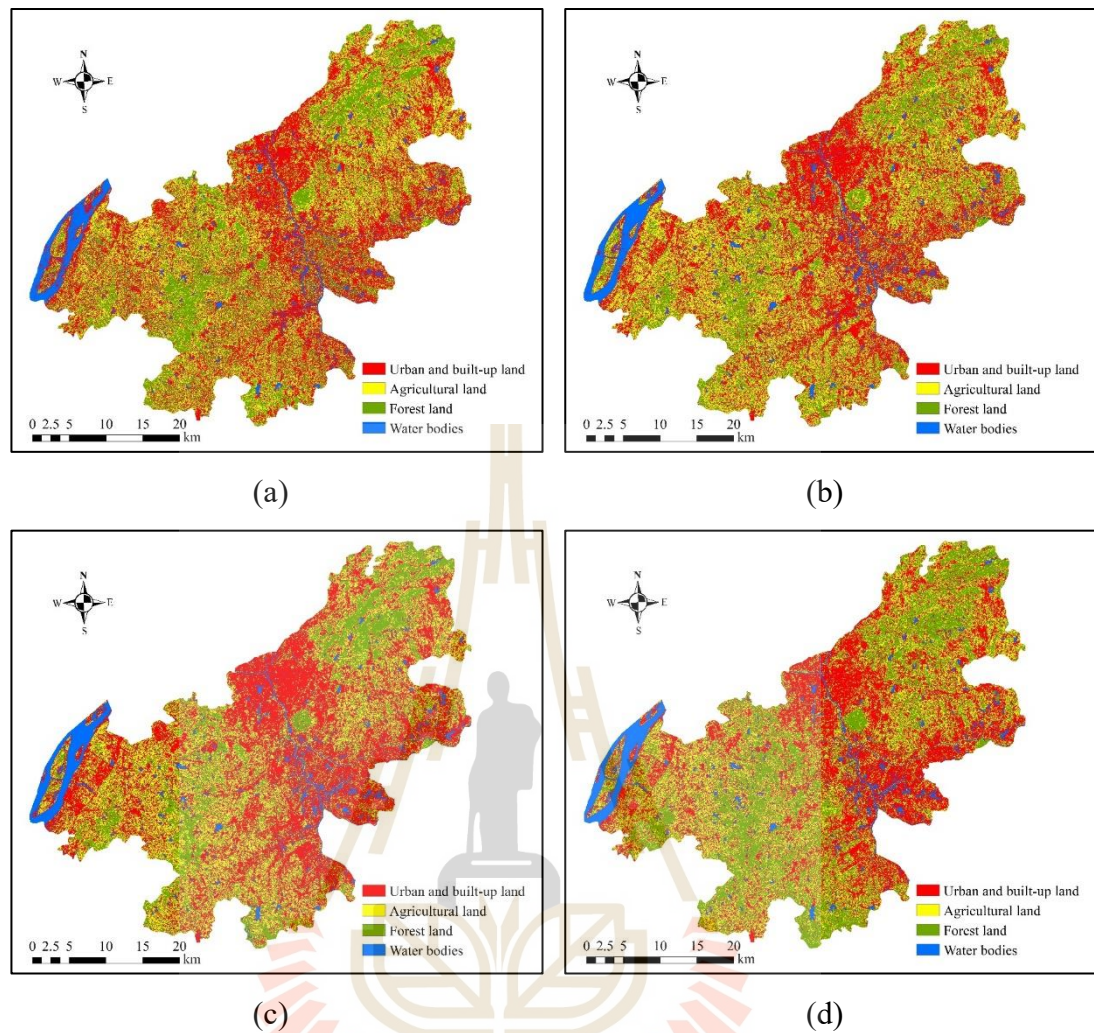


Figure 6.9 LULC classification map by K-Means with amplitude coefficient of NIR band: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) 18 May 2017.

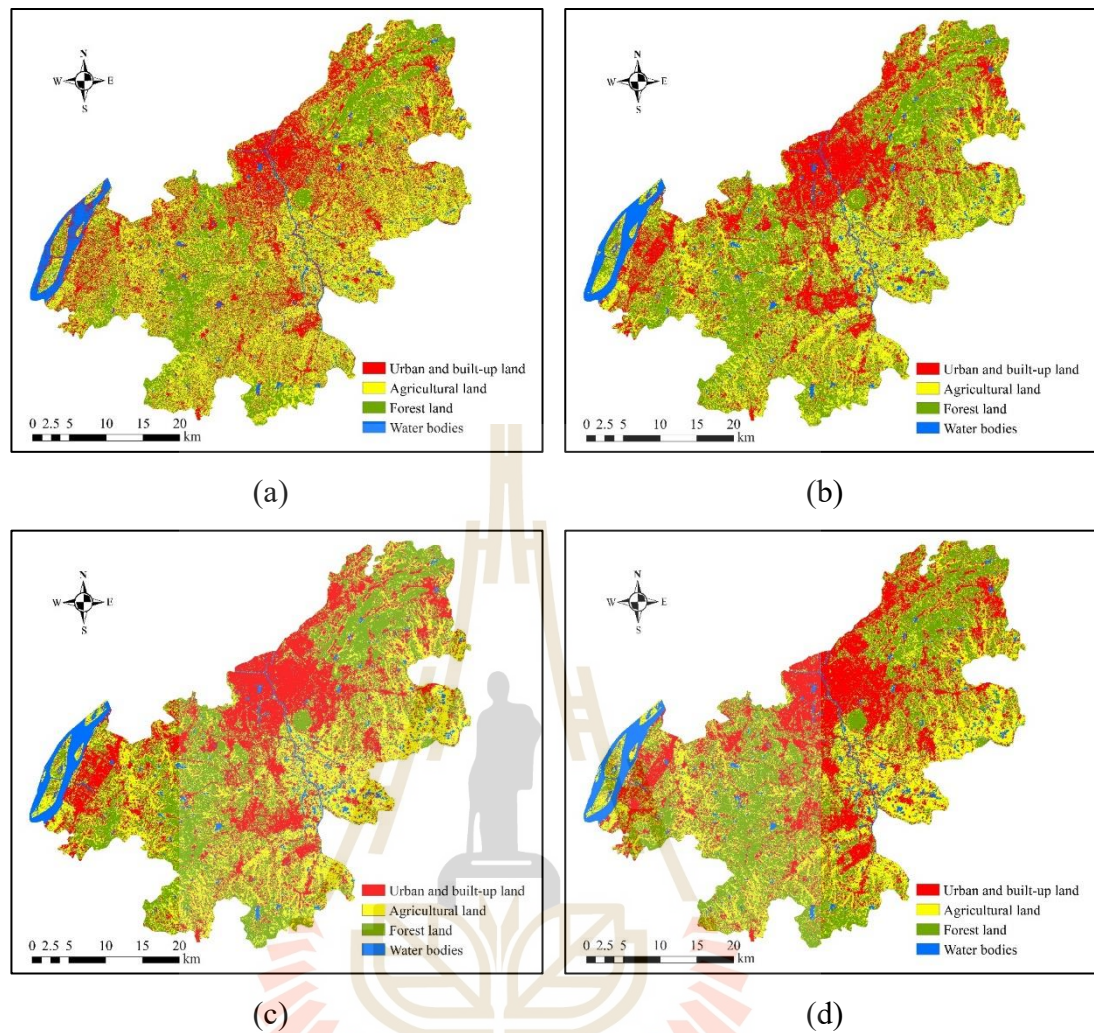


Figure 6.10 LULC classification map by K-Means with amplitude coefficients of TOP3 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) 18 May 2017.

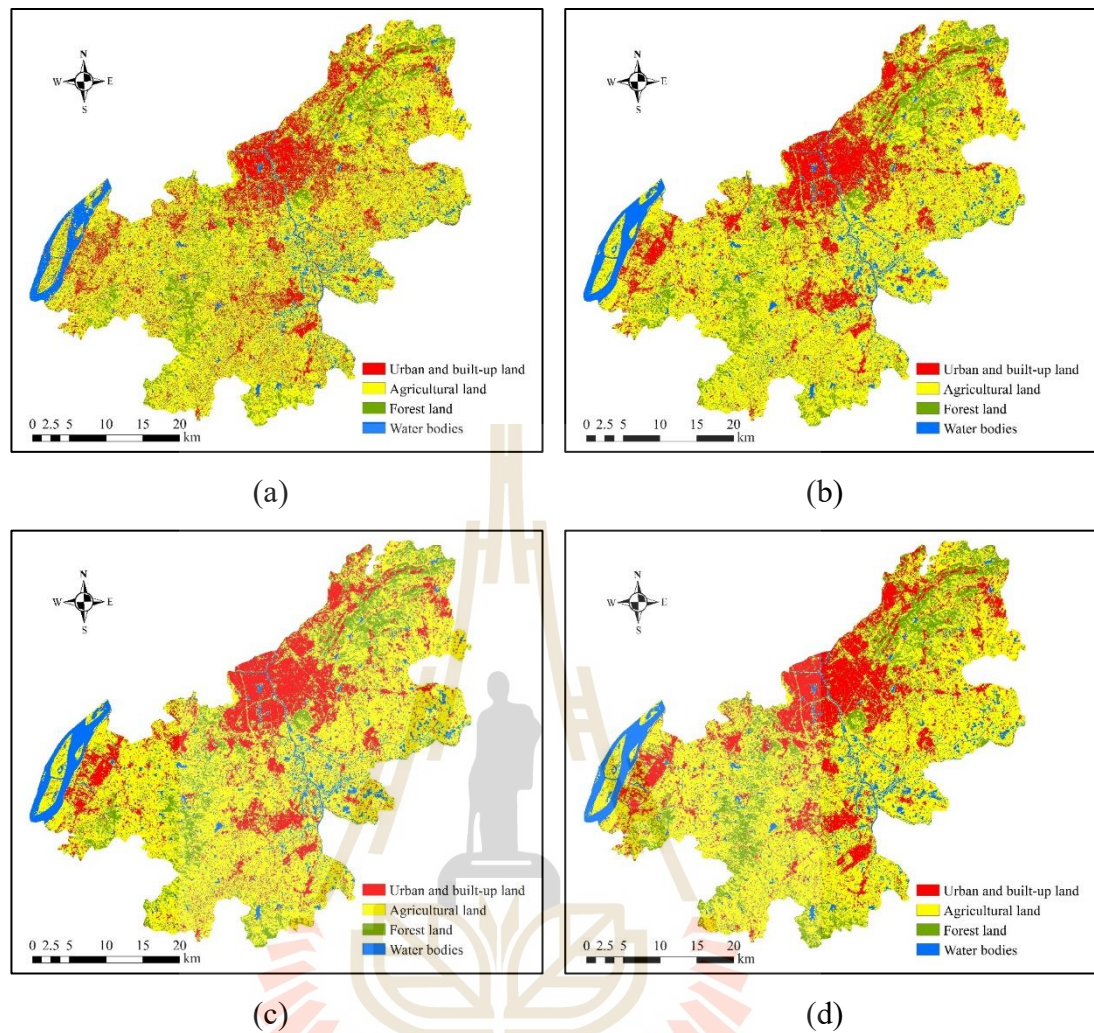


Figure 6.11 LULC classification map by K-Means with amplitude coefficients of TOP6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) 18 May 2017.

Besides, the amplitude coefficient feature of ORI6 spectral bands from Landsat data (BLUE, GREEN, RED, NIR, SWIR1, and SWIR2) was applied to classify LULC using K-Means algorithm and compared the result with TOP6 spectral features. Figure 6.12 demonstrates the LULC classification map in 2000, 2006, 2011, and 2017 using the K-Means algorithm based on the amplitude coefficient of ORI6 spectral bands.

Area and percentage of LULC classification in 2000, 2006, 2011, and 2017 using the K-Means algorithm from the various spectral feature(s) are summarized in Tables 6.11 to 6.14, respectively. The comparison of the proportional LULC type areas that were classified from the various spectral feature(s) in four different years is displayed in Figures 6.13 to 6.16, respectively.

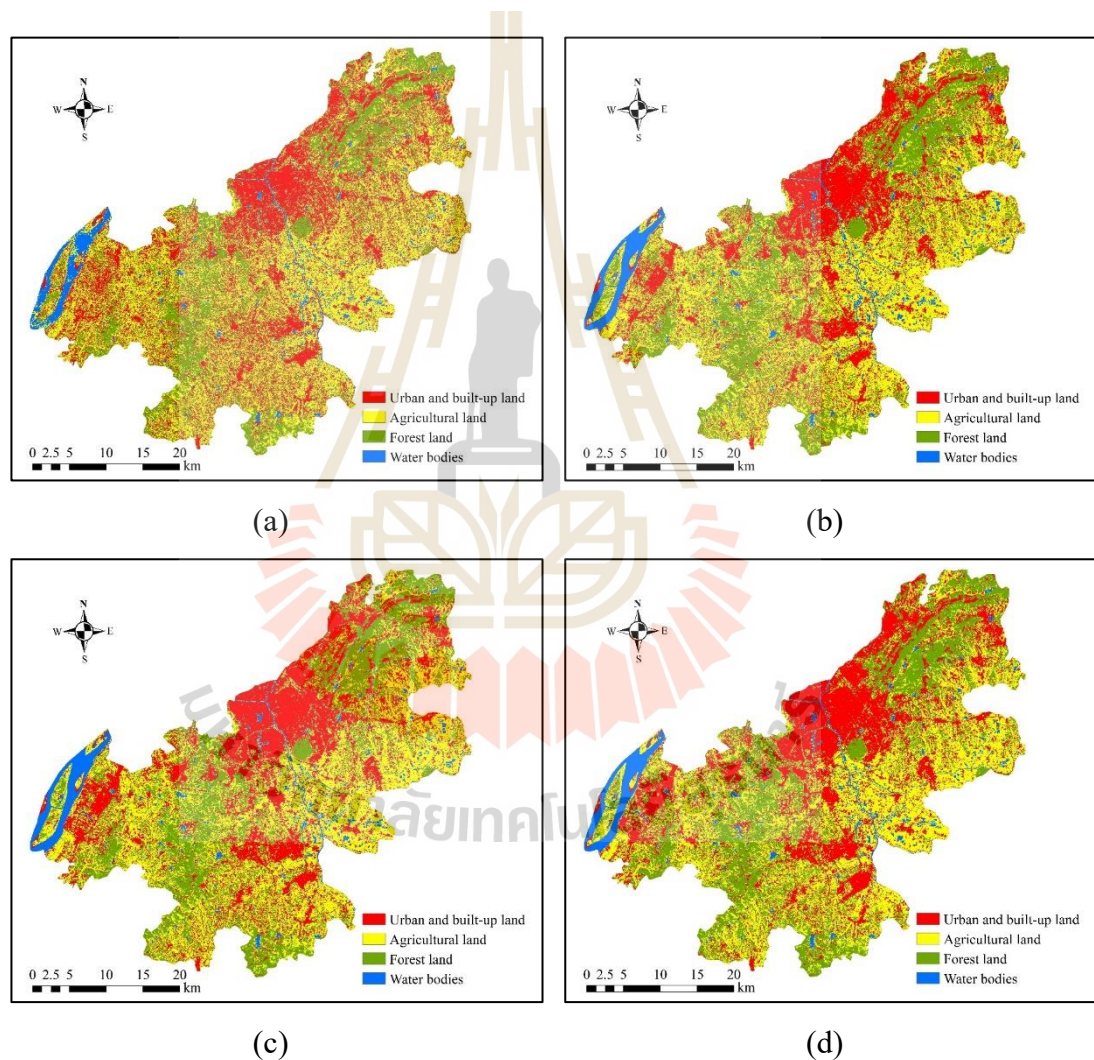


Figure 6.12 LULC classification map by K-Means with amplitude coefficients of ORI6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) 18 May 2017.

Table 6.11 Area and percentage of LULC classification in 2000.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	607.50	38.39	393.77	24.88	289.24	18.28	491.08	31.04
A	455.20	28.77	622.54	39.34	953.13	60.23	619.15	39.13
F	428.99	27.11	477.53	30.18	208.31	13.16	387.32	24.48
W	90.66	5.73	88.50	5.59	131.68	8.32	84.80	5.36

Table 6.12 Area and percentage of LULC classification in 2006.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	608.33	38.44	399.4	25.24	290.67	18.37	428.08	27.05
A	490.53	31.00	538.29	34.02	957.93	60.54	622.30	39.33
F	386.49	24.43	549.44	34.72	203.26	12.85	446.59	28.22
W	97.00	6.13	95.22	6.02	130.49	8.25	85.37	5.40

Table 6.13 Area and percentage of LULC classification in 2011.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	691.71	43.71	472.26	29.85	325.78	20.59	472.61	29.87
A	434.12	27.44	520.48	32.89	944.61	59.70	603.17	38.12
F	345.98	21.87	491.07	31.03	185.57	11.73	420.75	26.59
W	110.54	6.99	98.53	6.23	126.39	7.99	85.81	5.42

Table 6.14 Area and percentage of LULC classification in 2017.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	575.64	36.38	467.13	29.52	326.79	20.65	473.01	29.89
A	406.08	25.66	498.53	31.51	919.52	58.11	585.12	36.98
F	507.49	32.07	528.31	33.39	220.05	13.91	440.87	27.86
W	93.13	5.89	88.39	5.59	115.98	7.33	83.36	5.27

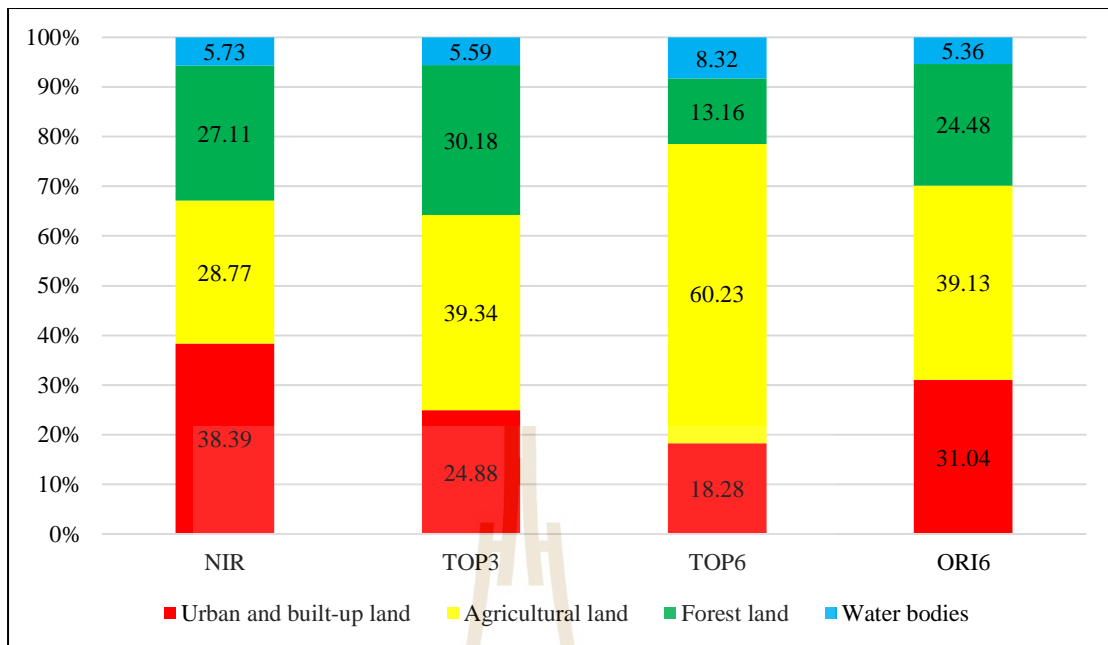


Figure 6.13 Comparison of the proportional LULC area in 2000 using the K-Means algorithm with various spectral features.

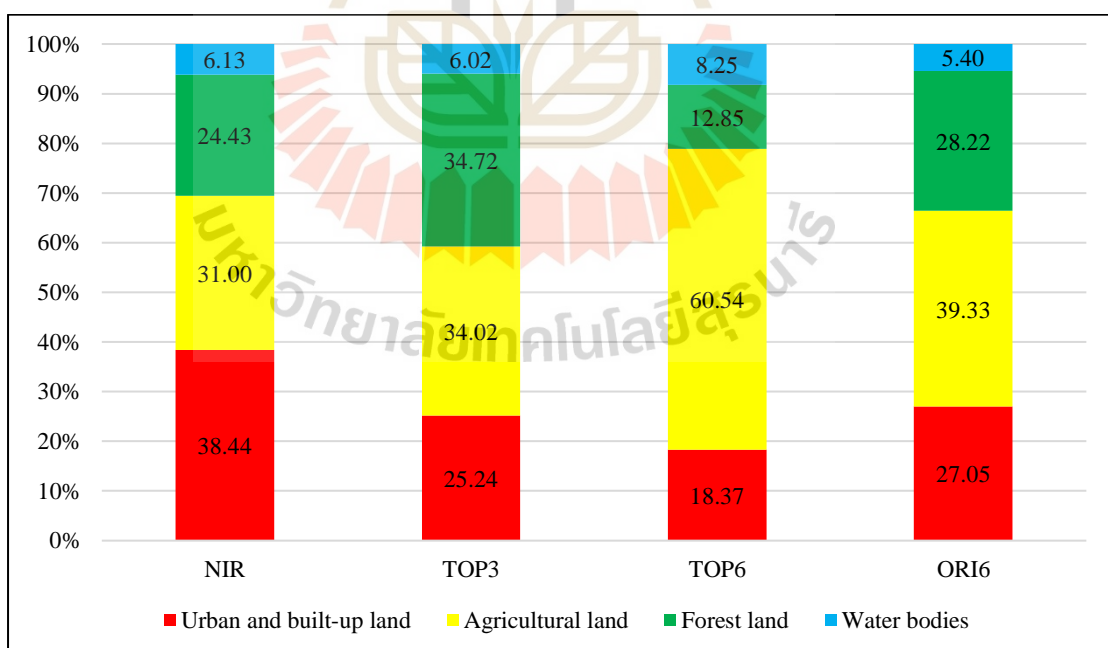


Figure 6.14 Comparison of the proportional LULC area in 2006 using the K-Means algorithm with various spectral features.

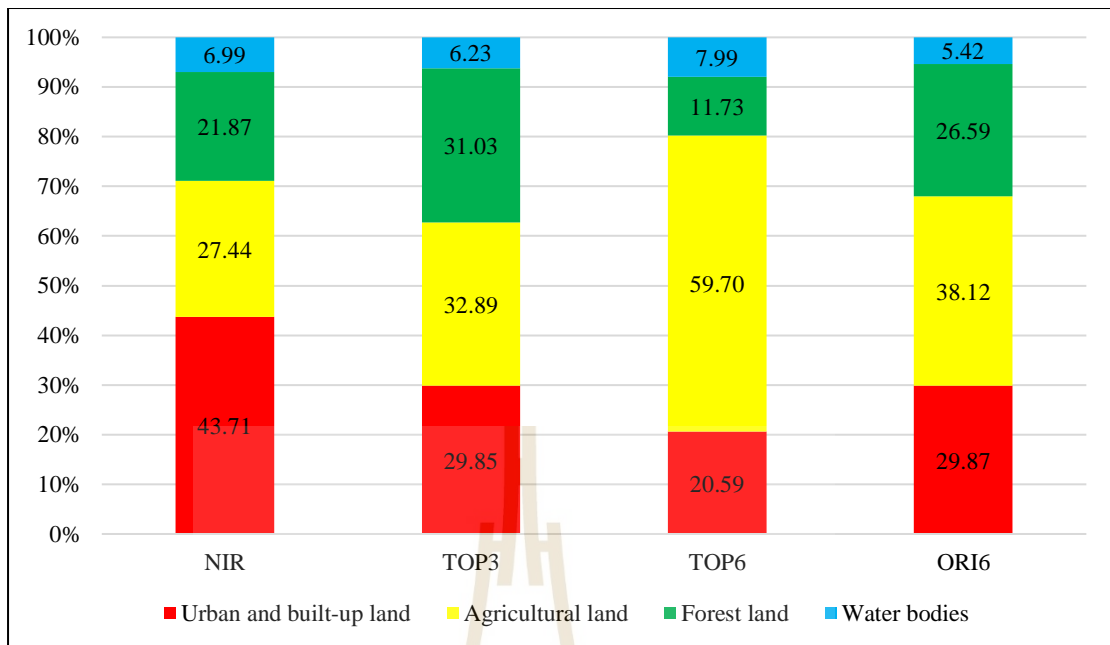


Figure 6.15 Comparison of the proportional LULC area in 2011 using the K-Means algorithm with various spectral features.

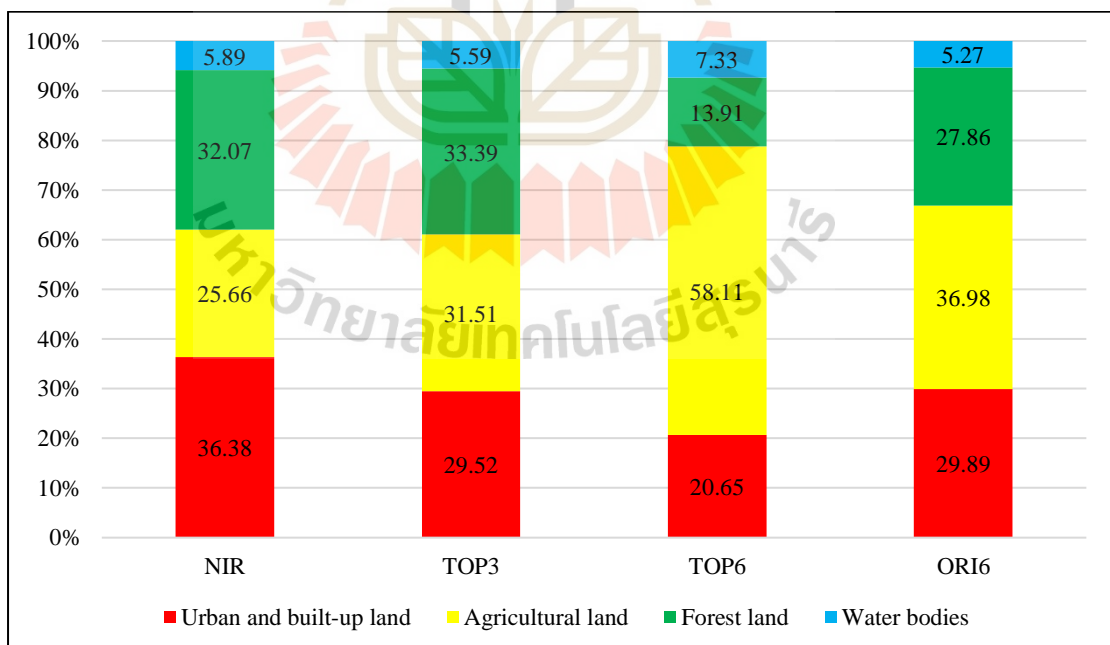


Figure 6.16 Comparison of the proportional LULC area in 2017 using the K-Means algorithm with various spectral features.

By comparing area and percentage among LULC types according to spectral feature, it reveals that area and percentage of agricultural land which were extracted based on the TOP6 spectral features are somewhat high when they are compared with other spectral features in every year. The percent of agricultural land varies from 58.11% in 2017 to 60.54% in 2006. Similarly, the classified water bodies based on TOP6 spectral features are also high when they are compared with other spectral features in every year. The percent of water bodies varies from 8.32% in 2000 to 7.33% in 2017.

On the contrary, the classified forest land based on TOP6 spectral features is rather low when they are compared with other spectral features in every year. The percent of forest land varies from 11.73% in 2011 to 13.91% in 2017. Likewise, the classified urban and built-up land based on the TOP6 spectral features is also rather low when it is compared with other spectral features in every year. The percent of urban and built-up land varies from 18.28% in 2000 to 20.65% in 2017.

These findings show the impact (influence) of the existing number of the natural clustering based on values of spectral reflectance and indices among spectral feature(s) combination on LULC classification using the K-Means algorithm. The derived spectral classes from ten initial clusters (classes) not obviously represent specific LULC types (U, A, F, and W). Based on this reason, areas of urban and built-up land, which include developed land and bareland, was dramatically increased in 2011, and it was dramatically decreased in 2017.

6.2.3 LULC classification by ISODATA algorithm

The ISODATA unsupervised classification calculates class means evenly distributed in the data space then iteratively clusters the remaining pixels using minimum distance techniques. Each iteration recalculates means and reclassifies pixels for the new means. Iterative class splitting, merging, and deleting is done based on input threshold parameters. All pixels are classified to the nearest class unless a standard deviation or distance threshold is specified, in which case some pixels may be unclassified if they do not meet the selected criteria. This process continues until the number of pixels in each class changes by less than the selected pixel change threshold, or the maximum number of iterations is reached (Jensen, 2017).

Similar to K-Means algorithm, Figures 6.17 to 6.19 demonstrate the LULC classification map in 2000, 2006, 2011 and 2017 using ISODATA algorithm based on amplitude coefficient of NIR TOP3 (NIR, EVI, and TCB), and TOP6 (NIR, EVI, TCB, BLUE, GREEN, and PC1) spectral feature(s), respectively.

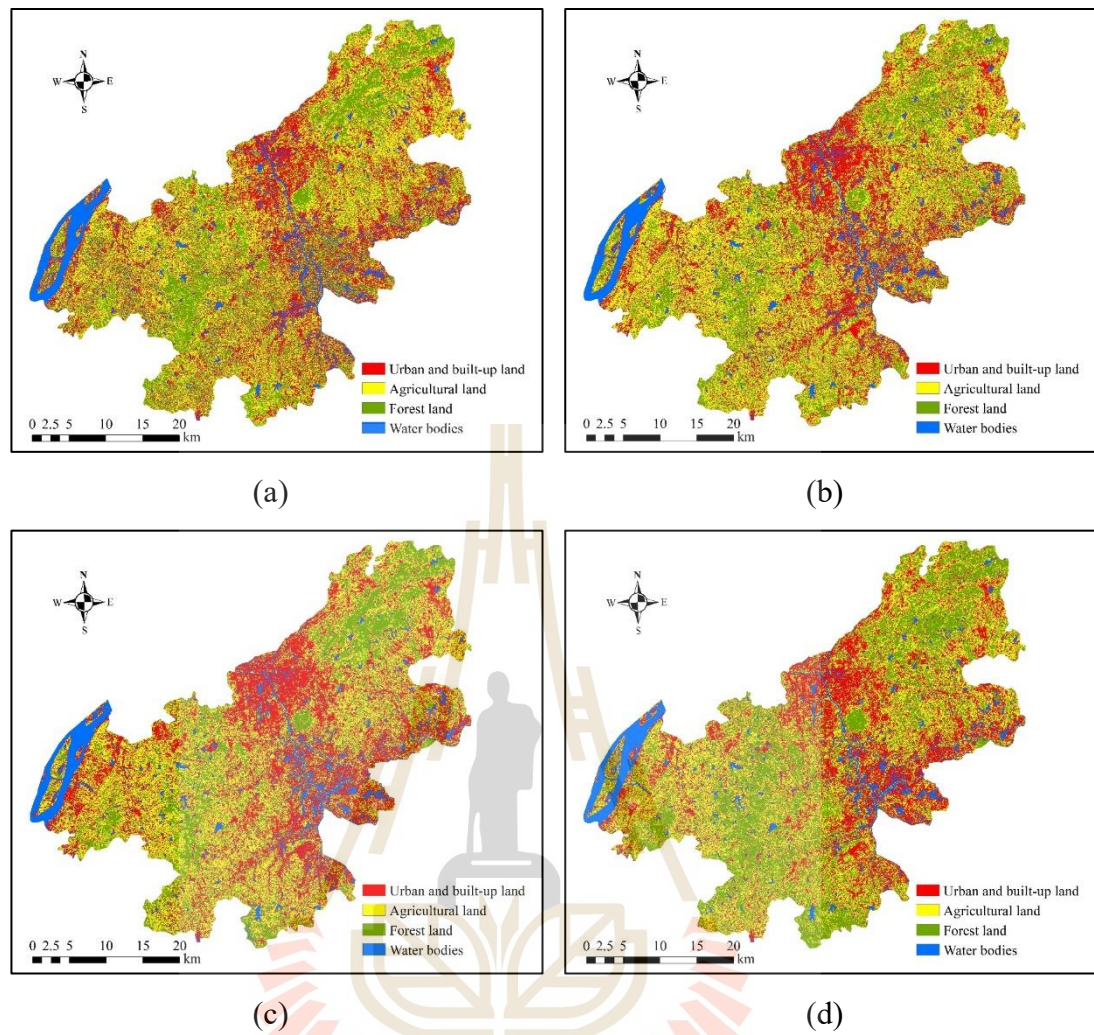


Figure 6.17 LULC classification map by ISODATA with amplitude coefficient of NIR band: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.

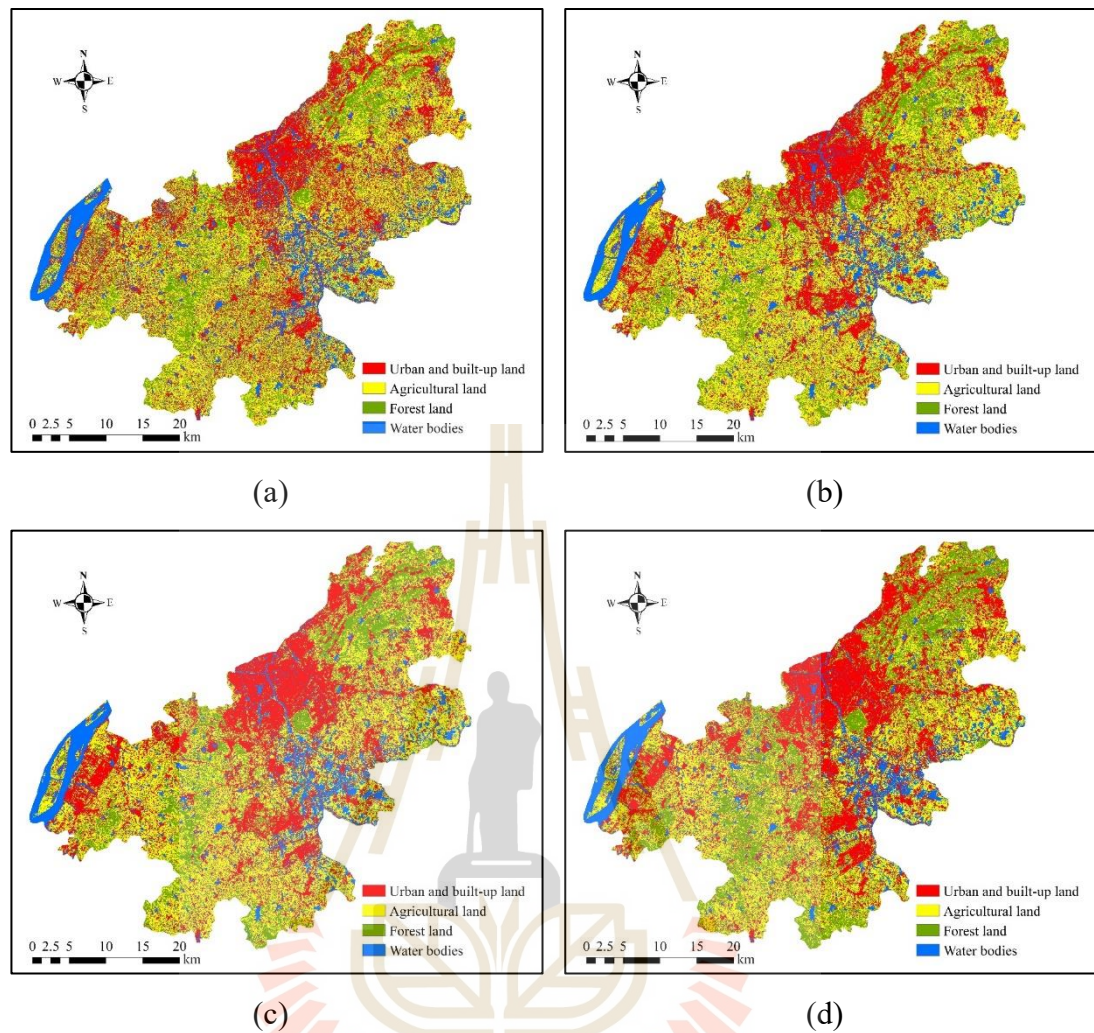


Figure 6.18 LULC classification map by ISODATA with amplitude coefficients of TOP3 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.

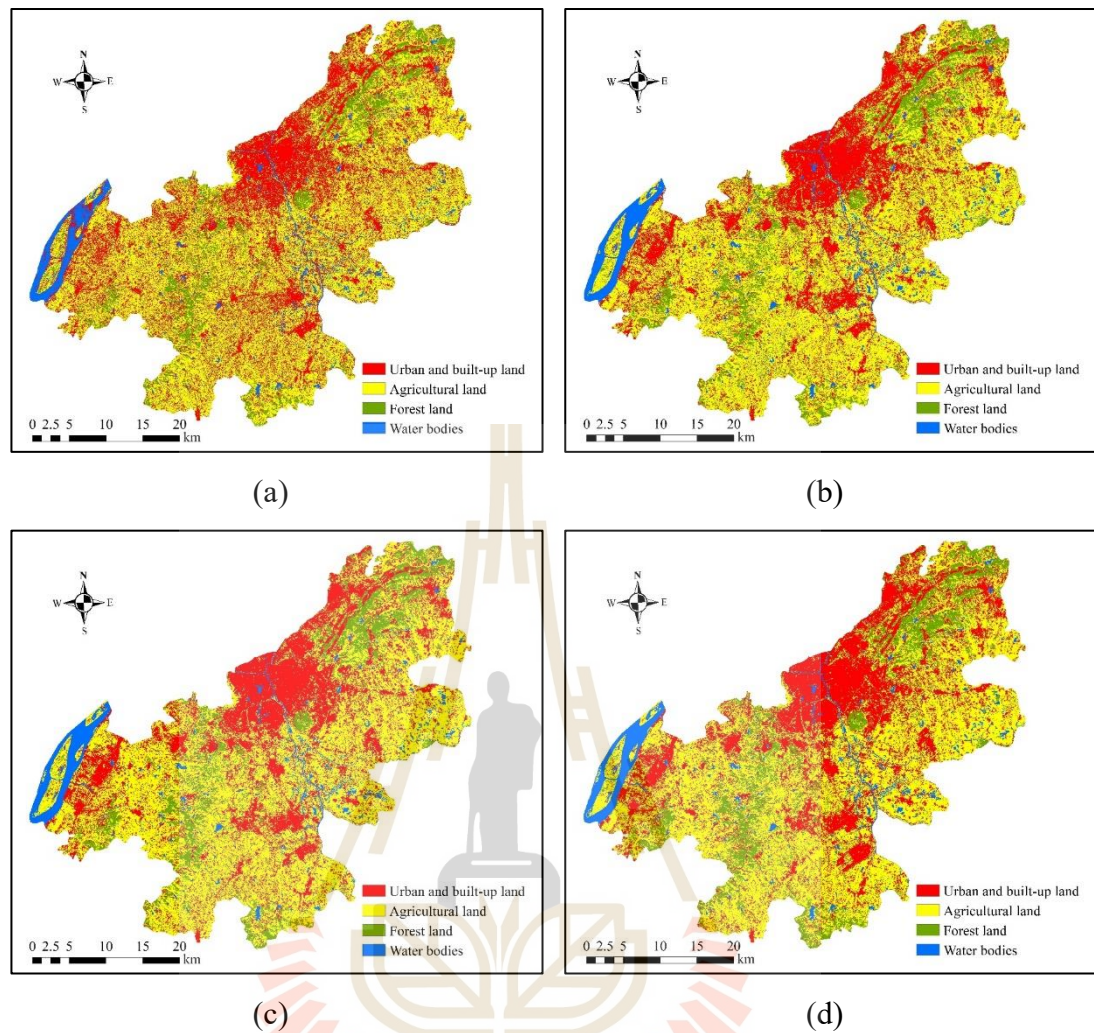


Figure 6.19 LULC classification map by ISODATA with amplitude coefficients of TOP6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.

Like the K-Means algorithm, the amplitude coefficient feature of ORI6 spectral bands from Landsat data was also applied to classify LULC using the ISODATA algorithm and compared the result with TOP6 spectral features. Figure 6.20 shows the LULC classification maps in 2000, 2006, 2011, and 2017 using the ISODATA algorithm based on the amplitude coefficient of ORI6 spectral bands.

Area and percentage of LULC classification in 2000, 2006, 2011, and 2017 using the ISODATA algorithm from various spectral features are summarized in Tables 6.15 to 6.18, respectively. The comparison of the proportional LULC type areas that were classified using the ISODATA algorithm from various spectral features in different four years is displayed in Figures 6.21 to 6.24, respectively.

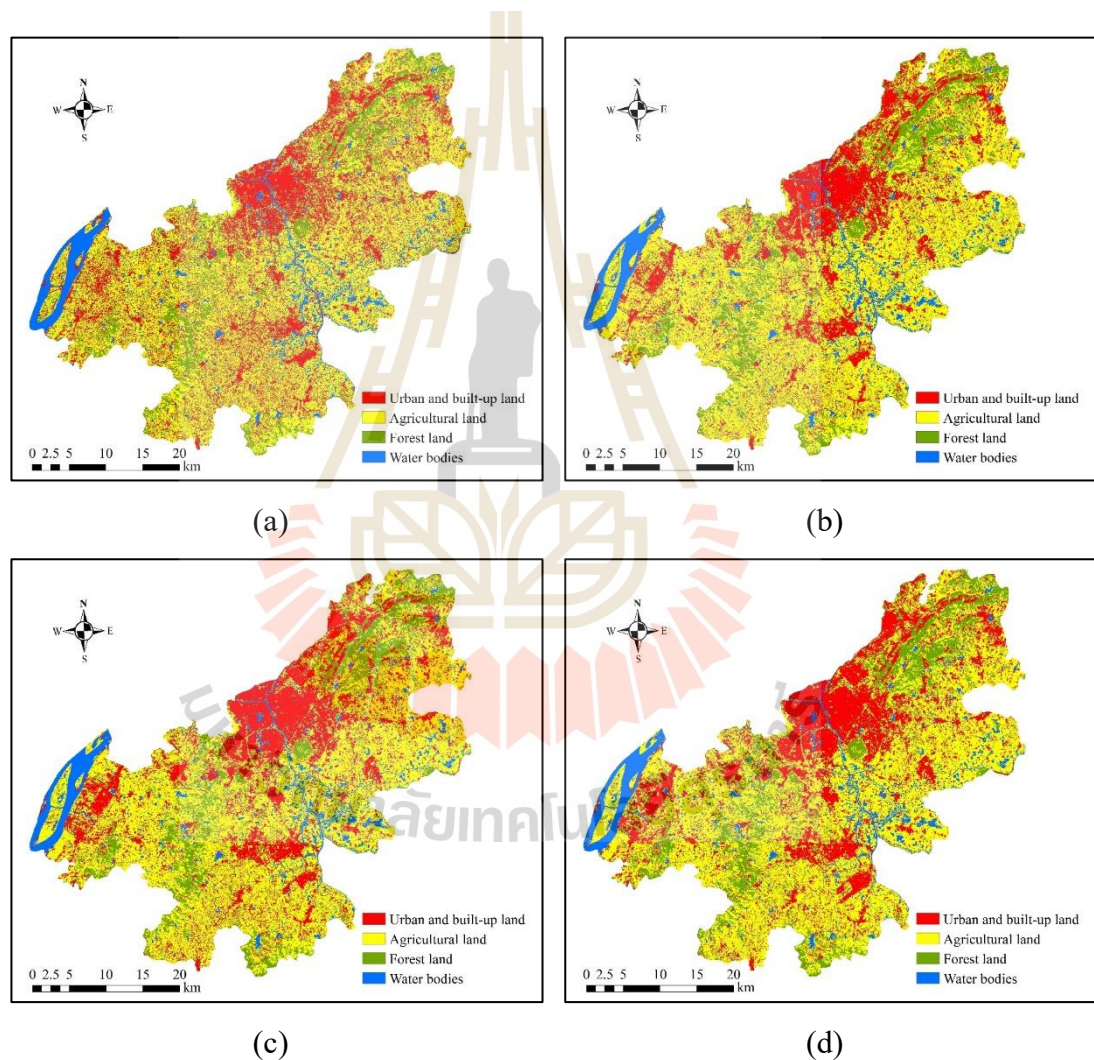


Figure 6.20 LULC classification map by ISODATA with amplitude coefficients of ORI6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.

Table 6.15 Area and percentage of LULC classification in 2000.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	436.55	27.59	444.38	28.08	499.27	31.55	413.35	26.12
A	577.41	36.49	604.55	38.21	803.19	50.76	827.69	52.31
F	416.62	26.33	292.41	18.48	191.59	12.11	181.39	11.46
W	151.76	9.59	241.01	15.23	88.30	5.58	159.91	10.11

Table 6.16 Area and percentage of LULC classification in 2006.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	440.20	27.82	443.43	28.02	442.61	27.97	358.44	22.65
A	626.78	39.61	637.91	40.31	830.28	52.47	877.08	55.43
F	373.91	23.63	299.99	18.96	213.61	13.50	213.05	13.46
W	141.46	8.94	201.01	12.70	95.84	6.06	133.78	8.45

Table 6.17 Area and percentage of LULC classification in 2011.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	514.08	32.49	509.91	32.22	480.78	30.38	411.00	25.97
A	566.63	35.81	582.82	36.83	813.64	51.42	841.67	53.19
F	335.60	21.21	273.19	17.26	193.48	12.23	195.87	12.38
W	166.03	10.49	216.44	13.68	94.44	5.97	133.81	8.46

Table 6.18 Area and percentage of LULC classification in 2017.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	426.34	26.94	504.37	31.87	483.89	30.58	414.21	26.18
A	525.94	33.24	517.59	32.71	792.64	50.09	831.28	52.53
F	496.72	31.39	373.45	23.60	218.26	13.79	211.47	13.36
W	133.34	8.43	186.94	11.81	87.56	5.53	125.39	7.92

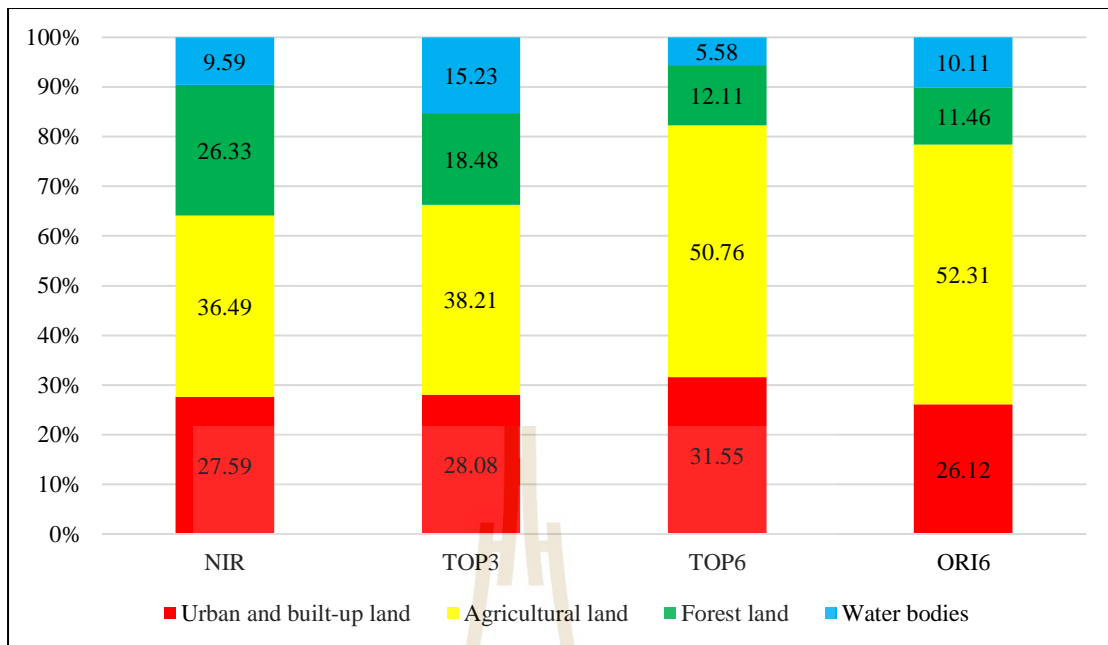


Figure 6.21 Comparison of the proportional LULC area in 2000 using the ISODATA algorithm with various spectral features.

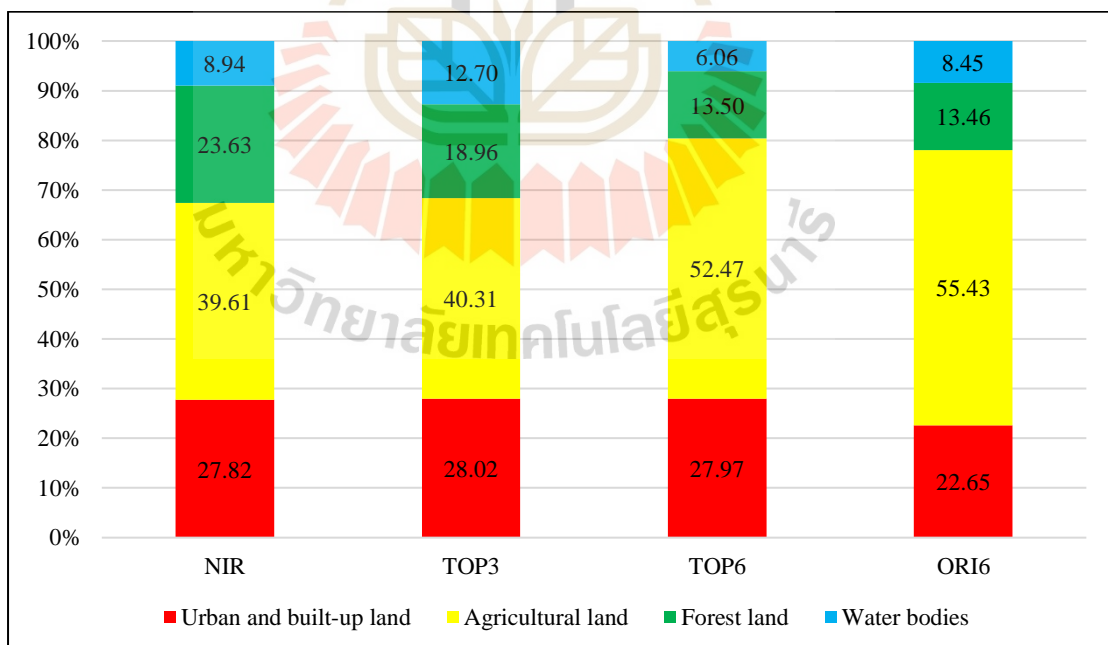


Figure 6.22 Comparison of the proportional LULC area in 2006 using the ISODATA algorithm with various spectral features.

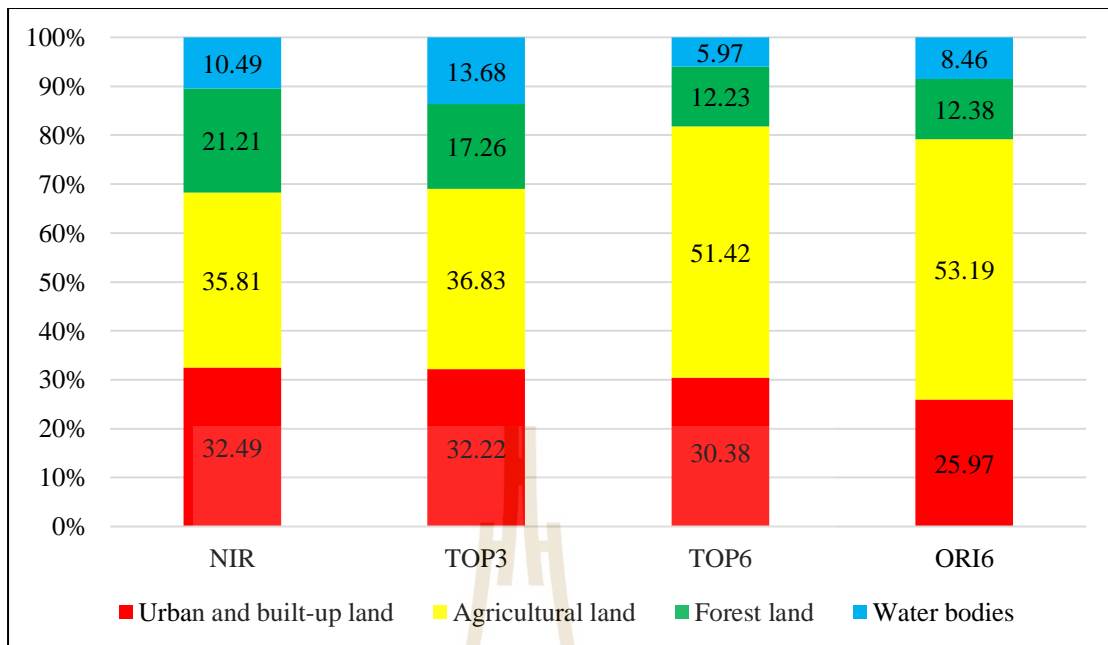


Figure 6.23 Comparison of the proportional LULC area in 2011 using the ISODATA algorithm with various spectral features.

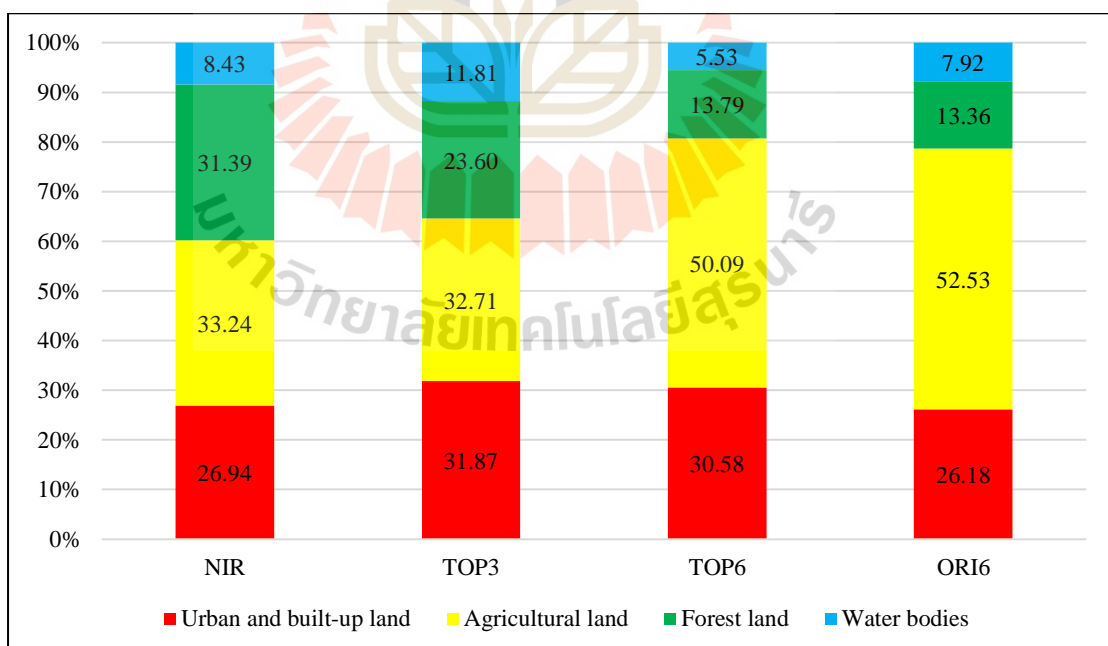


Figure 6.24 Comparison of the proportional LULC area in 2017 using the ISODATA algorithm with various spectral features.

By comparison of area and percentage among LULC types according to spectral feature, it reveals that area and percentage of agricultural land which were extracted based on TOP6 and ORI6 spectral features are somewhat high when they are compared with other spectral features in every year. The percent of agricultural land with TOP6 spectral features varies from 50.09% in 2017 to 52.47% in 2006, and the percent of agricultural land with ORI6 spectral features varies from 52.31% in 2000 to 55.43% in 2006.

On the contrary, the percent of classified forest land based on NIR or TOP3 spectral features is rather high when it is compared with other spectral features in every year. The percent of forest land with the NIR band varies from 21.21% in 2011 to 31.39% in 2017, and the percent of forest land with TOP3 spectral features varies from 17.26% in 2011 to 23.60% in 2017. Likewise, the percent of classified water bodies based on NIR and TOP3 spectral features is rather high when it is compared with other spectral features in every year. The percent of water bodies with the NIR band varies from 8.43% in 2017 to 10.49% in 2011, and the percent of water bodies with TOP3 spectral features varies from 11.81% in 2017 to 15.23% in 2000.

In the meantime, area and the percentage of urban and built-up land which was extracted based on ORI6 spectral features are rather low when they are compared with other spectral features in every year. The percent of urban and built-up land with ORI6 spectral features varies from 22.65% in 2006 to 26.18% in 2017.

Like the K-Means algorithm, these findings also show the impact of the existing number of the natural clustering based on values of spectral reflectance and indices among spectral feature(s) combination on LULC classification using ISODATA algorithm. The derived spectral classes from one to fifteen initial clusters

(classes) are not obviously represented for specific LULC types (U, A, F, and W). Based on this reason, areas of urban and built-up land, which include developed land and bareland, was dramatically increased in 2011, and it was dramatically decreased in 2017.

6.3 Accuracy assessment

The accuracy assessment consists of two aspects, one is whether LULC changes can be detected, and the other is whether the LULC classification before and after the change is correct. In this study, standard thematic accuracy assessment was assessed using overall accuracy (OA), producer's accuracy (PA), users' accuracy (UA), and Kappa hat coefficient based on error matrix between classified LULC data and ground reference information data (Congalton and Green, 2009).

For thematic accuracy assessment, the number of sample size was estimated based on the multinomial distribution (Congalton and Green, 2009; Tortora, 1978) using Eq. 6.1 and sample points were allocated for thematic accuracy assessment using a stratified random sampling technique.

$$N = \frac{B\Pi_i(1-\Pi_i)}{b_i^2} \quad (6.1)$$

Where, Π_i is the portion of a population in the i^{th} class out of k classes that has the proportion closest to 50%, b_i is the desired precision (e.g., 5%) for the class, B is the upper $(\alpha/k) * 100$ percentile of the chi-square (χ^2), distribution with 1 degree of freedom, k is the number of classes.

In this study, 636 sample points based on the multinomial distribution with the desired precision of 5%, and a level of confidence of 95% was applied to access thematic accuracy assessment.

In the meantime, pan-sharpened images with Gram-Schmidt pan-sharpening algorithm (Figure 6.25) as primary reference data and Google Earth images (Figure 6.26) as the second reference data were used as ground reference information for accuracy assessment of LULC change detection and classification. In general, the spatial resolution of Google Earth image is higher Landsat image, but usually, the acquisition time of these images is inconsistent with Landsat image. Single-use of Google Earth image as an accuracy verification reference image may bring some errors. The pan-sharpened image comes from the fusion of Landsat high spatial resolution panchromatic image, and a low spatial resolution multispectral image may be required. The acquisition time is consistent with the LULC classification map to be verified, but the spatial resolution is lower than the Google Earth image.

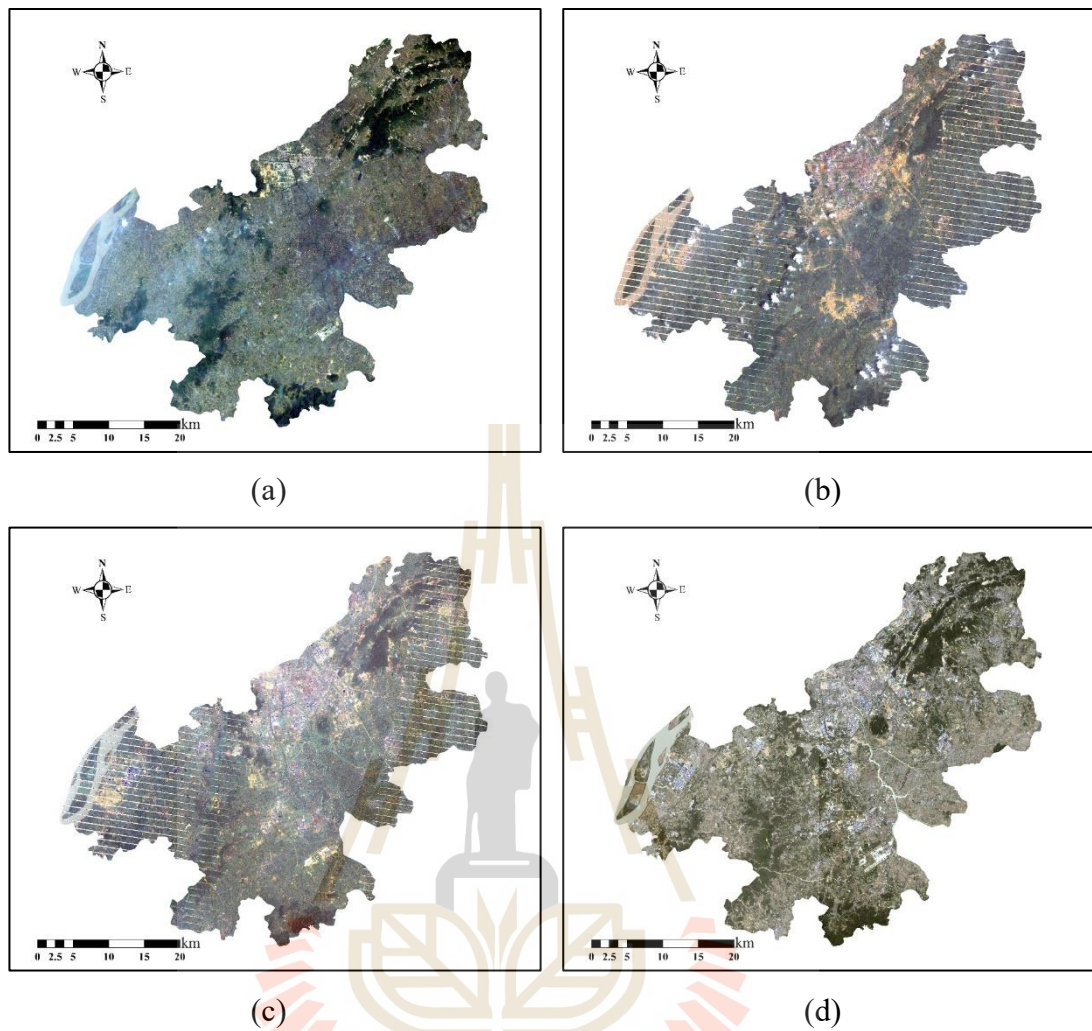


Figure 6.25 Pan-sharpened image from Landsat data in 4 different years: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.

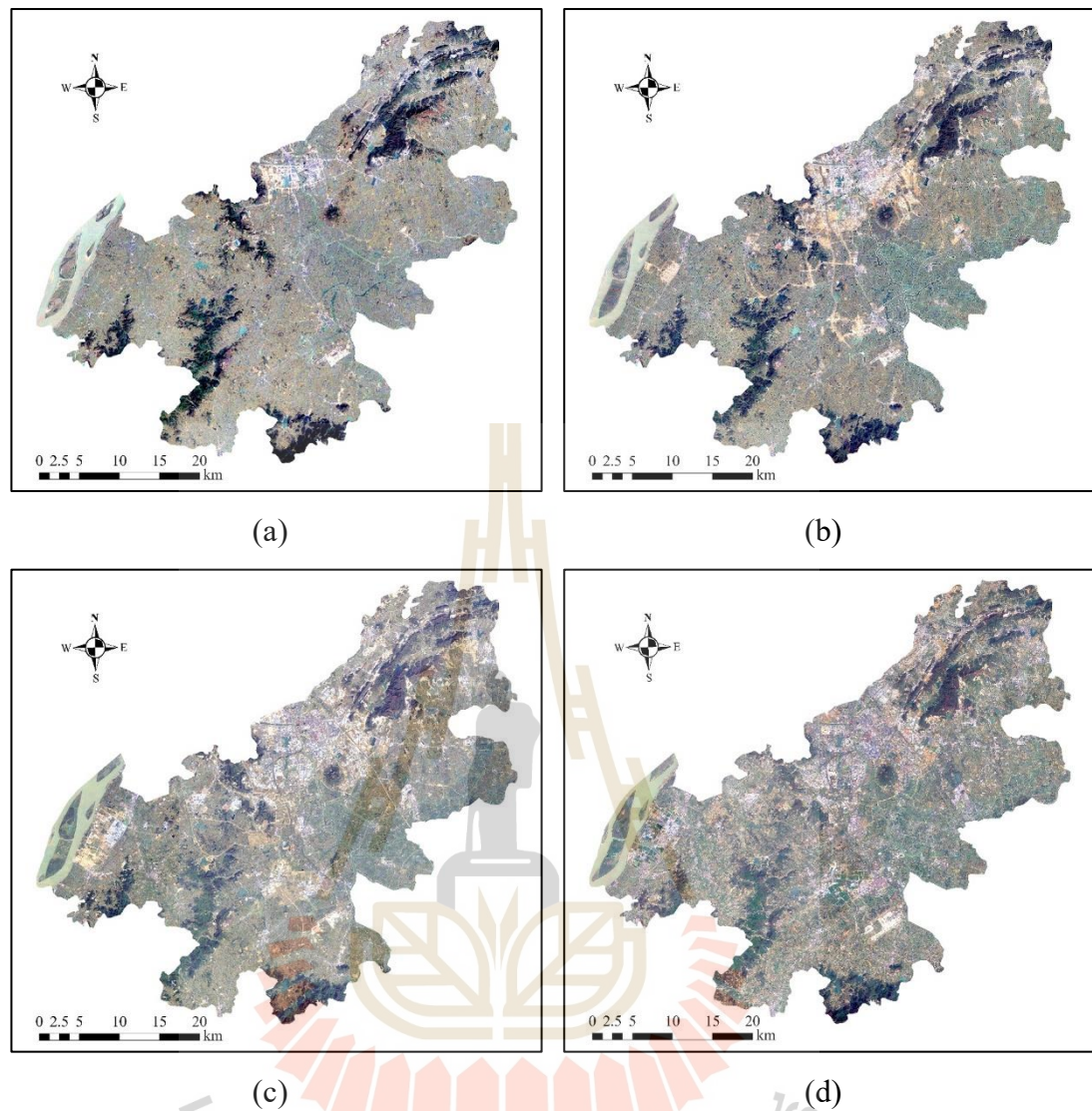


Figure 6.26 Images from Google Earth in 4 different years: (a) Date 2000, (b) Date 2006, (c) Date 2011, and (d) Date 2017.

Besides, the pairwise Z-test was conducted to examine the significant difference of thematic accuracy based on kappa hat coefficient values among various methods (Congalton and Green, 2009) as:

$$Z = \frac{|Khat_1 - Khat_2|}{\sqrt{\text{var}(Khat_1) + \text{var}(Khat_2)}} \quad (6.2)$$

and variance of KHAT is calculated by:

$$\widehat{\text{var}}(\widehat{K}) = \frac{1}{n} \left\{ \frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2-\theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4-4\theta_2)^2}{(1-\theta_2)^4} \right\} \quad (6.3)$$

when:

$$\begin{aligned} \theta_1 &= \frac{1}{n^2} \sum_{i=1}^k n_{ii}, & \theta_2 &= \frac{1}{n^2} \sum_{i=1}^k n_{i+n_i} \\ \theta_3 &= \frac{1}{n^2} \sum_{i=1}^k n_{ii}(n_{i+n_i}), & \theta_4 &= \frac{1}{n^3} \sum_{i=1}^k \sum_{j=1}^k n_{ij}(n_{j+n_i})^2 \end{aligned}$$

Z is standardized and normally distributed. Given the null hypothesis H0: (K1 - K2) = 0, and the alternative H1: (K1 - K2) ≠ 0, H0 is rejected if $Z \geq Z_{\alpha/2}$.

Finally, the optimum method among unsupervised and supervised methods will be identified using overall accuracy, Kappa hat coefficient, and pairwise Z-test.

The result of the accuracy assessment in two aspects is described and discussed in the following sections.

6.3.1 Accuracy assessment for change detection

Refer to area and percentage of LULC classification using the K-Means and ISODATA (in Tables 6.11 to 6.14 and Tables 6.15 to 6.18) in the previous section; they have shown possible variations in LULC for all cells in the entire study area when different spectral features are used. Two different classification algorithms determine the extent about LULC change (i.e., change and no change), and they provide from-to change information (i.e., one LULC type change to another LULC type). The combination of the LULC classification map using two different algorithms with different spectral features on different dates cannot directly applicable for change detection accuracy assessment. Therefore, in order to simplify the calculation, overlay analysis was applied here to identify stable and unstable pixels from LULC maps in

four different years (2000, 2006, 2011, and 2017). If areas are stable pixels, they are unchanged, but if areas are unstable pixels, they are changed.

Figure 6.27 and Figure 6.28 show the result of LULC change detection maps (stable LULC type and LULC change area) with different spectral features by K-Means and ISODATA algorithms, respectively.

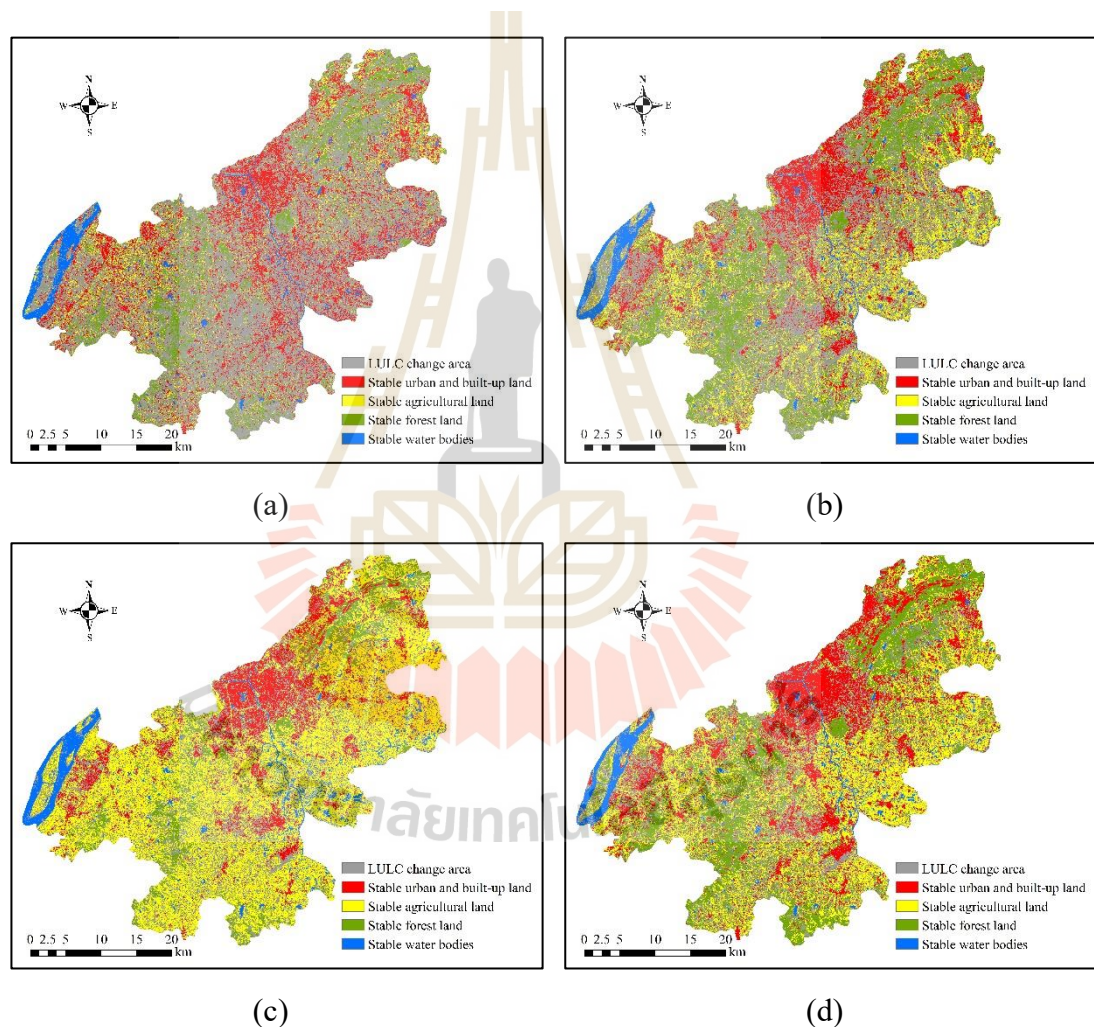


Figure 6.27 LULC change detection by the K-Means algorithm: (a) Using the NIR band, (b) Using TOP3 spectral features, (c) Using TOP6 spectral features, and (d) Using ORI6 spectral features.

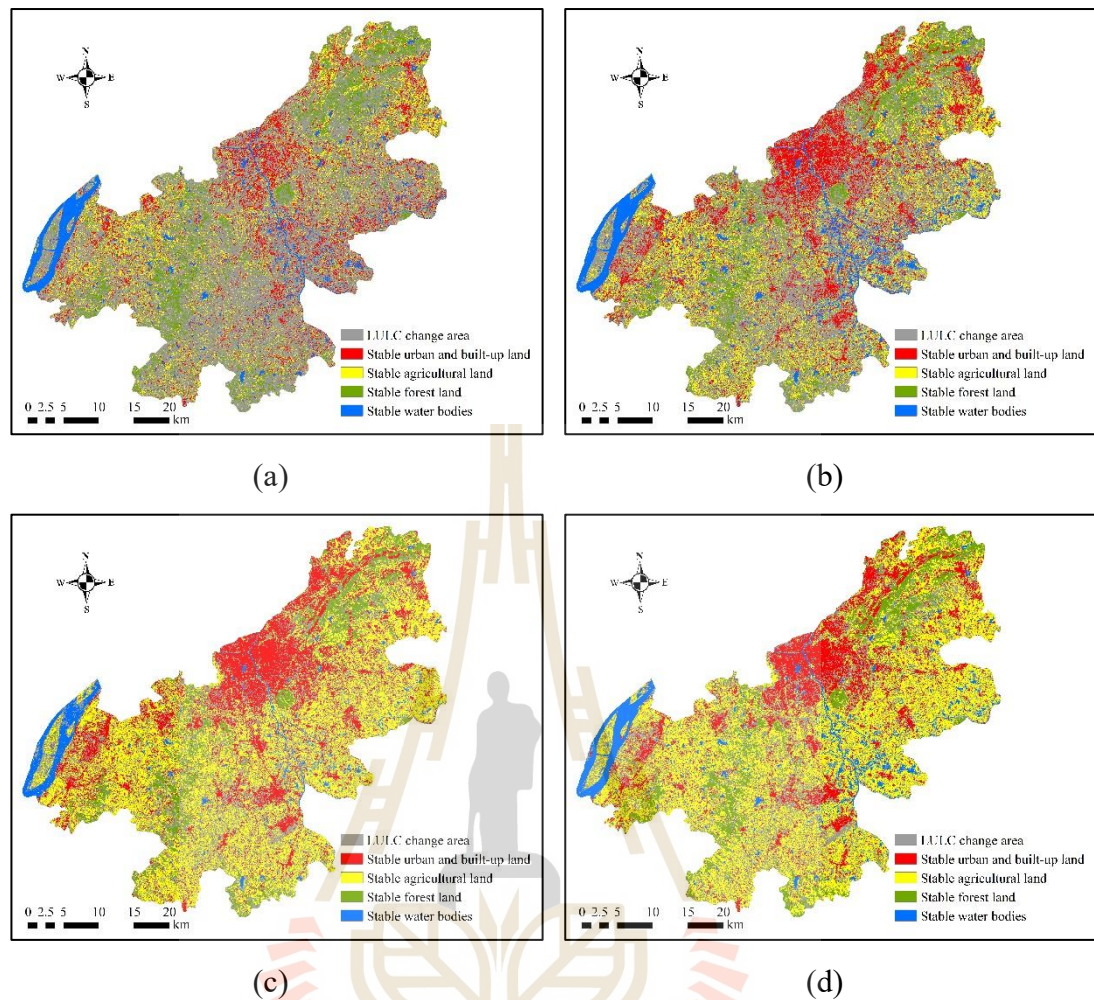


Figure 6.28 LULC change detection by the ISODATA algorithm: (a) Using the NIR band, (b) Using TOP3 spectral features, (c) Using TOP6 spectral features, and (d) Using ORI6 spectral features.

As the algorithms are capable of detecting change at a high temporal frequency, it is difficult to find reference data that can thoroughly assess its accuracy both spatially and temporally. There are no independent datasets available that have both more excellent spatial resolution and higher temporal frequency than Landsat images over the periods of Landsat series. To know where and when land cover change occurs, the primary source for reference data is the Landsat images themselves (Cohen,

Yang, and Kennedy, 2010). High spatial resolution images from Google Earth can help manual interpretation of LULC classes for accuracy assessment; however, the high spatial resolution images cannot provide the same temporal frequency as Landsat data, but the high spatial resolution images from the Google Earth are conducive to determining LULC change at more prolonged periods.

To detect LULC changes, whether different algorithms with different spectral features at a particular location, the four stable LULC and change areas in Figures 6.27 to 6.28 were reclassified into two groups: changed and unchanged areas, for change detection accuracy assessment.

A random stratified sampling scheme was here used for assessing the accuracy of change detection with four different spectral features using K-Means and ISODATA algorithms. Herein, a total of 502 reference pixels based on multinomial distribution with the desired precision of 5% and a level of confidence of 95% were applied for accuracy assessment.

Tables 6.19 to 6.22 show the results of the accuracy assessment of change detection between 2000 and 2017. Generally, the accuracy evaluation results show that for the same spectral feature(s) combinations, the K-Means and ISODATA algorithms provide a small difference in producer's accuracy, user's accuracy, and overall accuracy except for Kappa hat coefficient.

Table 6.19 The accuracy assessment of change detection using the NIR feature.

NIR	K-Means			ISODATA		
	Changed pixels	Stable pixels	Row Total	Changed pixels	Stable pixels	Row Total
Changed pixels	175	124	299	199	109	308
Stable pixels	26	177	203	17	177	194
Column Total	201	301	502	216	286	502
Producer's accuracy (%)	87.06	58.80		92.13	61.89	
User's accuracy (%)	58.53	87.19		64.61	91.24	
Overall accuracy (%)		70.12			74.90	
Kappa hat (%)		42.43			51.34	

Table 6.20 The accuracy assessment of change detection using TOP3 spectral features.

TOP3	K-Means			ISODATA		
	Changed pixels	Stable pixels	Row Total	Changed pixels	Stable pixels	Row Total
Changed pixels	169	68	237	176	73	249
Stable pixels	32	233	265	30	223	253
Column Total	201	301	502	206	296	502
Producer's accuracy (%)	84.08	77.41		85.44	75.34	
User's accuracy (%)	71.31	87.92		70.68	88.14	
Overall accuracy (%)		80.08			79.48	
Kappa hat (%)		59.71			58.91	

Table 6.21 The accuracy assessment of change detection using TOP6 spectral features.

TOP6	K-Means			ISODATA		
	Changed pixels	Stable pixels	Row Total	Changed pixels	Stable pixels	Row Total
Changed pixels	111	39	150	111	43	154
Stable pixels	69	283	352	60	288	348
Column Total	180	322	502	171	331	502
Producer's accuracy (%)	61.67	87.89		64.91	87.01	
User's accuracy (%)	74.00	80.40		72.08	82.76	
Overall accuracy (%)		78.49			79.48	
Kappa hat (%)		51.45			53.20	

Table 6.22 The accuracy assessment of change detection using ORI6 spectral features.

ORI6	K-Means			ISODATA		
	Changed pixels	Stable pixels	Row Total	Changed pixels	Stable pixels	Row Total
Changed pixels	107	45	152	116	32	148
Stable pixels	74	276	350	65	289	354
Column Total	181	321	502	181	321	502
Producer's accuracy (%)	59.12	85.98		64.09	90.03	
User's accuracy (%)	70.39	78.86		78.38	81.64	
Overall accuracy (%)		76.29			80.68	
Kappa hat (%)		46.73			56.36	

By comparison of overall accuracy according to spectral feature, LULC change detection using the NIR feature provides the least accurate (K-Means is 70.12%, and ISODATA is 74.90%). In contrast, the accuracy of TOP3 spectral features by K-Means is 80.08%, and ISODATA is 79.48%, TOP6 spectral features by K-Means is 78.49%, and ISODATA is 79.48%, and ORI6 spectral features by K-Means is 76.29%, and ISODATA is 80.68%. These multiple spectral features (TOP3, TOP6, and ORI6) can provide similar results with high accuracy, and multiple spectral features can provide accuracy higher than a single spectral feature (NIR). This finding is an expected result because a single spectral feature is hard to discriminate subtle class such as healthy and disturbed vegetation, since the average value is similar, while this distinction is easily made from multispectral spectral features (Chuvienco, 2016).

By comparison of accuracy according to the algorithm with different feature(s), the ISODATA algorithm with NIR can provide an overall accuracy of 74.90% and Kappa hat coefficient of 51.34% higher than K-Means algorithm which provides overall accuracy of 70.12% and Kappa hat coefficient of 42.43%. Besides, the

accuracy assessment for change detection by the ISODATA algorithm with the NIR show producer's accuracy of 92.13% and user's accuracy of 64.61% for changed pixels, and producer's accuracy of 61.89% and user's accuracy of 91.24% for stable pixels.

Likewise, the ISODATA algorithm with TOP6 spectral features can also provide overall accuracy (79.48%) and Kappa hat coefficient (53.20%) higher than the K-Means algorithm, which provides overall accuracy of 78.49% and Kappa hat coefficient of 51.45%. Also, the accuracy assessment for change detection by ISODATA algorithm with TOP6 spectral features show producer's accuracy of 64.91% and user's accuracy of 72.08% for changed pixels, and producer's accuracy of 87.01% and user's accuracy of 82.76% for stable pixels.

Similarly, the ISODATA algorithm with ORI6 spectral features can also provide an overall accuracy of 80.68% and Kappa hat coefficient of 56.36% higher than the K-Means algorithm, which provides overall accuracy of 76.29% and Kappa hat coefficient of 46.73%. Besides, the accuracy assessment for change detection by ISODATA algorithm with ORI6 spectral features show producer's accuracy of 64.09% and user's accuracy of 78.38% for changed pixels, and producer's accuracy of 90.03% and user's accuracy of 81.64% for stable pixels.

On the contrary, the K-Means algorithm with TOP3 spectral features can provide an overall accuracy of 80.08% and Kappa hat coefficient of 59.71% higher than the ISODATA algorithm, which provides overall accuracy of 79.48% and Kappa hat coefficient of 58.91%. However, both accuracy values from both algorithms are not significantly different.

In addition, the accuracy assessment for change detection by K-Means algorithm with TOP3 spectral features show producer's accuracy of 84.08% and user's

accuracy of 71.31% for changed pixels, and producer's accuracy of 77.41% and user's accuracy of 87.92% for stable pixels. In the meantime, the accuracy assessment for change detection by ISODATA algorithm with TOP3 spectral features show producer's accuracy of 85.44% and user's accuracy of 70.68% for changed pixels, and producer's accuracy of 75.34% and user's accuracy of 88.14% for stable pixels. In fact, the overall accuracy, producer's accuracy, user's accuracy, and Kappa hat coefficient of the ISODATA algorithm are slightly lower than the K-Means algorithm.

The relative lower user's accuracy indicates more commission errors than omission errors in detected changes. A higher threshold ($RMSE > 2$) or using more spectral feature combination (TOP3, TOP6, ORI6 spectral features) may better balance the commission and omission errors. As it can be seen from Tables 6.19 to 6.22, regardless of which algorithm is used, as the number of spectral features for change detection increases, the producer's accuracy of the changed pixels demonstrates a downward trend, while the user's accuracy of the changed pixels shows an upward trend. So, it can be concluded that when a single feature is used, the changed pixel has a high producer's accuracy and low user's accuracy, and when a plurality of features is used, the changed pixel has low producer's accuracy and high user's accuracy.

Commission errors are mainly due to overfitting, which means that LULC changes that have not occurred are considered to have undergone LULC changes. The major causes of this phenomenon are: (1) spectral characteristics are too sensitive for LULC changes, (2) clouds missed three or more times consecutively, and (3) very small threshold.

Using point X in Figure 6.2 as an example, this point (in 2017) comes from the new runway of Nanjing Lukou International Airport. By comparing with

Google Earth images in four different years, this point LULC consists of three stages: before the runway's completion, during construction, and after construction. Figures 6.3 to 6.6 shows the results of different feature combinations for the detection of this LULC change. From time-series HA model plots, it can be seen that these points can be divided into three parts, before 2013, around 2013, and after 2013. The NIR feature divides the harmonic plots into eight parts, and TOP3 spectral features also divide them into eight parts, while TOP6 spectral features and ORI6 spectral bands are only divided into four and six parts, respectively. This result means that the fewer the number of features is, the more likely the model can overfit. Of course, if the shape of the fitted harmonic function curve is similar, when the next step is based on the K-Means and ISODATA algorithms, the different parts will be reclassified into one class. However, this still increases the possibility of misclassifying stable pixels into changing pixels; that is, the fewer the number of features used for change detection is, the higher the commission error exists.

The second reason is the lack of continuous multi-phase data for some pixels. For example, in the summer, the continuous multi-period images of the same location cover by cloud and the data loss caused by the Landsat7 strip are also considered as a cloud pixel. Table 6.23 shows the proportion of all the pixels in the study area in 2013. It can be seen that during the period from April 21, 2013, to September 4, 2013, the cloud coverage ratio of the multi-stage image is very high. So, some pixels can be possibly covered by clouds for a long time, and these pixels might be misjudged as having a LULC change during the change detection.

Table 6.23 Cloud cover of Landsat in the study area.

Date	Cloud	Date	Cloud	Date	Cloud	Date	Cloud
20130107	0.82	20130413	0.15	20130726	0.61	20131014	0.00
20130123	0.67	20130421	0.70	20130811	0.00	20131022	0.50
20130224	0.09	20130616	0.86	20130819	0.87	20131107	0.06
20130312	0.09	20130702	0.86	20130827	0.81	20131115	0.03
20130328	0.31	20130710	0.15	20130904	0.78	20131123	0.71
20130407	0.00	20130718	0.32	20130920	0.49	20131201	0.10

The last reason for false detection of change is a very small threshold. Since the change detection algorithm uses the RMSE value from HA model estimation for defining the threshold for land cover change, when the RMSE value is small, a very slight change can be caused by the atmosphere or other factors. This problem is common in the same type of LULC with different intensity changes, such as change of water content over agricultural land, turbidity and clearness in the water, and changes in the degree of vegetation density.

The relative lower producer's accuracy indicates more omission errors than commission errors in detected changes. A low threshold or using a single spectral feature may better balance the commission and omission errors.

The omission errors are mostly due to some reasons: (1) partially changed pixels; (2) change occurs too early before the model is initialized; (3) very large threshold; and (4) mutual interference of multiple spectral features.

The partially changed pixels are always difficult to detect, as the magnitude of change is mostly dependent on the proportion of change within that pixel. As it can be observed from Figures 6.10 to 6.13, a large number of scattered points are not fit during the model initialization phase; it means if a change occurs at the beginning

of model initialization, the algorithm is unable to detect any change as there are not enough observations to initialize the time-series model.

For threshold, some possible LULC changes may be ignored if the threshold is too large. By comparison among Figures 6.3 to 6.6, it can be observed that when LULC change detection is performed using more spectral features, some possible LULC changes are interfered by multiple spectral features, resulting in these changes not being detected.

In general, in the case of change detection, when using fewer spectral features, the producer of the changed pixel has higher accuracy, and the user accuracy is lower, and it is easy to misjudge the stable pixel as the changed pixel, and the overall accuracy of the change detection is low. When more spectral features are used, the producer of the changed pixel has lower accuracy, and the user accuracy is higher, and it is easy to misjudge the changed pixel as a stable pixel, and the overall accuracy of the change detection is high. The overall accuracy of TOP3, TOP6, and ORI6 spectral features is relatively close. When considering the computational efficiency of change detection, the time used for TOP3 spectral features is about three times of NIR band, and the time used for TOP6 and ORI6 spectral features is about six times of NIR band. Therefore, it can be seen from the two aspects of detection accuracy and time efficiency that TOP3 spectral features are a better combination of spectral features for change detection.

6.3.2 Accuracy assessment for LULC classification

Results of accuracy assessment of LULC classification including producer's accuracy, user's accuracy, overall accuracy and Kappa hat coefficient by different classification algorithms with different spectral features in four different years are displayed in Tables 6.24 to 6.27 and Figures 6.29 to 6.31. The derived overall accuracy is compared and described in three aspects: algorithm, spectral features, and years.

It is easy to see from Figure 6.29 that when TOP6 spectral features were used in 2000 and TOP3 or TOP6 spectral features were used in 2006 for LULC classification, the overall accuracy of the K-Means algorithm is slightly higher than the ISODATA algorithm. Under other conditions, the accuracy of the ISODATA algorithm is higher than the K-Means algorithm. Considering the problems of pan-sharpened images and Google Earth images themselves as reference data for accuracy verification, and the result by human error from visual interpretation. Nevertheless, it can be concluded that the overall accuracy of LULC maps from the ISODATA algorithm is slightly higher than the K-Means algorithm.

Table 6.24 Error matrix and accuracy assessment of LULC Classification in 2000.

K_MEANS						ISODATA				
NIR	U	A	F	W	Row Total	U	A	F	W	Row Total
U	98	110	18	9	235	85	67	12	12	176
A	11	145	23	1	180	17	185	30	0	232
F	4	74	92	1	171	4	74	87	2	167
W	0	3	0	47	50	4	5	0	52	61
Column Total	113	332	133	58	636	110	331	129	66	636
Producer's accuracy (%)	86.73	43.67	69.17	81.03		77.27	55.89	67.44	78.79	
User's accuracy (%)	41.70	80.56	53.80	94.00		48.30	79.74	52.10	85.25	
Overall accuracy (%)			60.06					64.31		
Kappa hat (%)			44.78					48.94		
TOP3	U	A	F	W	Row Total	U	A	F	W	Row Total
U	97	56	2	3	158	107	68	3	1	179
A	15	194	26	5	240	1	202	40	0	243
F	4	71	112	1	188	2	34	81	0	117
W	0	2	0	48	50	4	10	0	83	97
Column Total	116	323	140	57	636	114	314	124	84	636
Producer's accuracy (%)	83.62	60.06	80.00	84.21		93.86	64.33	65.32	98.81	
User's accuracy (%)	61.39	80.83	59.57	96.00		59.78	83.13	69.23	85.57	
Overall accuracy (%)			70.91					74.37		
Kappa hat (%)			57.90					63.64		
TOP6	U	A	F	W	Row Total	U	A	F	W	Row Total
U	76	39	1	0	116	90	86	8	12	196
A	1	324	57	1	383	10	250	40	3	303
F	0	24	60	0	84	0	13	74	0	87
W	1	3	0	49	53	0	1	0	49	50
Column Total	78	390	118	50	636	100	350	122	64	636
Producer's accuracy (%)	97.44	83.08	50.85	98.00		90.00	71.43	60.66	76.56	
User's accuracy (%)	65.52	84.60	71.43	92.45		45.92	82.51	85.06	98.00	
Overall accuracy (%)			80.03					72.80		
Kappa hat (%)			65.41					58.49		
ORI6	U	A	F	W	Row Total	U	A	F	W	Row Total
U	88	89	9	7	193	88	66	11	1	166
A	0	215	19	3	237	8	278	43	4	333
F	2	46	108	0	156	0	10	63	0	73
W	0	1	0	49	50	1	2	0	61	64
Column Total	90	351	136	59	636	97	356	117	66	636
Producer's accuracy (%)	97.78	61.25	79.41	83.05		90.72	78.09	53.85	92.42	
User's accuracy (%)	45.60	90.72	69.23	98.00		53.01	83.48	86.30	95.31	
Overall accuracy (%)			72.33					77.04		
Kappa hat (%)			59.99					63.88		

Table 6.25 Error matrix and accuracy assessment of LULC Classification in 2006.

	K_MEANS					ISODATA				
NIR	U	A	F	W	Row Total	U	A	F	W	Row Total
U	110	93	14	20	237	104	53	11	9	177
A	15	165	12	1	193	20	203	27	2	252
F	3	48	103	2	156	0	53	97	0	150
W	2	1	0	47	50	4	2	0	51	57
Column Total	130	307	129	70	636	128	311	135	62	636
Producer's accuracy (%)	84.62	53.75	79.84	67.14		81.25	65.27	71.85	82.26	
User's accuracy (%)	46.41	85.49	66.03	94.00		58.76	80.56	64.67	89.47	
Overall accuracy (%)			66.82					71.54		
Kappa hat (%)			53.85					58.84		
TOP3	U	A	F	W	Row Total	U	A	F	W	Row Total
U	117	35	3	5	160	141	30	5	2	178
A	7	172	25	7	211	4	215	37	0	256
F	5	70	139	1	215	2	30	89	0	121
W	1	3	0	46	50	5	3	1	72	81
Column Total	130	280	167	59	636	152	278	132	74	636
Producer's accuracy (%)	90.00	61.43	83.23	77.97		92.76	77.34	67.42	97.30	
User's accuracy (%)	73.13	81.52	64.65	92.00		79.21	83.98	73.55	88.89	
Overall accuracy (%)			74.53					81.29		
Kappa hat (%)			63.94					73.38		
TOP6	U	A	F	W	Row Total	U	A	F	W	Row Total
U	102	10	4	1	117	117	42	13	4	176
A	19	317	45	4	385	8	268	37	4	317
F	1	15	66	0	82	1	13	79	0	93
W	0	3	0	49	52	0	0	0	50	50
Column Total	122	345	115	54	636	126	323	129	58	636
Producer's accuracy (%)	83.61	91.88	57.39	90.74		92.86	82.97	61.24	86.21	
User's accuracy (%)	87.18	82.34	80.49	94.23		66.48	84.54	84.95	100	
Overall accuracy (%)			83.96					80.82		
Kappa hat (%)			73.54					70.72		
ORI6	U	A	F	W	Row Total	U	A	F	W	Row Total
U	133	31	1	5	170	124	19	0	1	144
A	8	199	24	8	239	8	302	40	2	352
F	1	61	115	0	177	0	13	72	1	86
W	0	0	0	50	50	1	1	0	52	54
Column Total	142	291	140	63	636	133	335	112	56	636
Producer's accuracy (%)	93.66	68.38	82.14	79.37		93.23	90.15	64.29	92.86	
User's accuracy (%)	78.24	83.26	64.97	100		86.11	85.80	83.72	96.30	
Overall accuracy (%)			78.14					86.48		
Kappa hat (%)			68.75					78.53		

Table 6.26 Error matrix and accuracy assessment of LULC Classification in 2011.

K_MEANS						ISODATA				
NIR	U	A	F	W	Row Total	U	A	F	W	Row Total
U	142	94	13	23	272	117	64	12	13	206
A	29	113	31	1	174	25	169	30	4	228
F	1	45	94	0	140	2	25	107	1	135
W	1	2	0	47	50	2	5	1	59	67
Column Total	173	254	138	71	636	146	263	150	77	636
Producer's accuracy (%)	82.08	44.49	68.12	66.20		80.14	64.26	71.33	76.62	
User's accuracy (%)	52.21	64.94	67.14	94.00		56.80	74.12	79.26	88.06	
Overall accuracy (%)			62.26					71.07		
Kappa hat (%)			47.43					59.51		
TOP3	U	A	F	W	Row Total	U	A	F	W	Row Total
U	155	23	3	6	187	165	36	4	0	205
A	5	169	21	10	205	4	200	30	0	234
F	3	61	129	1	194	1	19	90	0	110
W	0	1	0	49	50	2	7	0	78	87
Column Total	163	254	153	66	636	172	262	124	78	636
Producer's accuracy (%)	95.09	66.54	84.31	74.24		95.93	76.34	72.58	100	
User's accuracy (%)	82.89	82.44	66.49	98.00		80.49	85.47	81.82	89.66	
Overall accuracy (%)			78.93					83.81		
Kappa hat (%)			70.51					77.21		
TOP6	U	A	F	W	Row Total	U	A	F	W	Row Total
U	114	16	1	0	131	140	35	9	6	190
A	38	291	44	7	380	11	255	42	2	310
F	1	13	60	0	74	2	15	69	0	86
W	0	4	1	46	51	0	0	0	50	50
Column Total	153	324	106	53	636	153	305	120	58	636
Producer's accuracy (%)	74.51	89.81	56.60	86.79		91.50	83.61	57.50	86.21	
User's accuracy (%)	87.02	76.58	81.08	90.20		73.68	82.26	80.23	100	
Overall accuracy (%)			80.35					80.82		
Kappa hat (%)			68.30					71.01		
ORI6	U	A	F	W	Row Total	U	A	F	W	Row Total
U	141	38	5	2	186	140	23	1	1	165
A	9	185	20	18	232	14	273	44	7	338
F	3	42	123	0	168	0	14	65	0	79
W	0	0	0	50	50	0	2	1	51	54
Column Total	153	265	148	70	636	154	312	111	59	636
Producer's accuracy (%)	92.16	69.81	83.11	71.43		90.91	87.50	58.56	86.44	
User's accuracy (%)	75.81	79.74	73.21	100		84.85	80.77	82.28	94.44	
Overall accuracy (%)			78.46					83.18		
Kappa hat (%)			69.55					73.99		

Table 6.27 Error matrix and accuracy assessment of LULC Classification in 2017.

K_MEANS						ISODATA				
NIR	U	A	F	W	Row Total	U	A	F	W	Row Total
U	148	59	4	13	224	126	33	5	7	171
A	20	128	14	1	163	39	149	21	2	211
F	9	65	123	2	199	6	56	136	2	200
W	1	5	0	44	50	3	5	0	46	54
Column Total	178	257	141	60	636	174	243	162	57	636
Producer's accuracy (%)	83.15	49.81	87.23	73.33		72.41	61.32	83.95	80.70	
User's accuracy (%)	66.07	78.53	61.81	88.00		73.68	70.62	68.00	85.19	
Overall accuracy (%)			69.65					71.86		
Kappa hat (%)			57.92					60.47		
TOP3	U	A	F	W	Row Total	U	A	F	W	Row Total
U	157	19	6	2	184	174	23	3	3	203
A	10	162	13	11	196	12	168	24	4	208
F	3	69	131	3	206	5	41	104	0	150
W	0	3	0	47	50	6	4	2	63	75
Column Total	170	253	150	63	636	197	236	133	70	636
Producer's accuracy (%)	92.35	64.03	87.33	74.60		88.32	71.19	78.20	90.00	
User's accuracy (%)	85.33	82.65	63.59	94.00		85.71	80.77	69.33	84.00	
Overall accuracy (%)			78.14					80.03		
Kappa hat (%)			69.47					72.17		
TOP6	U	A	F	W	Row Total	U	A	F	W	Row Total
U	112	18	2	0	132	143	28	16	3	190
A	41	267	41	14	363	20	242	32	6	300
F	5	12	74	0	91	2	15	77	2	96
W	0	6	0	44	50	0	2	0	48	50
Column Total	158	303	117	58	636	165	287	125	59	636
Producer's accuracy (%)	70.89	88.12	63.25	75.86		86.67	84.32	61.60	81.36	
User's accuracy (%)	84.85	73.55	81.32	88.00		75.26	80.67	80.21	96.00	
Overall accuracy (%)			78.14					80.19		
Kappa hat (%)			66.01					70.55		
ORI6	U	A	F	W	Row Total	U	A	F	W	Row Total
U	157	24	4	1	186	141	24	2	0	167
A	24	170	19	12	225	30	246	44	14	334
F	5	44	125	1	175	2	9	74	0	85
W	2	6	0	42	50	2	2	0	46	50
Column Total	188	244	148	56	636	175	281	120	60	636
Producer's accuracy (%)	83.51	69.67	84.46	75.00		80.57	87.54	61.67	76.67	
User's accuracy (%)	84.41	75.56	71.43	84.00		84.43	73.65	87.06	92.00	
Overall accuracy (%)			77.67					79.72		
Kappa hat (%)			68.41					69.41		

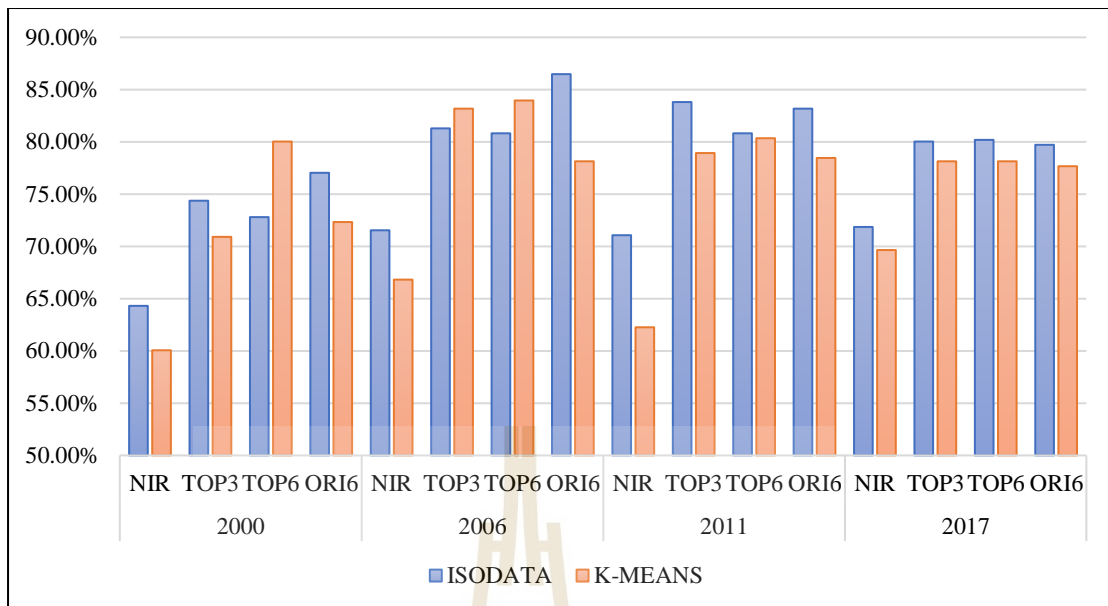


Figure 6.29 Comparison of the overall accuracy of the two algorithms.

Besides, it can also be observed from Figure 6.30 that the overall accuracy is lower at approximately 70% when using a single spectral feature under different years and different algorithms conditions, while the overall accuracy is higher when using multiple spectral features, approximately 80%. This finding is consistent with accuracy assessment for change detection as mentioned above that the multiple spectral features can provide higher accuracy than a single spectral feature.

Furthermore, when using different numbers of spectral features for LULC classification, the minimum memory required for TOP3 spectral features is about three times of NIR band, and the minimum memory required for TOP6 and ORI6 spectral features is about six times of NIR band. In the same computer configuration, TOP3 spectral features take about three times of NIR, and TOP6 and ORI6 spectral features take about six times of NIR band.

By considering accuracy, minimum memory requirement, and processing time for LULC classification, TOP3 spectral features can be selected as a suitable spectral feature for LULC change detection and classification.

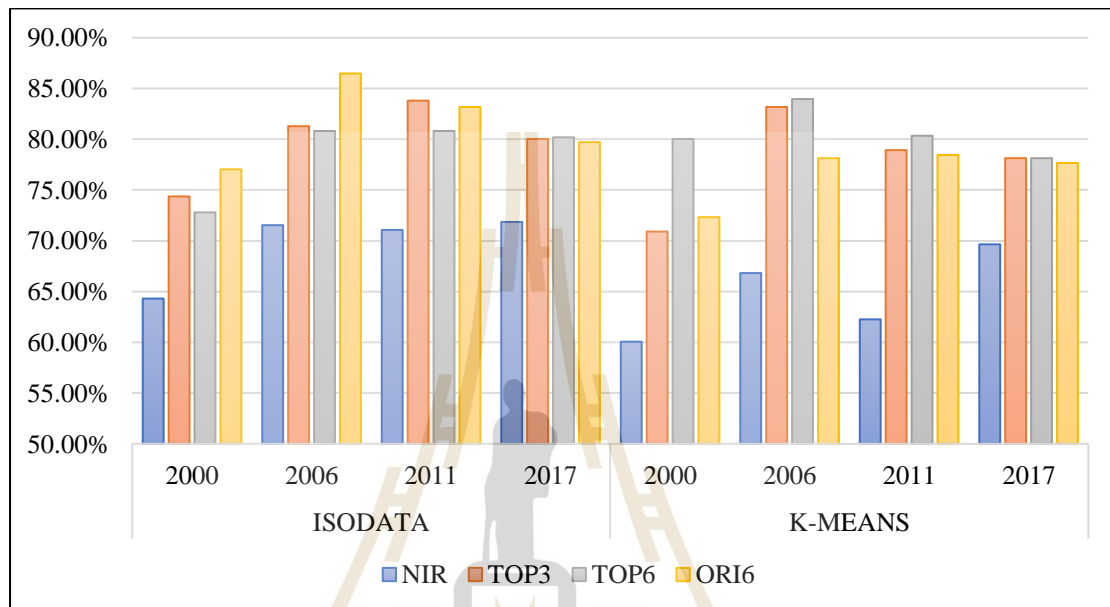


Figure 6.30 Comparison of the overall accuracy of four spectral features.

Figure 6.31 displays a comparison of the overall accuracy in four different years. The overall accuracy of the LULC classification map in 2000 by K-Means and ISODATA algorithms with different spectral features is rather low, while the overall accuracy of 2006, 2011, and 2017 is high. The lower accuracy of the LULC classification map in 2000 may be related to more errors accumulated during the initialization phase of the model during the change detection (see Figures 6.3 to 6.6), where a large number of points are not fitted during the model initialization phase.

Since the model initialization phase requires 12 clean pixels, and the Landsat image acquisition period is eight days (combined with Landsat 7 and Landsat

8), considering the possible effects of clouds and stripes (Landsat 7), it is recommended that when studying the change of LULC in a certain period based on time-series data, the Landsat data time range used needs to be one year earlier than the research period.

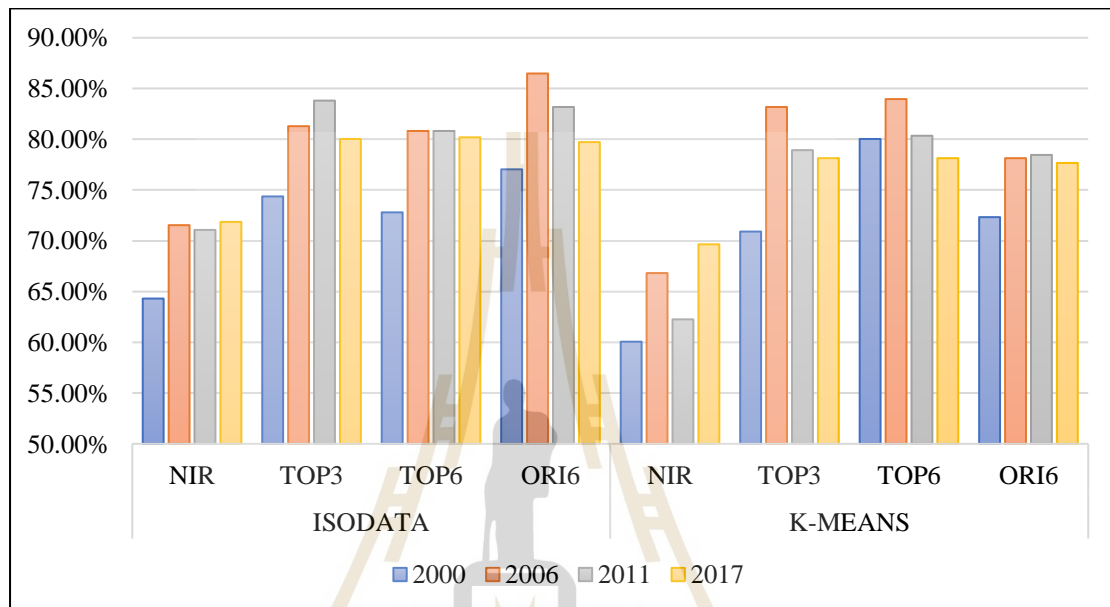


Figure 6.31 Comparison of the overall accuracy of four years.

For analyzing the producer's and user's accuracies of LULC classification maps obtained in different spectral features combinations, different algorithms, and different years, it is difficult and unnecessary to calculate the producer's and user's accuracies of the LULC classification map for each specific date. Because the interpretation of both accuracies results from the time-series of LULC classification maps (every eight days or 16 days) of both methods are time-consuming. What is important is the average producer's accuracy and user's accuracy of the LULC classification maps obtained by different spectral features and different algorithms. Therefore, the producer's accuracy and user's accuracy of the entire study period (2000-

2017) are represented by an average value of producer's accuracy and user's accuracy of the LULC classification map from one selected date of four years. According to the previous analysis, due to the initialization of the model, the accuracy of the LULC classification map in 2000 is reduced. The average producer's accuracy and user's accuracy of each LULC type are firstly separately calculated from the LULC classification maps of 2006, 2011 and 2017 and the derived average producer's accuracy and user's accuracy from four different LULC types are further applied to calculate average producer's accuracy and user's accuracy again.

Tables 6.28 to 6.29 demonstrate an average producer's accuracy and user's accuracy of each LULC type and four different LULC types based on the different spectral feature(s) of K-Means and ISODATA algorithms, respectively.

Table 6.28 Average producer's accuracy and user's accuracy using K-Means algorithms.

Spectral feature(s)	Accuracy	The average value of each LULC type				Average of 4 LULC type
		U	A	F	W	
NIR	Producer's (%)	83.28	49.35	78.40	68.89	69.98
	User's (%)	54.90	76.32	64.99	92.00	72.05
TOP3	Producer's (%)	92.48	64.00	84.96	75.60	79.26
	User's (%)	80.45	82.20	64.91	94.67	80.56
TOP6	Producer's (%)	76.34	89.94	59.08	84.46	77.45
	User's (%)	86.35	77.49	80.96	90.81	83.90
ORI6	Producer's (%)	89.78	69.29	83.24	75.27	79.39
	User's (%)	79.49	79.52	69.87	94.67	80.89

Table 6.29 Average producer's accuracy and average's user accuracy using ISODATA algorithms.

Spectral feature(s)	Accuracy	The average value of each LULC type				Average of 4 LULC type
		U	A	F	W	
NIR	Producer's (%)	77.93	63.62	75.71	79.86	74.28
	User's (%)	63.08	75.10	70.64	87.57	74.10
TOP3	Producer's (%)	92.34	74.96	72.73	95.77	83.95
	User's (%)	81.80	83.41	74.90	87.52	81.91
TOP6	Producer's (%)	90.34	83.63	60.11	84.59	79.67
	User's (%)	71.81	82.49	81.80	98.67	83.69
ORI6	Producer's (%)	88.24	88.40	61.51	85.32	80.87
	User's (%)	85.13	80.07	84.35	94.25	85.95

Figures 6.32 to 6.33 show the producer's accuracy and user's accuracy based on Tables 6.28 to 6.29. It can be seen from the figures, in terms of the algorithm, the average producer's accuracy and user's accuracy of the four LULC types obtained by the ISODATA algorithm are superior to the K-Means algorithm. In terms of spectral feature combinations, the results obtained using TOP3, TOP6, and ORI6 spectral features are superior to the NIR band.

Besides, the average producer's accuracy and user's accuracy obtained from each LULC type using K-Means and ISODATA algorithm displays that TOP3, TOP6, and ORI6 spectral features can be applied to classify each LULC type superior to NIR spectral feature.

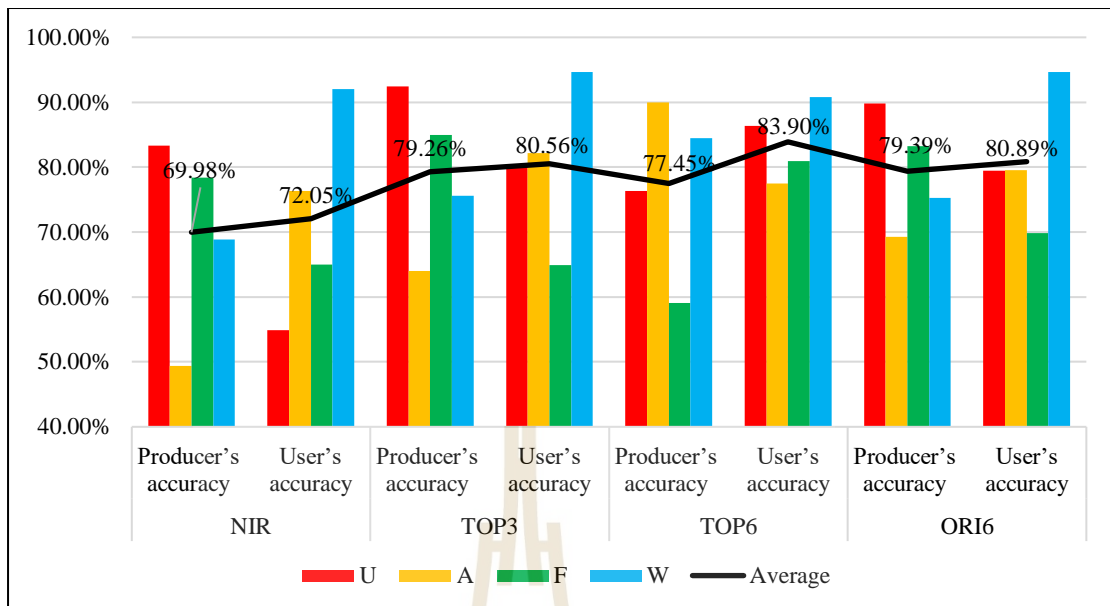


Figure 6.32 Producer's accuracy and user's accuracy based on the K-Means algorithm.

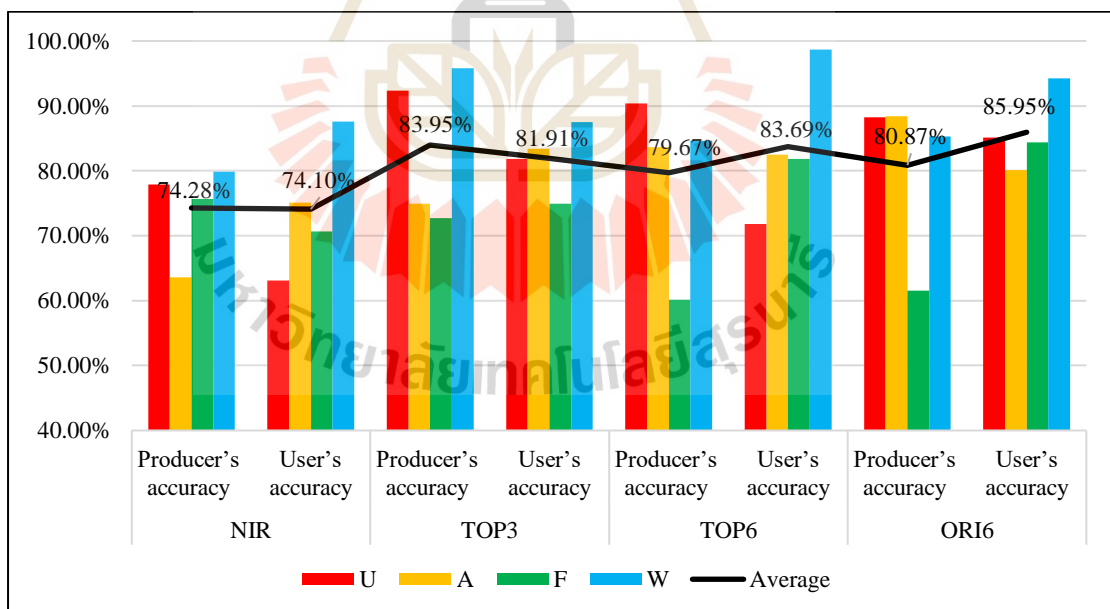


Figure 6.33 Producer's accuracy and user's accuracy based on the ISODATA.

Furthermore, the pairwise Z-test is conducted to examine the significant difference of accuracy from two algorithms using four spectral feature(s) combination based on Kappa analysis using Eq 6.2 and Eq 6.3 as results shown in Tables 6.30 to 6.32.

Table 6.30 shows the result of the pairwise Z-test of accuracy for LULC classification between the K-Means and the ISODATA algorithm from a variety spectral feature(s). It can be seen that most of the Z-values are less than 1.96 (except for a few cases). These results indicate that most of the accuracies for LULC classification using both algorithms from a variety spectral feature(s) are not significantly different. However, the accuracy of the LULC classification in 2011 based on NIR using the K-Means algorithm is a significant difference from the ISODATA algorithm (See Table 6.26). Likewise, the accuracy of LULC classification in 2006 and 20011 based on TOP3 using the K-Means algorithm is a significant difference from the ISODATA algorithm (See Tables 6.25 and 6.26), and the accuracy of LULC classification in 2011 based on ORI6 using K-Means algorithm is a significant difference from ISODATA algorithm (See Tables 6.25).

From Tables 6.31 and 6.32, all Z-values of the pairwise test of accuracy between NIR and TOP3, NIR and TOP6, and NIR and ORI6 using a different algorithm (K-Means or ISODATA) are higher than 1.96. These results indicate that the accuracies of LULC classification using K-Means or ISODATA algorithm based on the NIR spectral feature is a significant difference from the accuracies of LULC classification with TOP3, TOP6, and ORI6 spectral features.

In contrast, whether using the K-Means algorithm or ISODATA algorithm, the values of pairwise Z-test of accuracy among TOP3, TOP6, and ORI6

spectral features are almost less than 1.96. These results indicate that the accuracies of LULC classification with TOP3, TOP6, and ORI6 spectral features are insignificantly different, except for accuracy between TOP3 and TOP6 in 2011 and between TOP6 and ORI6 in 2006.

Table 6.30 Pairwise Z-test between the K-Means and ISODATA algorithm.

YEAR	Z-values			
	NIR	TOP3	TOP6	ORI6
2000	1.0419	1.5585	1.7661	1.0434
2006	1.3298	2.8203 *	0.8250	3.0258 *
2011	3.1927 *	2.1722 *	0.7728	1.3479
2017	0.7066	0.8418	1.2944	0.2938

Note * It is a significant difference of accuracy based on the Kappa hat coefficient.

Table 6.31 Pairwise Z-test between different spectral features using the K-Means algorithm.

YEAR	Z-values					
	NIR_TOP3	NIR_TOP6	NIR_ORI6	TOP3_TOP6	TOP3_ORI6	TOP6_ORI6
2000	3.3456 *	5.1618 *	3.9365 *	1.9318 *	0.5572	1.4153
2006	2.7681 *	5.4728 *	4.1683 *	2.7572 *	1.3903	1.4073
2011	6.4156 *	5.5043 *	6.0944 *	0.6404	0.2913	0.3605
2017	3.3550 *	2.2247 *	3.0156 *	0.9978	0.3202	0.6856

Note * It is a significant difference of accuracy based on the Kappa hat coefficient.

Table 6.32 Pairwise Z-test between different spectral features using the ISODATA algorithm.

YEAR	Z-values					
	NIR_TOP3	NIR_TOP6	NIR_ORI6	TOP3_TOP6	TOP3_ORI6	TOP6_ORI6
2000	3.9029 *	2.4318 *	3.8618 *	1.3864	0.0652	1.4114
2006	4.2033 *	3.3254 *	5.7451 *	0.8108	1.6536	2.4082 *
2011	5.3569 *	3.2845 *	4.1824 *	1.9657 *	1.0345	0.8983
2017	3.4552 *	2.8973 *	2.5337 *	0.4992	0.8359	0.3351

Note * It is a significant difference of accuracy based on the Kappa hat coefficient.

SUMMARY

By comparison of overall accuracy according to spectral features, LULC change detection using the NIR feature provided the least accurate. In contrast, the multiple spectral features (TOP3, TOP6, and ORI6) could provide similar results with high accuracy, and the multiple spectral features provided accuracy higher than a single spectral feature (NIR). It is an expected result because a single spectral feature is hard to discriminate subtle class such as healthy and disturbed vegetation, since the average value is similar, while this distinction is easily made from multispectral spectral features.

According to the performance of the clustering algorithm for LULC classification, it can be observed that the accuracy of thematic maps derived by the ISODATA algorithm was slightly higher than the K-Means algorithm. Overall accuracy and Kappa hat coefficient of the classified LULC maps in four years with ORI6 spectral features varied from 77.04% to 86.48% and 63.88% to 78.53%, respectively.

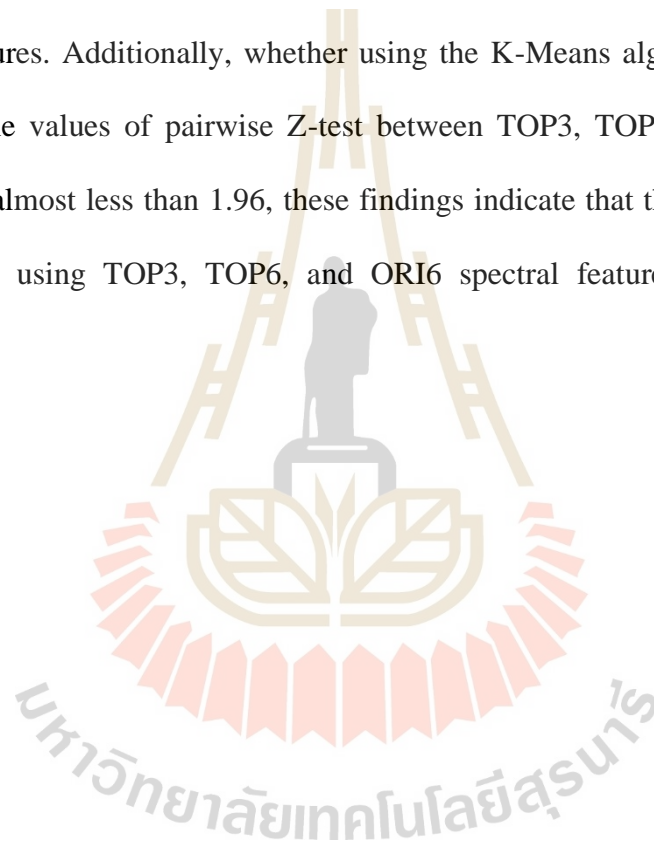
Besides, by considering spectral feature application to LULC classification, when a single spectral feature was applied to classify LULC maps in different years

with two algorithms, the overall accuracy of LULC maps was approximately 70%. On the contrary, when multiple spectral features were applied to classify LULC maps in different years with two algorithms, the overall accuracy of LULC maps was approximately 80%. As a result, it was confirmed that the multiple spectral features could provide higher accuracy than a single spectral feature. Additionally, by considering minimum memory requirement for LULC classification, the required minimum memory for LULC classification with TOP3 spectral features was about three times of NIR band, and the required minimum memory for LULC classification with TOP6 and ORI6 spectral features was about six times of NIR band. Likewise, with the same standard of computer configuration, the processing time for LULC classification with TOP3 spectral features was about three times of NIR while the processing time for LULC classification with TOP6 and ORI6 spectral features was about six times of NIR band. Subsequently, by considering accuracy, minimum memory requirement, and processing time for LULC classification, TOP3 spectral features can be selected as a suitable spectral feature for LULC change detection and classification.

Furthermore, the overall accuracy of the LULC classification map in 2000 by K-Means and ISODATA algorithms with different spectral features is rather low, while the overall accuracy of 2006, 2011, and 2017 is high. The lower accuracy of the LULC classification map in 2000 may be related to more errors accumulated during the initialization phase of the model during the change detection, where a large number of points are not fitted during the model initialization phase. Since the model initialization phase requires 12 clean pixels, and the Landsat image acquisition period is eight days (combined with Landsat 7 and Landsat 8), considering the possible effects of clouds and stripes (Landsat 7), it is recommended that when studying the change of LULC in

a certain period based on time-series data, the Landsat data time range used needs to be one year earlier than the research period.

In addition, the pairwise Z-test indicates that the accuracy of the LULC classification from both algorithms is not significantly different. The pairwise Z-test indicating that the accuracy of LULC classification using the NIR spectral feature is different from the accuracy of LULC classification using TOP3, TOP6, and ORI6 spectral features. Additionally, whether using the K-Means algorithm or ISODATA algorithm, the values of pairwise Z-test between TOP3, TOP6, and ORI6 spectral features are almost less than 1.96, these findings indicate that the accuracy of LULC classification using TOP3, TOP6, and ORI6 spectral features are insignificantly different.



CHAPTER VII

TIME-SERIES LAND USE AND LAND COVER

CLASSIFICATION USING A SUPERVISED METHOD

In order to increase the speed of change monitoring and classification, a nearly-real-time LULC classification method using HA with a minimum spectral distance algorithm was developed in this study. The workflow of time-series LULC classification using HA with a minimum spectral distance algorithm is schematically displayed in Figure 7.1. The detail of three major tasks and their result of this component included (1) harmonic function curve conversion and standard harmonic curve construction, (2) spectral distance measurement and probability calculation, (3) time-series LULC classification and (4) accuracy assessment were separately described and discussed in the following section.

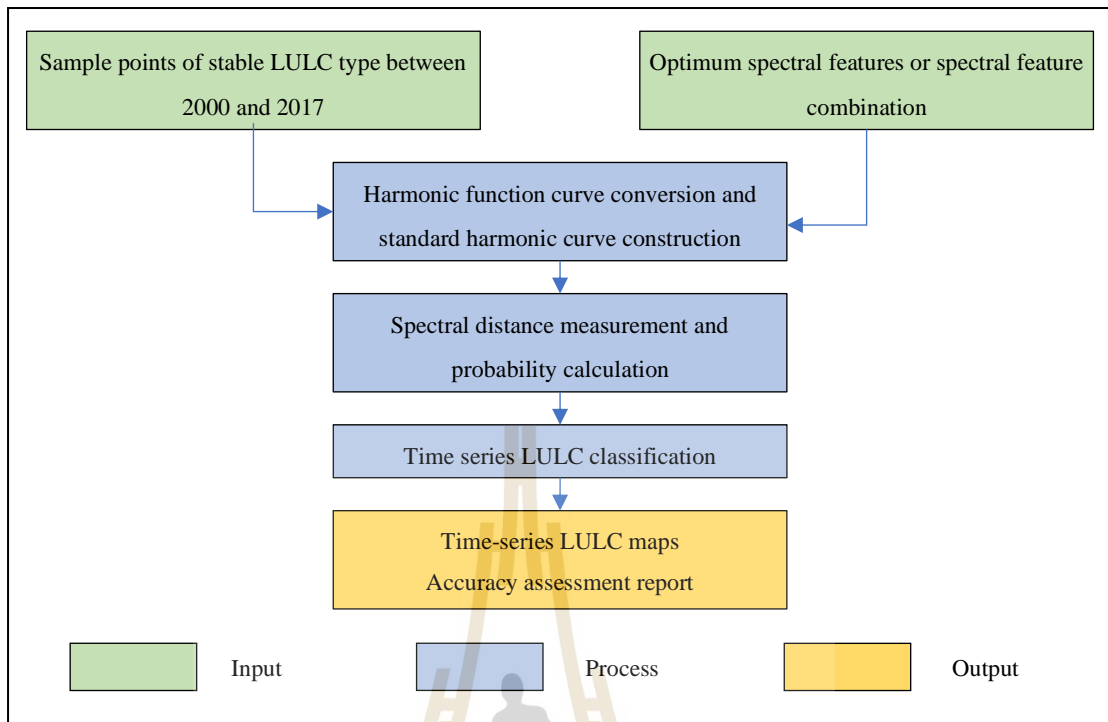


Figure 7.1 Workflow of Component 4: Time-series LULC classification using supervised method.

7.1 Harmonic function curve transformation and standard harmonic curve construction

The time-series curve of each data point is expressed as the sum of a series of cosine or sine waves; each wave determines by a different amplitude and phase (Leica, 2005). These continuous amplitude and phase are summed to produce a compound curve (Jakubauskas et al., 2002). Since the cosine of the sine wave is a typical periodic change curve, therefore, HA can be used to simulate the periodic change of spectral reflectivity.

To obtain the harmonic curve characteristic of different LULC types, the selected stable pixels for each LULC type between 2000 and 2017 were first

transformed into a spectral harmonic curve using Eq. 7.1 modified from Zhu and Woodcock (2014b).

$$y_i = a_i + b_i t + A_i \cos\left(\frac{2\pi}{T} t - \varphi_i\right) \quad (7.1)$$

where, t is the Julian date, i is the i^{th} Landsat band, T is the number of days per year ($T=365$), a_i is the coefficient of intercept value, b_i is the coefficient of slope value, A_i is the coefficient of amplitude value, φ_i is the coefficient of phase value, y is the reconstructed reflectance value at Julian date t . In this study, four spectral features combinations: NIR, TOP3, TOP6, and ORI6, which contain nine spectral features, namely blue, green, red, NIR, SWIR1, SWIR2, EVI, TCB, and PC1 were further applied to transform harmonic function curve for time-series LULC classification.

After that, the median values of six spectral feature coefficients from all sample points were extracted, and the standard harmonic function curve of each LULC type from six bands was constructed using Eq. 7.1 to simplify the character of each LULC type for spectral distance measurement with any unclassified pixel.

7.1.1 Spectral harmonic function curve of LULC types

The spectral harmonic function curves of four LULC types were transformed from sample points of stable pixels between 2000 and 2017 by using different spectral features display in Figure 7.2. Since the harmonic function curves of each LULC type come from nearly 10,000 sample points, the harmonic function curves of each LULC type from each spectral band appear like a ribbon form, and the harmonic function curves of different LULC types overlap among LULC types. For example, under the TCB feature, the harmonic curve of the water body overlaps with forest land, while the harmonic curve of the water body overlaps with another LULC type under

the GREEN feature. So, it is not easy to justify the similarity among LULC types of the unclassified pixel using spectral polyline and the reference harmonic function curve. Therefore, it is necessary to find a standard harmonic function curve that can represent different LULC types in various bands.



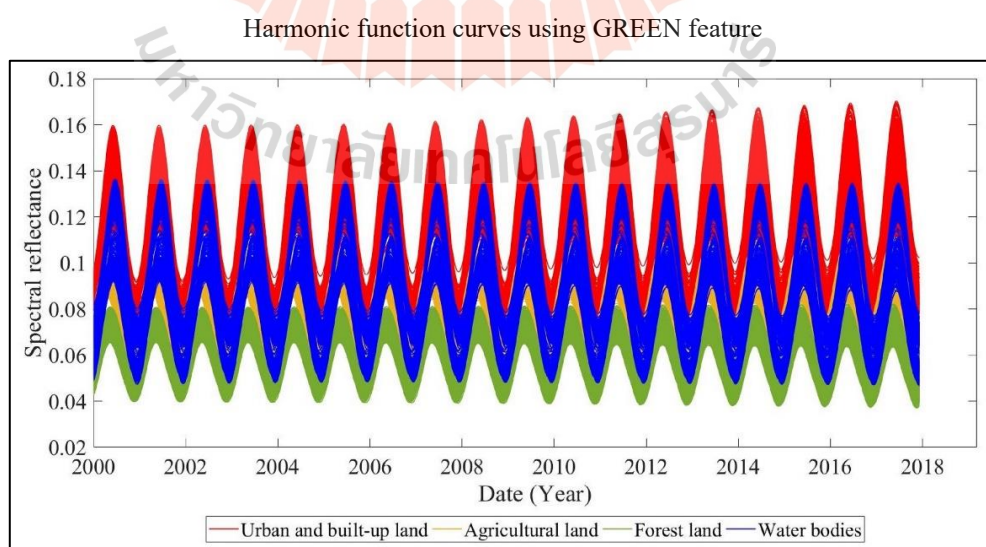
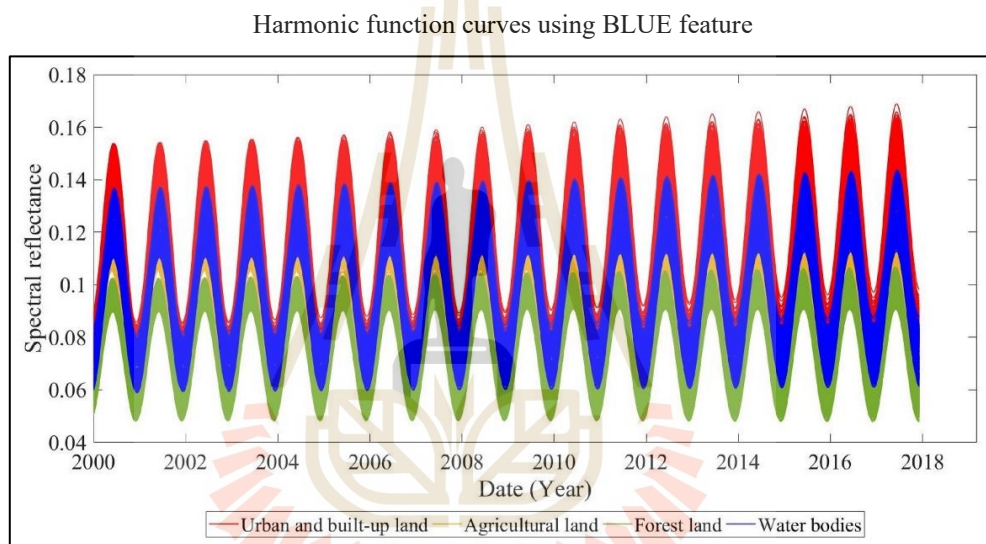
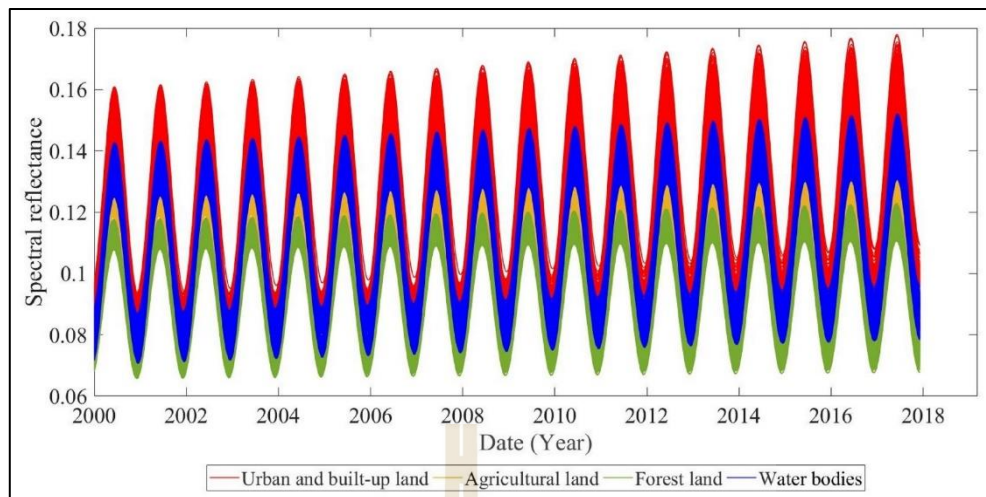


Figure 7.2 Harmonic function curves of four LULC types.

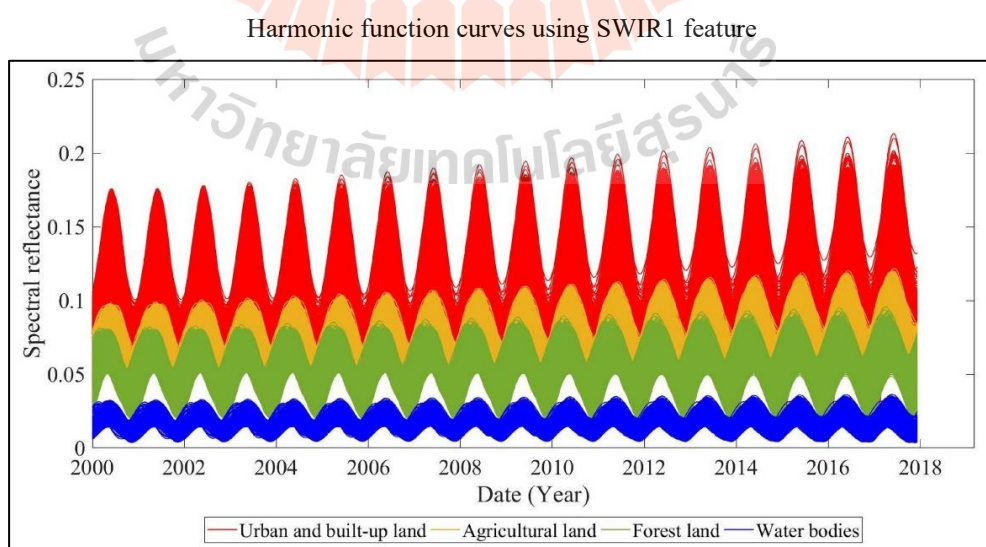
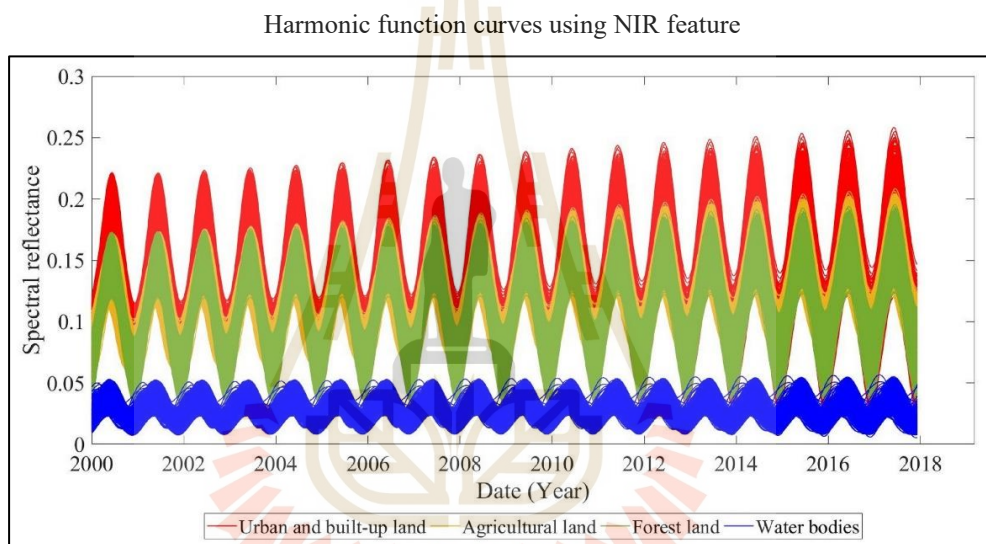
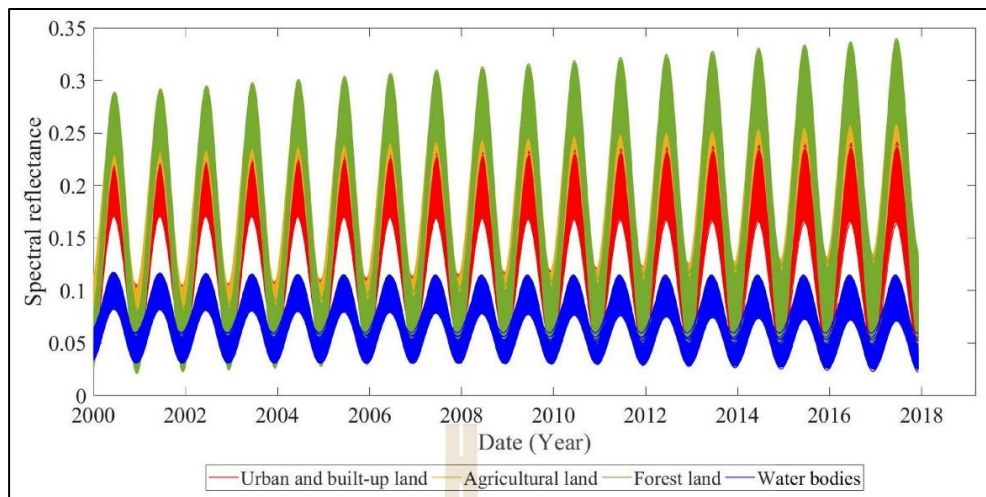


Figure 7.2 (Continued).

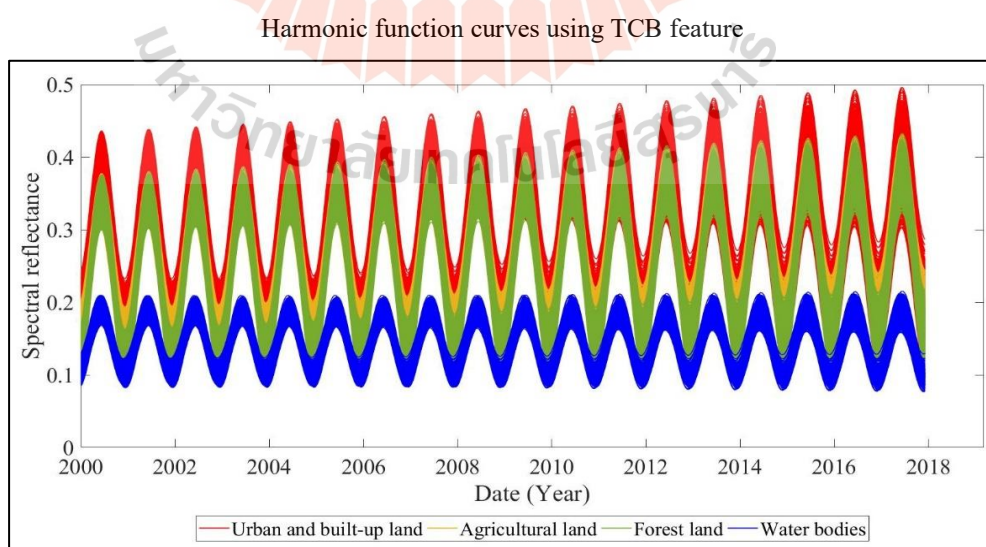
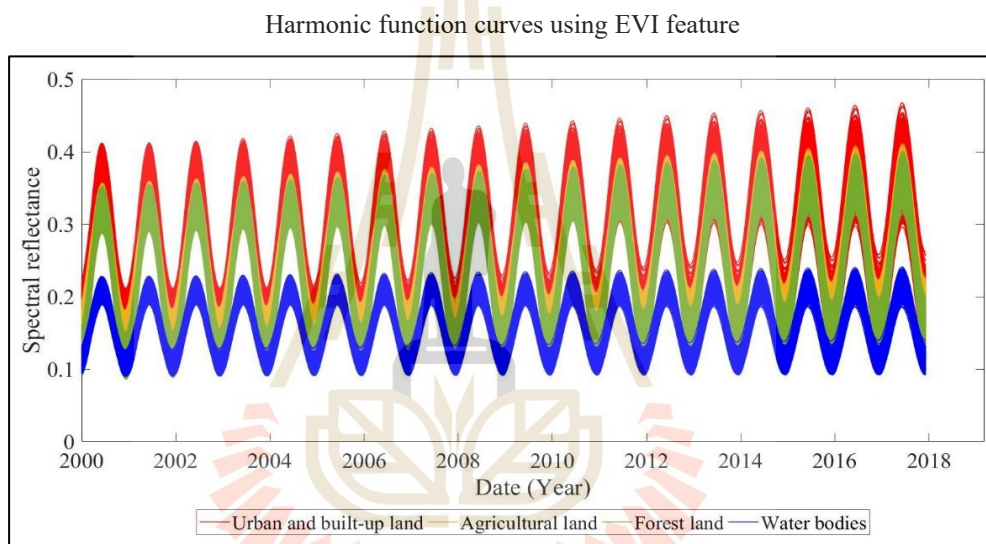
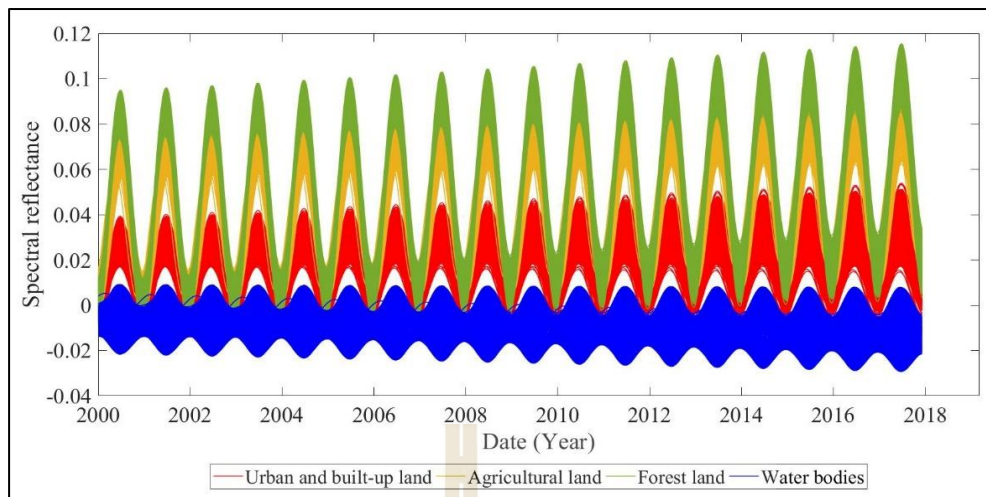


Figure 7.2 (Continued).

7.1.2 Standard harmonic function curve of each LULC type

Table 7.1 shows the median value of nine spectral feature coefficients for four LULC types, while Figure 7.3 presents the harmonic function curve of nine spectral features for four LULC types based on Table 7.1. As a result, it revealed that there is no overlapped strip among LULC types, and it is easy to differentiate the harmonic function curve of LULC types. Importantly, these standard harmonic function curves will be applied to measure the spectral distance between each LULC type and an unclassified pixel and to calculate a probability of an unclassified pixel of being specific LULC type (U, A, F, and W) at a particular point of time.

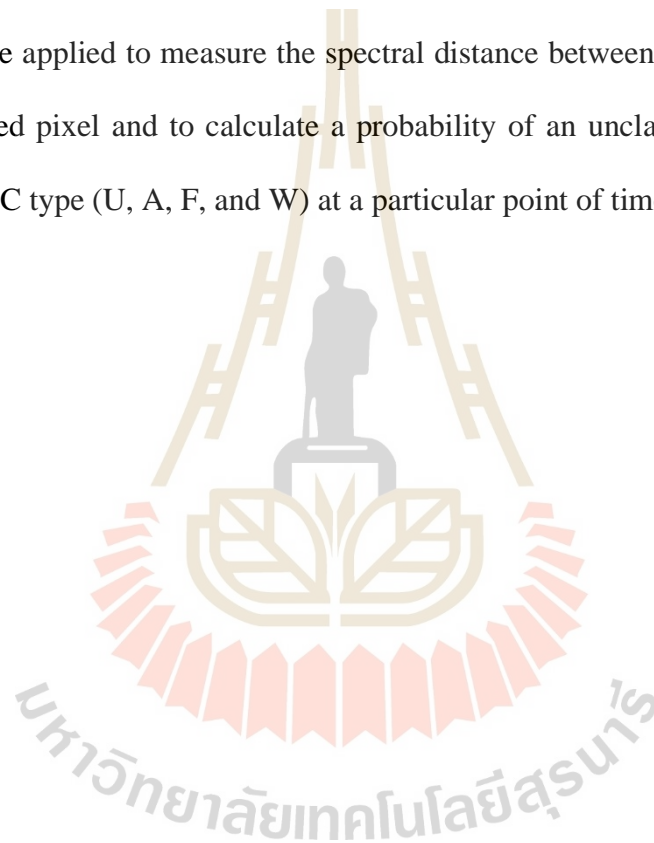
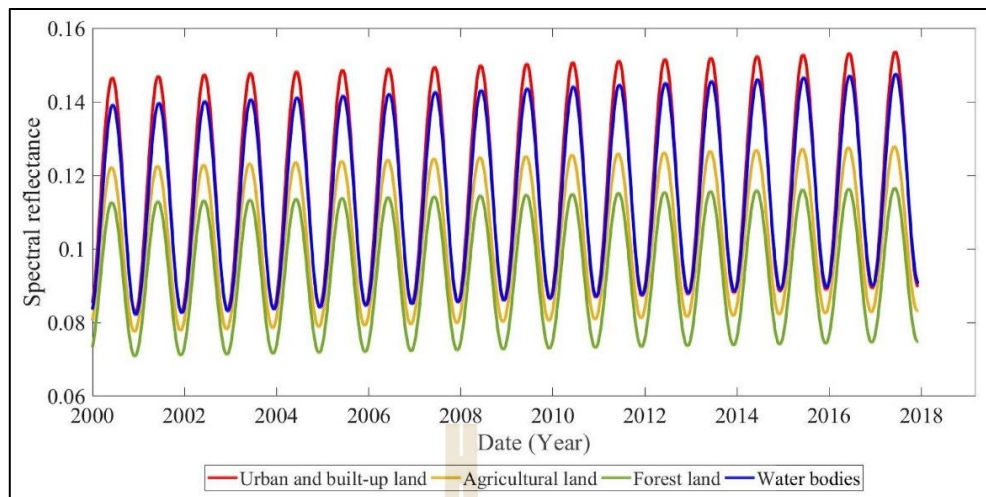
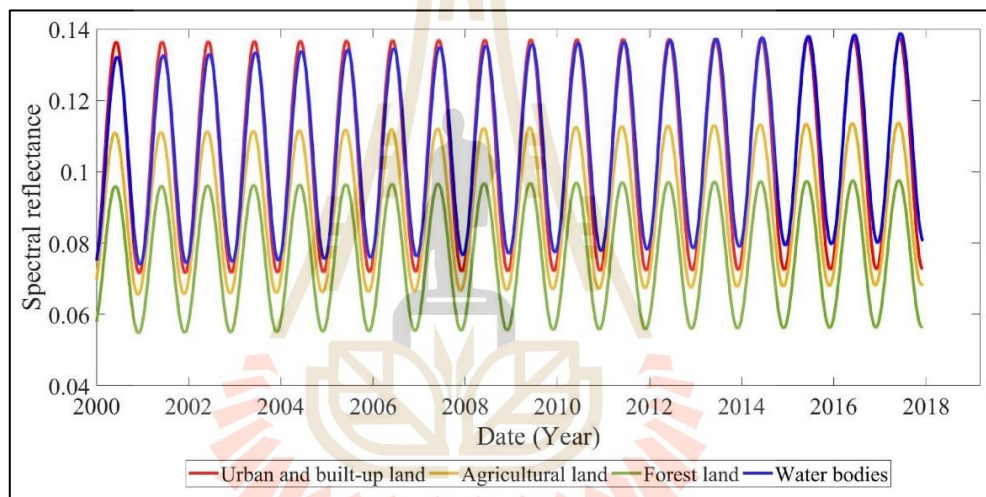


Table 7.1 Median of coefficients for four LULC types.

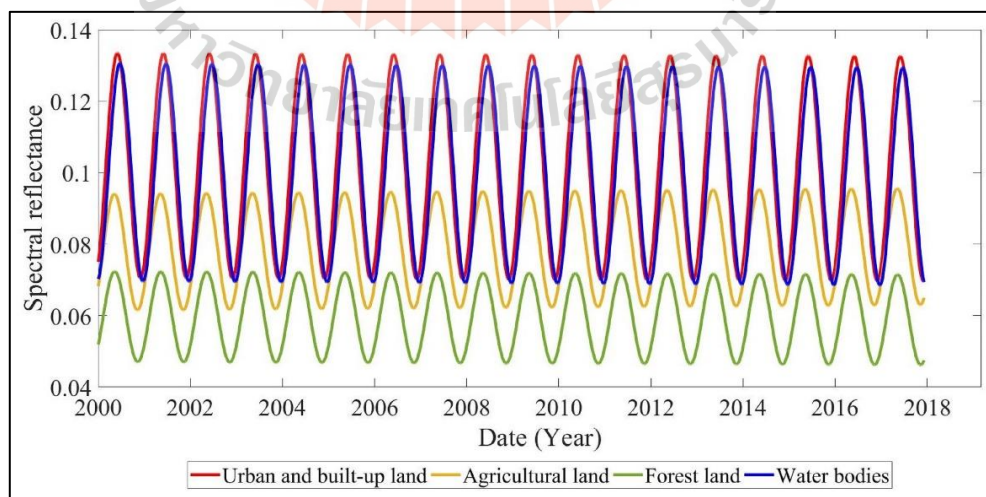
Spectral features	LULC	I	S (E-04)	A	P
BLUE	U	0.1143	0.0114	-0.0321	0.4298
	A	0.0996	0.0092	-0.0225	0.5379
	F	0.0916	0.0063	-0.0210	0.4925
	W	0.1103	0.0136	-0.0287	0.3571
GREEN	U	0.1039	0.0020	-0.0325	0.4572
	A	0.0881	0.0045	-0.0228	0.6110
	F	0.0753	0.0027	-0.0206	0.5527
	W	0.1028	0.0105	-0.0293	0.3148
RED	U	0.1020	-0.0014	-0.0314	0.5205
	A	0.0777	0.0024	-0.0163	0.9323
	F	0.0597	-0.0014	-0.0126	0.8894
	W	0.1001	-0.0019	-0.0305	0.1774
NIR	U	0.1368	0.0161	-0.0626	0.3530
	A	0.1613	0.0515	-0.0836	0.2996
	F	0.1548	0.0619	-0.1045	0.3088
	W	0.0685	-0.0214	-0.0250	0.2702
SWIR1	U	0.1250	0.0102	-0.0512	0.4175
	A	0.1103	0.0387	-0.0360	0.4323
	F	0.1032	0.0230	-0.0426	0.4420
	W	0.0198	0.0003	-0.0089	0.7526
SWIR2	U	0.0947	0.0083	-0.0364	0.5518
	A	0.0641	0.0200	-0.0166	0.9975
	F	0.0500	0.0088	-0.0151	1.0717
	W	0.0122	0.0002	-0.0059	0.8611
EVI	U	0.0140	0.0081	-0.0133	0.1604
	A	0.0355	0.0213	-0.0304	0.1211
	F	0.0423	0.0264	-0.0418	0.1953
	W	-0.0140	-0.0117	0.0027	0.0725
TCB	U	0.2587	0.0246	-0.0971	0.4188
	A	0.2425	0.0633	-0.0884	0.4245
	F	0.2195	0.0612	-0.1010	0.4022
	W	0.1614	0.0092	-0.0515	0.3117
PC1	U	0.2710	0.0267	-0.1002	0.4375
	A	0.2543	0.0666	-0.0919	0.4213
	F	0.2314	0.0598	-0.1068	0.4022
	W	0.1452	-0.0039	-0.0430	0.3665



Standard reference harmonic waves with BLUE feature

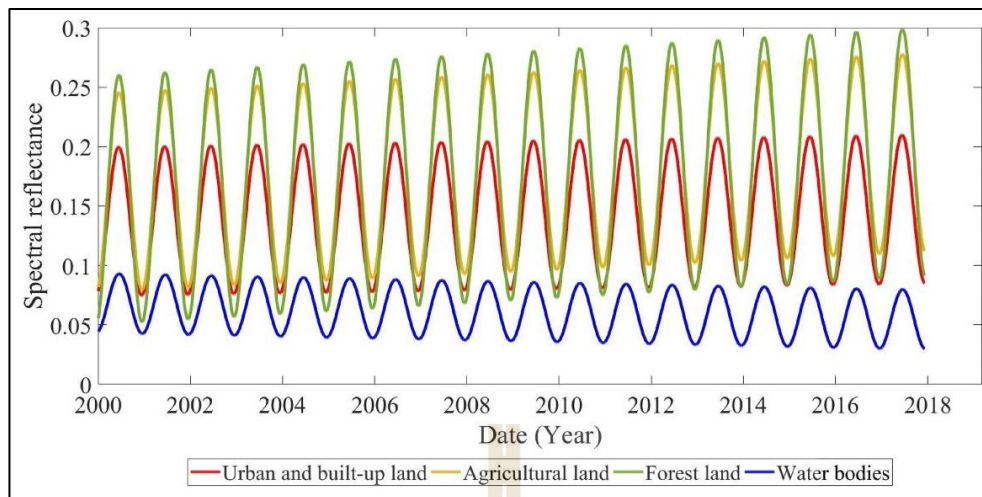


Standard reference harmonic waves with GREEN feature

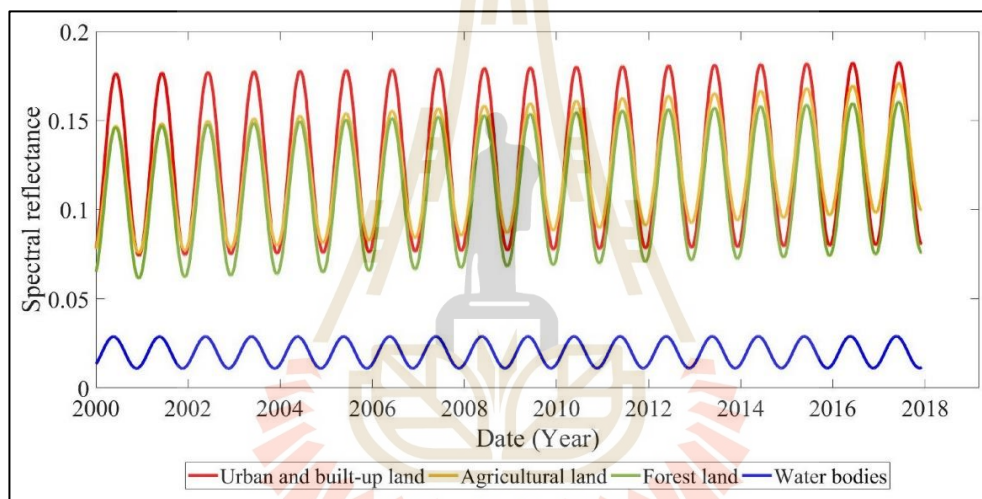


Standard reference harmonic waves with RED feature

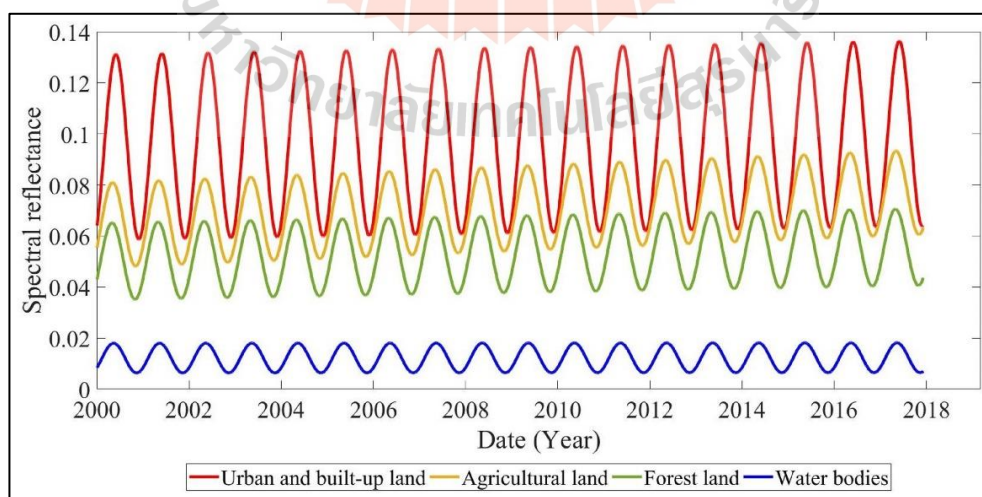
Figure 7.3 Standard reference harmonic waves of four LULC types.



Standard reference harmonic waves with NIR feature

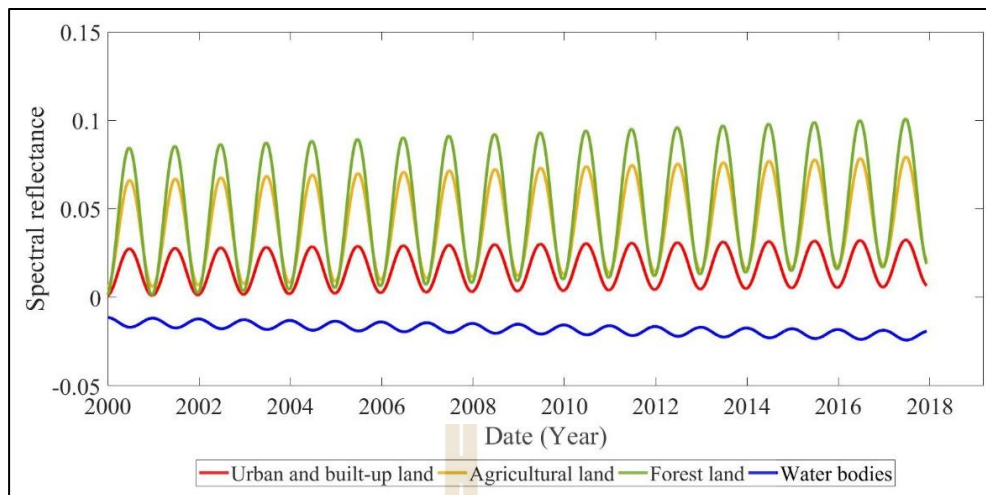


Standard reference harmonic waves with SWIR1 feature

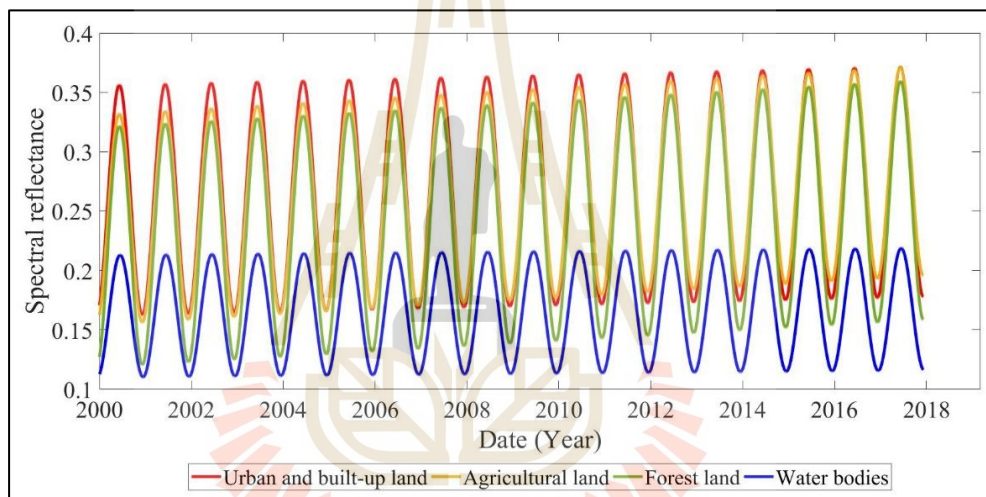


Standard reference harmonic waves with SWIR2 feature

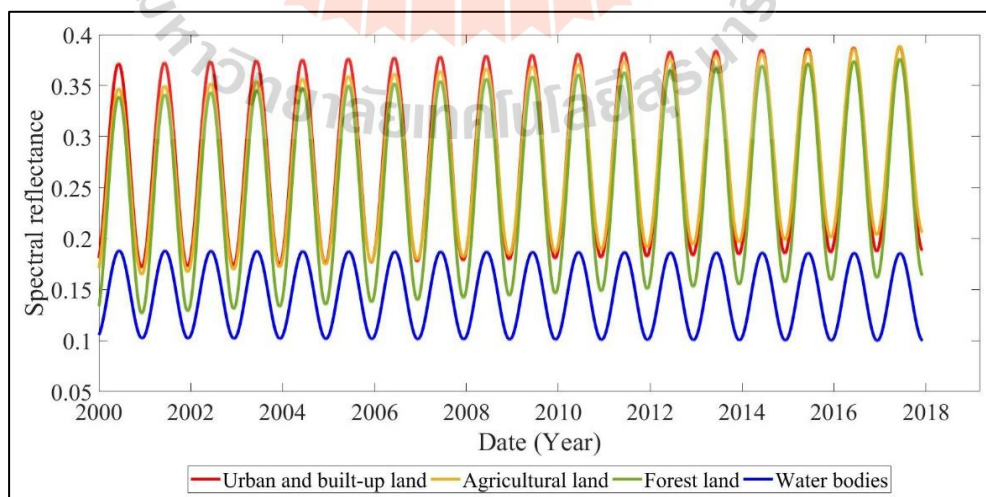
Figure 7.3 (Continued).



Standard reference harmonic waves with EVI feature



Standard reference harmonic waves with TCB feature



Standard reference harmonic waves with PC1 feature

Figure 7.3 (Continued).

7.2 Spectral distance measurement and probability calculation

In principle, the spectral distance can be measured by simple subtraction of spectral reflectance value between the standard harmonic function curves of each LULC type (U, A, F, and W) and an unclassified pixel at a specific time point. However, spectral distance values of four LULC types to an unclassified pixel have different range values from time to time, and they cannot be directly applied to compare among different LULC categories for a specific time point due to a different scale. Consequently, it is necessary to normalize spectral distance values. With this step, the spectral distance value of each LULC type will rescale into a range of [0,1].

In this study, Eq. 7.2 and Eq. 7.3 were first applied to calculate the maximum and minimum spectral distance between the standard harmonic function curve of each LULC type (U, A, F, and W) and an unclassified pixel at a specific time point. Then, maximum and minimum values from any LULC type at the same specific time point were further used to calculate the normalized spectral distance between the standard harmonic function curve of each LULC type (U, A, F, and W) and an unclassified pixel at the same specific time point using Eq. 7.4 to Eq. 7.7, respectively.

$$MAX_{ij} = \max(|R_{Uij} - R_{Xij}|, |R_{Aij} - R_{Xij}|, |R_{Fij} - R_{Xij}|, |R_{Wij} - R_{Xij}|) \quad (7.2)$$

$$MIN_{ij} = \min(|R_{Uij} - R_{Xij}|, |R_{Aij} - R_{Xij}|, |R_{Fij} - R_{Xij}|, |R_{Wij} - R_{Xij}|) \quad (7.3)$$

$$Normalized\ spectral\ distance_{Uij} = \frac{|R_{Uij} - R_{Xij}| - MIN_{ij}}{MAX_{ij} - MIN_{ij}} \quad (7.4)$$

$$Normalized\ spectral\ distance_{Aij} = \frac{|R_{Aij} - R_{Xij}| - MIN_{ij}}{MAX_{ij} - MIN_{ij}} \quad (7.5)$$

$$Normalized\ spectral\ distance_{Fij} = \frac{|R_{Fij} - R_{Xij}| - MIN_{ij}}{MAX_{ij} - MIN_{ij}} \quad (7.6)$$

$$\text{Normalized spectral distance}_{Wij} = \frac{|R_{Wij} - R_{Xij}| - \text{MIN}_{ij}}{\text{MAX}_{ij} - \text{MIN}_{ij}} \quad (7.7)$$

where, R_{Uij} , R_{Aij} , R_{Fij} , R_{Wij} are the spectral reflectance from standard harmonic function curve of U, A, F, and W, respectively, at the time point i using band j . R_{Xij} is the spectral reflectance of an unclassified pixel at the time point i using band j .

As mentioned earlier, the domain value of the normalized spectral distance is 0 to 1. The closer the spectral distance is, the higher the probability is. The lowest normalized spectral distance between any LULC type and an unclassified pixel (i.e., a spectral distance equals 0) will provide the highest likelihood of being a specific LULC type (i.e., probability equals 1). Therefore, the normalized spectral distance of an unclassified pixel to any LULC type was applied to calculate a probability of an unclassified pixel of being specific LULC type (U, A, F, and W) using Eq. 7.8 to Eq. 7.11, respectively.

$$\text{Probability}_{Uij} = 1 - \text{Normalized spectral distance}_{Uij} \quad (7.8)$$

$$\text{Probability}_{Aij} = 1 - \text{Normalized spectral distance}_{Aij} \quad (7.9)$$

$$\text{Probability}_{Fij} = 1 - \text{Normalized spectral distance}_{Fij} \quad (7.10)$$

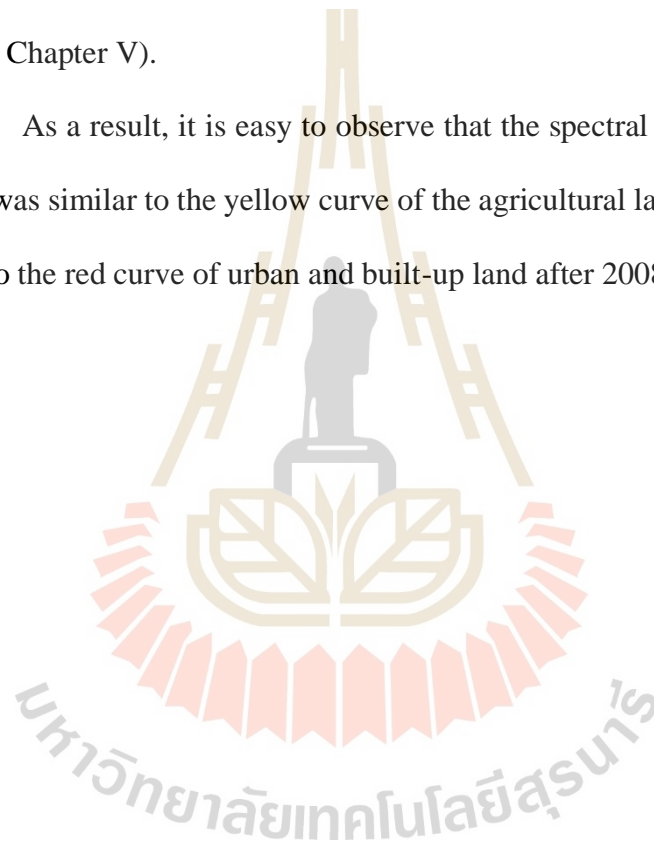
$$\text{Probability}_{Wij} = 1 - \text{Normalized spectral distance}_{Wij} \quad (7.11)$$

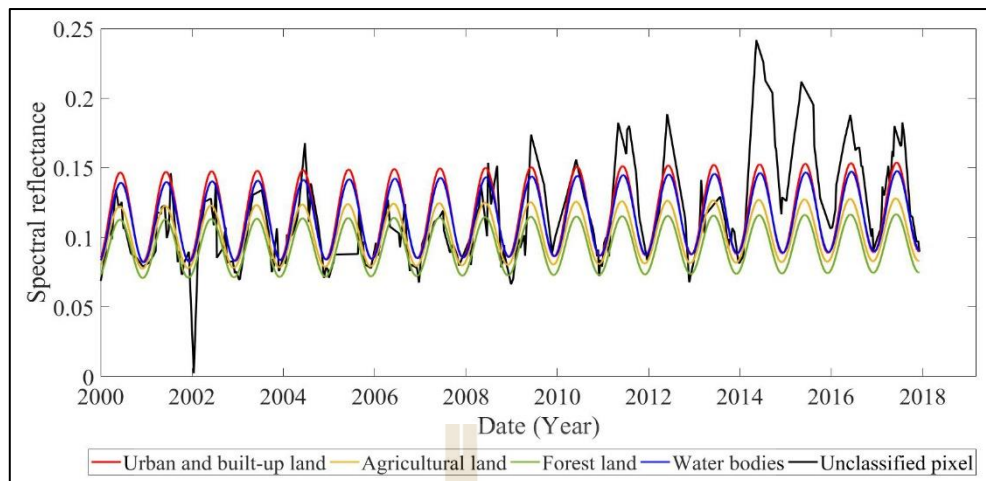
where, Probability_{Uij} , Probability_{Aij} , Probability_{Fij} , and Probability_{Wij} are the probability of an unclassified pixel of being the U, A, F, and W, respectively, at the time point i using band j .

7.2.1 Spectral distance measurement

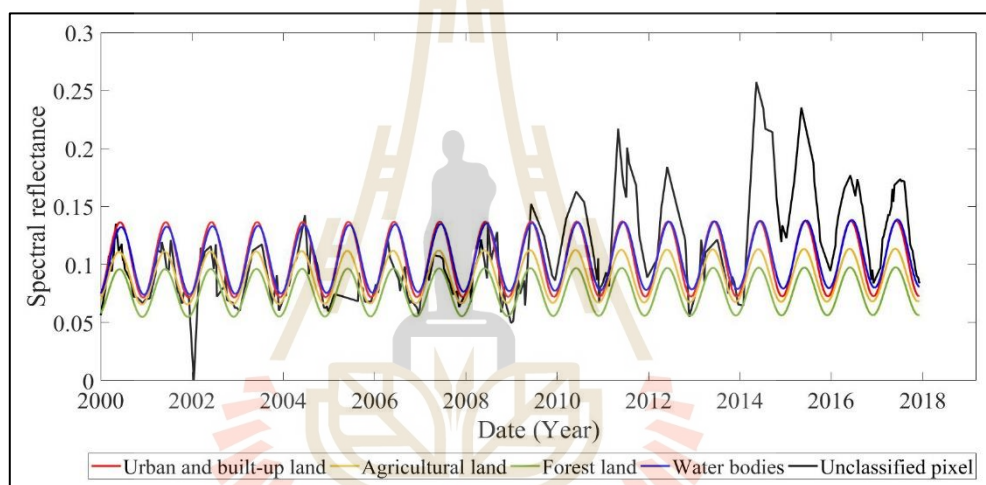
Figure 7.4 shows an example of the four reference harmonic waves and polyline of unclassified point X with different features from spectral similarity observation. In this figure, red, yellow, green, and blue curves are reference harmonic waves of urban and built-up land, agricultural land, forest land, and water bodies, respectively, while black polyline was connected to reflectance value of pixel X location (See Chapter V).

As a result, it is easy to observe that the spectral polyline (black color) of this pixel was similar to the yellow curve of the agricultural land before 2008, and it was similar to the red curve of urban and built-up land after 2008.

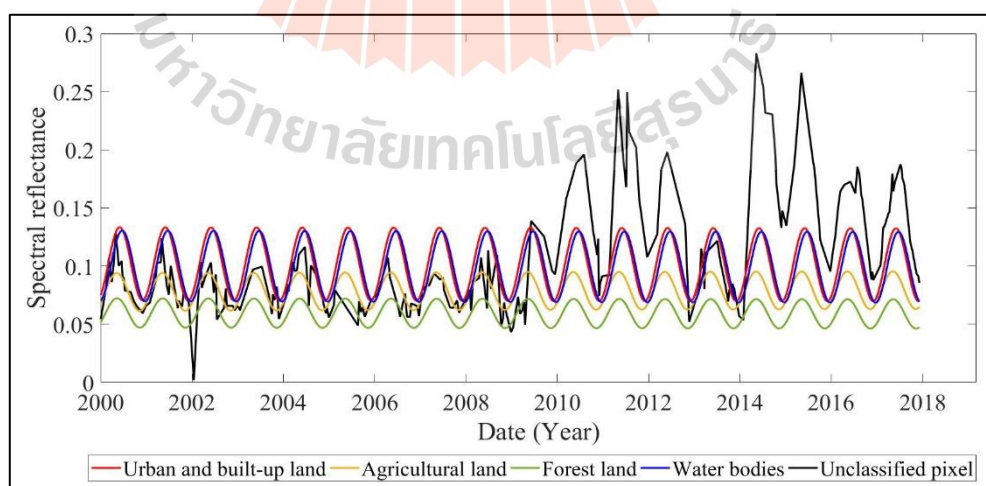




(a) BLUE



(b) GREEN



(c) RED

Figure 7.4 Spectral distance between the standard harmonic function curves and polyline of the unclassified pixels between 2000 and 2017.

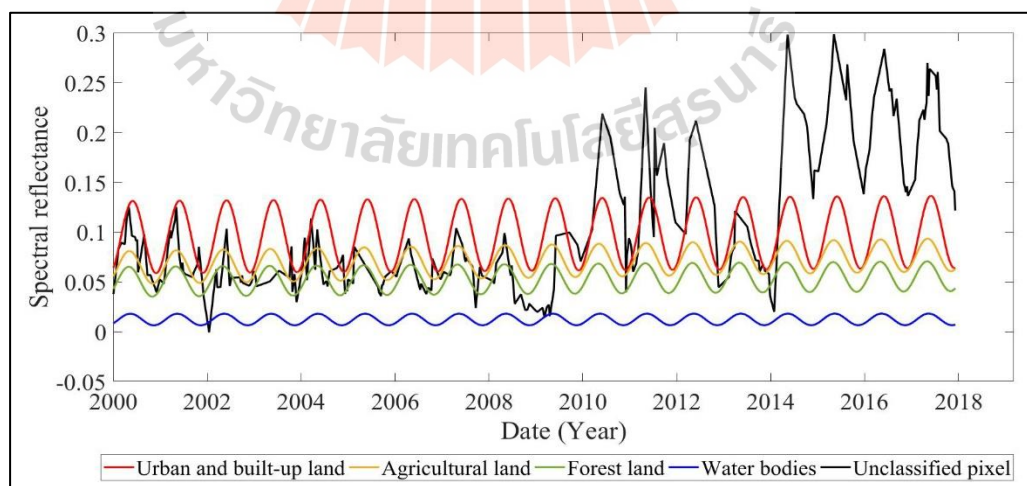
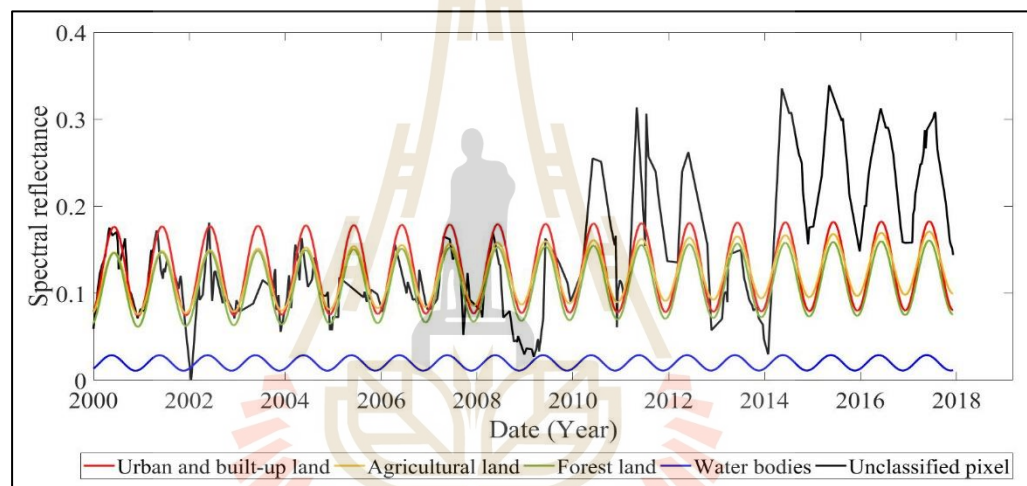
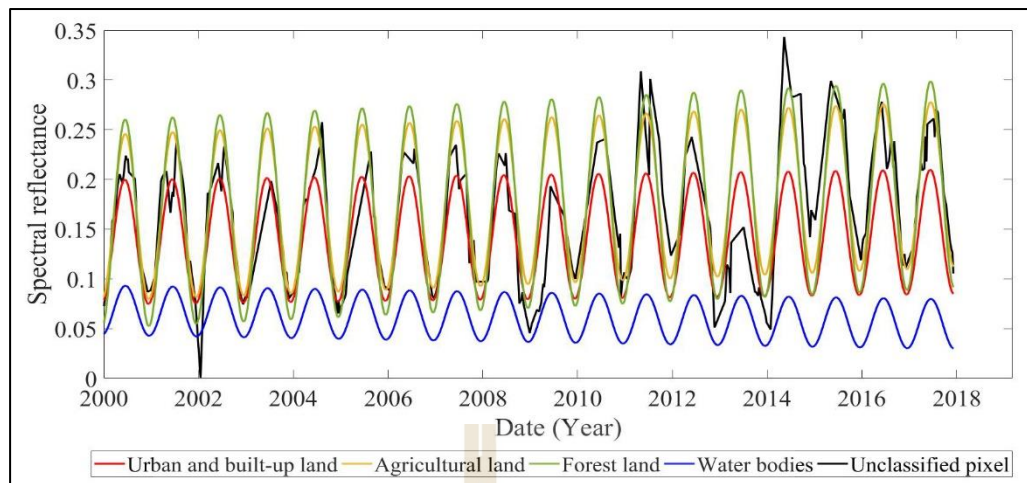
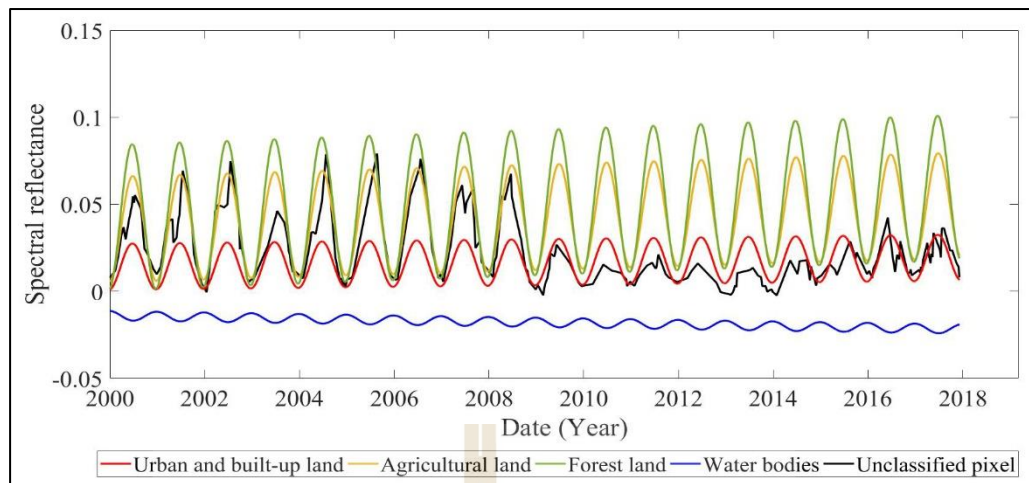
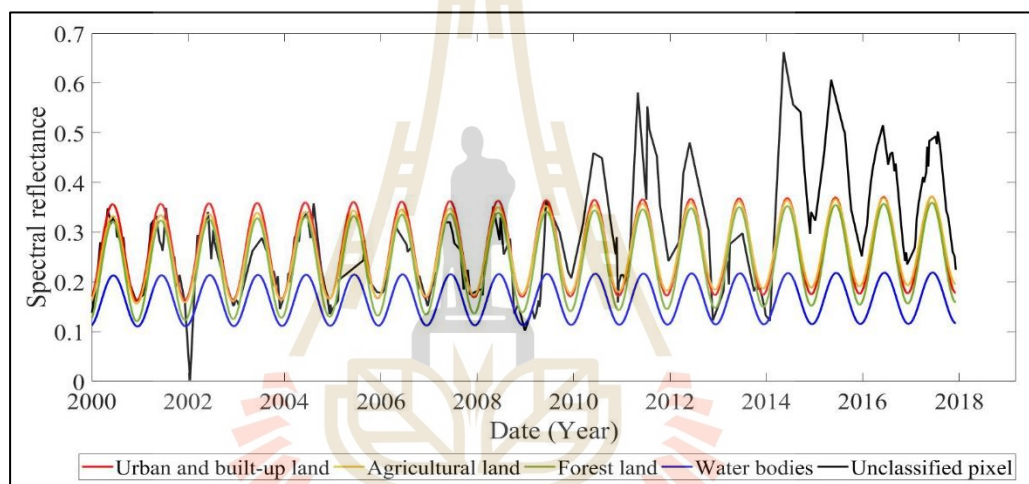


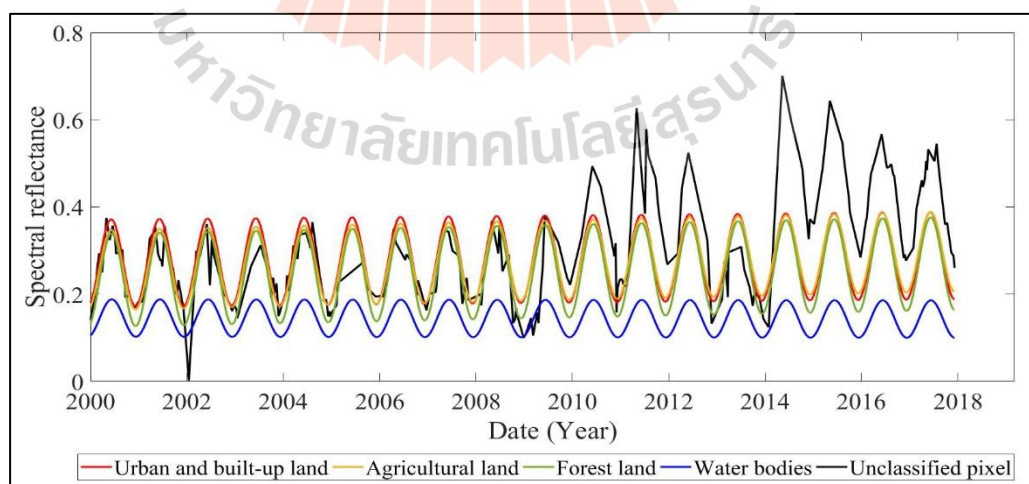
Figure 7.4 (Continued).



(g) EVI



(h) TCB

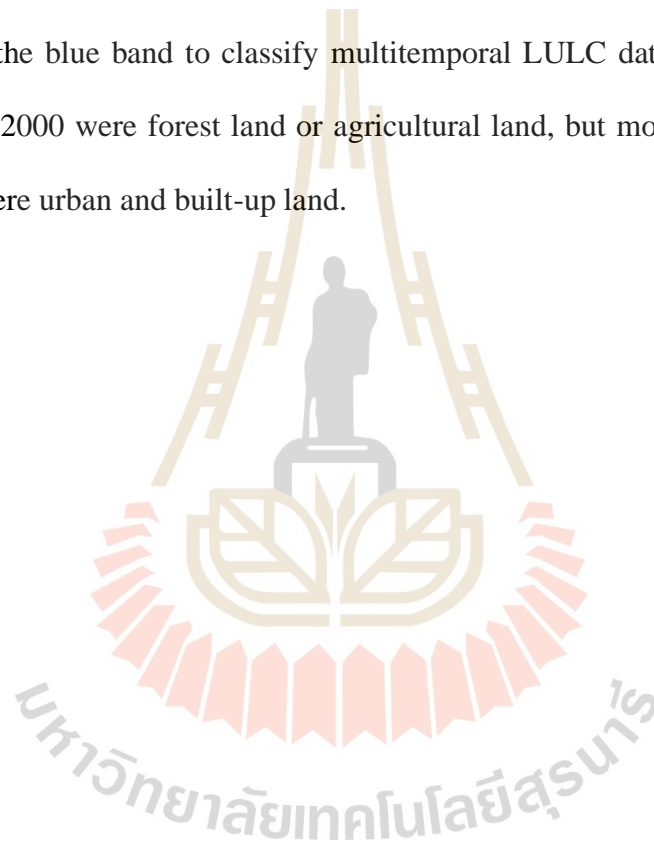


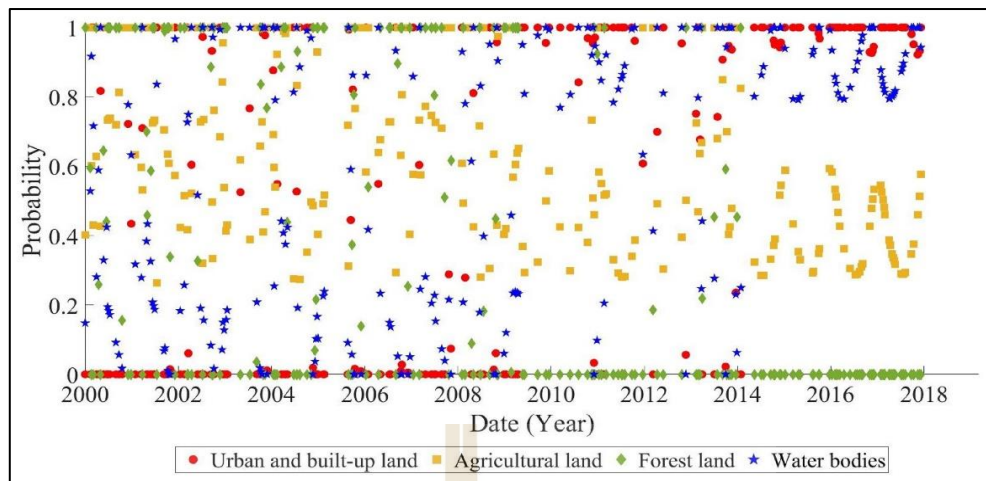
(i) PC1

Figure 7.4 (Continued).

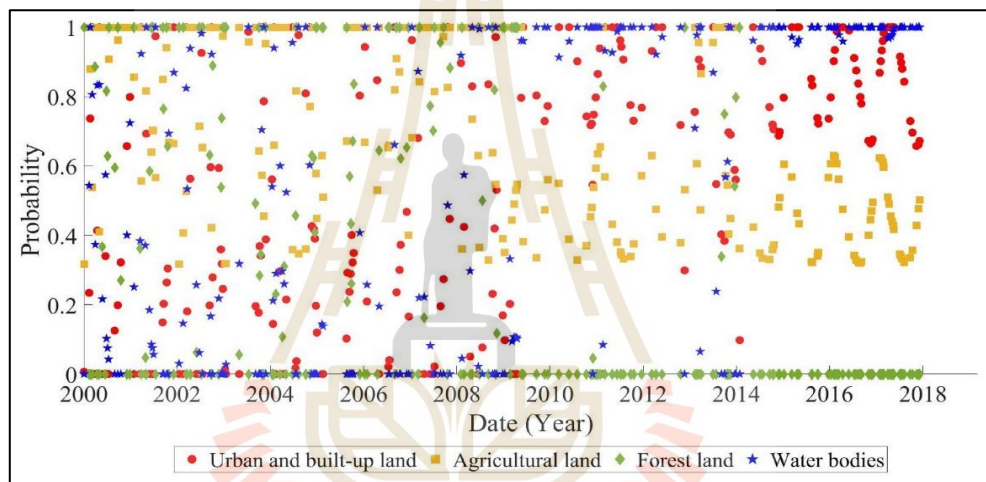
7.2.2 Probability of an unclassified pixel of being specific LULC type

The probability of an unclassified pixel being a specific LULC type (U, A, F, and W) between 2000 and 2017, based on the nine spectral features, is displayed in Figure 7.5. This figure shows the variation in the probability of an unclassified pixel being a specific LULC type between 2000 and 2017 and indicates the probability of an unclassified pixel being a specific LULC type at a particular time point. For example, if we select the blue band to classify multitemporal LULC data, most of the LULC types before 2000 were forest land or agricultural land, but most of the LULC types after 2008 were urban and built-up land.

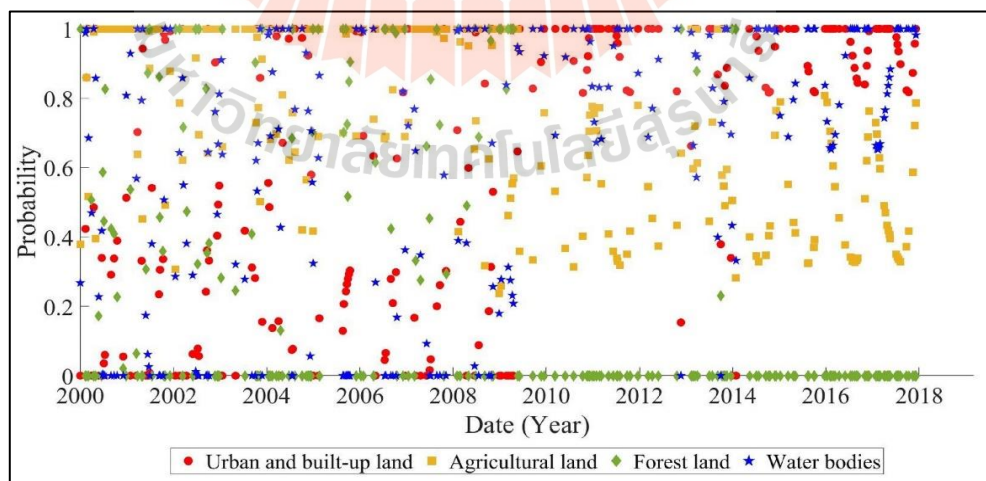




(a) BLUE

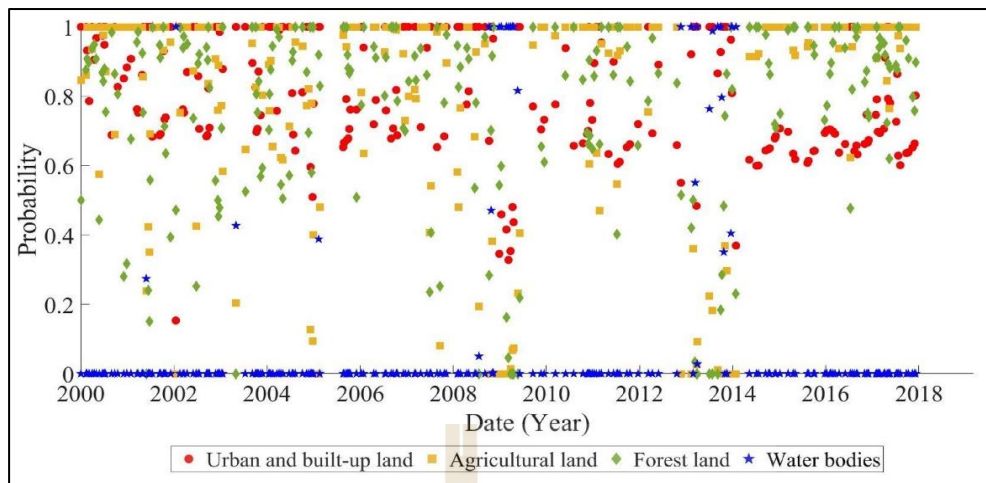


(b) GREEN

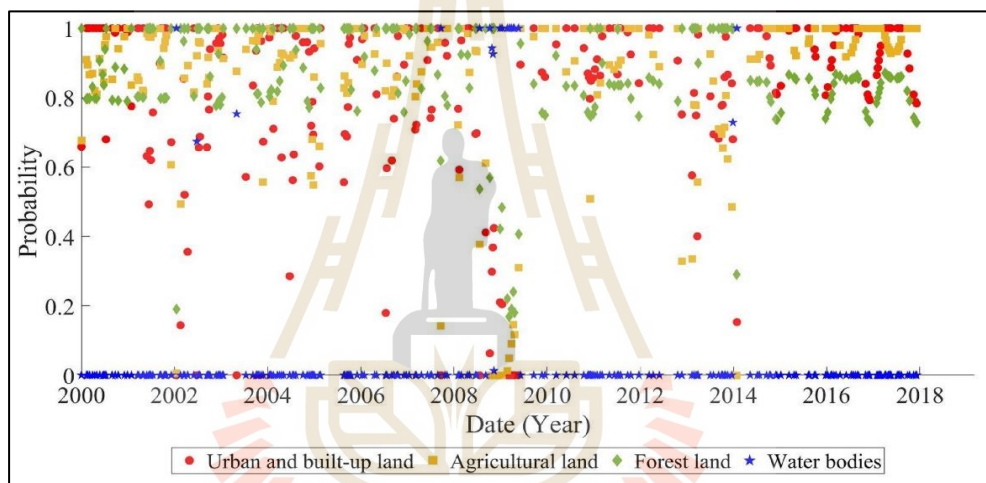


(c) RED

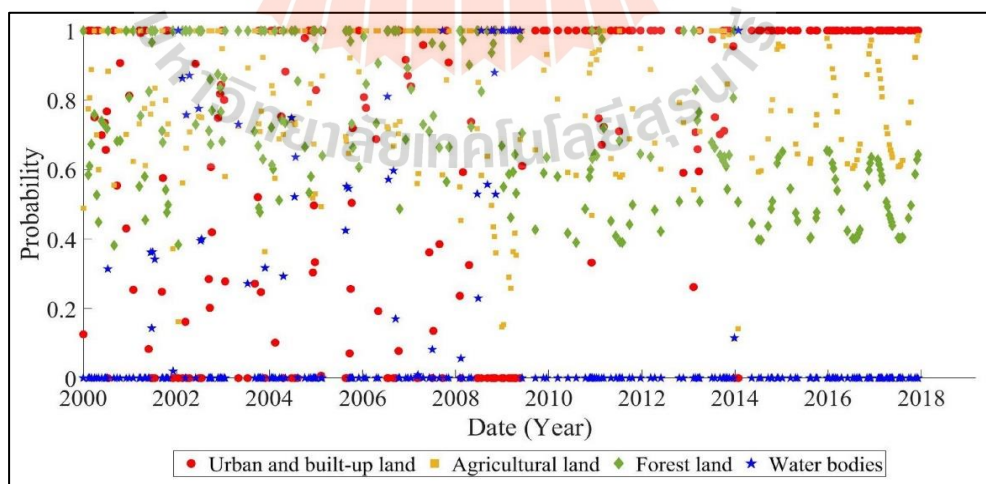
Figure 7.5 Probability of an unclassified pixel for each LULC type between 2000 and 2017 in each spectral feature.



(d) NIR

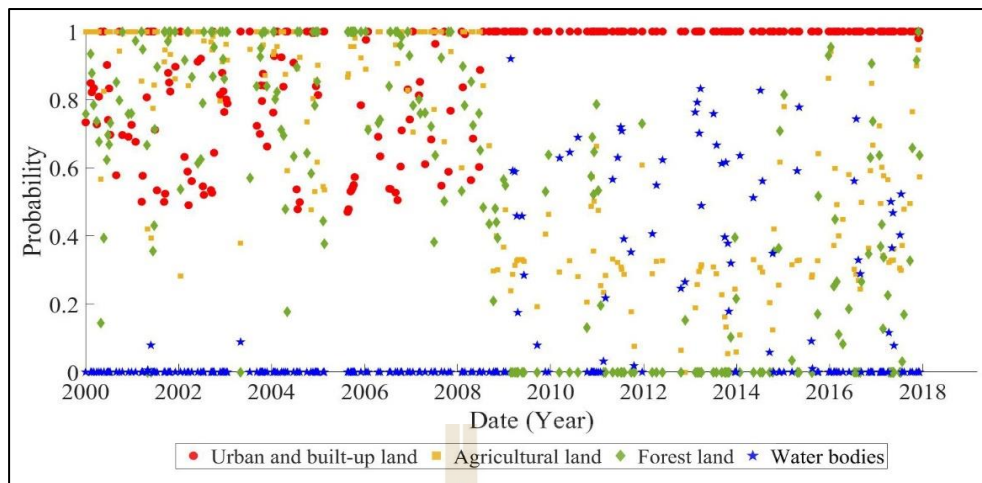


(e) SWIR1

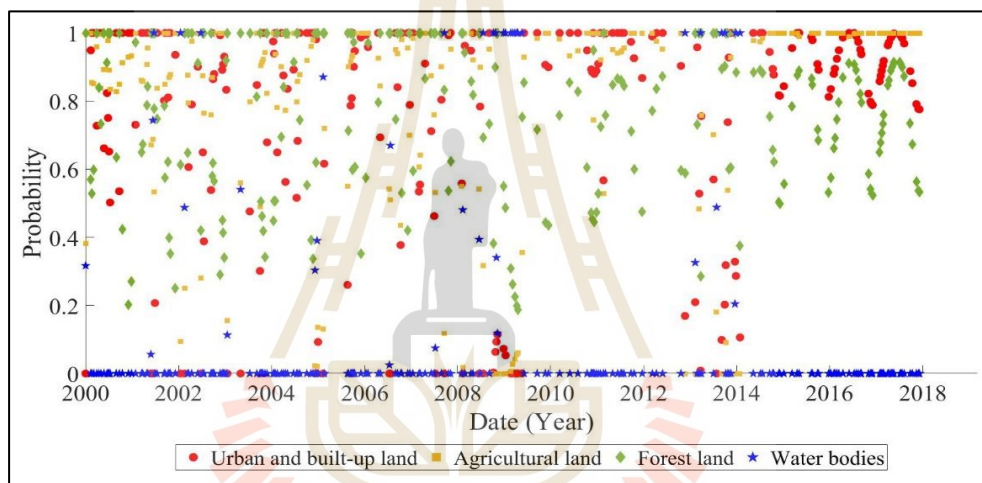


(f) SWIR2

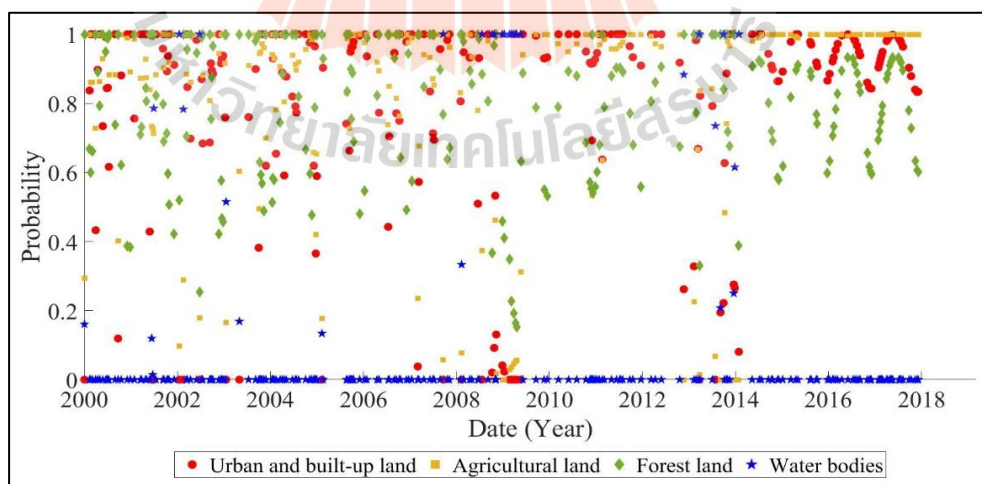
Figure 7.5 (Continued).



(g) EVI



(h) TCB



(i) PC1

Figure 7.5 (Continued).

7.3 Time-series LULC classification

In principle, the calculated probability of an unclassified pixel of being specific LULC type from Eq. 7.8 to Eq. 7.11 at a specific time point from one spectral band can be separately applied for multitemporal LULC classification.

In this study, four spectral features combinations: NIR, TOP3, TOP6, and ORI6, which contain 1, 3, 6, and 6 spectral bands, respectively, were applied for multitemporal LULC classification. So, average probabilities of an unclassified pixel of being specific LULC type (U, A, F, and W) by a number of the band (n) were first separately calculated using Eq. 7.8 to Eq. 7.11. Then, average probabilities of an unclassified pixel of being specific LULC type (U, A, F, and W) from four LULC types were compared to identify the highest value and a corresponding LULC type that provides the highest probability was then assigned to an unclassified pixel at a specific time point using Eq. 7.12 to Eq. 7.15.

$$\text{Average probability}_{Ui} = \sum_{j=1}^n \text{Probability}_{Uij} / n \quad (7.12)$$

$$\text{Average probability}_{Ai} = \sum_{j=1}^n \text{Probability}_{Aij} / n \quad (7.13)$$

$$\text{Average probability}_{Fi} = \sum_{j=1}^n \text{Probability}_{Fij} / n \quad (7.14)$$

$$\text{Average probability}_{Wi} = \sum_{j=1}^n \text{Probability}_{Wij} / n \quad (7.15)$$

where, *Average probability*_{Ui}, *Average probability*_{Ai}, *Average probability*_{Fi}, *Average probability*_{Wi} are the probability of an unclassified pixel of being the U, A, F, and W, respectively, at the time point i using n bands.

After multitemporal LULC classification based on average probability, post-classification processing was applied to remove the unexpected errors since the differences in acquisition and atmospheric conditions of time-series datasets that can create spectral shift (Tuia, Persello, and Bruzzone, 2016), so, the LULC type may be misclassified at a specific time point. In this study, mode function with a moving window with $1 * 9$ size of the spatio-temporal filtering under the MATLAB software was operated to eliminate the unexpected errors. Herein, the center pixel was replaced by the value that occurs most frequently in the window.

7.3.1 Average probability of an unclassified pixel being a specific LULC type

The average probability of an unclassified pixel being a specific LULC type (U, A, F, and W) between 2000 and 2017, based on the nine spectral bands, is presented in Figure 7.6.

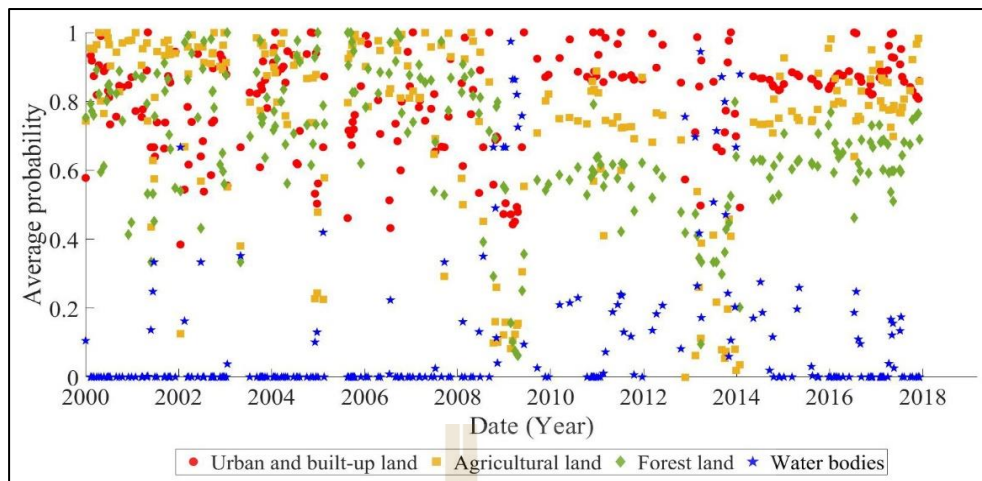
Figure 7.6(a) shows the variation of the average probability of an unclassified pixel for each LULC type from the three spectral bands (TOP3) between 2000 and 2017 at a specific time point.

Figure 7.6(b) shows the variation of the average probability of an unclassified pixel for each LULC type from the six spectral bands (TOP6) between 2000 and 2017 at a specific time point.

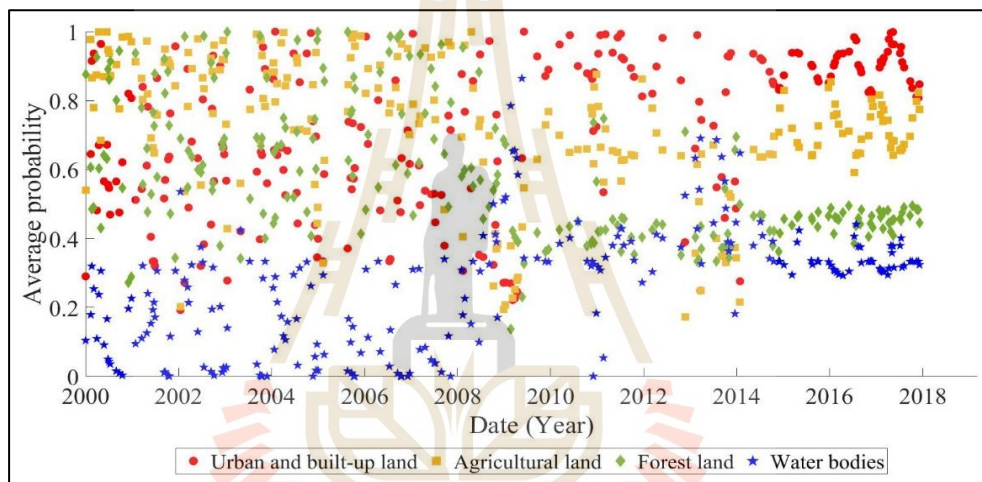
Figure 7.6(c) shows the variation of the average probability of an unclassified pixel for each LULC type from the six spectral bands (ORI6) between 2000 and 2017 at a specific time point.

As a result, the unclassified pixel belongs to different LULC types at different time points. Before 2008, this pixel alternatively shows a high average probability to appear in forest land and agricultural land, but after 2008, this pixel displays a high average probability to appear in urban and built-up land. This finding indicates that the LULC type of this pixel can be classified at a specific time point according to the highest average probability value of the corresponding LULC type.

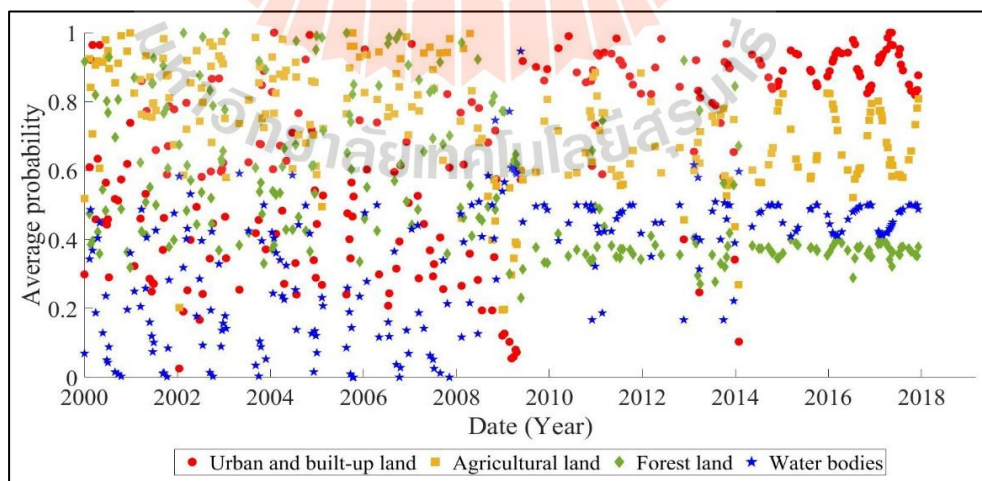




(a) TOP3 features



(b) TOP6 features



(c) ORI6 features

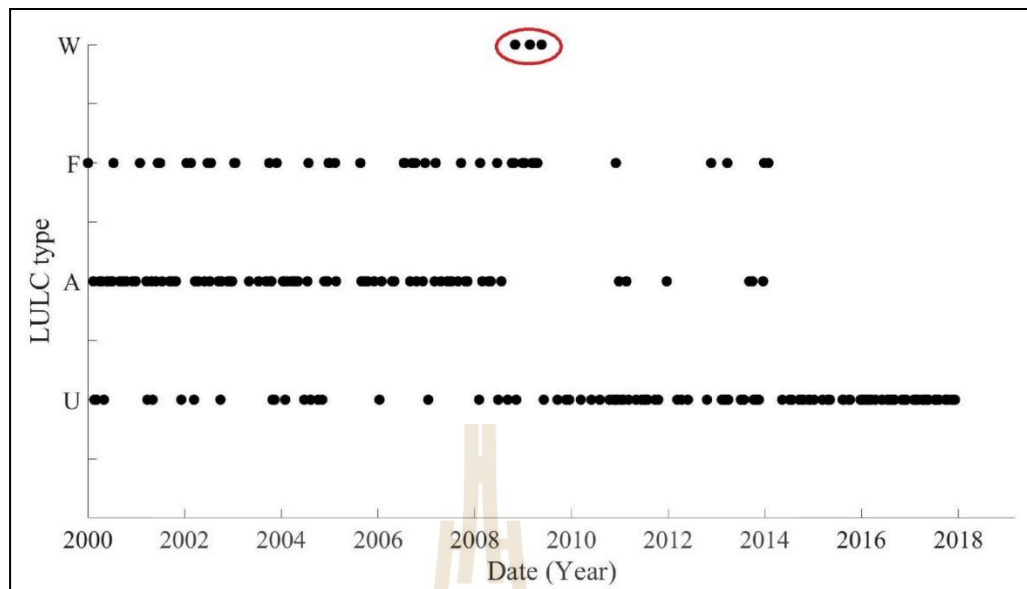
Figure 7.6 Average probability of the unclassified pixel for each LULC type between 2000 and 2017 from TOP3, TOP6, and ORI6 features.

7.3.2 Supervised LULC classification after post-classification processing

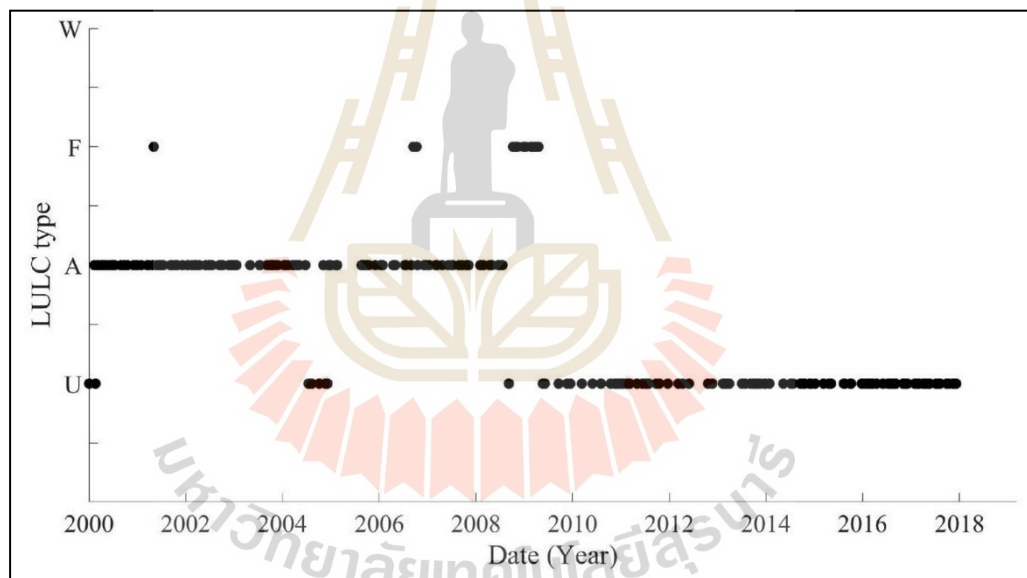
Figure 7.7 shows an example of a multitemporal LULC classification of one unclassified pixel between 2000 and 2017.

Figure 7.7(a) shows the LULC type of the unclassified pixel between 2000 and 2017 at specific dates of the 388 Landsat data before post-classification processing. As a result, the LULC type of the pixel alternately belongs to forest land, agricultural land, and urban and built-up land between 2000 and 2008. It also frequently belongs to agricultural land during this period. In contrast, this pixel belongs to urban and built-up land after 2008. However, this pixel belongs to water bodies at three dates (as shown by the red oval). Based on our prior knowledge, it is difficult for the LULC type to undergo two sharp changes in a short period. The LULC type of this pixel at the three-time points indicates unexpected errors like salt and pepper noise (Jensen, 2015).

On the contrary, Figure 7.7(b) shows the final results after post-classification processing for the unclassified pixel; this figure reveals that the unexpected errors from water bodies in Figure 7.7(a) have been removed. Accordingly, the classified LULC type of the unclassified pixel was agricultural land before 2008 and was urban and built-up land after 2008.



(a)



(b)

Figure 7.7 Multitemporal LULC classification of one unclassified pixel between 2000 and 2017: (a) before post-classification and (b) after post-classification.

7.3.3 Result of supervised LULC classification

Results of time-series LULC classification using HA and a minimum spectral distance algorithm based on single and multiple spectral features from a selected date in 2000, 2006, 2011, and 2017 are presented in Figures 7.8 to 7.11.

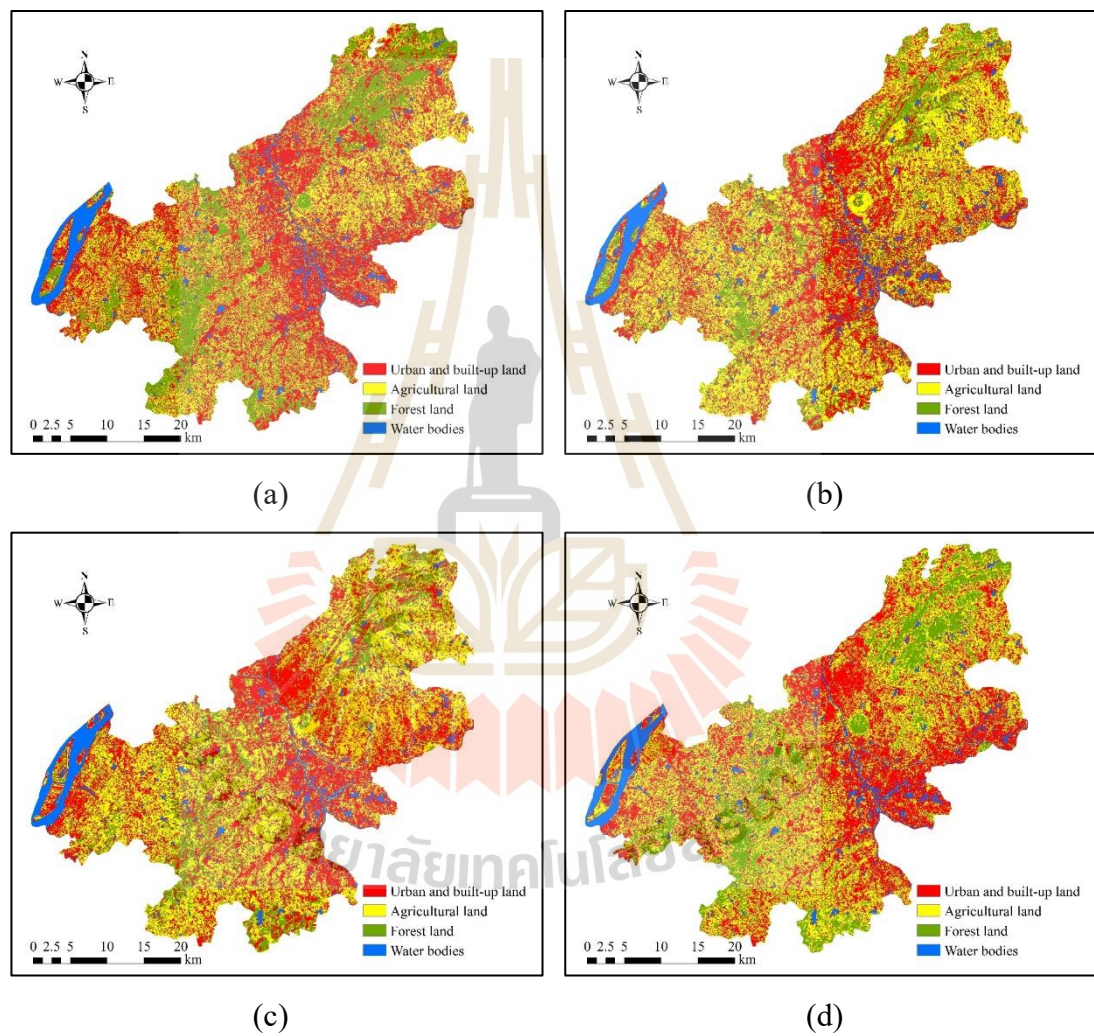


Figure 7.8 LULC classification map by the supervised algorithm with NIR band: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.

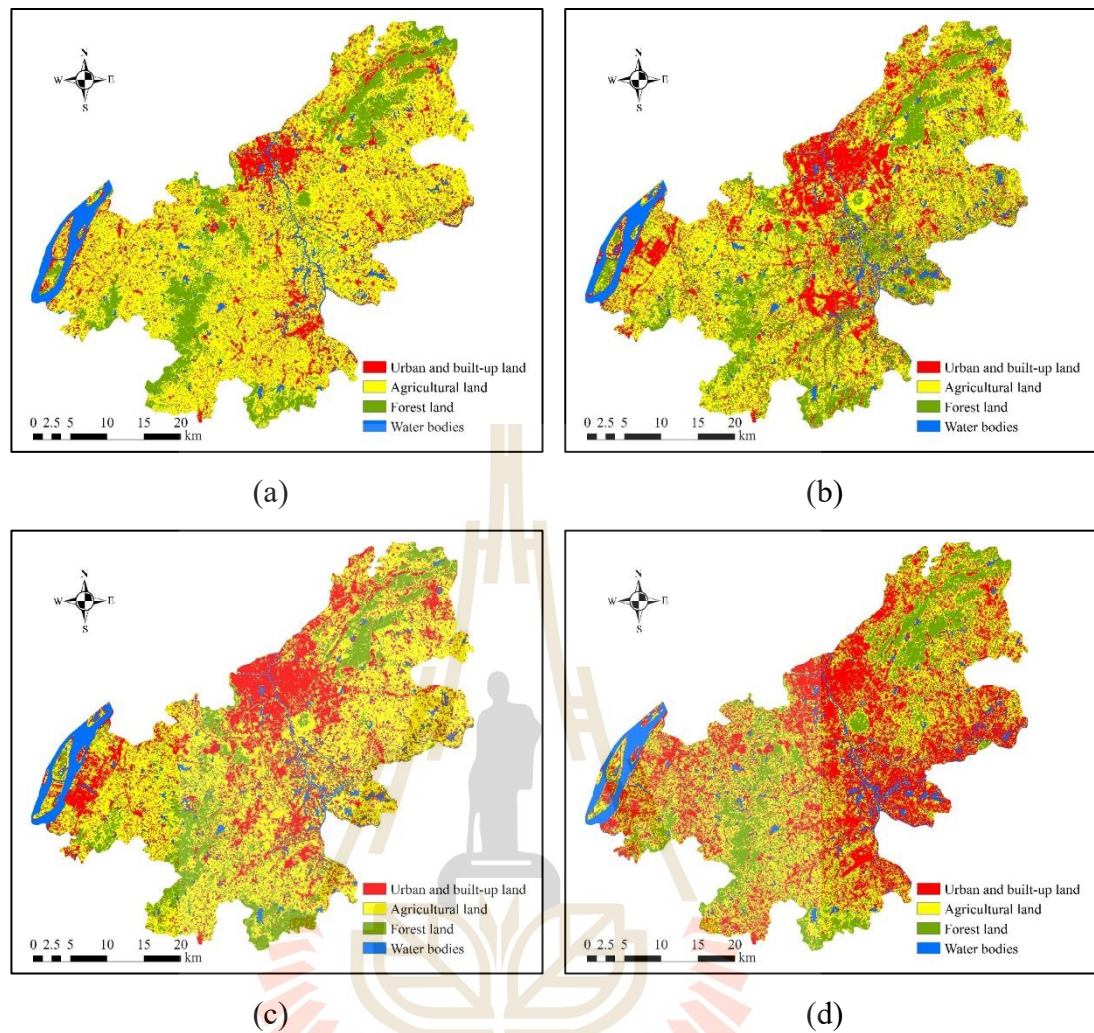


Figure 7.9 LULC classification map by the supervised algorithm with TOP3 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.

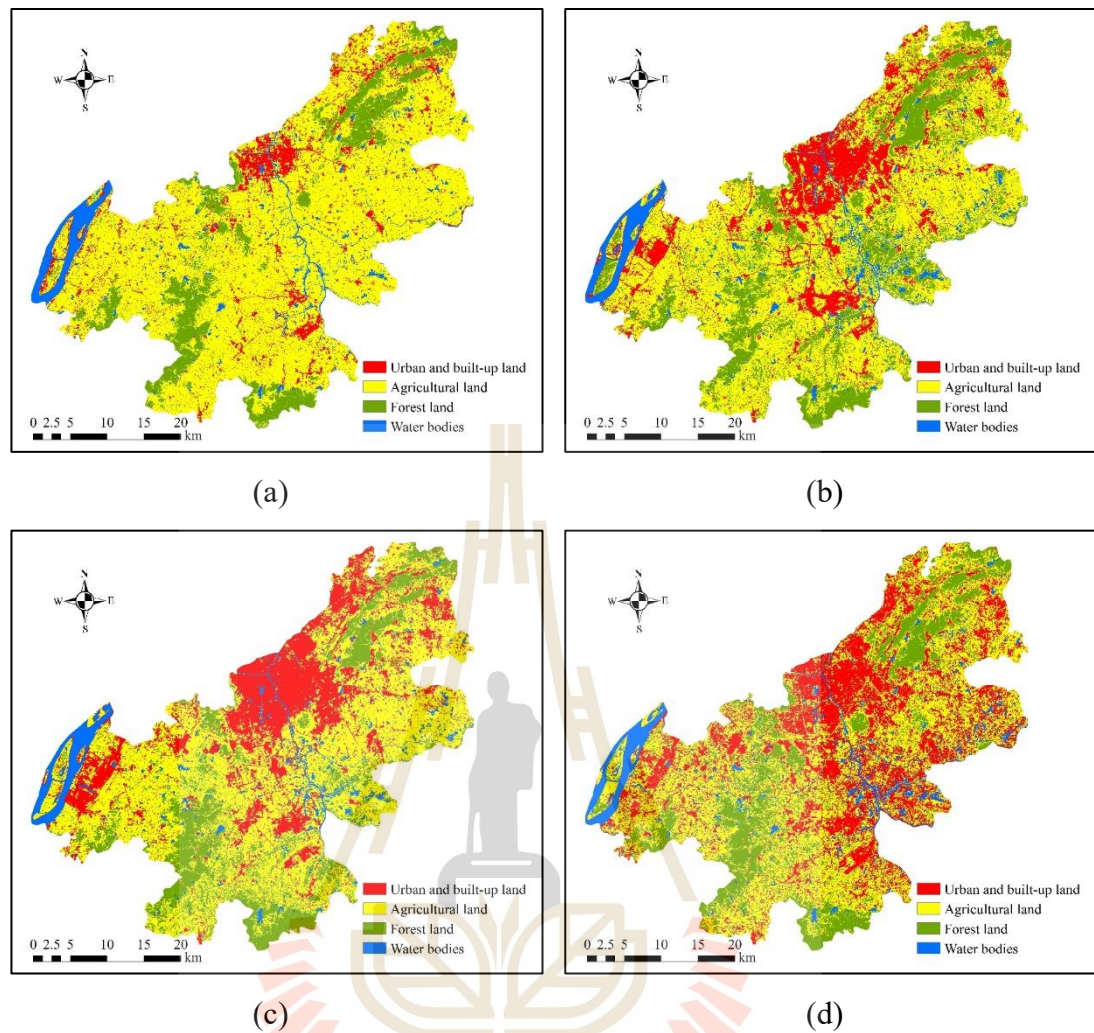


Figure 7.10 LULC classification map by the supervised algorithm with TOP6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.

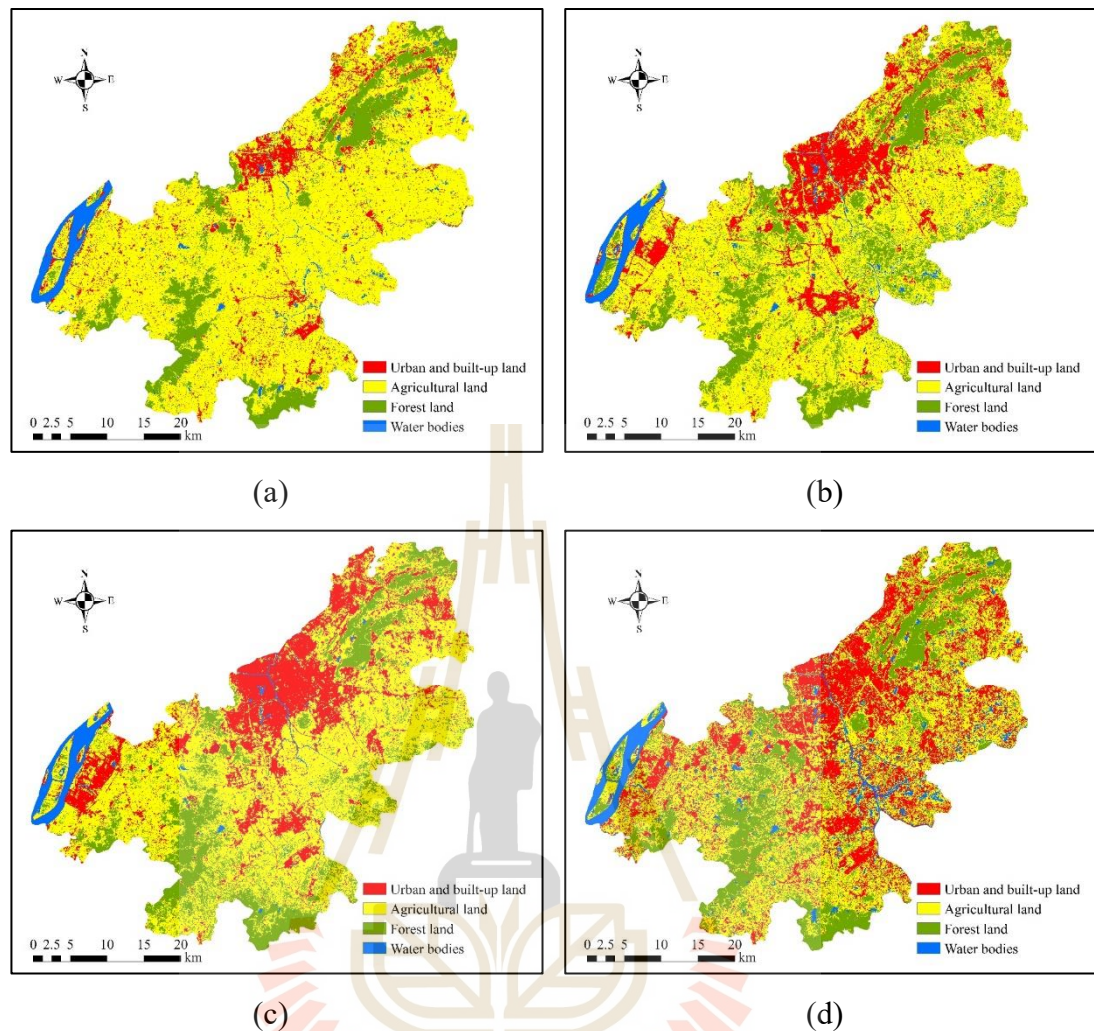


Figure 7.11 LULC classification map by the supervised algorithm with ORI6 spectral features: (a) Date 12 June 2000, (b) Date 31 July 2006, (c) Date 29 July 2011, and (d) Date 18 May 2017.

In detail, Figure 7.8 shows the supervised LULC classification maps in 2000, 2006, 2011, and 2017 using the NIR band. Meanwhile, Figure 7.9 shows the supervised LULC classification maps in 2000, 2006, 2011, and 2017 using the TOP3 spectral features (NIR, EVI, and TCB). In the meantime, Figure 7.10 and Figure 7.11 presents the supervised LULC classification maps in 2000, 2006, 2011 and 2017 using the TOP6 spectral features (NIR, EVI, TCB, BLUE, GREEN, and PC1) and ORI6

spectral bands from Landsat data (BLUE, GREEN, RED, NIR, SWIR1, and SWIR2), respectively.

Meanwhile, area and percentage of supervised LULC classification in 2000, 2006, 2011, and 2017 according to single and multiple spectral features were summarized in Tables 7.2 to 7.5. In the meantime, the proportional LULC areas using supervised classification with single and multiple spectral features in 2000, 2006, 2011, and 2017 are comparatively displayed in Figures 7.12 to 7.15, respectively.

By comparing area and percentage according to spectral feature, it reveals that areas and percentages of urban and built-up land, which was extracted based on the NIR band are somewhat high when they are compared with other spectral features in every year except 2017. The percent of urban and built-up land using NIR band decreased from 41.14% in 2000 to 36.89% in 2017. This finding showed an unexpected result when it was compared with the report of MOHURD in 2017. On the contrary, areas and percentages of urban and built-up land between 2000 and 2017 using multi-spectral features (TOP3, TOP6, and ORI6) have been continuously increased, and they are consistent with the report of MOHURD in 2017. Additionally, areas and percentages of agricultural land based on NIR are rather low when it is compared with other spectral features in every year. In the meantime, areas and percentages of forest land and water bodies using the NIR band are similar to other spectral features in every year.

Table 7.2 Area and percentage of supervised LULC classification in 2000.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	650.95	41.14	245.72	15.53	149.78	9.47	147.02	9.29
A	478.49	30.24	944.82	59.71	1116.99	70.59	1141.86	72.17
F	346.21	21.88	282.03	17.82	214.13	13.53	225.51	14.25
W	106.62	6.74	109.72	6.93	101.38	6.41	67.89	4.29

Table 7.3 Area and percentage of supervised LULC classification in 2006.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	558.79	35.32	362.27	22.90	271.40	17.15	258.45	16.33
A	620.02	39.19	719.82	45.49	841.49	53.18	909.45	57.48
F	298.72	18.88	389.07	24.59	350.04	22.12	349.66	22.10
W	104.81	6.62	111.19	7.03	119.42	7.55	64.79	4.09

Table 7.4 Area and percentage of supervised LULC classification in 2011.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	574.50	36.31	430.17	27.19	371.79	23.50	353.24	22.32
A	617.10	39.00	757.23	47.86	812.66	51.36	864.94	54.66
F	293.90	18.57	293.37	18.54	292.43	18.48	308.56	19.50
W	96.85	6.12	101.57	6.42	105.47	6.67	55.62	3.51

Table 7.5 Area and percentage of supervised LULC classification in 2017.

LULC type	NIR		TOP3		TOP6		ORI6	
	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%	Area (km ²)	%
U	583.62	36.89	594.83	37.59	468.26	29.59	443.62	28.04
A	596.67	37.71	590.11	37.30	698.59	44.15	689.74	43.59
F	299.15	18.91	286.91	18.13	288.29	18.22	337.84	21.35
W	102.90	6.50	110.49	6.98	127.21	8.04	111.14	7.02

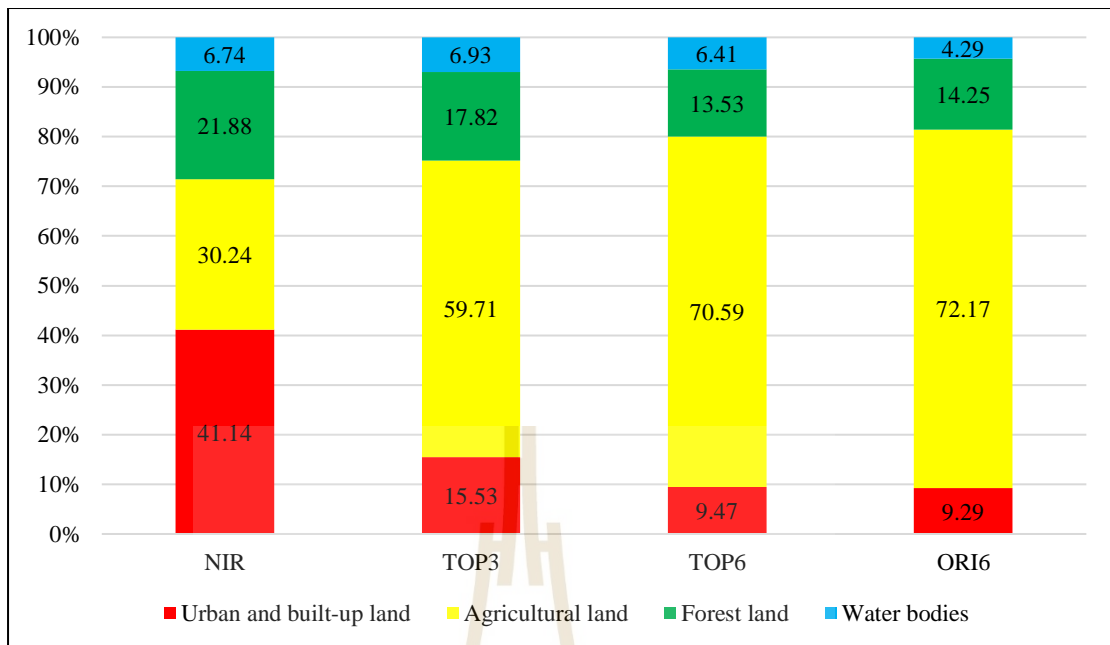


Figure 7.12 Comparison of the proportional LULC area in 2000 using supervised classification with various spectral features.

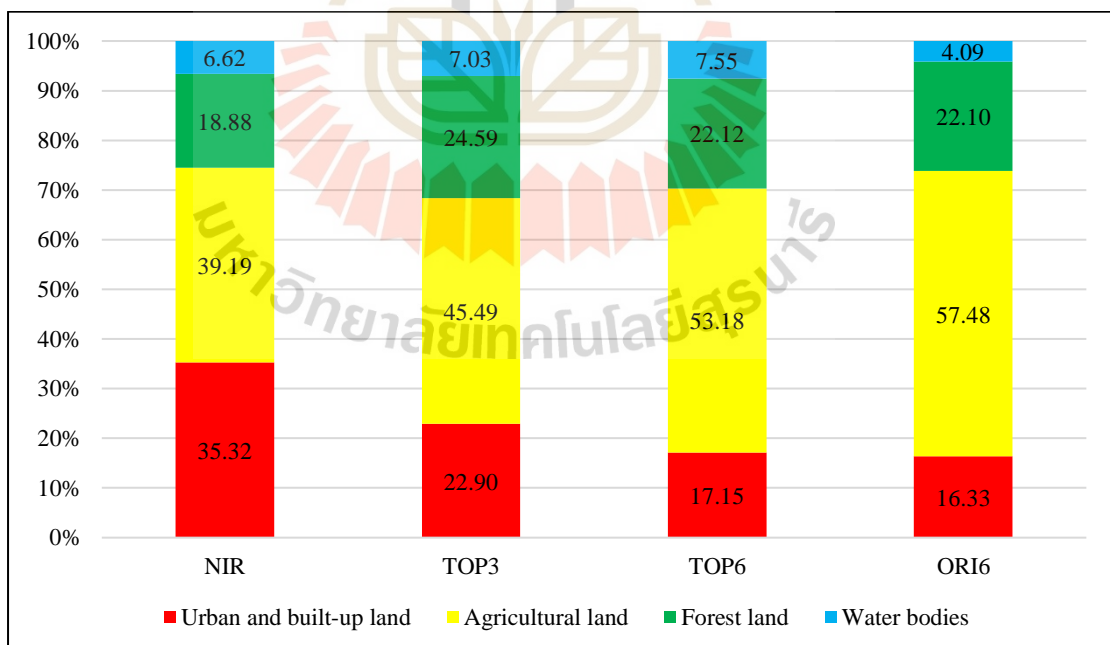


Figure 7.13 Comparison of the proportional LULC area in 2006 using supervised classification with various spectral features.

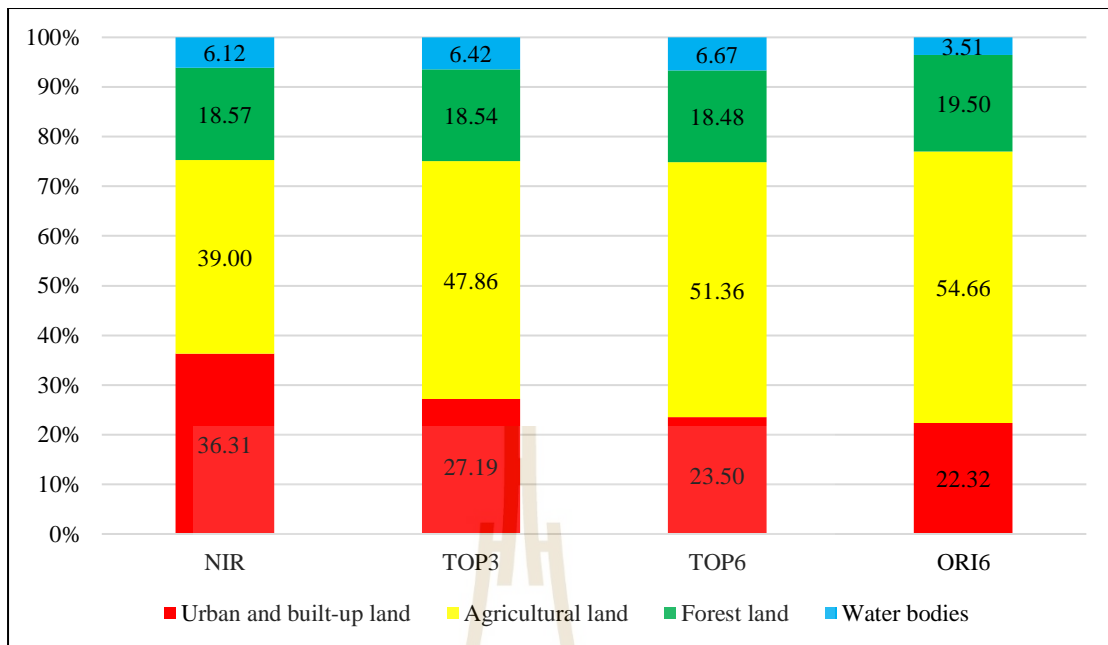


Figure 7.14 Comparison of the proportional LULC area in 2011 using supervised classification with various spectral features.

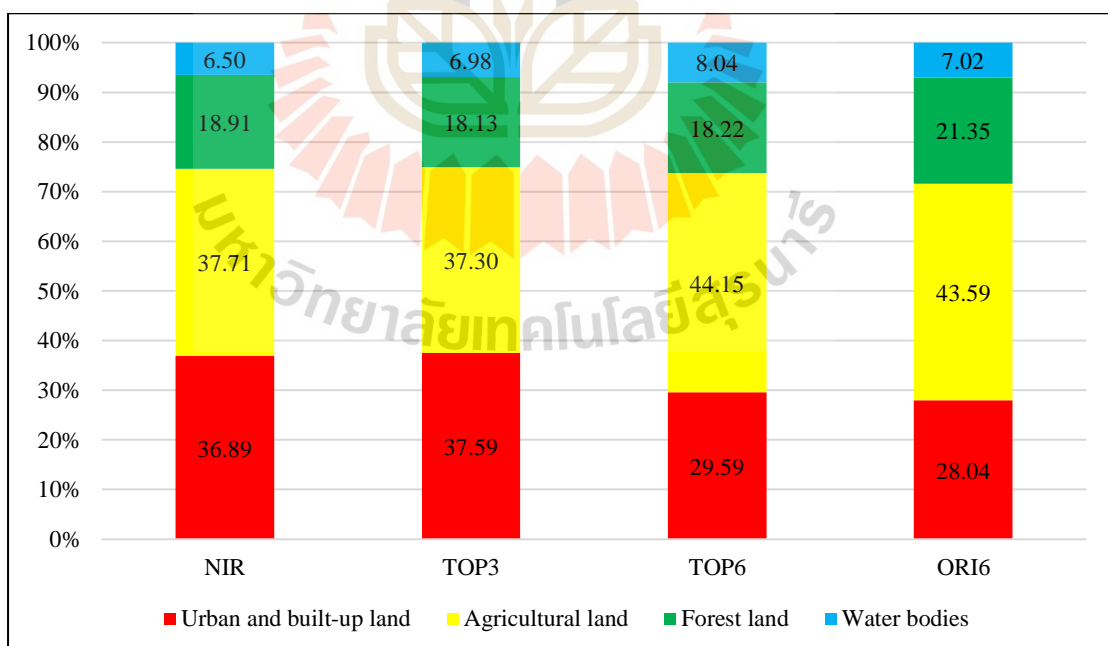


Figure 7.15 Comparison of the proportional LULC area in 2017 using supervised classification with various spectral features.

7.4 Accuracy assessment

Similar to the unsupervised classification approach, accuracy assessment of the supervised classification approach consists of two aspects, one is whether LULC changes can be detected, and the other is whether the LULC classification before and after the change is correct. Likewise, standard thematic accuracy assessment was assessed using overall accuracy (OA), producer's accuracy (PA), users' accuracy (UA), and Kappa hat coefficient based on error matrix between classified LULC data and ground reference information data (Congalton and Green, 2009).

7.4.1 Accuracy assessment for change detection

To detect LULC changes, whether different algorithms with different spectral features a particular location, the four stable LULC types and LULC change areas were reclassified into two groups: changed and unchanged areas, for change detection accuracy assessment. Figure 7.16 shows the result of LULC change detection maps (stable LULC type and LULC change area) with different spectral features by supervised classification algorithms.

Same as Chapter VI, a stratified random sampling scheme was here used for assessing the accuracy of change detection with four different spectral features using supervised classification algorithms. Herein, 502 reference pixels are based on multinomial distribution with the desired precision of 5% and a level of confidence of 95%. Additionally, high spatial resolution images from the Google Earth in 2000, 2006, 2011, and 2017 as ground reference were used to determine if the LULC type at a particular location has changed.

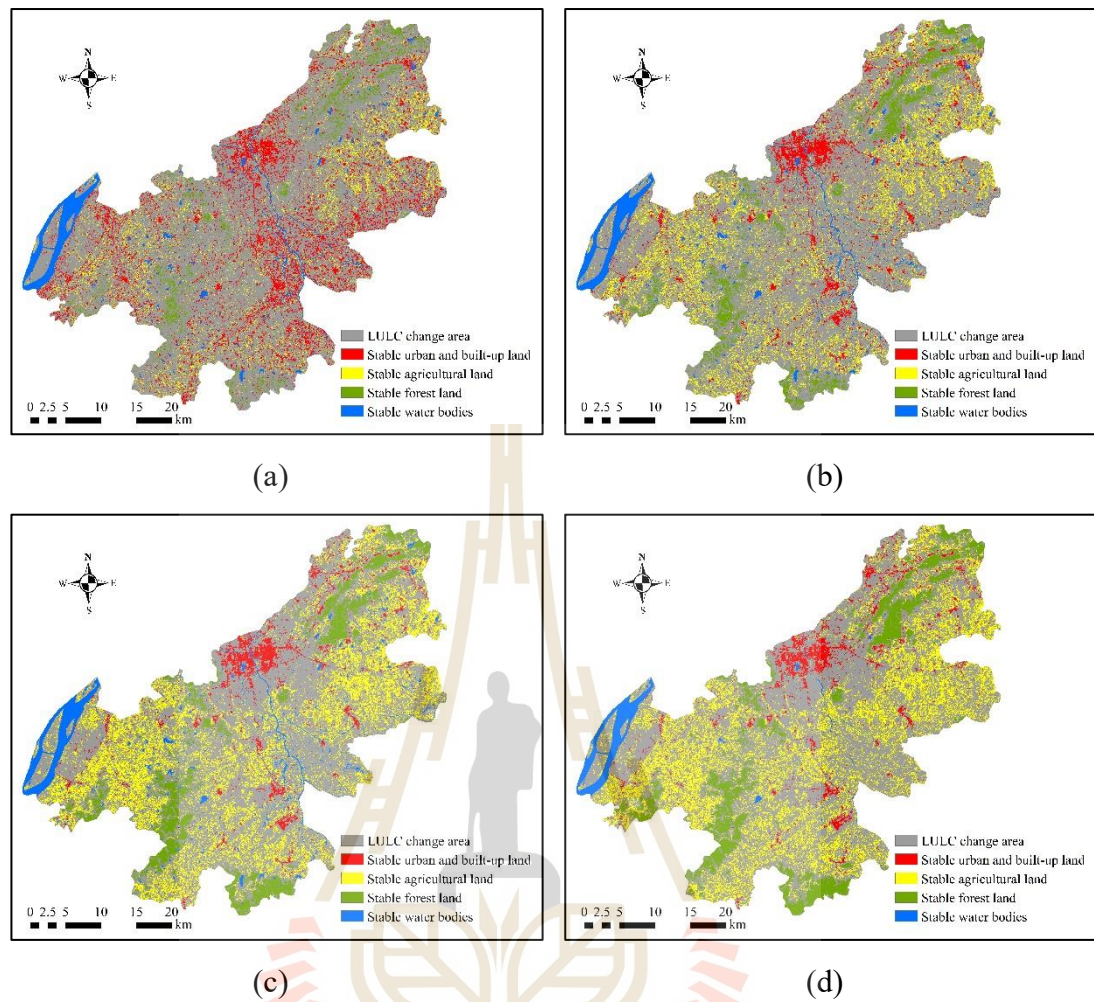


Figure 7.16 LULC change detection by the supervised algorithm: (a) Using NIR band, (b) Using TOP3 spectral features, (c) Using TOP6 spectral features, and (d) Using ORI6 spectral features.

Tables 7.6 to 7.9 show the results of the accuracy assessment of change detection.

Table 7.6 The accuracy assessment of change detection using the NIR feature.

NIR	Changed pixels	Stable pixels	Row Total
Changed pixels	256	115	371
Stable pixels	12	119	131
Column Total	268	234	502
Producer's accuracy (%)	95.52	50.85	
User's accuracy (%)	69.00	90.84	
Overall accuracy (%)		74.70	
Kappa hat (%)		47.71	

Table 7.7 The accuracy assessment of change detection using TOP3 spectral features.

TOP3	Changed pixels	Stable pixels	Row Total
Changed pixels	269	84	353
Stable pixels	8	141	149
Column Total	277	225	502
Producer's accuracy (%)	97.11	62.67	
User's accuracy (%)	76.20	94.63	
Overall accuracy (%)		81.67	
Kappa hat (%)		61.74	

Table 7.8 The accuracy assessment of change detection using TOP6 spectral features.

TOP6	Changed pixels	Stable pixels	Row Total
Changed pixels	247	55	302
Stable pixels	10	190	200
Column Total	257	245	502
Producer's accuracy (%)	96.11	77.55	
User's accuracy (%)	81.79	95.00	
Overall accuracy (%)		87.05	
Kappa hat (%)		73.98	

Table 7.9 The accuracy assessment of change detection using ORI6 spectral features.

ORI6	Changed pixels	Stable pixels	Row Total
Changed pixels	247	53	300
Stable pixels	6	196	202
Column Total	253	249	502
Producer's accuracy (%)	97.63	78.71	
User's accuracy (%)	82.33	97.03	
Overall accuracy (%)		88.25	
Kappa hat (%)		76.46	

By comparing overall accuracy according to spectral feature, LULC change detection using NIR provides the least accuracy (74.70%). In contrast, the accuracy of TOP3, TOP6, and ORI6 spectral features for change detection are 81.67%, 87.05%, and 88.25%, respectively. These combinations of spectral features can provide similar results with high accuracy.

Similar to the overall accuracy, the Kappa hat coefficient of NIR is quite low (47.71%), it is slightly higher when using TOP3 spectral features, and it is higher when using TOP6 and ORI6 spectral features with values of 73.98%, 76.46%, respectively. In other words, when applying more spectral features, the Kappa hat coefficient is higher.

Interestingly, for the detection of changed pixels, all spectral feature combinations can achieve higher producer's accuracy, 95.52%, 97.11%, 96.11%, and 97.63%, respectively. For the detection of stable pixels, all spectral feature combinations can achieve higher user's accuracy, 90.84%, 94.63%, 95.00%, 97.03%, respectively. For the changed pixels, the user's accuracy is lower than the producer's accuracy, and the user's accuracy of the different spectral feature combinations are

69.00%, 76.20%, 81.79%, and 82.33%, respectively. For stable pixels, the producer's accuracy is lower than the user's accuracy, and the producer's accuracy of different spectral feature combinations is 69.00%, 76.20%, 81.79%, and 82.33%, respectively. In consequence, it can be concluded that more spectral features can provide higher user's accuracy for changed pixels and the producer's accuracy for stable pixels.

In general, the relative lower user's accuracy indicates more commission errors than omission errors in change detection. The significant causes of commission errors are: (1) spectral characteristics are too sensitive for LULC changes, and (2) the existing clouds occur three or more times consecutively.

Likewise, the relative lower producer's accuracy indicates more omission errors than commission errors in change detection. The omission errors mostly result from some reasons: (1) partially changed pixels and (2) mutual interference of multiple spectral features.

In general, in the case of change detection, when using fewer spectral features, the producer's accuracy of the changed pixel has higher accuracy, and the user's accuracy is lower, and it is easy to misjudge the stable pixel as the changed pixel, and the overall accuracy and Kappa hat coefficient of the change detection is low. When more spectral features are used, the producer's accuracy of the changed pixel also has higher accuracy, and the user's accuracy is higher, and the overall accuracy and Kappa hat coefficient of the change detection is high.

The overall accuracy of TOP3, TOP6, and ORI6 spectral features is relatively close. When considering the computational efficiency of change detection, the time used for TOP3 spectral features is about three times of NIR band, and the time used for TOP6 and ORI6 spectral features is about six times of NIR band. Therefore, it

can be seen from the two aspects of change detection accuracy and time efficiency that TOP3 (NIR, EVI, and TCB) is the best combination of spectral features for change detection.

Although the TOP6 and ORI6 spectral features take twice as long as TOP3 spectral features since the supervised classification algorithm is very fast compared to the unsupervised classification algorithm, then the difference in efficiency can be ignored, and the accuracy of TOP6 and ORI6 spectral features are higher than TOP3 spectral features. Therefore, TOP6 and ORI6 spectral features are suitable to apply for LULC classification in the actual use.

7.4.2 Accuracy assessment for LULC classification

The classified LULC maps in four different years (2000, 2006, 2011, and 2017) were further assessed their accuracies using 636 sample points based on the multinomial distribution with the desired precision of 5% and a level of confidence of 95% using the stratified random sampling technique. Therefore, pan-sharpened images (Figure 6.25) as primary reference data and very high resolution satellite images from Google Earth (Figure 6.26) as the second reference data were again used for accuracy assessment of LULC classification in this study.

Results of accuracy assessment of LULC classification with different spectral features in four different years, which include producer's accuracy, user's accuracy, overall accuracy, and Kappa hat coefficient are reported in Tables 7.10 to 7.13 and displayed in Figures 7.17 to 7.18. The derived overall accuracy was compared and described in two aspects: spectral features and years.

Table 7.10 Accuracy assessment of LULC classification in 2000.

NIR	U	A	F	W	Row Total
U	80	155	17	3	255
A	5	174	11	0	190
F	1	38	102	0	141
W	0	0	0	50	50
Column Total	86	367	130	53	636
Producer's accuracy (%)	93.02	47.41	78.46	94.34	
User's accuracy (%)	31.37	91.58	72.34	100.00	
Overall accuracy (%)			63.84		
Kappa hat (%)			49.88		
TOP3	U	A	F	W	Row Total
U	74	25	2	1	102
A	11	344	13	0	368
F	0	16	100	0	116
W	1	1	0	48	50
Column Total	86	386	115	49	636
Producer's accuracy (%)	86.05	89.12	86.96	97.96	
User's accuracy (%)	72.55	93.48	86.21	96.00	
Overall accuracy (%)			88.99		
Kappa hat (%)			81.29		
TOP6	U	A	F	W	Row Total
U	59	8	1	0	68
A	4	411	11	0	426
F	0	1	91	0	92
W	0	0	0	50	50
Column Total	63	420	103	50	636
Producer's accuracy (%)	93.65	97.86	88.35	100.00	
User's accuracy (%)	86.76	96.48	98.91	100.00	
Overall accuracy (%)			96.07		
Kappa hat (%)			92.40		
ORI6	U	A	F	W	Row Total
U	69	6	0	1	76
A	2	403	2	0	407
F	0	2	98	3	103
W	0	0	0	50	50
Column Total	71	411	100	54	636
Producer's accuracy (%)	97.18	98.05	98.00	92.59	
User's accuracy (%)	90.79	99.02	95.15	100.00	
Overall accuracy (%)			97.48		
Kappa hat (%)			95.35		

Table 7.11 Accuracy assessment of LULC Classification in 2006.

NIR	U	A	F	W	Row Total
U	102	97	17	4	220
A	14	204	25	0	243
F	4	48	71	0	123
W	0	0	0	50	50
Column Total	120	349	113	54	636
Producer's accuracy (%)	85.00	58.45	62.83	92.59	
User's accuracy (%)	46.36	83.95	57.72	100.00	
Overall accuracy (%)			67.14		
Kappa hat (%)			51.96		
TOP3	U	A	F	W	Row Total
U	113	23	2	8	146
A	9	267	6	1	283
F	0	74	82	1	157
W	1	0	0	49	50
Column Total	123	364	90	59	636
Producer's accuracy (%)	91.87	73.35	91.11	83.05	
User's accuracy (%)	77.40	94.35	52.23	98.00	
Overall accuracy (%)			80.35		
Kappa hat (%)			70.16		
TOP6	U	A	F	W	Row Total
U	99	8	0	3	110
A	3	324	8	0	335
F	0	52	89	0	141
W	0	0	0	50	50
Column Total	102	384	97	53	636
Producer's accuracy (%)	97.06	84.38	91.75	94.34	
User's accuracy (%)	90.00	96.72	63.12	100.00	
Overall accuracy (%)			88.36		
Kappa hat (%)			81.05		
ORI6	U	A	F	W	Row Total
U	109	3	2	0	114
A	1	319	4	4	328
F	0	45	96	3	144
W	0	0	0	50	50
Column Total	110	367	102	57	636
Producer's accuracy (%)	99.09	86.92	94.12	87.72	
User's accuracy (%)	95.61	97.26	66.67	100.00	
Overall accuracy (%)			90.25		
Kappa hat (%)			84.48		

Table 7.12 Accuracy assessment of LULC classification in 2011.

NIR	U	A	F	W	Row Total
U	112	88	19	5	224
A	20	202	18	0	240
F	27	17	78	0	122
W	2	1	0	47	50
Column Total	161	308	115	52	636
Producer's accuracy (%)	69.57	65.58	67.83	90.38	
User's accuracy (%)	50.00	84.17	63.93	94.00	
Overall accuracy (%)			69.03		
Kappa hat (%)			54.91		
TOP3	U	A	F	W	Row Total
U	139	30	2	1	172
A	7	274	12	0	293
F	0	22	99	0	121
W	0	0	0	50	50
Column Total	146	326	113	51	636
Producer's accuracy (%)	95.21	84.05	87.61	98.04	
User's accuracy (%)	80.81	93.52	81.82	100.00	
Overall accuracy (%)			88.36		
Kappa hat (%)			82.42		
TOP6	U	A	F	W	Row Total
U	140	9	0	1	150
A	6	288	22	0	316
F	0	27	93	0	120
W	0	1	0	49	50
Column Total	146	325	115	50	636
Producer's accuracy (%)	95.89	88.62	80.87	98.00	
User's accuracy (%)	93.33	91.14	77.50	98.00	
Overall accuracy (%)			89.62		
Kappa hat (%)			84.08		
ORI6	U	A	F	W	Row Total
U	137	6	1	1	145
A	13	287	8	2	310
F	0	17	106	8	131
W	0	0	0	50	50
Column Total	150	310	115	61	636
Producer's accuracy (%)	91.33	92.58	92.17	81.97	
User's accuracy (%)	94.48	92.58	80.92	100.00	
Overall accuracy (%)			91.19		
Kappa hat (%)			86.74		

Table 7.13 Accuracy assessment of LULC classification in 2017.

NIR	U	A	F	W	Row Total
U	139	70	8	12	229
A	50	163	20	1	234
F	4	17	102	0	123
W	0	2	0	48	50
Column Total	193	252	130	61	636
Producer's accuracy (%)	72.02	64.68	78.46	78.69	
User's accuracy (%)	60.70	69.66	82.93	96.00	
Overall accuracy (%)			71.07		
Kappa hat (%)			58.54		
TOP3	U	A	F	W	Row Total
U	173	50	3	9	235
A	27	181	25	0	233
F	0	8	109	1	118
W	0	0	0	50	50
Column Total	200	239	137	60	636
Producer's accuracy (%)	86.50	75.73	79.56	83.33	
User's accuracy (%)	73.62	77.68	92.37	100.00	
Overall accuracy (%)			80.66		
Kappa hat (%)			72.32		
TOP6	U	A	F	W	Row Total
U	163	18	0	7	188
A	23	245	10	3	281
F	1	6	109	0	116
W	0	1	0	50	51
Column Total	187	270	119	60	636
Producer's accuracy (%)	87.17	90.74	91.60	83.33	
User's accuracy (%)	86.70	87.19	93.97	98.04	
Overall accuracy (%)			89.15		
Kappa hat (%)			84.13		
ORI6	U	A	F	W	Row Total
U	142	31	0	4	177
A	31	228	7	6	272
F	1	4	130	2	137
W	0	1	0	49	50
Column Total	174	264	137	61	636
Producer's accuracy (%)	81.61	86.36	94.89	80.33	
User's accuracy (%)	80.23	83.82	94.89	98.00	
Overall accuracy (%)			86.32		
Kappa hat (%)			80.24		

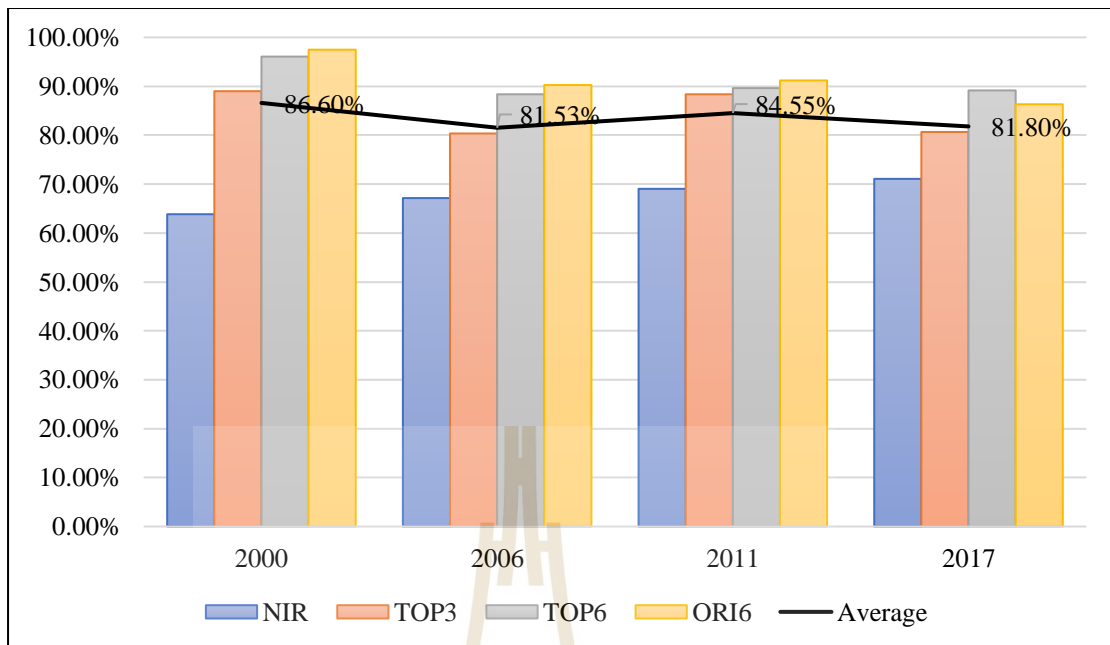


Figure 7.17 Comparison of the overall accuracy of four spectral features.

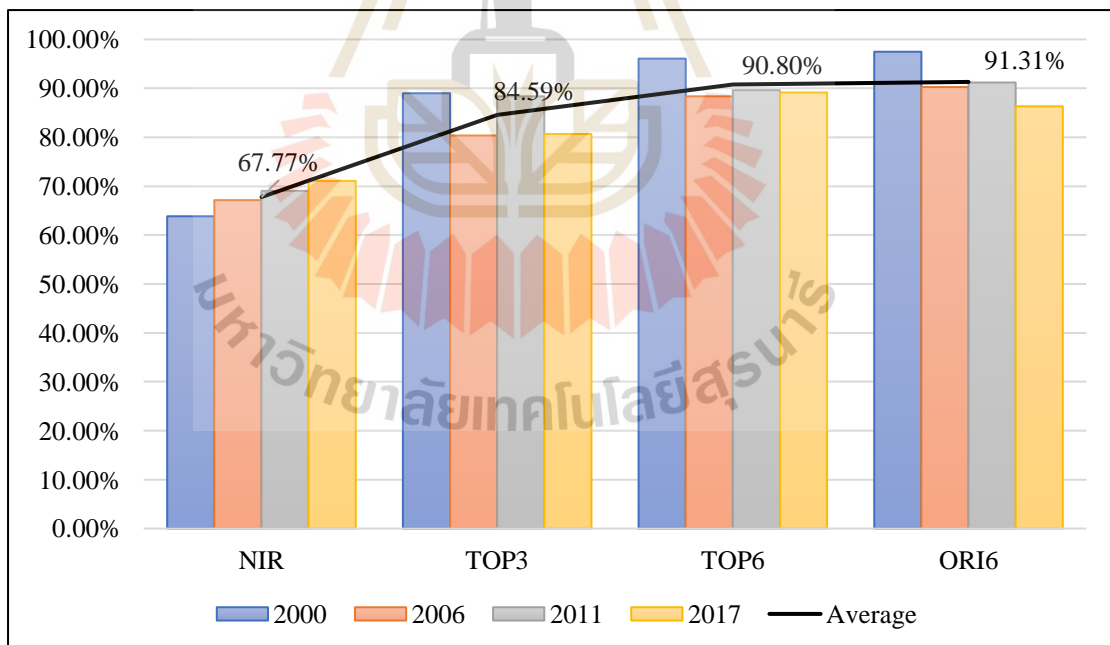


Figure 7.18 Comparison of the overall accuracy of four years.

As shown in Figure 7.17, an average overall accuracy of LULC classification from different spectral features (NIR, TOP3, TOP6, and ORI6) in 2000, 2006, 2011 and 2017 are approximately the same with the value of 86.60%, 81.53%, 84.55%, 81.80%, respectively. Regardless of the impact of the year, the result of LULC classification with TOP6 and ORI6 spectral features can provide very high accuracy, followed by TOP3 spectral features. In contrast, the result of LULC classification with the NIR spectral feature provides the lowest overall accuracy. So, it can be concluded that as the number of spectral features applied to LULC classification increases, the overall accuracy of the LULC classification result also increases.

Besides, an average overall accuracy of four different years from NIR, TOP3, TOP6 and ORI6 spectral feature in the Figure 7.18 is 67.77%, 84.59%, 90.80%, and 91.31%, respectively. As a result, it infers that when the number of spectral features increases, the overall accuracy increases.

Furthermore, when we apply the same spectral feature(s) for different years, the overall accuracies are roughly the same. For NIR spectral feature, overall accuracy slightly increases over time. One possible reason for this observation might be the rapid urban development in the study area by converting agricultural land to urban and built-up land, and some agricultural land without crops are misclassified to urban and built-up land, so the overall accuracy slowly increases.

For TOP3, TOP6, and ORI6, spectral features the overall accuracy is slightly reduced. The possible reason may be that in the early stage of urban development (2000), urban and built-up land is relatively concentrated, and most of the construction locates in the core area of the city. The pixels are mostly pure, and the LULC type is easy to be identified. However, in the later stage of urban development

(2017), other LULC types are converted into urban and built-up land, and the developed areas are more disperse, resulting in a large number of mixed pixels so, the LULC types are challenging to be identified.

Moreover, Table 7.14 reports the average producer's accuracy and user's accuracy of LULC classification maps with different spectral features, and Figure 7.19 shows the average producer's accuracy and user's accuracy based on Table 7.14.

Table 7.14 Average producer's accuracy and average user's accuracy using supervised classification.

Spectral feature(s)	Accuracy	The average value of each LULC type				Average of 4 LULC type
		U	A	F	W	
NIR	Producer's (%)	79.90	59.03	71.90	89.00	74.96
	User's (%)	47.11	82.34	69.23	97.50	74.04
TOP3	Producer's (%)	89.91	80.56	86.31	90.60	86.84
	User's (%)	76.10	89.76	78.16	98.50	85.63
TOP6	Producer's (%)	93.44	90.40	88.14	93.92	91.48
	User's (%)	89.20	92.88	83.38	99.01	91.12
ORI6	Producer's (%)	92.30	90.98	94.80	85.65	90.93
	User's (%)	90.28	93.17	84.41	99.50	91.84

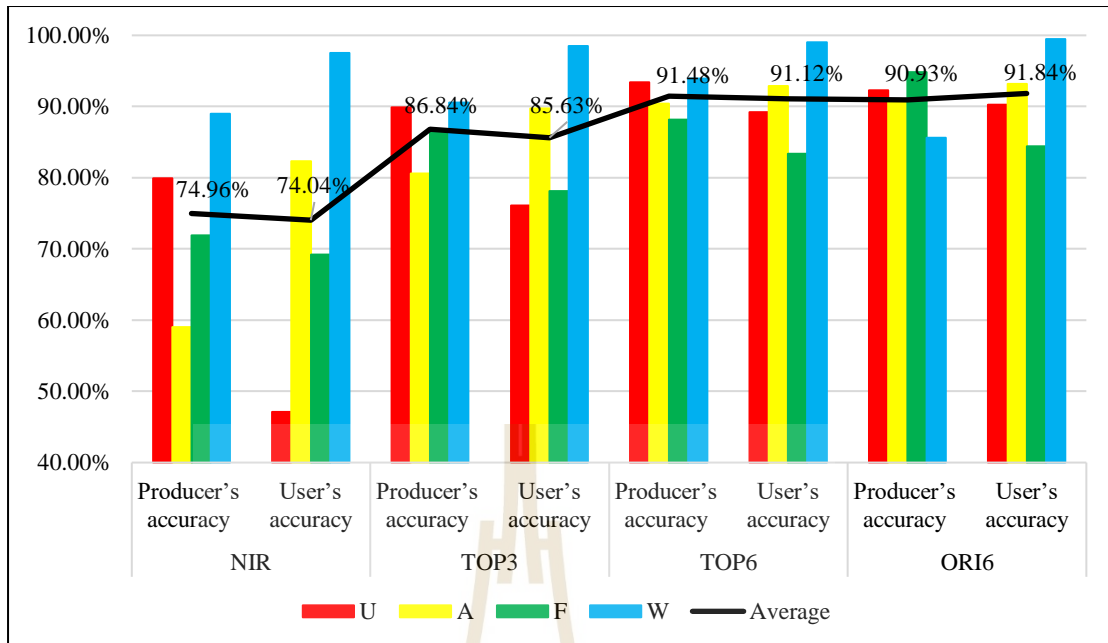


Figure 7.19 Producer's accuracy and user's accuracy.

As shown in Figure 7.19, it reveals that the producer's and user's accuracy of water bodies from different spectral features are very high (more than 85%) since standard reference harmonic wave from optimum spectral features are significantly different from others LULC curves (See Figure 7.3). Likewise, the producer's and user's accuracy for all LULC types from TOP6 and ORI6 spectral features are very high (greater than 83%), especially, user's accuracy of water bodies is close to 100%. In the meantime, for TOP3 spectral features, producer's accuracy varies from 80.56% for agricultural land to 90.60% for water bodies while the user's accuracy varies from 76.10% for urban and built-up land to 98.50% for water bodies. In contrast, for NIR spectral feature, the producer's accuracy varies from 59.03% for agricultural land to 89.00% for water bodies while the user's accuracy varies from 47.11% for urban and built-up land to 97.50% for water bodies. This finding infers that much

confusion arises when classifying the three LULC types (urban and built-up land, agricultural land, and forest land) with the NIR band.

Furthermore, the overall derived accuracy for multitemporal LULC classification with TOP6 and ORI6 in this study is comparable with that of other studies that applied time-series Landsat datasets to classify multitemporal LULC maps with a specific algorithm. For example, Gebhardt et al. (2014) applied 135 Landsat scenes to classify a series of seven maps of Mexico between 1993 and 2008 by using a decision tree algorithm, which provided an overall accuracy of about 76%. Zhu and Woodcock (2014b) applied their developed CCDC algorithm with random forests (RF) classifier to classify multitemporal land cover maps from time-series Landsat datasets (1982-2011) in coastal New England, United States, achieving an overall accuracy of about 90%. Gounaridis, Symeonakis, Choriantopoulos, and Koukoulas (2018) applied an RF classifier with time-series Landsat datasets (1991-2016) to classify LULC maps in Attica, Greece and attained an overall accuracy varying from 90.5% to 93.5%. Lu et al. (2018) applied Landsat images from 1990 to 2015 at five-year intervals to classify impervious surface and non-impervious surface areas using a linear spectral mixture analysis of the six selected metropolises in the coastal and inland metropolises in 2015, achieving an overall accuracy varying from 94% to 95%. Mi et al. (2019) applied an RF classifier with time-series Landsat datasets (1987-2017) to detect the LULC changes in a mining area, achieving an average overall accuracy of about 84%. Buitre, Zhang, and Lin (2019) applied a support vector machine with time-series Landsat datasets (1987-2016) to classify mangroves, non-mangroves, seawater, and clouds in the Philippines and attained an average overall accuracy of about 84%.

Similar to unsupervised classification, pairwise Z-test is conducted to examine the significant difference of accuracy from four spectral feature(s) combination based on Kappa analysis using Eq 6.2 and Eq 6.3 as results shown in Table 7.15.

As a result, it can be seen that the values of the pairwise Z-test between the NIR spectral feature and the other three spectral features (TOP3, TOP6, ORI6) are higher than 1.96, which indicates that the classification accuracy of NIR is significantly different from the others. Also, the values of pairwise Z-test between TOP3 and the other two spectral features (TOP6 and ORI6) are almost higher than 1.96. This finding shows that the classification accuracy of TOP3 is also significantly different from the other two spectral features. Conversely, all of the pairwise Z-test values between TOP6 and ORI6 spectral features are less than 1.96, indicating that the classification accuracy of TOP6 and ORI6 is insignificantly different.

Table 7.15 Pairwise Z-test between different spectral features.

	NIR_TOP3	NIR_TOP6	NIR_ORI6	TOP3_TOP6	TOP3_ORI6	TOP6_ORI6
2000	8.8971 *	13.3437 *	14.9653 *	4.2725 *	5.8159 *	1.5648
2006	4.9034 *	8.2716 *	9.5748 *	3.4145 *	4.6869 *	1.2201
2011	8.1973 *	8.7941 *	9.8944 *	0.6141	1.6748	1.0521
2017	3.9157 *	7.9089 *	6.5023 *	4.0413 *	2.6100 *	1.4404

Note * It is a significant difference of accuracy based on the Kappa hat coefficient.

SUMMARY

In this study, a new supervised classification method for multitemporal land use and land cover classification was successfully developed using harmonic analysis with a minimum spectral distance algorithm under the MATLAB environment through

systematic Landsat selection and time-series spectral reflectance reconstruction by converting the space domain to the time domain.

By comparing the overall accuracy of change detection according to spectral feature, LULC change detection using NIR provided the least accuracy (74.70%). In contrast, the accuracy of TOP3, TOP6, and ORI6 spectral features were 81.67%, 87.05%, and 88.25%, respectively. The combination of spectral features (TOP3, TOP6, and ORI6) provided similar and high accuracy. The overall accuracy of TOP3, TOP6, and ORI6 spectral features was relatively close. When considering the computational efficiency of change detection, the required processing time with TOP3 spectral features was about three times of NIR band, and the required processing time with TOP6 and ORI6 spectral features was about six times of NIR band. Therefore, it can be seen from the two aspects of change detection accuracy and time efficiency that TOP3 (NIR, EVI, and TCB) is the best combination of spectral features for change detection. Of course, although the TOP6 and ORI6 spectral features take twice as long as TOP3 spectral features since the supervised classification algorithm is very fast compared to the unsupervised classification algorithm, then the difference in efficiency can be ignored, and the accuracy of TOP6 and ORI6 spectral features are higher than TOP3 spectral features. Consequently, TOP6 and ORI6 spectral features are suitable to apply for LULC change detection in the actual use.

The average overall accuracy of LULC classification from different spectral features (NIR, TOP3, TOP6, and ORI6) in 2000, 2006, 2011, and 2017 were approximately the same with the value of 86.60%, 81.53%, 84.55%, 81.80%, respectively. Regardless of the impact of the year, the result of LULC classification with TOP6 and ORI6 spectral features provided very high accuracy, followed by TOP3

spectral features. In contrast, the result of LULC classification with the NIR spectral feature provided the lowest overall accuracy. So, it can be concluded that as the number of spectral features applied to LULC classification increases, the overall accuracy of the LULC classification result also increases. Besides, an average overall accuracy of four different years from NIR, TOP3, TOP6, and ORI6 spectral feature was 67.77%, 84.59%, 90.80%, and 91.31%, respectively. As a result, it infers that when the number of spectral features increases, the overall accuracy increases.

The values of the pairwise Z-test between the NIR spectral feature and the other three spectral features (TOP3, TOP6, ORI6) were higher than 1.96, which indicates that the classification accuracy of NIR is significantly different from the others. Also, the values of pairwise Z-test between TOP3 and the other two spectral features (TOP6 and ORI6) were almost higher than 1.96. This finding shows that the classification accuracy of TOP3 is also significantly different from the other two spectral features. Conversely, all of the pairwise Z-test values between TOP6 and ORI6 spectral features were less than 1.96, indicating that the classification accuracy of TOP6 and ORI6 is insignificantly different.

Therefore, it can be concluded that the newly developed supervised classification method using harmonic analysis with a minimum spectral distance algorithm can be efficiently used to classify and map multitemporal land use and land cover and to detect its changes from time-series Landsat datasets with high reliability for the information. However, this developed classification method should be examined in other areas to determine its spatial and temporal transferability. Nevertheless, the presented workflow of the research methodology can be used as a guideline for software developers for (semi) automatic land use and land cover classification and mapping.

Furthermore, it can be stated that the derived overall accuracy and Kappa hat coefficient of classified LULC maps in 2000, 2006, 2011, and 2017 based on the developed supervised classification method (harmonic analysis with a minimum spectral distance algorithm) are obviously higher than the derived thematic accuracies using a traditional unsupervised method (ISODATA algorithm) (Table 7.16). Subsequently, it can be concluded that the newly developed supervised classification method with multi-spectral features can be efficiently applied to classify multitemporal LULC maps from time-series Landsat datasets with high reliability for the information.

Table 7.16 Comparison of accuracy between unsupervised (ISODATA) and developed supervised classification methods.

Spectral features	Classification Method	Accuracy	LULC map in			
			2000	2006	2011	2017
NIR	ISODATA algorithm	OA	64.31	71.54	71.07	71.86
		Kappa	48.94	58.84	59.51	60.47
	Developed algorithm	OA	63.84	67.14	69.03	71.07
		Kappa	49.88	51.96	54.91	58.54
TOP3	ISODATA algorithm	OA	74.37	81.29	83.81	80.03
		Kappa	63.64	73.38	77.21	72.17
	Developed algorithm	OA	88.99	80.35	88.36	80.66
		Kappa	81.29	70.16	82.42	72.32
TOP6	ISODATA algorithm	OA	72.8	80.82	80.82	80.19
		Kappa	58.49	70.72	71.01	70.55
	Developed algorithm	OA	96.07	88.36	89.62	89.15
		Kappa	92.4	81.05	84.08	84.13
ORI6	ISODATA algorithm	OA	77.04	86.48	83.18	79.72
		Kappa	63.88	78.53	73.99	69.41
	Developed algorithm	OA	97.48	90.25	91.19	86.32
		Kappa	95.35	84.48	86.74	80.24

CHAPTER VIII

IMPACT STUDY OF LAND USE AND LAND COVER CHANGE ON LAND SURFACE TEMPERATURE

Land surface temperature (LST) is an essential parameter in the climate system. The dynamics of LST can furthermore be used as an indicator of climate variability and change (Arnfield, 2003; Bastiaanssen et al., 1998; Hansen et al., 2010; Kogan, 2001; Su, 2002; Voogt and Oke, 2003; Weng, 2009; Weng et al., 2004). LST is regularly measured from satellite sensors on-broad with moderately spatial scale and high temporal resolution, such as Landsat. In general, the derived LST data from thermal infrared (TIR) bands of a satellite is a crucial variable to understand the impacts of urbanization induced LULC change (Zhu and Woodcock, 2014b). A temporal analysis of thermal landscapes, therefore, requires the consideration of time-varying thermal characteristics. One way to avoid non-stationarity in modeling the temporal thermal landscape patterns is to divide time-series observations into individual segments that correspond to different land covers (Zhu and Woodcock, 2014b). As such, consistent time-series land surface temperature (TSLST) datasets are called for in revealing the urban thermal dynamics caused by land cover conversions (Weng et al., 2014). Nonetheless, at present, TSLST datasets at medium spatial resolution with regular temporal frequency are not available.

Under this chapter, the Landsat data and MODIS product were first used to extract LST for each independent scene using the SC algorithm. Later, a corrected and

uncorrected LST cubes were constructed to indicate which pixel's LST values need to be recalculated. Finally, the incorrect LST data were replaced by the newly calculated LST data by the HA model, while the corrected LST data remain, thereby a time-series LST dataset was retrieval.

Figure 8.1 shows the overview workflow and linkage of the impact study of LULC on LST, which consisted of four components. Details of the workflow of each component include input, process, and output are described separately in each section.

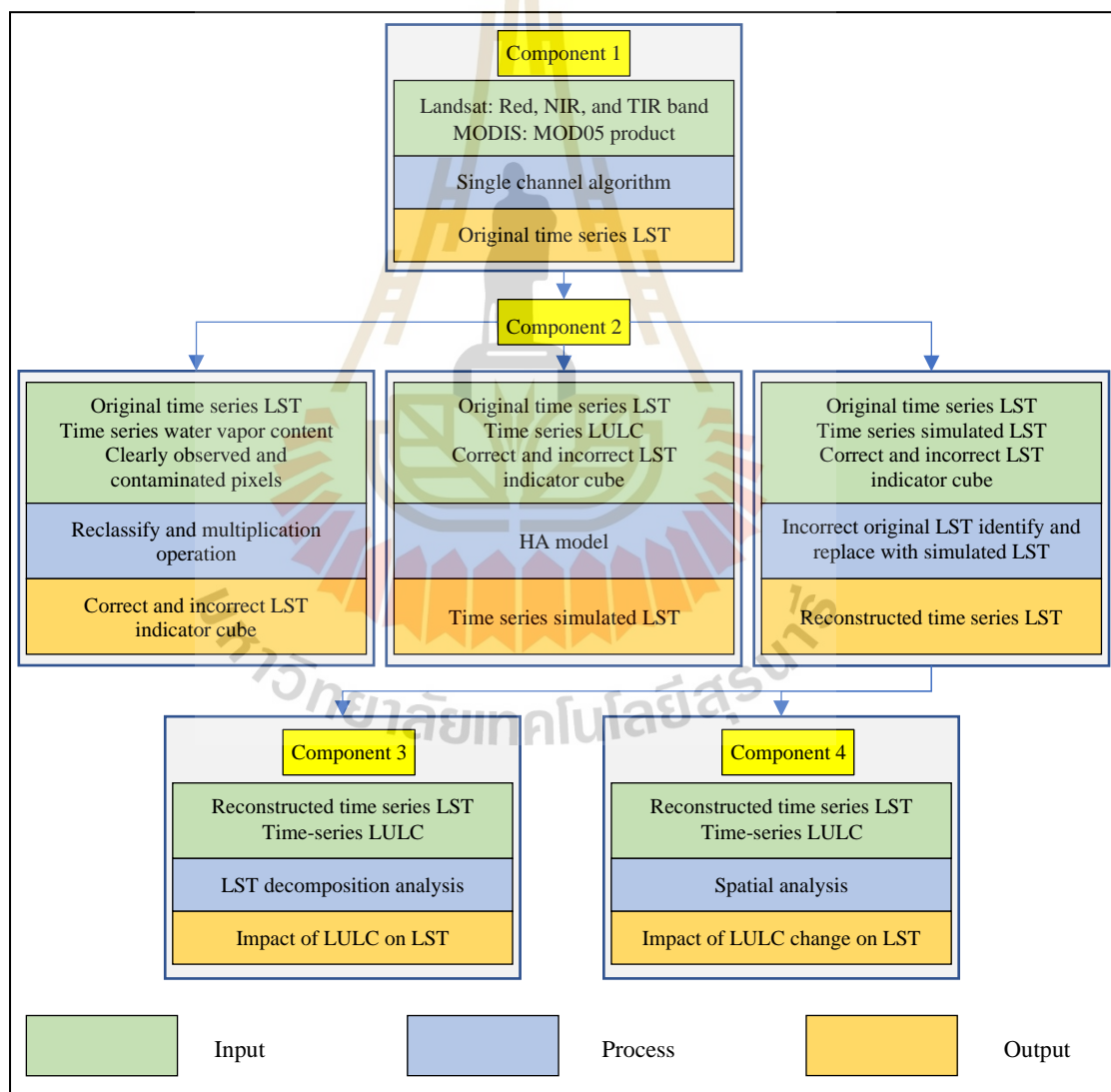


Figure 8.1 Workflow of Component 5: Time-series LST retrieves and time-series LST reconstruction.

8.1 Time-series LST retrieve using Single Channel algorithm

The LST is a highly variable quantity of the Earth's surface, in both space and time. Its temporal variability mainly results from the annual and daily cycles of solar irradiation, which are further influenced by cloud cover and general weather situations. Spatial variability is governed by surface characteristics like albedo, emissivity, soil moisture, the heat capacity of the surface soil layers, and topography (Bonan et al., 1992; Foley et al., 2005; Lee et al., 2015; Mahmood et al., 2014; McPherson, 2007; Pielke, 2005).

The USGS (2017) pointed out that band 11 of Landsat 8 provides the more considerable uncertainty data, users should work with band 10 data of Landsat 8 as a single spectral band (like TM of Landsat 5, ETM of Landsat 7) and should not attempt a split-window correction using both TIRs bands 10 and 11 of Landsat 8. Consequently, the SC algorithm was applied to retrieve LST in this study (Figure 8.2).

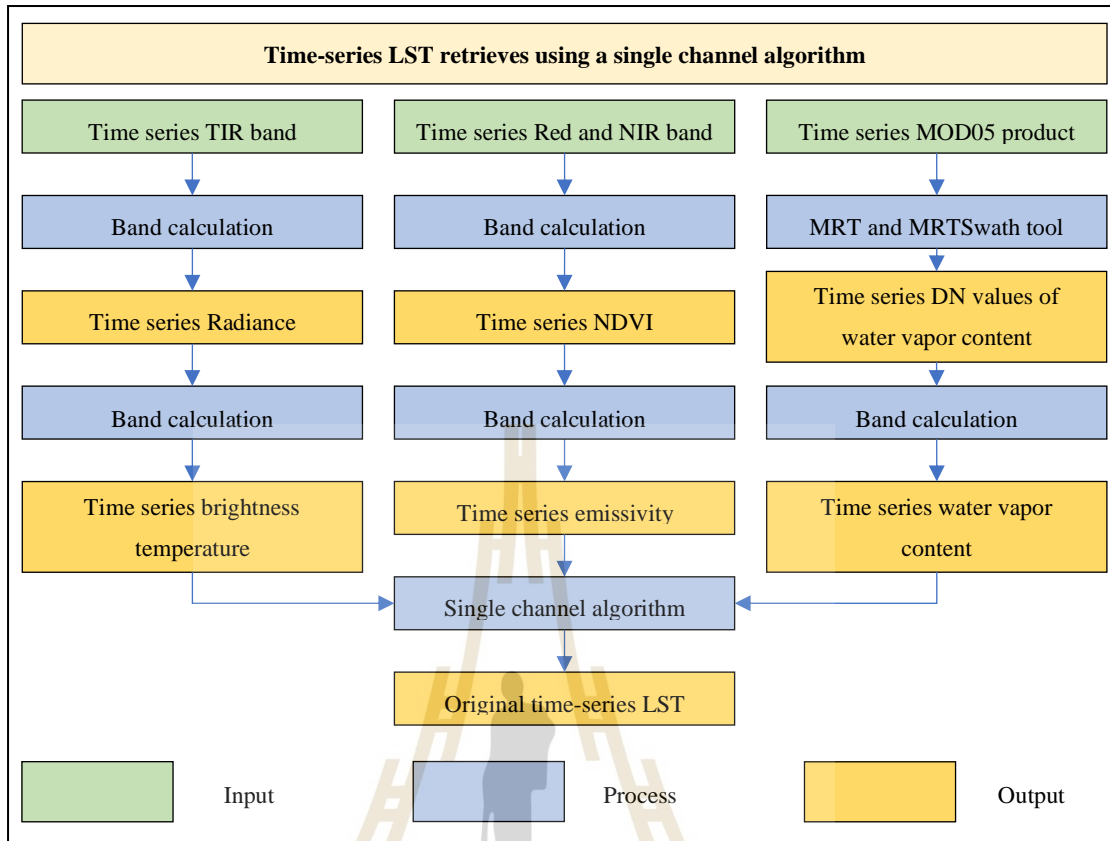


Figure 8.2 Workflow of time-series LST retrieves using a single channel algorithm.

The SC algorithm based on Jiménez-Muñoz and Sobrino (2003) was chosen for LST retrieve from Landsat series data using the following equation:

$$T_s = \gamma \left[\frac{1}{\varepsilon} (\varphi_1 L_{sen} + \varphi_2) + \varphi_3 \right] + \delta \quad (8.1)$$

where, T_s is land surface temperature, L_{sen} is the at-sensor radiance, ε is the surface emissivity, and (γ, δ) are two parameters given by:

$$\gamma \approx \frac{T_{sen}^2}{b_\gamma L_{sen}}, \quad \delta \approx T_{sen} - \frac{T_{sen}^2}{b_\gamma} \quad (8.2)$$

where, T_{sen} is the at-sensor Brightness Temperature (BT); $b_\gamma = c_2/\lambda$; and φ_1, φ_2 , and φ_3 are so-called atmospheric functions, given by:

$$\varphi_1 = 1/\tau; \quad \varphi_2 = -L_d - L_u/\tau; \quad \varphi_3 = L_d \quad (8.3)$$

where τ is the total atmospheric transmissivity, L_d is the down-welling radiance,

L_u is the up-welling atmospheric radiance.

The practical approach proposed in the SC algorithm consists of the approximation of the atmospheric functions defined in Eq. 8.4 versus the atmospheric water vapor content w from a second-order polynomial fit, expressed in matrix notation as follows ($\varphi = c w$):

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \begin{bmatrix} w^2 \\ w \\ 1 \end{bmatrix} \quad (8.4)$$

where coefficients c_{ij} are obtained by simulation.

Alternatively, if atmospheric parameters τ , L_u , and L_d are known, the atmospheric functions can be calculated from Eq. 8.3, thus avoiding the empirical relationship versus w .

The atmospheric functions φ_1 , φ_2 , and φ_3 for Landsat series data can be obtained as a function of the total atmospheric water vapor content (w) using the following equations.

For band 6 of Landsat 5 and Landsat 7 data:

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix} = \begin{bmatrix} 0.14714 & -0.15583 & 1.1234 \\ -1.1836 & -0.37607 & -0.52894 \\ 0.04554 & 1.8719 & -0.39071 \end{bmatrix} \begin{bmatrix} w^2 \\ w \\ 1 \end{bmatrix} \quad (8.5)$$

(Jiménez-Muñoz et al., 2009; Sobrino et al., 2004)

For band 10 of Landsat 8 data:

$$\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \end{bmatrix} = \begin{bmatrix} 0.04019 & 0.02916 & 1.01523 \\ -0.38333 & -1.50294 & 0.20324 \\ 0.00918 & 1.36072 & -0.27514 \end{bmatrix} \begin{bmatrix} w^2 \\ w \\ 1 \end{bmatrix} \quad (8.6)$$

(Jiménez-Muñoz et al., 2014).

The required three parameters, including brightness temperature, emissivity, and water vapor content, were calculated as follows.

8.1.1 Brightness temperature

The brightness temperature was calculated in two steps: (1) DNs to radiance, and (2) radiance to brightness temperature.

Step 1. DNs to Radiance

DNs of the thermal infrared band (band 6 of Landsat 5 and Landsat 7, and band 10 of Landsat 8) of all selected 388 Landsat images were processed in units of absolute radiance using 32-bit floating-point calculations. These values were then converted to 16-bit integer values in the finished Level-1 product. These values can then be converted to spectral radiance using the radiance scaling factors provided in the metadata file:

$$L_{\lambda} = M_{\lambda} * Q_{cal} + A_{\lambda} \quad (8.7)$$

where, L_{λ} is spectral radiance ($W/(m^2 \cdot sr \cdot \mu m)$), M_{λ} is radiance multiplicative scaling factor for the band from the metadata, A_{λ} is radiance additive scaling factor for the band from the metadata, Q_{cal} is the L1 pixel value in DN.

Step2. Radiance to brightness temperature

The TIRs data were converted from spectral radiance to brightness temperature, which is the effective temperature viewed by the satellite under an assumption of unity emissivity. The conversion formula is as follows:

$$T = K2 / \ln \left(\frac{K1}{L_{\lambda}} + 1 \right) \quad (8.8)$$

where T is brightness temperature in Kelvin (K), L_{λ} is spectral radiance ($W/(m^2 \cdot sr \cdot \mu m)$), $K1$, and $K2$ is thermal conversion constant for the band from the metadata. The derived outputs are further applied for LST retrieve.

An example of radiance and brightness temperature images from Landsat 5 and Landsat 7 is displayed in Figures 8.3 and 8.4, respectively.

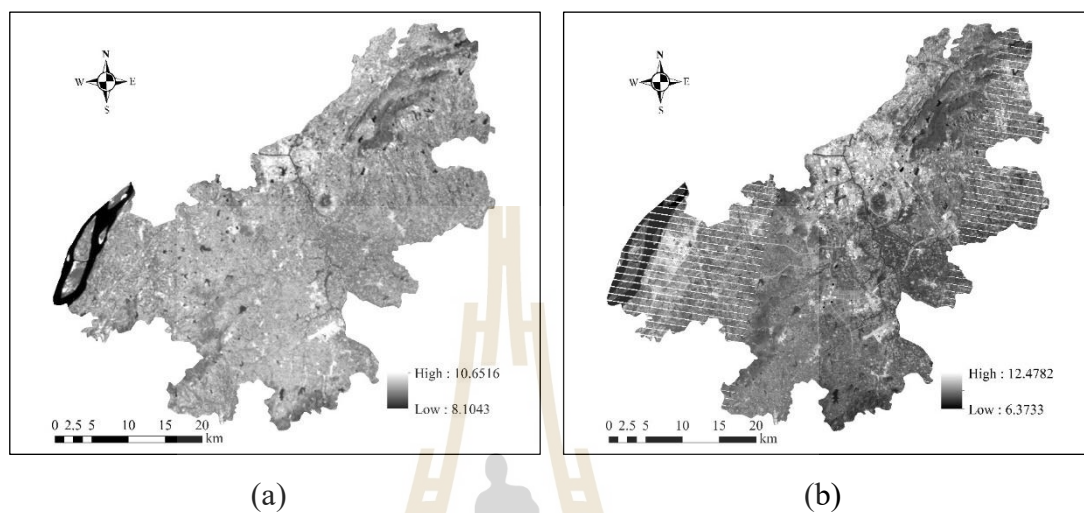


Figure 8.3 Spatial distribution of spectral radiance data: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.

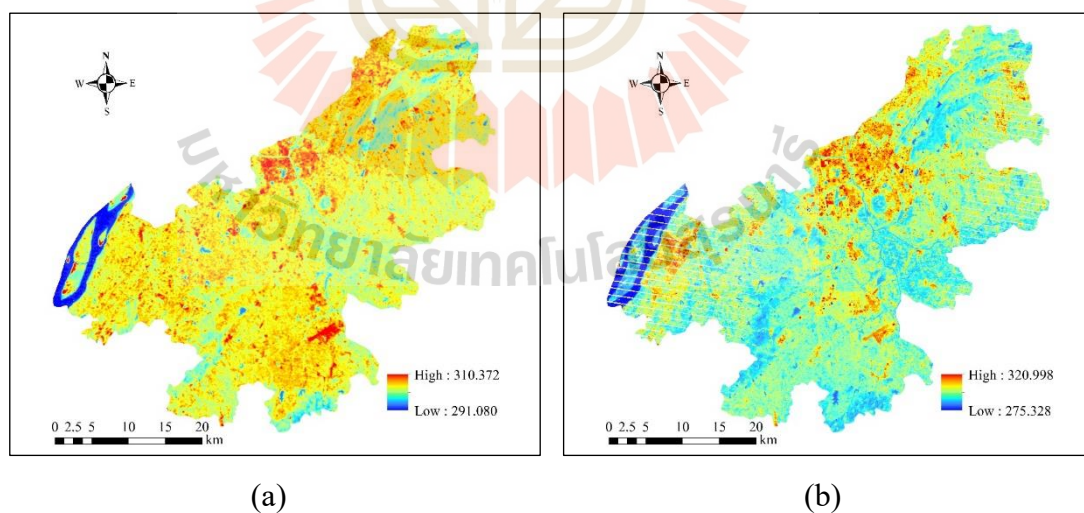


Figure 8.4 Spatial distribution of brightness temperature data: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.

8.1.2 Emissivity extraction

In principle, the emissivity of an object is mainly determined by its thermo-physical characteristics. For the ground surface, the components composing the surface are the main factors determining the ground emissivity. At present, many effective methods have been approved to estimate the emissivity for LST retrieval. Since the emissivity is variable with the wavelength, the normalized difference vegetation index (NDVI) threshold method (Sobrino et al., 2008) can be used to estimate the emissivity of different land surfaces in the 10-12 μm range. Additionally, the spectral range of band 6 of Landsat 5/7 and band 10 of Landsat 8 is suitable in this range. At this wavelength range, the emissivity could be modeled as follows:

$$\varepsilon_{\lambda} = \begin{cases} \varepsilon_w & NDVI \leq 0 \\ \varepsilon_{s\lambda} & 0 < NDVI < NDVI_s \\ \varepsilon_{v\lambda} \cdot P_v + \varepsilon_{s\lambda} \cdot (1 - P_v) + C_{\lambda} & NDVI_s \leq NDVI \leq NDVI_v \\ \varepsilon_{v\lambda} \cdot P_v + C_{\lambda} & NDVI > NDVI_v \end{cases} \quad (8.9)$$

subject to:

$$C_{\lambda} = (1 - \varepsilon_{s\lambda}) \cdot \varepsilon_{v\lambda} \cdot F' \cdot (1 - P_v) \quad (8.10)$$

where, ε_{λ} is the band emissivity, $\varepsilon_{v\lambda}$ and $\varepsilon_{s\lambda}$ are respectively the vegetation and soil emissivity, P_v is the proportion of vegetation, C is a term due to surface roughness ($C = 0$ for a flat surface), $NDVI_v$ and $NDVI_s$ are the $NDVI$ for a fully vegetated pixel and a soil one, respectively, and F' is a geometrical factor ranging between zero and one.

Generally, the vegetation cover fraction at the pixel scale can be computed from its $NDVI$ as follows (Sobrino et al., 2004):

$$P_v = \left(\frac{NDVI - NDVI_s}{NDVI_v - NDVI_s} \right)^2 \quad (8.11)$$

Over particular areas, $NDVI_v$ and $NDVI_s$ values can be extracted from the $NDVI$ histogram. Values of $NDVI_v = 0.5$ and $NDVI_s = 0.2$ were proposed to apply in global conditions (Sobrino et al., 2008).

Meanwhile, the TIR band 10 of Landsat 8 is different from band 6 of Landsat 5 and 7. It is worth noting that most satellite sensors measure the outgoing radiation with a finite spectral-bandwidth, and the channel-effective quantities of interest, therefore a weighted average are expressed by Li et al. (2013) as:

$$X = \frac{\int_{\lambda_1}^{\lambda_2} f(\lambda)X(\lambda)d\lambda}{\int_{\lambda_1}^{\lambda_2} f(\lambda)d\lambda} \quad (8.12)$$

where, X is a weighted average value of emissivity in the thermal infrared band, $X(\lambda)$ is various spectral quantity considered as emissivity, λ_1 and λ_2 are the lower and upper boundaries of the wavelength in TIR channel, and $f(\lambda)$ is the spectral response function.

An example of the $NDVI$ images is presented in Figures 8.5, while an example of emissivity data is shown in Figure 8.6.

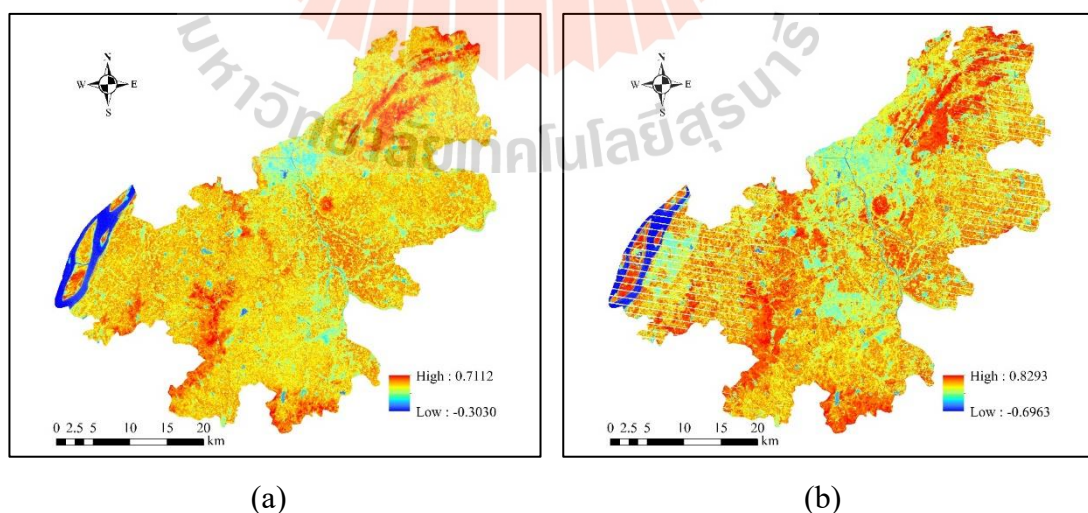


Figure 8.5 Spatial distribution of $NDVI$ data: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.

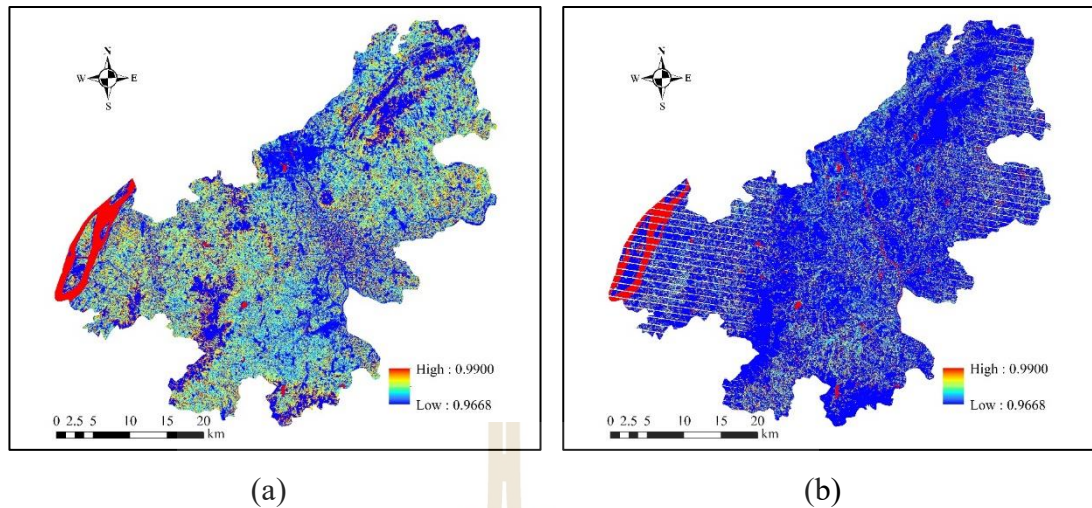


Figure 8.6 Spatial distribution of emissivity data: (a) 3 May 2000 and (b) 4 May 2009.

8.1.3 Water vapor content

Since Landsat itself does not have a band that can be used to detect water vapor content. So, in this study, the MOD05_L2 product on the same day as Landsat was used to calculate water vapor content. The MOD05_L2 product of MODIS/Terra on the same day with the final selection of Landsat (5, 7, and 8) data were first downloaded from LAADS DAAC (<https://ladsweb.modaps.eosdis.nasa.gov/>).

MODIS Reprojection Tool (MRT) and MRTSwath tool were used for reprojection and conversion of MODIS product to DN values of water vapor content.

To retrieve the real value of water vapor content, the DN values were divided by the scaling factor (Scaling factor = 1000).

$$w = \text{DN}/1000 \quad (8.13)$$

where w is water vapor content in vertical column water vapor amounts (cm), DN is the pixel value.

The MOD05_L2 of MODIS/Terra at the same with the final selection of Landsat data with a total of 388 images was downloaded reprojection and conversion. An Example of water vapor content is presented in Figures 8.7.

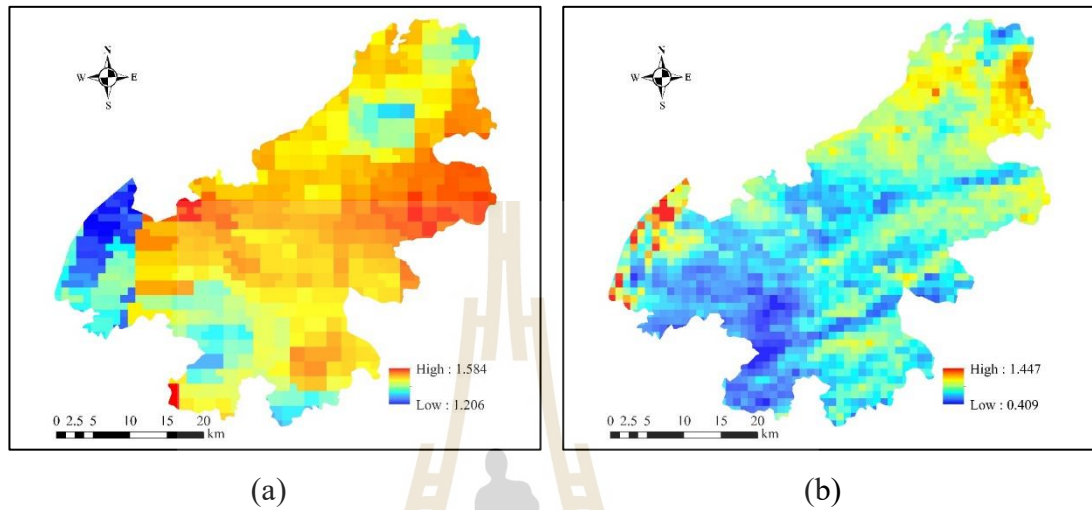


Figure 8.7 Spatial distribution of water vapor content from the MODIS product: (a) Date 3 May 2000 and (b) Date 4 May 2009.

Finally, the brightness temperature, the emissivity, and the water vapor content were used to invert the LST using Eq 8.1 as a result shown in Figures 8.8.

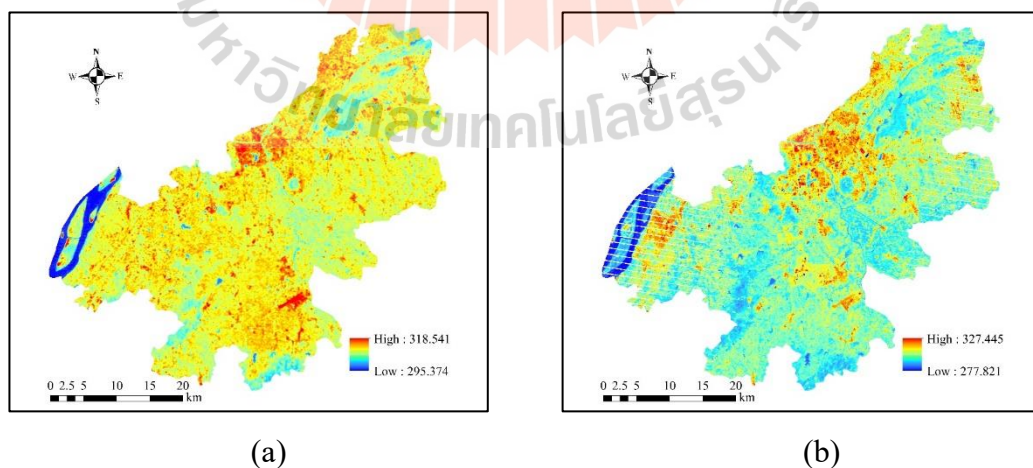


Figure 8.8 Spatial distribution of land surface temperature data: (a) Landsat 5, date 3 May 2000, and (b) Landsat 7, date 4 May 2009.

8.2 Time-series LST reconstruction

LST related research of the past decades often focused on the analysis of a single scene. Due to the complexity of land surface and radiative transmission, LST inversion is a very complicated process, and various factors (data, algorithms, processing, etc.) introduce errors into single scene LST (Weng et al., 2014).

So, the original time-series LST constructed from these single scene LSTs contains two kinds of values: correct and incorrect LST values. These incorrect LST values can be corrected based on correct LST values. Furthermore, the corrected incorrect LST values and correct LST values were used to reconstruct the new time-series LST.

In practice, new time-series LST reconstruction consisted of the essential four steps: (1) identification of correct and incorrect LST values from original time-series LST, (2) simulated LST recalculation using HA model, (3) time-series LST reconstruction combination from original and simulated time-series LST, and (4) LST reconstruction accuracy verification.

8.2.1 Construction of spatiotemporal cube for correct and incorrect LST value

The contaminated pixels (e.g., cloud, cloud shadow, Landsat 7 gap, etc.) of Landsat time-series data not only affect the classification of LULC but also lead to incorrect LST values. Therefore, the cube of the clearly observed and contaminated pixel, which calculated in the previous chapter, was used here to indicate the correct and incorrect LST values.

Radiative transfer equation (RTE) is the best option for LST retrieval from one thermal band since it does not involve additional approximations. However, this technique requires an accurate knowledge of atmospheric parameters such as transmissivity and atmospheric upwelling and downwelling radiances, which is not always possible. The SC algorithm has solved this problem by fitted atmospheric parameters versus more accessible parameters, but these relationships involve a strong approximation. Meanwhile, these relationships are not valid for a full range of water vapor content (w) values but only for low/moderate w values ($w < 3 \text{ cm}$). When w values lower than 3 cm are selected, the SC algorithm provides RMSEs less than 2 K, while w values higher than 3 cm are considered, the SC algorithm provides RMSEs higher than 5 K (Jiménez-Muñoz et al., 2009; Jiménez-Muñoz and Sobrino, 2003; Jiménez-Muñoz et al., 2014). The poor (incorrect) LST values retrieve by SC algorithm can identify base on w (w higher or lower than 3 cm) values obtain from the MOD05 product.

Besides the above two main reasons for produce incorrect LST, some other occasional errors (e.g., data processing errors) can also cause the LST values to be outliers (very high or very low). In this study, the LST outlier was detected base on the mean ± 3 standard deviation (SD). Any LST values higher than mean plus 3SD or less than mean minus 3SD are identified as an outlier (Shiffler, 1988).

So, a spatiotemporal cube of correct and incorrect LST values was constructed based on three types of error sources as a summary in the following sections.

(1) Spatiotemporal data error cube. The spatiotemporal data error cube of clearly observed and contaminated pixels, which was applied for unsupervised and supervised LULC classification, was directly used to indicate incorrect LST values

due to contaminated pixels in Landsat data. Herein, LST values that were calculated from clearly observed pixels are considered as corrected LST with a value of zero (0), while LST values that were calculated from contaminated pixels are considered as incorrect LST with a value of one (1). The value of spatiotemporal data error cube consists of 0 or 1.

(2) Spatiotemporal algorithm cube. The time-series water vapor content with a total of 388 scenes (2D array) was converted from space to time using reshape function of MATLAB as “spatiotemporal water vapor content cube”, and then this cube was reclassified using threshold value by water vapor content depth of 3 cm as “spatiotemporal algorithm cube”. Herein, LST values that were calculated from pixels with water vapor content less than or equal three centimeters are considered as correct LST with value of zero (0), while LST values that were calculated from pixels with water vapor content higher than three centimeters are considered as incorrect LST with value of one (1). The value of spatiotemporal algorithm cube consists of 0 or 1.

(3) Spatiotemporal LST outlier cube. Only the correct LST values of the original LST cube that indicates by data error and algorithm cubes were used to calculate the mean and standard deviation. Then, the combined cube was applied to reclassify and construct a “spatiotemporal LST outlier cube”. In this study, LST values within mean ± 3 SD are here considered as corrected LST with a value of zero (0); others are considered as an outlier with a value of one (1). The value of spatiotemporal LST outlier cube consists of 0 or 1.

Figure 8.9 shows a detailed flowchart of the construction of the spatiotemporal cube for correct and incorrect LST values in this study.

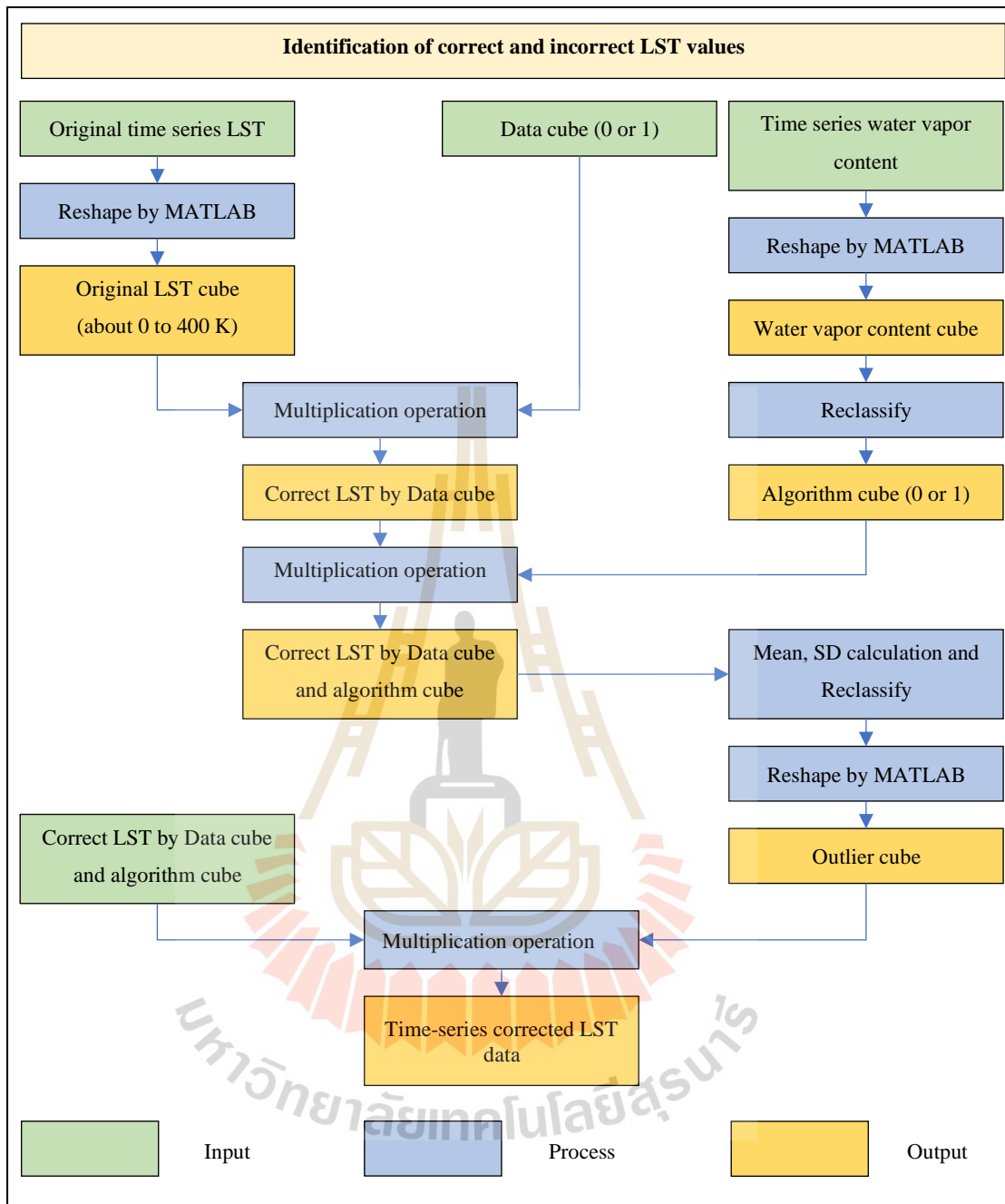


Figure 8.9 Workflow of Section 8.2.1: Identification of correct and incorrect LST values.

Figure 8.10 shows the histogram distribution of LST data from different sources.

Figure 8.10(a) displays the histogram distribution of original LST values after retrieval by SC algorithm, LST values of clearly observed pixels are around 300 K. On the contrary, the LST values of contaminated pixels are 0, these LST values (0 K) are identified as incorrect LST values and are ignored under data error cube.

Figure 8.10(b) shows the histogram distribution of LST values after correcting by spatiotemporal data error cube. Most of the LST values are around 300 K. However, some LST values are around 350 K, and they are inconsistent with the actual temperature of Nanjing. Later, these higher LST values, which are caused by high water vapor content under the SC algorithm, were corrected using spatiotemporal algorithm cube based on threshold setting of water vapor content depth, as mentioned earlier.

Figure 8.10(c) displays the histogram distribution of LST values after correcting by algorithm cubes. Most of the LST values are around 300 K, but there are still some outliers around 180 K and 350 K, these outliers are identified by spatiotemporal LST outlier cube. In this study, LST values that are higher than $294.12+3*13.3$ or less than $294.12-3*13.3$ are considered here as an outlier.

Figure 8.10(d) shows the distribution of LST values after correcting by LST outlier cubes. All of the remaining LST values are around 300 K. Basic statistic data of LST values in different stages of processing, as displayed in Figure 8.10, is summarized in Table 8.1.

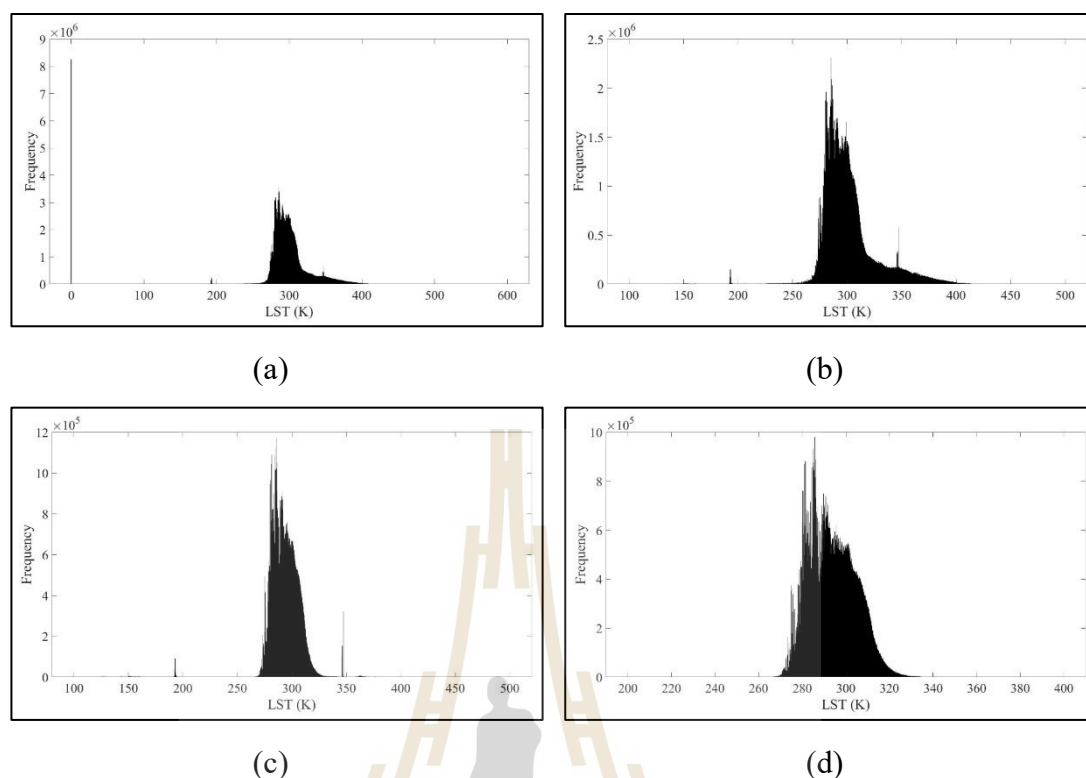


Figure 8.10 Histogram distribution of LST values in different stages of processing: (a) original LST values by SC algorithm, (b) after applying data error cube, (c) after applying algorithm cube, and (d) after applying LST outlier cube.

Table 8.1 The mean and standard deviation of LST data in different processing steps.

Processing stage	Mean	SD
Original LST values by SC algorithm (Figure 8.10(a))	302.0	25.6
LST values after applying data error cube (Figure 8.10(b))	303.6	24.8
LST values after applying algorithm cube (Figure 8.10(c))	294.1	13.3
LST values after applying the LST outlier cube (Figure 8.10(d))	294.2	10.7

In order to ensure the integrity of the original LST cube (1844 * 2000 * 388) for data analysis in MATLAB, no LST values were deleted from this cube, and only three spatiotemporal cubes (data error cube, algorithm cube, and LST outlier cube) were applied to construct time-series corrected LST data.

8.2.2 Simulated LST recalculation using the HA model

Refer to the previous section, incorrect LST values need to be identified and recalculated. Here, the original time-series LST cube (value in the cube is LST), spatiotemporal cube for correct and incorrect LST (multiplication of data error cube, algorithm cube, and LST outlier cube) with value of 0 or 1, and LULC cube (value in the cube is 1 to 4; U = 1, A = 2, F = 3, and W = 4) were used to recalculate simulated LST using HA model.

The original time-series LST of each location is divided into a LULC time-series homogeneous segment (period), in which the harmonic terms of LST were constant in each segment. Then the selected correct LST values according to spatiotemporal cube for correct and incorrect LST were used to fit the harmonic function curve (calculate harmonic parameters), and the simulated LST values can be recalculated at any time according to the harmonic parameters.

Figure 8.11 shows a detailed flowchart for simulated LST recalculation using the HA model. Brief information for time-series recalculation of essential five steps is described in the following sections.

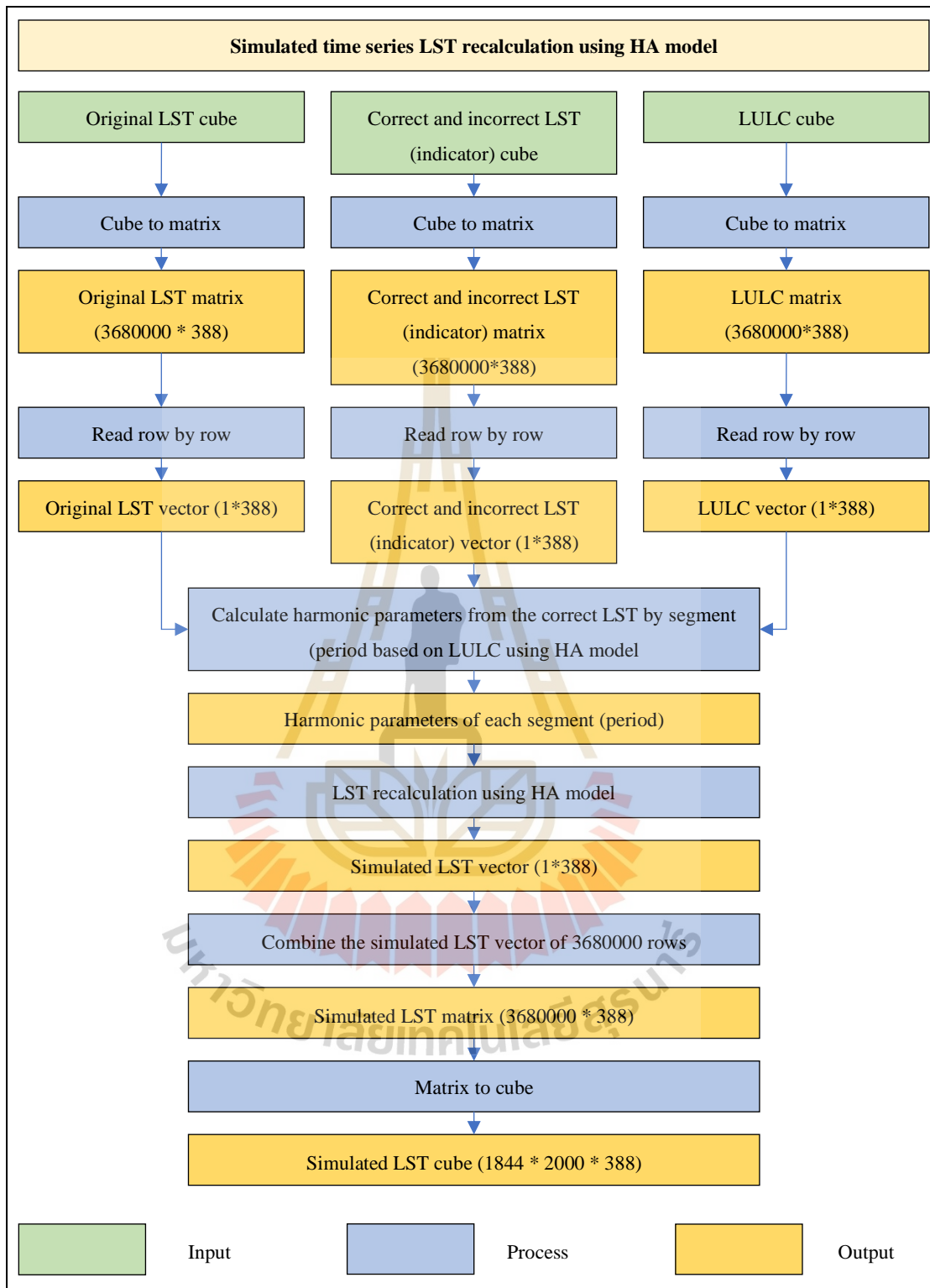


Figure 8.11 Workflow of Section 8.2.2: Simulated LST recalculation using the HA model.

Step 1. The original LST cube (3D array, $1844 * 2000 * 388$) was reshaped to a matrix (2D array, $3680000 * 388$) under MATLAB. Likewise, correct and incorrect (indicator) LST cube and LULC cube were also reshaped to a matrix. The values from these three matrices were read row by row as vector (1D array, $1*388$).

Step 2. The original LST vector was here divided into the LST segment based on the homogeneity of the LULC type in the LULC vector. Then the correct LST values were identified and selected from the correct and incorrect LST (indicator) vector.

Step 3. The selected correct segment of LST values was then used to fit the harmonic curve and calculate harmonic parameters (4 harmonic parameters per LST segment).

Step 4. The harmonic parameters were used to recalculate the simulated LST by the HA model.

Step 5. Steps 2 to 4 were here repeated 3,680,000 times. These results were recorded and saved as a matrix of $3680000 * 388$, which was then be converted into a simulated LST cube ($1844 * 2000 * 388$) using MATLAB.

Figure 8.12 shows the primary calculation process and the results of one row of data from the three matrices (original LST, correct and incorrect LST, and LULC vectors).

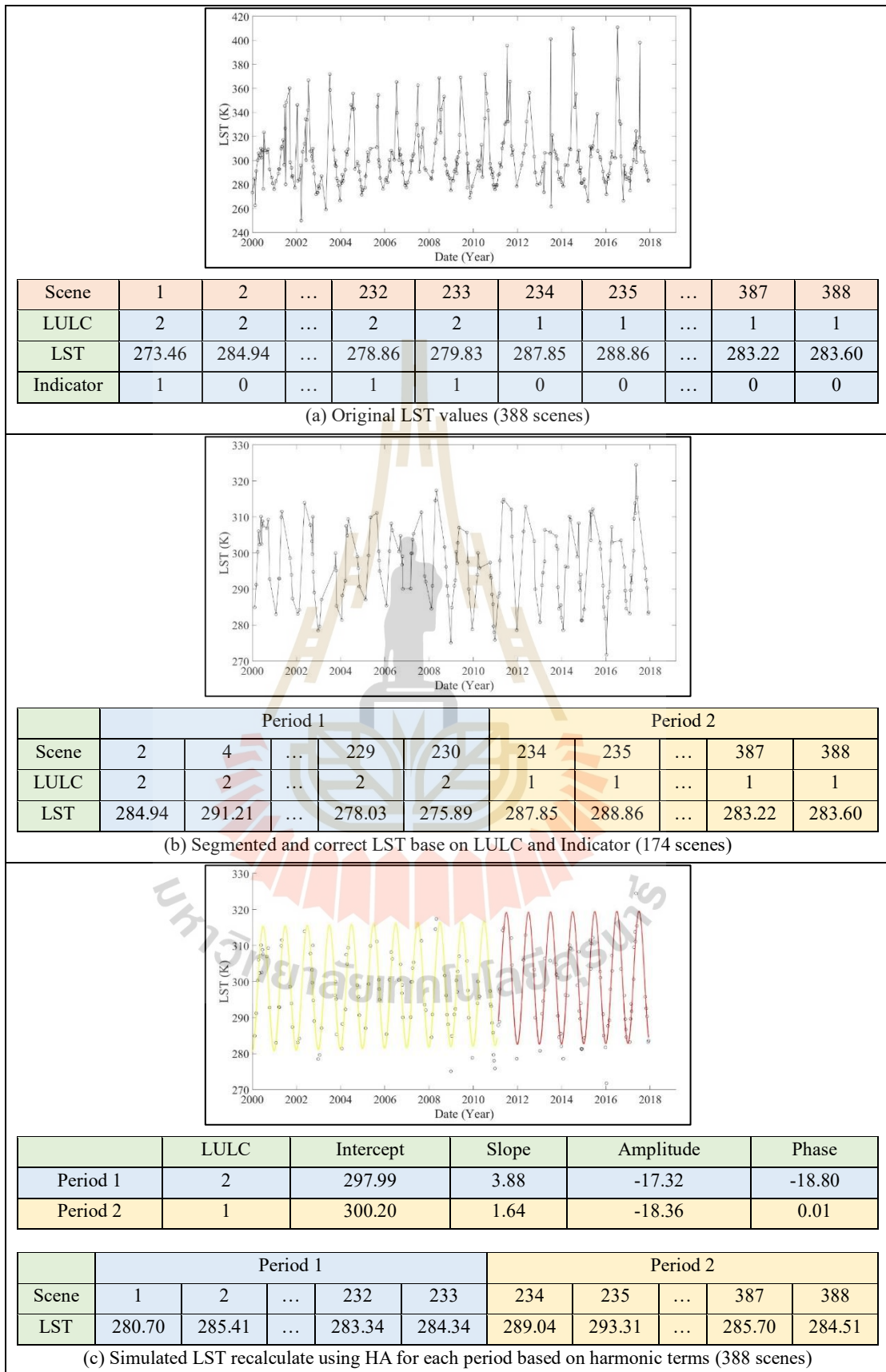


Figure 8.12 Main calculation process and results of one row of data.

Figure 8.12(a) shows input vector data of one pixel of 388 scenes read from the original LST matrix, correct and incorrect LST (indicator) matrix, and LULC matrix. As a result, the LULC vector indicates about spatiotemporal LULC change for the whole study period (388 dates) implies how LULC change, and when LULC take place. The LST vector shows a variation of LST values in one location for the whole study period (388 dates). The indicator of the correct and incorrect LST with a value of 0 (correct) and 1 (incorrect) shows which values of the LULC vector and LST vector were used for further data analysis. Meanwhile, the LST vector (388 scenes) are plotted and shown in the figure. It can be seen that most of the LST values oscillate above and below 300 K, but some of the values are very high or very low.

Figure 8.12(b) shows that the original LST vector is divided into two segments (periods) based on the LULC vector, and only correct LST values are selected base on correct and incorrect LST values indicator vector. Moreover, only 174 scenes remain, as shown in the figure. As a result, the LST value fluctuates 18 times with the season (corresponding to the Landsat data used from 2000 to 2017).

Figure 8.12(c) shows the harmonic function parameters of two periods, which represent the stability of the LULC type. Meanwhile, the simulated LST of 388 scenes, which were recalculated from parameters, is plotted as a harmonic function curve, as shown in the figure.

8.2.3 Time-series LST reconstruction

In this section, the original LST cube and simulated LST cube were combined according to the correct and incorrect LST indicator cube. The values of the reconstructed LST cube were recalculated using the equation as follow:

$$\text{Reconstructed LST} = \begin{cases} \text{Original LST}, & \text{LST indicator} = 0 \\ \text{Simulated LST}, & \text{LST indicator} = 1 \end{cases} \quad (8.14)$$

Figure 8.13 shows the detailed flowchart for the time-series LST reconstruction, while Figure 8.14 presents a schematic diagram of LST reconstruction based on the correct and incorrect LST indicator, the original LST, and the simulated LST with a 2D matrix.

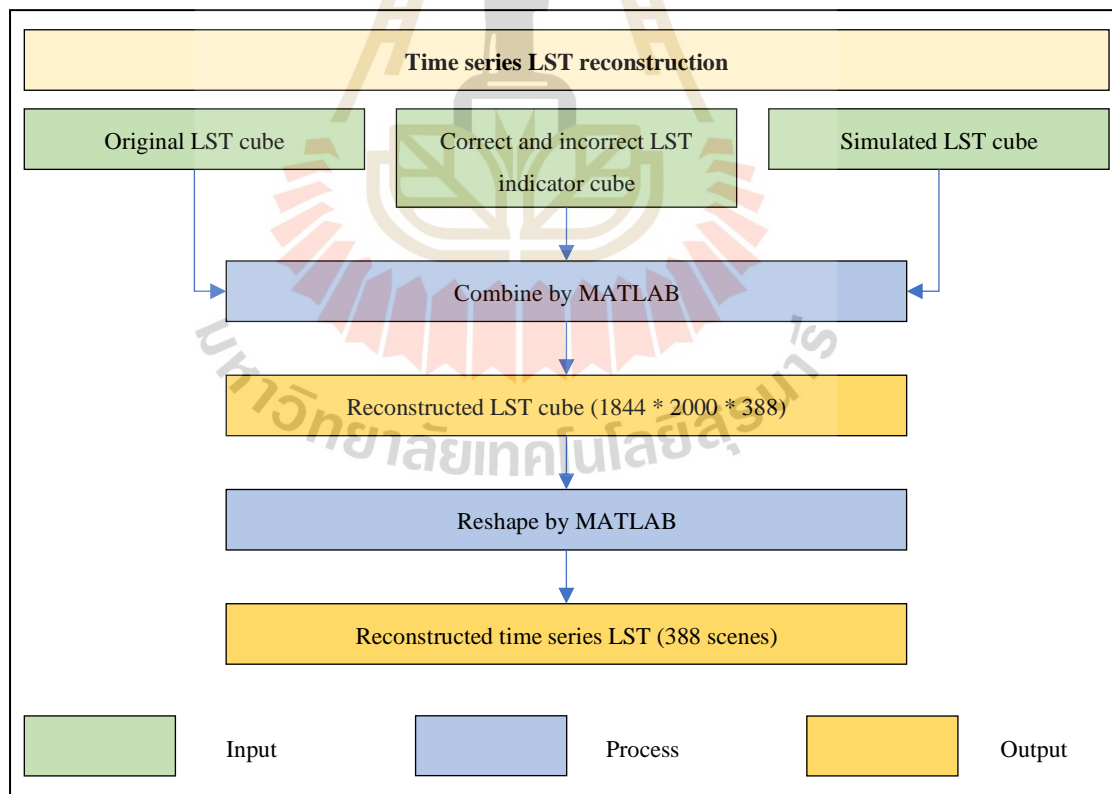


Figure 8.13 Workflow of Section 8.2.3: Time-series LST reconstruction.

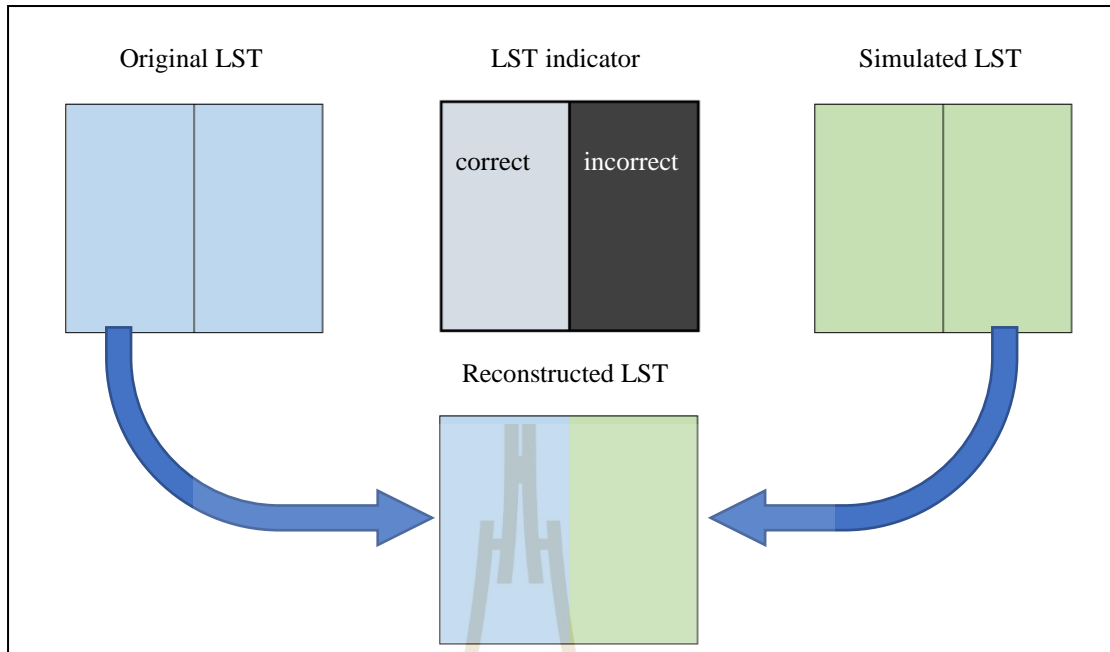


Figure 8.14 Schematic diagram of LST reconstruction based on the LST indicator, the original LST, and the simulated LST.

As shown in Figure 8.14, if the LST indicator shows the original LST values are correct (gray color in the left side of LST indicator matrix), the reconstructed LST values are taken from original LST (blue color in the left side of original LST matrix); while the LST indicator shows the original LST values are incorrect (black color in the right side of LST indicator matrix), the reconstructed LST values are taken from simulated LST (green color in the right side of simulated LST matrix). Based on this method, a new LST cube with a dimension of $1844 * 2000 * 388$ was first reconstructed and then reshaped from time to space to create a time-series LST dataset of 388 dates.

Figure 8.15 shows the results of reconstructed LST of one pixel of 388 LST data, as applied in Figure 8.12. This figure is here only used to illustrate the result of LST reconstruction for a single row (a vector) in the matrix of $3680000 * 388$.

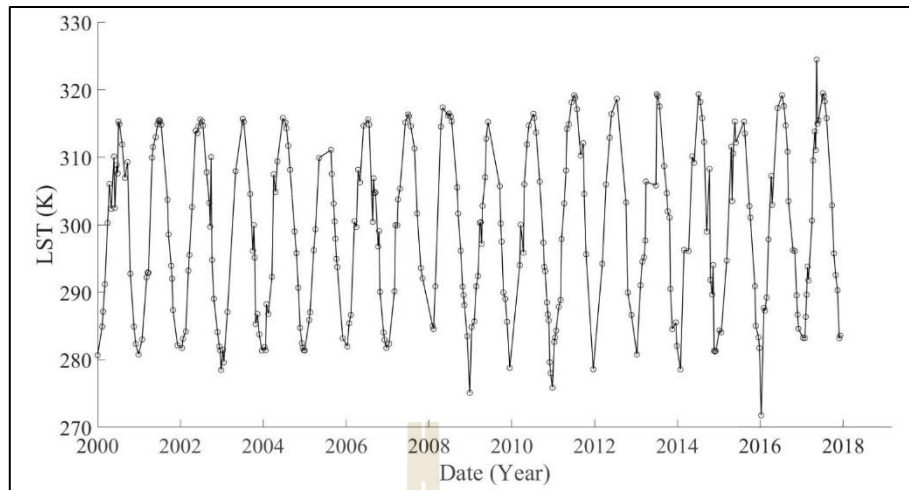


Figure 8.15 Reconstructed LST data.

In actual operation, the reconstruction of LST was directly based on these three cubes (original LST cube, simulated LST cube, and correct and incorrect LST indication cube). In this way, the time-series LST dataset was automatically reconstructed under MATLAB. Figure 16 shows one scene from the time-series LST dataset (388 scenes), date 30 October 2010, in different processing steps. Meanwhile, basic statistical data of LST from different processing steps are reported in Table 8.2.

Figure 16(a) shows one scene from the original LST data. The LST values range from Mean $- 3$ SD (288.0 K, 14.8 °C) to Mean $+ 3$ SD (301.6 K, 28.5 °C). On the left side of this figure (white stripe), some LST values are missing due to the Landsat 7 gap. At the upper right of this figure (blue patch), due to the influence of the atmosphere, the LST values are significantly lower than the surrounding area.

Figure 16(b) shows one scene from the simulated LST, the LST values range from Mean $- 3$ SD (290.4 K, 17.2 °C) to Mean $+ 3$ SD (299.1 K, 25.9 °C). The LST values in this figure are calculated from four harmonic terms (intercept, slope, amplitude, and phase).

Figure 16(c) shows one scene from the reconstructed LST. The LST values range from Mean $- 3$ SD (289.1 K, 15.9 °C) to Mean $+ 3$ SD (300.7 K, 27.6 °C). The LST values of the white stripe and blue patch are replaced with simulated LST.

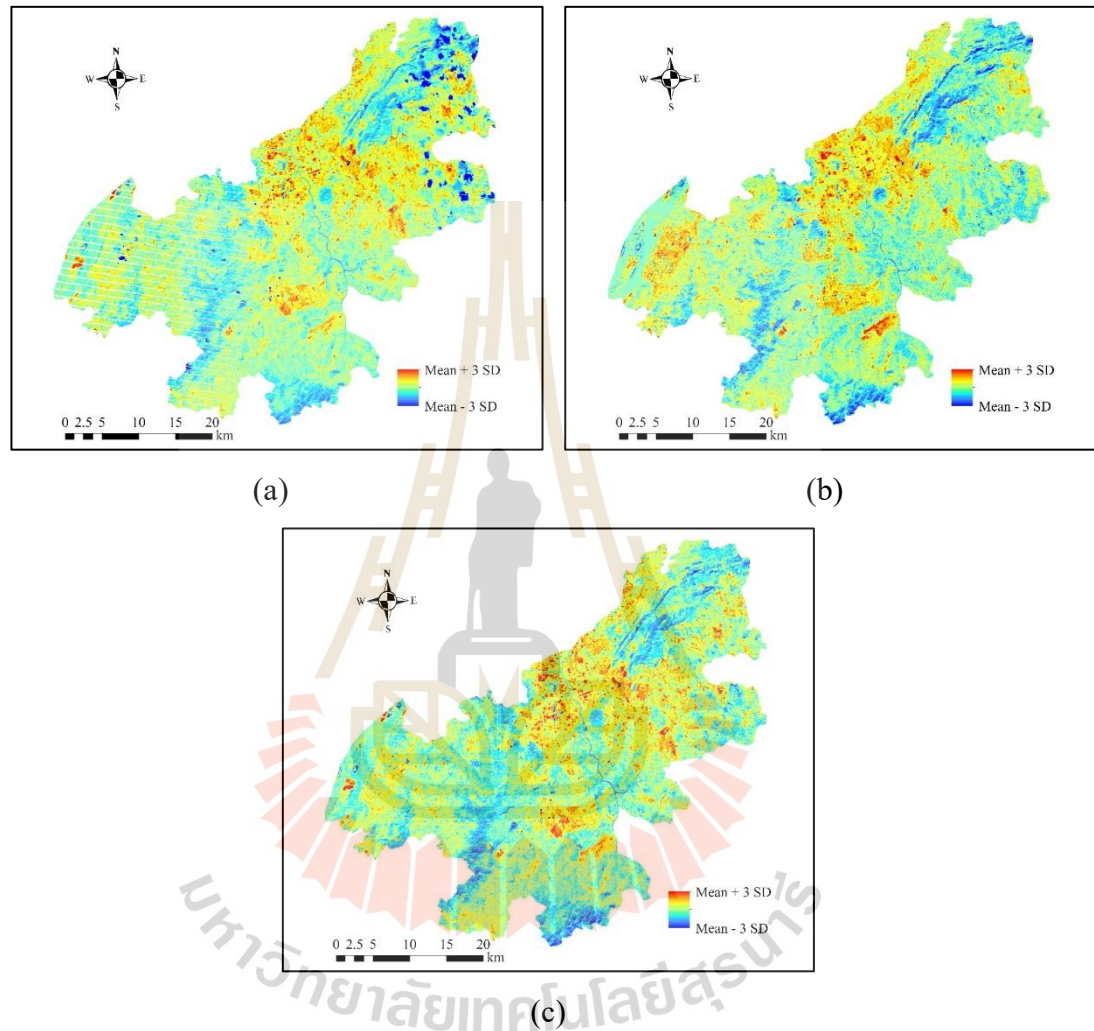


Figure 8.16 Reconstruction of LST data of Landsat 7, date 30 October 2010: (a) LST from original LST, (b) LST from simulated LST, and (c) LST from reconstructed LST.

Table 8.2 The mean and SD of LST data.

	Mean	SD	Mean - 3SD	Mean + 3SD
Original LST (K)	294.8	2.3	288.0	301.6
Simulated LST (K)	294.7	1.5	290.4	299.1
Reconstructed LST (K)	294.9	1.9	289.1	300.7

8.2.4 Verification of simulated LST data

The reconstructed LST value comes from two parts, one from the original real LST inversion value (when the indicator shows that the original LST value is correct), and the other comes from the simulated LST value based on HA recalculation (when the indicator shows that the original LST value is incorrect). In this study, the calculated LST values using the SC algorithm with depth of water vapor content less or equal three centimeters are acceptance as suggested by Jiménez-Muñoz et al. (2009). Meanwhile, the simulated LST values with the depth of water vapor content higher three centimeters or contaminated pixels or outliers are required to verify whether the calculated values are consistent with the actual ground temperature values. However, in actual operation, it is difficult to obtain the precise ground temperature of this part because of the atmospheric influence (cloud, water vapor) and other conditions, which makes it difficult to verify the accuracy of this simulated value directly.

In this study, mean error (ME) and mean absolute error (MAE), which frequently used to measure the difference between two continuous variables (Cort and Kenji, 2005; Nijbroek et al., 2018), was applied to verify the accuracy of simulated LST. In practice, the simulated LST data were compared with estimated LST using the SC algorithm. Figure 8.17 shows a schematic diagram of simulated LST data verification with a 2D matrix.

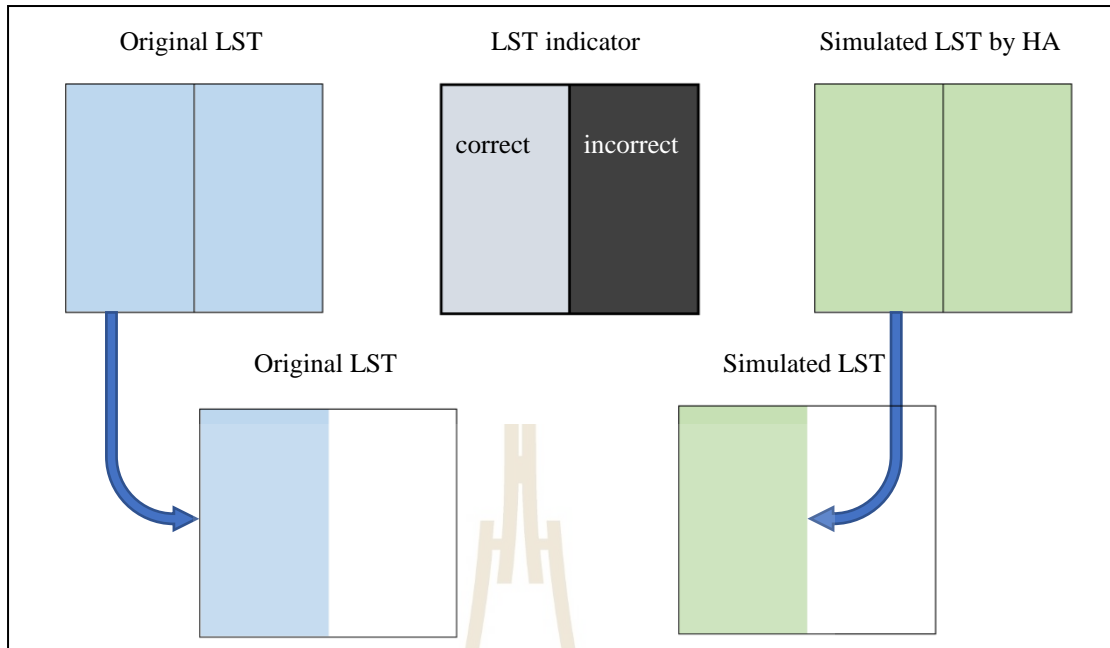


Figure 8.17 Schematic diagram of LST reconstruction accuracy verification.

As shown in Figure 8.17, the left box of the LST indicator indicates correct data from the original LST, while the right box LST indicator represents incorrect data from the original LST. In theory, it is necessary to verify the reliability of replacing the original LST values (blue color on the right box of the original LST) with incorrect values by the simulated LST values (green color on the right box of simulated LST by HA). However, the verification of the simulated LST values (right box of simulated LST by HA) that are replaced on the incorrect original LST (blue color on the right box of the original LST) is impossible, as mentioned earlier. Therefore, the simulated LST values on the left box of the simulated LST by HA were verified with the correct original LST on the left box by using the following equations (Cort and Kenji, 2005; Nijbroek et al., 2018).

$$ME = \frac{\sum_{i=1}^n (y_i - x_i)}{n} \quad (8.15)$$

$$AE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (8.16)$$

where x_i is the correct original LST values, y_i is the corresponded simulated LST values, and n is the total number of correct LST pixels of all 388 dates.

In this study, the mean error (ME) value between the simulated LST recalculated from the HA model, and the original LST is 0.015 K while mean absolute error which ignores positive and negative offset is 3.520 K. As a result, the replacement of simulated LST value on the incorrect original LST value can be accepted since ME value is less than 1 K as suggested by Jiménez-Muñoz et al. (2009).

8.3 LST decomposition analysis and impact of LULC on LST

In this section, the reconstructed LST was decomposed into three parts: trend component (intercept and slope), seasonality component (amplitude and phase), and residual component according to LULC type of each pixel.

Firstly, the harmonic parameters (intercept, slope, amplitude, and phase) were reclassified and grouped into four groups accord to LULC types (U, A, F, and W).

Secondly, the harmonic parameters were also reclassified into 18 subgroups of the previous four LULC groups according to the persistent years (1 year to 18 years).

Finally, the harmonic parameters that relate to the conversion of agricultural land, forest land and water bodies into urban and built land were selected and categorized into three groups (from A to U, from F to U, and from W to U).

Figure 8.18 shows the detailed flowchart of LST decomposition analysis and the impact of LULC on LST. Brief information on LST decomposition and impact of LULC on LST is described and discussed in the following sections.

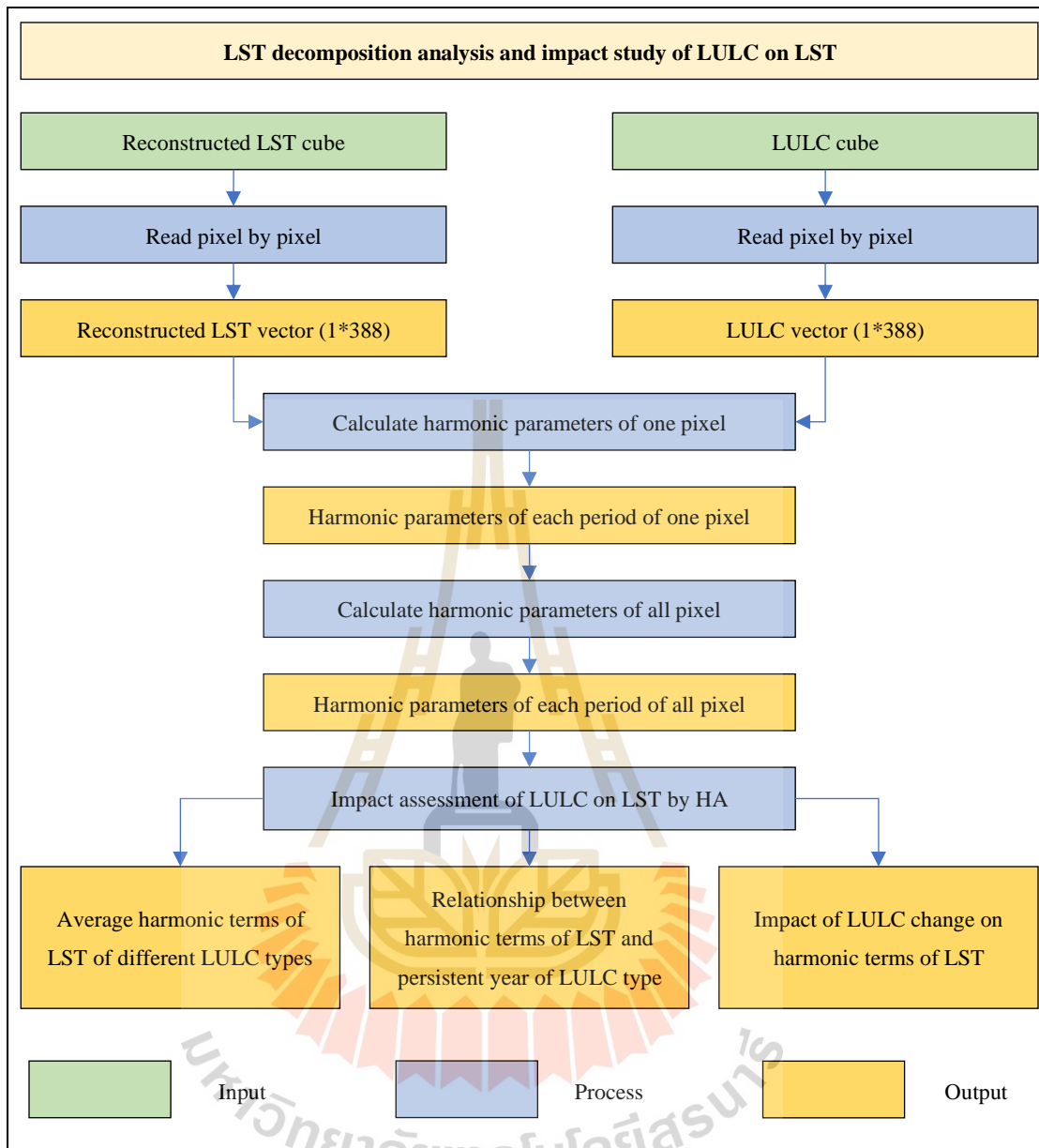


Figure 8.18 Workflow of Section 8.3: LST decomposition analysis and impact study of LULC on LST.

8.3.1 LST decomposition

The occurrence of LULC changes may induce non-stationarity in a TSLST dataset since surface thermal responses are different before and after land-cover conversions (Weng et al., 2014). To analyze the time-varying surface thermal characteristics, the time-series LST were first divided into temporally homogeneous segments (periods) in which the trend component value and the seasonality component value were constant. Then, a decomposition scheme of a time-series additive model comprising the seasonality component and the trend component (Eq. 8.17) was applied to both stable and unstable pixels. The LOcally wEighted regrESSion Smoother (LOESS) scheme was utilized for decomposition due to its ability to reduce the outlier effects (Cleveland, Cleveland, McRae, and Terpenning, 1990). The equations that were modified from (Fu and Weng, 2016) were applied to calculate these components as follows:

$$Y_t = T_t + S_t + \varepsilon \quad (8.17)$$

$$T_t = a + bt \quad (8.18)$$

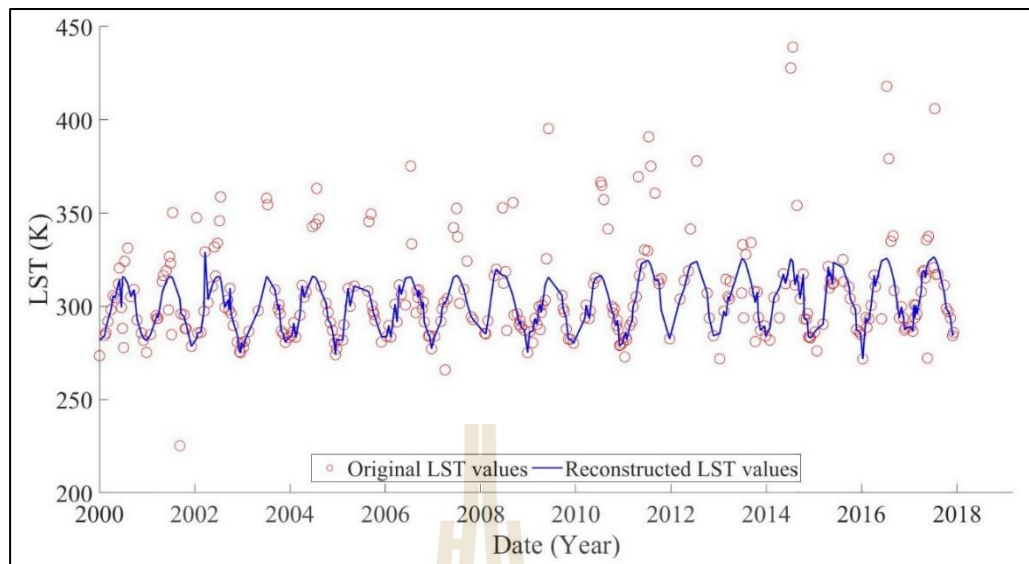
$$S_t = A \cos\left(\frac{2\pi}{T}t - \varphi\right) \quad (8.19)$$

where, Y_t is the time-series observations, T_t is the trend component, S_t is the seasonality component, and ε is the noise. a and b are the coefficients for the fitted linear trend, t is the Julian day, A and φ are the coefficient and periodic frequency for the harmonic item, $T=365$.

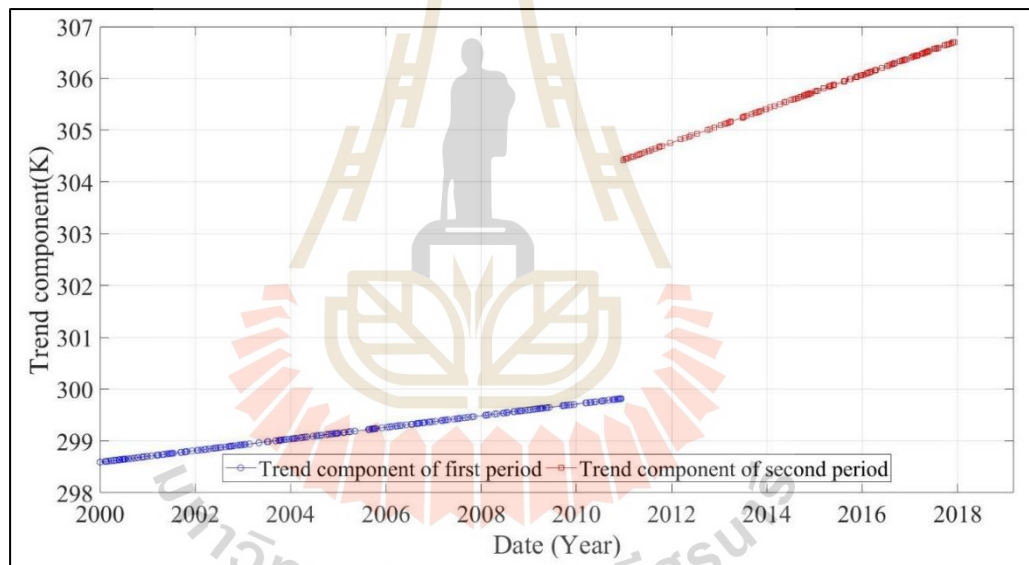
Meanwhile, a linear change was assumed to characterize the gradual change over the years. Eventually, the number of the harmonic terms (variable N) was adopted due to its ability to deal with the irregular spaced remote sensing time-series and to characterize the complex periodic patterns (Verbesselt, Hyndman, Newnham, et

al., 2010; Verbesselt, Hyndman, Zeileis, and Culvenor, 2010). The decomposition procedure was here applied to each pixel in the study area.

Figure 8.19 shows a result of the LST decomposition procedure using time-series LST at one point where land cover changed from agricultural land to urban and built-up land. Figure 8.19(a) compares a result of the original LST and reconstructive LST. Meanwhile, Figure 8.19(b-d) shows the trend component, seasonality component, and modeling residues (differences between reconstructed LST and trend component + seasonality component), respectively. As a result, the absolute mean residue for this point from all 388 scenes is only 2.4 K. The lower residue means that the error is smaller when reconstructed LST is only represented by the trend component (intercept and slope) and seasonality component (amplitude and phase). In other words, the error of representing reconstructed LST values on 388 dates of this pixel by eight harmonic terms (4 harmonic terms per period) is only 2.4K. In this study, it is possible to analyze the effect of LULC on LST (all pixels of all dates) with a small error by using the trend and seasonal components comprehensively instead of the reconstructed LST of all the pixels.

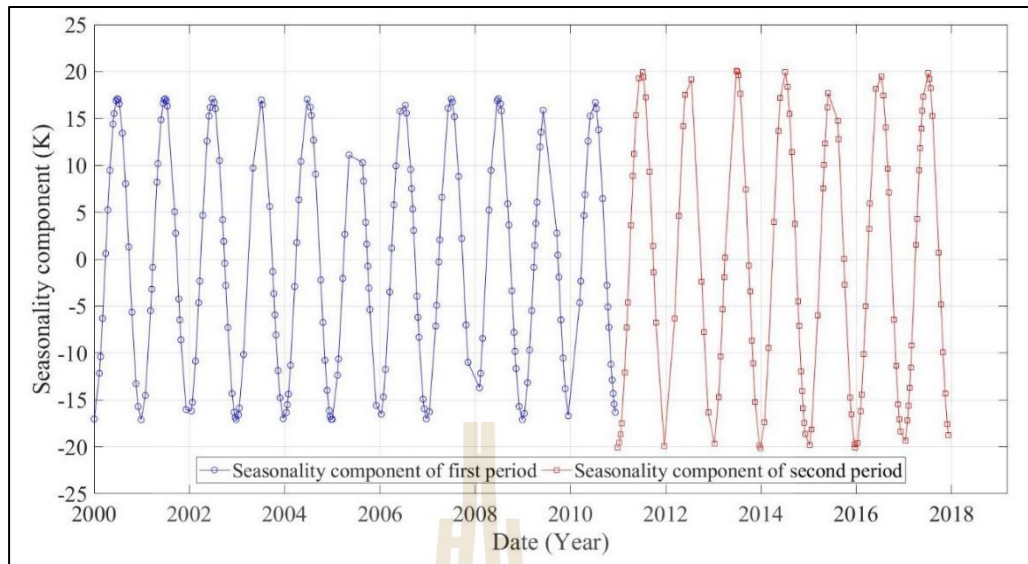


(a)

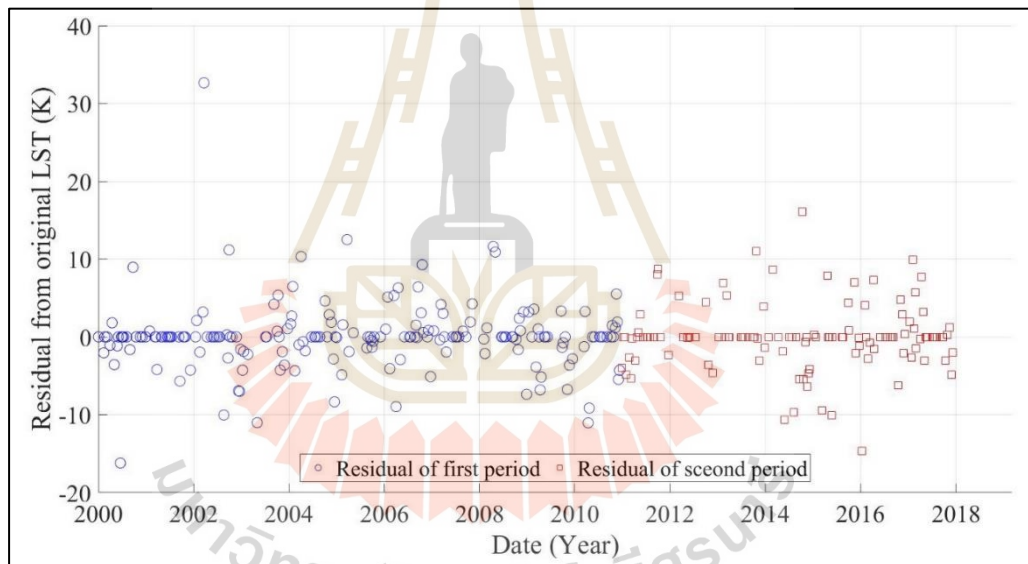


(b)

Figure 8.19 LST decomposition procedure using time-series LST at one point: (a) Original and reconstructed LST, (b) Trend component, (c) Seasonality component, and (d) Residue.



(c)



(d)

Figure 8.19 (Continued).

Table 8.3 shows an example of a final result of the harmonic parameters of each pixel of each period.

Table 8.3 Final result of harmonic parameters of some pixels from all periods.

Pixel No.	Period (Julian day)	Harmonic terms of LST				LULC type
		Intercept	Slope(E-4)	Amplitude	Phase	
(1, 2000)	1-6545	294.89	-0.99	-13.76	-0.21	4
...
(24,1211)	1-3976	299.26	-1.05	-17.05	6.29	2
	3977-6545	293.72	5.45	-14.25	0.02	3
...
(1844, 2000)	1-6545	298.95	-2.09	-17.12	-17.69	2

8.3.2 Impact of LULC on LST

The impact of LULC on LST was assessed based on the derived harmonic term of LST from 18 years. In this study, three aspects include (1) average value of harmonic terms of LST of different LULC types; (2) relationship between harmonic terms of LST and persistent year of LULC type; and (3) impact of LULC change on harmonic terms of LST were here investigated. Brief information on the method and results are separately describe and discussed in the following sections.

(1) Average value of harmonic terms of LST of different LULC types

In this study, the value of the harmonic term (intercept, slope, amplitude, and phase) of LST of each LULC type (U, A, F, W) were first counted based on the classified LULC data from 388 dates. Then, the number of each harmonic term of each LULC type were plotted as histogram distribution for comparison of each harmonic term value. Since the sample size and bin width of the histogram from different LULC types are different, it is difficult to compare them. Herein, standard histogram and

normalized histogram were presented. Besides, the average value of the harmonic term (intercept, slope, amplitude, and phase) of the LST of each LULC type (U, A, F, W) was also extracted and compared to their values.

The result of standard and normalized histogram distribution of intercept, slope, amplitude, and phase of LST for four different LULC types is displayed in Figure 8.20. Meanwhile, the average value of four harmonic terms of LST of four LULC types is presented in Figure 8.21.



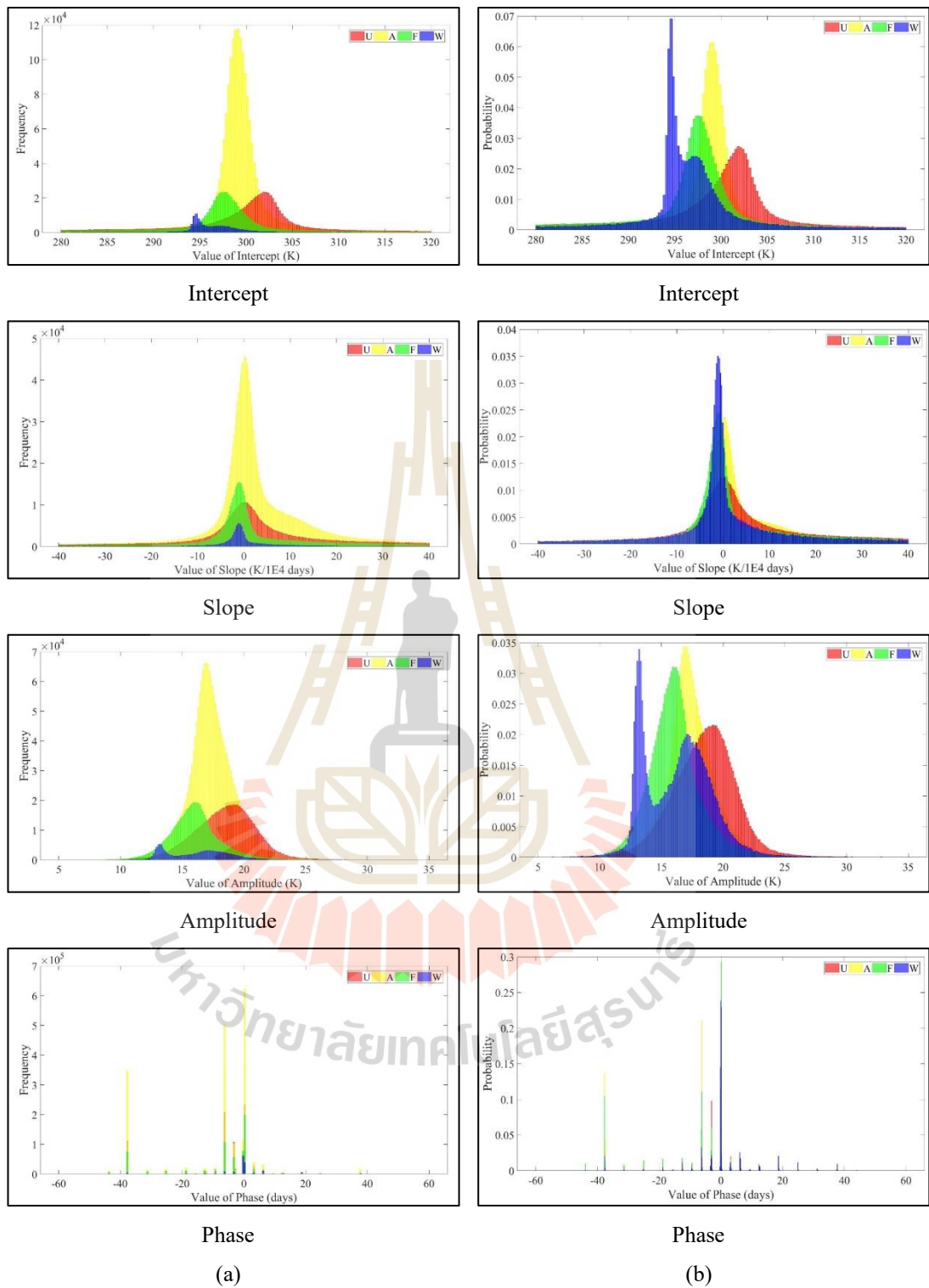


Figure 8.20 Distribution of frequency and probability of intercept, slope, amplitude, phase for four different LULC types: (a) original histogram distribution of harmonic term and (b) normalization of histogram distribution of harmonic term.

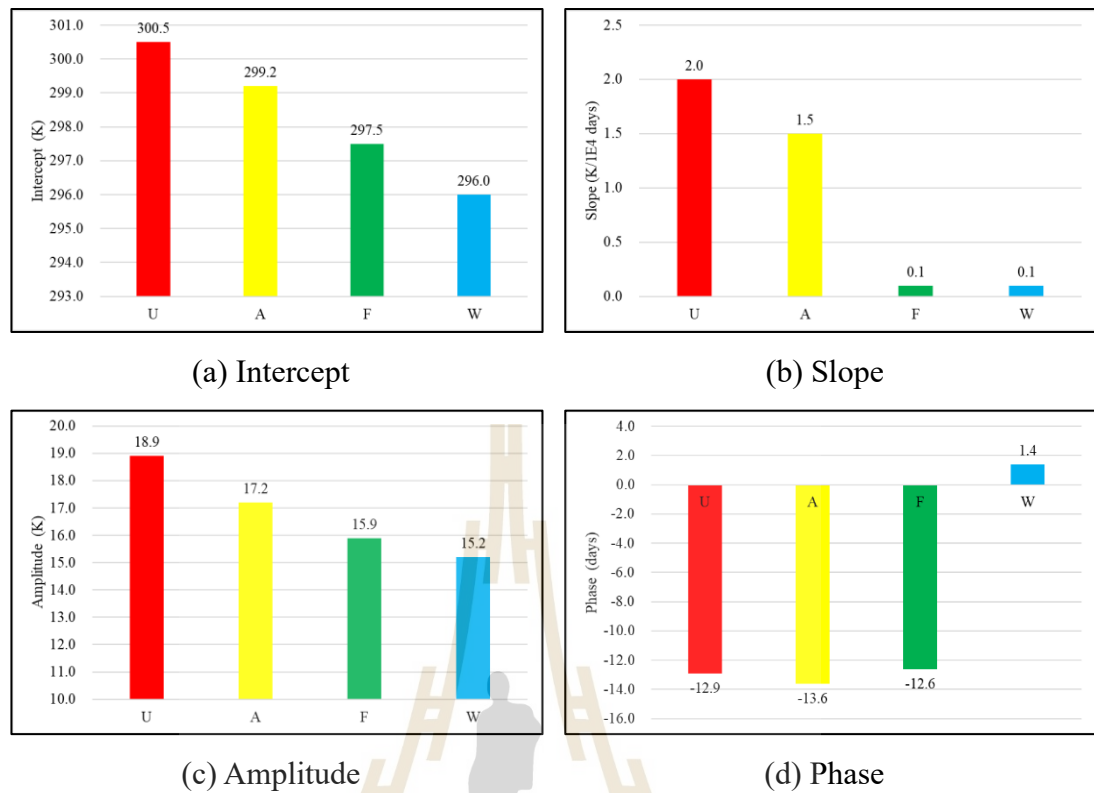


Figure 8.21 Comparison of four harmonic terms between four different LULC types.

As a result, in Figure 8.20 and Figure 8.21, intercept, slope and amplitude of urban and built-up land (U) have the highest value, followed by agricultural land (A), forest land (F), and water bodies (W) has the lowest value.

When we pay attention to the phase value, it was found that the phase value of the water body is only 1.4 days. This finding indicates that the temporal change of water bodies' temperature is very low; in other words, temperatures of water bodies in different periods are the same or similar. That is, extreme values (the maximum and minimum values per year) of LST values are reached at a similar time of each year, and the difference does not exceed two days. On the contrary, for non-water body areas, the time at which the LST values reaches the extreme value varies more than twelve days.

(2) Relationship between harmonic terms of LST and persistent year of LULC type

In this study, the relationship between harmonic terms of LST and persistent year of LULC type were examined by two approaches: simple chart and simple linear regression based on the extracted value of a persistent year of four harmonic terms of LST of LULC types. In detail, the simple linear regression analysis between the harmonic term of LST of each LULC type and persistent year in 18 years was here examined to identify the suitable harmonic term of LST for describing its relationship with the persistent year of each LULC type using correlation coefficient as suggested by Cohen (1988). In statistics, the correlation coefficient (r) measures the strength and direction of a linear relationship between two variables on a scatterplot (Table 8.4).

Table 8.4 Scale of the correlation coefficient.

Value of r	Meaning of r
-1	A perfect negative linear relationship
$-1 < r < -0.5$	A strong negative linear relationship
$-0.5 < r < -0.3$	A moderate negative relationship
$-0.3 < r < -0.1$	A weak negative linear relationship
0	No linear relationship
$0.1 < r < 0.3$	A weak positive linear relationship
$0.3 < r < 0.5$	A moderate positive relationship
$0.5 < r$	A strong positive linear relationship
1	A perfect positive linear relationship

Source: Modified from Cohen (1988).

The result of the extracted value of a persistent year of four harmonic terms of LST of LULC types is reported in Table 8.5 and displayed in Figure 8.22.

Table 8.5 The time-weighted average of different persistent time of four harmonic terms of four different LULC types.

LULC	Persistent time (years)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
U	Intercept	302.5	299.4	295.9	293.4	298.5	299.2	300.6	299.6	299.4	300.2	301.6	302.0	302.0	301.9	302.2	302.0	301.9	301.7
	Slope	-11.9	5.4	11.1	12.0	4.3	3.2	1.0	2.5	3.4	1.9	-0.4	-1.0	-1.0	-0.4	-0.7	0.1	0.0	0.8
	Amplitude	18.7	18.2	18.2	17.7	17.8	17.9	18.0	18.3	18.6	18.6	18.8	18.7	18.7	19.2	19.4	19.8	19.8	20.0
	Phase	-1.9	-2.2	-2.3	-2.9	-4.0	-5.4	-7.4	-8.4	-9.1	-9.7	-11.2	-13.0	-14.0	-15.5	-17.6	-19.0	-19.0	-22.0
A	Intercept	305.3	299.2	296.6	296.4	298.4	299.4	299.6	299.1	298.9	299.1	299.5	299.9	299.9	299.6	299.5	299.5	299.4	299.2
	Slope	-24.3	6.5	14.0	7.4	6.9	3.9	3.1	4.2	6.4	3.9	1.4	0.4	0.1	0.7	0.1	-0.9	-1.1	-0.1
	Amplitude	18.1	17.9	17.7	17.7	17.5	17.6	17.2	17.4	17.9	17.6	17.6	17.8	17.8	17.8	17.4	17.0	16.8	16.9
	Phase	-1.9	-1.6	-1.7	-1.6	-2.5	-3.7	-4.3	-4.9	-3.9	-6.8	-11.3	-16.1	-16.0	-17.3	-18.1	-18.8	-19.1	-17.4
F	Intercept	307.0	297.5	294.7	291.9	296.0	297.3	298.8	297.8	297.7	298.4	299.0	299.3	299.4	298.9	298.9	298.4	298.4	297.5
	Slope	-15.4	9.6	11.3	10.5	4.0	1.4	-1.1	1.0	1.1	-0.1	-1.9	-2.8	-3.3	-2.4	-2.5	-2.0	-2.1	-1.2
	Amplitude	16.9	16.7	16.5	15.7	15.5	15.5	15.6	16.1	16.0	16.2	16.0	16.2	16.0	16.1	16.2	16.5	16.3	15.8
	Phase	-2.1	-1.5	-1.7	-2.5	-3.2	-4.1	-5.2	-6.5	-6.6	-8.1	-10.1	-11.0	-10.1	-11.6	-12.3	-13.5	-14.8	-16.6
W	Intercept	299.4	297.6	294.3	297.2	297.0	297.3	297.3	297.6	297.4	297.4	297.2	297.7	297.7	297.1	297.0	297.2	297.4	295.1
	Slope	-2.2	7.9	10.2	2.4	3.6	1.9	1.6	0.7	0.9	0.3	-0.2	-0.6	-1.6	-0.9	-0.6	-1.5	-1.6	-1.2
	Amplitude	17.4	17.3	17.2	17.5	17.3	17.1	16.8	17.3	17.0	17.3	16.7	17.5	16.6	16.8	16.6	16.7	16.8	13.8
	Phase	-1.2	-0.5	-0.4	0.0	-0.6	-0.7	-0.6	-0.6	-1.4	-2.5	-0.3	-1.9	-2.9	-2.4	-2.8	-4.3	-7.5	3.7

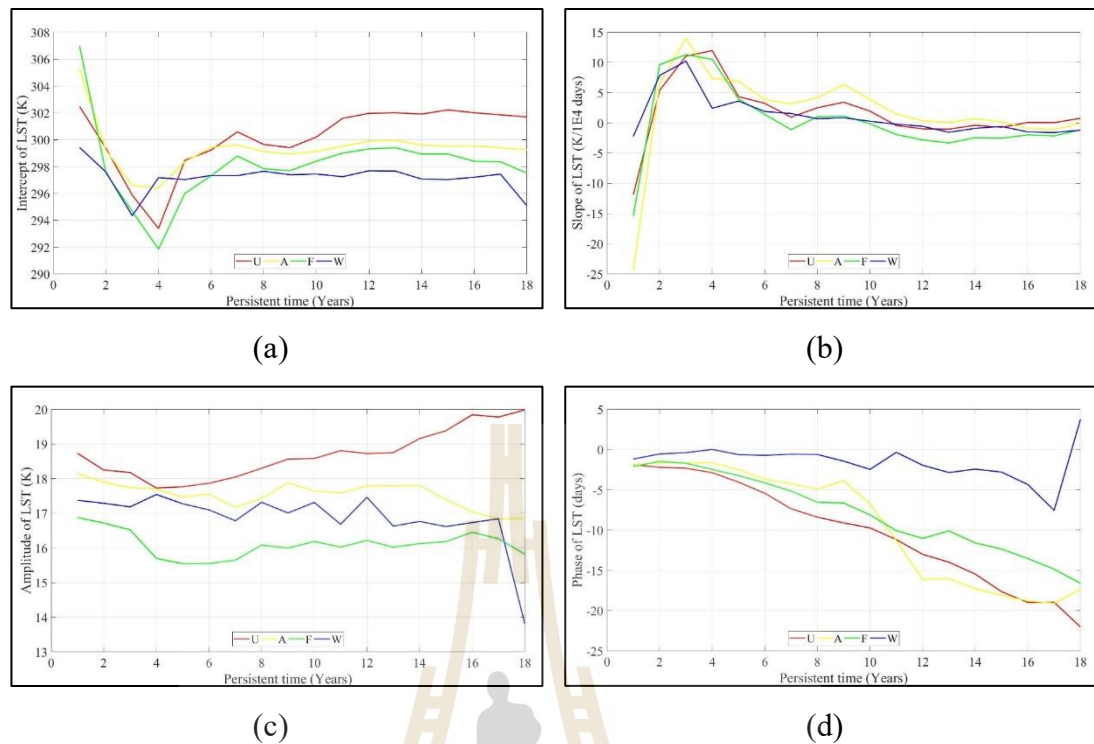


Figure 8.22 The relationship between persistent time and the value of four harmonic terms for four different LULC types: (a) intercept, (b) slope, (c) amplitude, and (d) phase.

As results in Table 8.5 and Figure 8.22, if we observed the relationship between intercept of LST and LULC type, the persistent time for stability of each LULC is different. In this study, the persistent years for a stability of intercept of LST of water bodies start after four years while the persistent years for a stability of intercept of LST of agricultural land start after seven years but the persistent years for a stability of intercept of LST of urban and built-up land and forest land start after 11 years. These findings indicate the development stages of each LULC type. For example, urban and built-up land require some years to complete the development stage with full complete landscaping.

Meanwhile, if we observed the relationship between the slope of LST and LULC type, the persistent time for stability of each LULC is different. Herein, the persistent year for the stability of slope of LST of water bodies starts after eight years, but the persistent years for the stability of slope of LST of urban and built-up land, agricultural land, and forest land start after 11 years.

Furthermore, if we observed the relationship between the amplitude of LST and LULC type, the persistent time for stability of each LULC is different. Herein, the persistent years for a stability of amplitude of LST of water bodies start after six years while the persistent years for a stability of amplitude of LST of forest land start after eight years but the persistent years for a stability of amplitude of LST of urban and built-up land and agricultural land cannot be observed here.

Additionally, if we observed the relationship between the phase of LST and LULC type, the value of the phase of LST is not stable as the persistent time increases. Herein, the phase of LST of urban and built-up land, agricultural land, forest land, and water bodies decrease while the persistent time increase.

Meanwhile, linear regression equations with r and R^2 values for the relationship between harmonic terms of LST of each LULC type and its persistent year of each LULC type are reported in Table 8.6.

Table 8.6 Linear regression of persistent years and values of four harmonic terms.

LULC type	Harmonic terms	Linear regression	
U	Intercept	$y = 0.2739x + 297.61$	$R^2 = 0.3647, r = 0.6039$
	Slope	$y = -0.2479x + 4.0348$	$R^2 = 0.0676, r = -0.2601$
	Amplitude	$y = 0.1111x + 17.635$	$R^2 = 0.7116, r = 0.8436$
	Phase	$y = -1.2102x + 1.2446$	$R^2 = 0.9826, r = -0.9913$
A	Intercept	$y = -0.0098x + 299.46$	$R^2 = 0.0009, r = -0.0296$
	Slope	$y = -0.0846x + 2.6104$	$R^2 = 0.0036, r = -0.0596$
	Amplitude	$y = -0.0436x + 17.954$	$R^2 = 0.4049, r = -0.6363$
	Phase	$y = -1.2678x + 2.7754$	$R^2 = 0.8890, r = -0.9429$
F	Intercept	$y = 0.0273x + 297.9$	$R^2 = 0.0025, r = 0.0505$
	Slope	$y = -0.3722x + 3.7583$	$R^2 = 0.1060, r = -0.3255$
	Amplitude	$y = -0.0092x + 16.194$	$R^2 = 0.0170, r = -0.1305$
	Phase	$y = -0.8956x + 0.6382$	$R^2 = 0.9762, r = -0.9880$
W	Intercept	$y = -0.0379x + 297.53$	$R^2 = 0.0374, r = -0.1935$
	Slope	$y = -0.4007x + 4.8657$	$R^2 = 0.4150, r = -0.6442$
	Amplitude	$y = -0.0924x + 17.751$	$R^2 = 0.3581, r = -0.5985$
	Phase	$y = -0.1469x - 0.1076$	$R^2 = 0.1212, r = -0.3481$

As a result, the relationship between four harmonic terms of LST for urban and built-up land and persistent year provide different strength and direction of a linear relationship. The amplitude and intercept of LST have a strong positive linear relationship with persistent year, the r value of 0.8436, and 0.6039, respectively, while the phase of LST shows a strong negative linear relationship with persistent year, r value of -0.9913. In contrast, the slope of LST shows a weak negative linear relationship with persistent year, r value of -0.2601. This finding implies that when a persistent year of urban and built-up land increases, the value of the intercept and amplitude of LST in urban and built-up land increases.

Meanwhile, the relationship between four harmonic terms of LST for agricultural land and persistent year also provide different strength and direction of a

linear relationship. The amplitude and phase of LST have a strong negative linear relationship with persistent years of agricultural land, r value of -0.6363, and -0.9429, respectively. This finding implies that when a persistent year of agricultural land increases, the value of the amplitude and phase of LST in agricultural land decreases. On the contrary, there is no linear relationship between intercept and slope of LST with a persistent year in agricultural land. The derived equations only provide the r value of -0.0296 and -0.0596, respectively.

For the meantime, the relationship between four harmonic terms of LST for forest land and persistent year also provide different strength and direction of a linear relationship. The phase of LST has a strong negative linear relationship with persistent years of forest land, r value of -0.9880. This finding implies that when a persistent year of forest land increases, the value of the phase of LST in forest land decrease. On the contrary, there is a moderate negative relationship between the slope of LST with a persistent year in forest land, a weak negative relationship between the amplitude of LST with a persistent year in forest land, and no linear relationship between intercept of LST with a persistent year in forest land. The derived equations only provide the r value of -0.3255, -0.1305, and 0.0505, respectively.

In the meantime, the relationship between four harmonic terms of LST for water bodies and persistent year also provide different strength and direction of a linear relationship. The slope and amplitude of LST have a strong negative linear relationship with persistent years of water bodies, r value of -0.6442, and -0.5985, respectively. This finding implies that when a persistent year of water bodies increases, the value of the slope and amplitude of LST in water bodies decreases. On the contrary, there is a moderate negative relationship between the phase of LST with a persistent

year in water bodies and a weak negative relationship between the intercept of LST with a persistent year in water bodies. The derived equations only provide the r value of -0.3481, -0.1935, respectively.

(3) Impact of LULC change on harmonic terms of LST

In this study, the pixels with the LULC change from A to U, F to U, and W to U were firstly selected under HA. Then the average value of harmonic terms (intercept, slope, amplitude, and phase) of LST of those selected pixels before and after LULC change were extracted and calculated the change. The measurement unit of change based on intercept, slope (speed of change), amplitude, and phase of LST are Kelvin, Kelvin/10,000 days, Kelvin, days, respectively.

Table 8.7 shows the result of the average value of harmonic terms of LST and its change of “from-to” change class (A to U, F to U, and W to U).

Table 8.7 Thermal signatures and thermal signatures change of the LULC change area.

Change types	Intercept (K)			Slope (K/1E4 days)			Amplitude (K)			Phase (days)		
	Before	After	change	Before	After	change	Before	After	change	Before	After	change
A to U	299.3	301.1	1.8	4.5	1.5	-3.0	17.5	19.6	2.1	-8.2	-11.7	-3.4(4)
F to U	298.1	299.4	1.2	3.4	2.2	-1.2	16.6	18.9	2.4	-4.7	-8.2	-3.4(4)
W to U	297.3	300.0	2.7	5.5	2.5	-3.0	16.2	18.7	2.5	0.9	-8.7	-9.6(8)

As a result, in Table 8.7, it suggests that the most significant LST change was observed in the conversion type of water bodies to urban and built-up land. The differences are 2.7 K, -3 K/1E4 days, 2.5 K, and -9.6 days in terms of intercept, slope, amplitude, and phase, respectively. In contrast, in terms of intercept and phase, the conversion of forest land to urban and built-up land provides the smallest difference,

but in terms of amplitude, the conversion of agricultural land to urban and built-up land provides the smallest difference.

In general, when other types of LULC are converted to urban and built land, their intercept and amplitude values will increase. This finding is similar to the values of the harmonic parameters shown in Figure 8.21; that is, the intercept and amplitude values of urban and built-up land are higher than agricultural land, forest land, and water bodies. This finding implies that the average temperature and annual temperature change in urban areas are higher than in other non-urban areas.

On the contrary, the slope value in Table 8.7 is calculated using only the pixels converted from other non-urban lands to urban land. The intercept and amplitude values of these pixels will suddenly increase with the conversion. However, vegetation is usually planted around these new urban and built-up land. Over time, although the LULC type of these areas has not changed (still urban and built-up land), the vegetation coverage will increase, which will slow down the increasing speed of LST.

8.4 Spatial analysis and impact of LULC change on LST

The impact of LULC change on LST was assessed using zonal statistics analysis based on the derived multitemporal LULC data and LST in 2000, 2006, 2011, and 2017 (Figure 8.23 and Figure 8.24). In this study, LULC change from three different periods (2000-2006, 2006-2011, and 2011-2017) were first conducted using a post-classification comparison change detection algorithm for extraction “from-to” change classes (e.g., A_{2000} to U_{2006} , F_{2000} to U_{2006} , and W_{2000} to U_{2006} , and so on). After that, the mean LST value of “from-to” change classes (e.g., A_{2000} to U_{2006} , F_{2000} to U_{2006} , and W_{2000} to U_{2006} , and so on) were extracted and calculated its change.

Table 8.8 shows the results of the LST change in three periods (2000-2006, 2006-2011, and 2011-2017), which relate to the conversion of other LULC types into urban and built-up land (A to U, F to U, and W to U).

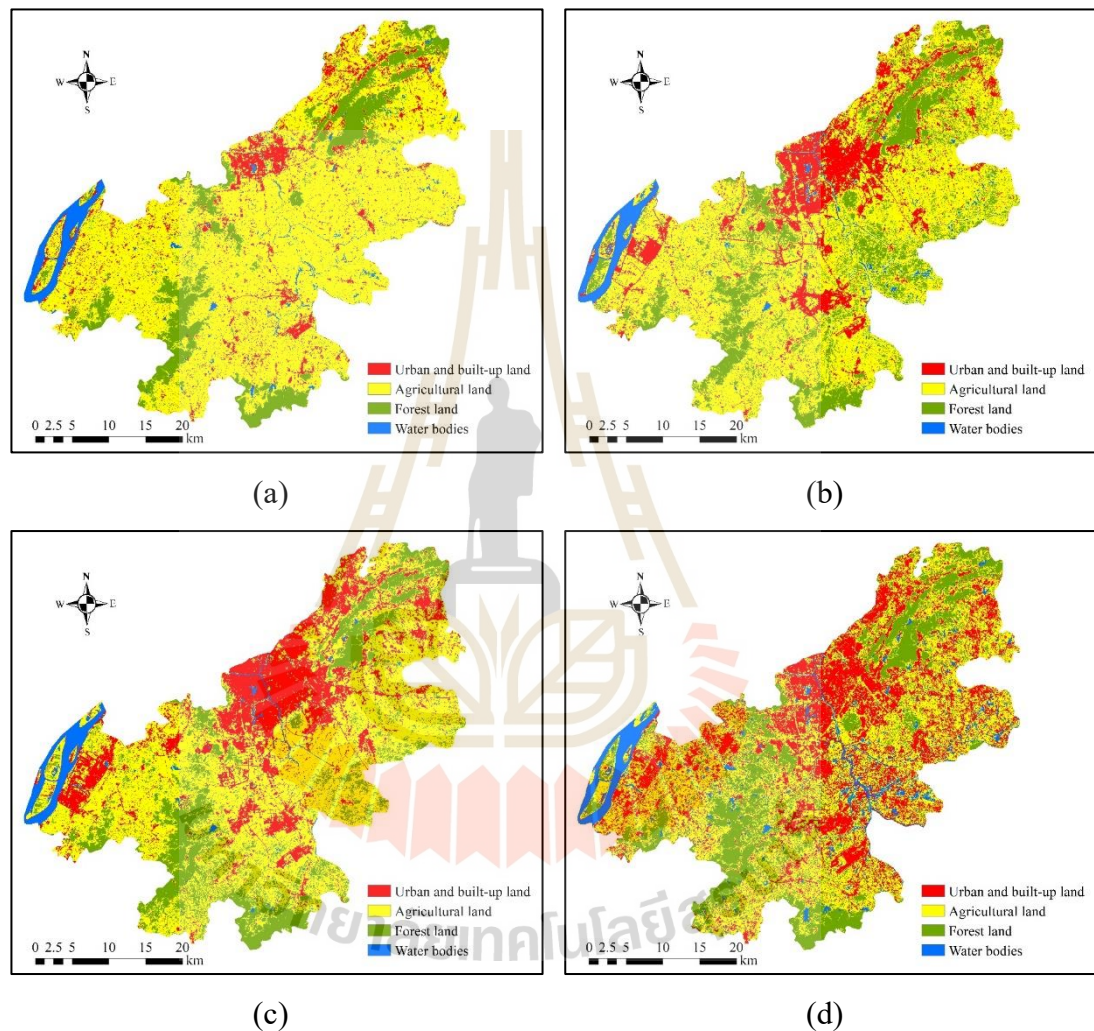


Figure 8.23 LULC map of four years: (a) date 30 July 2000, (b) date 31 July 2006, (c) date 29 July 2011, (d) date 29 July 2017.

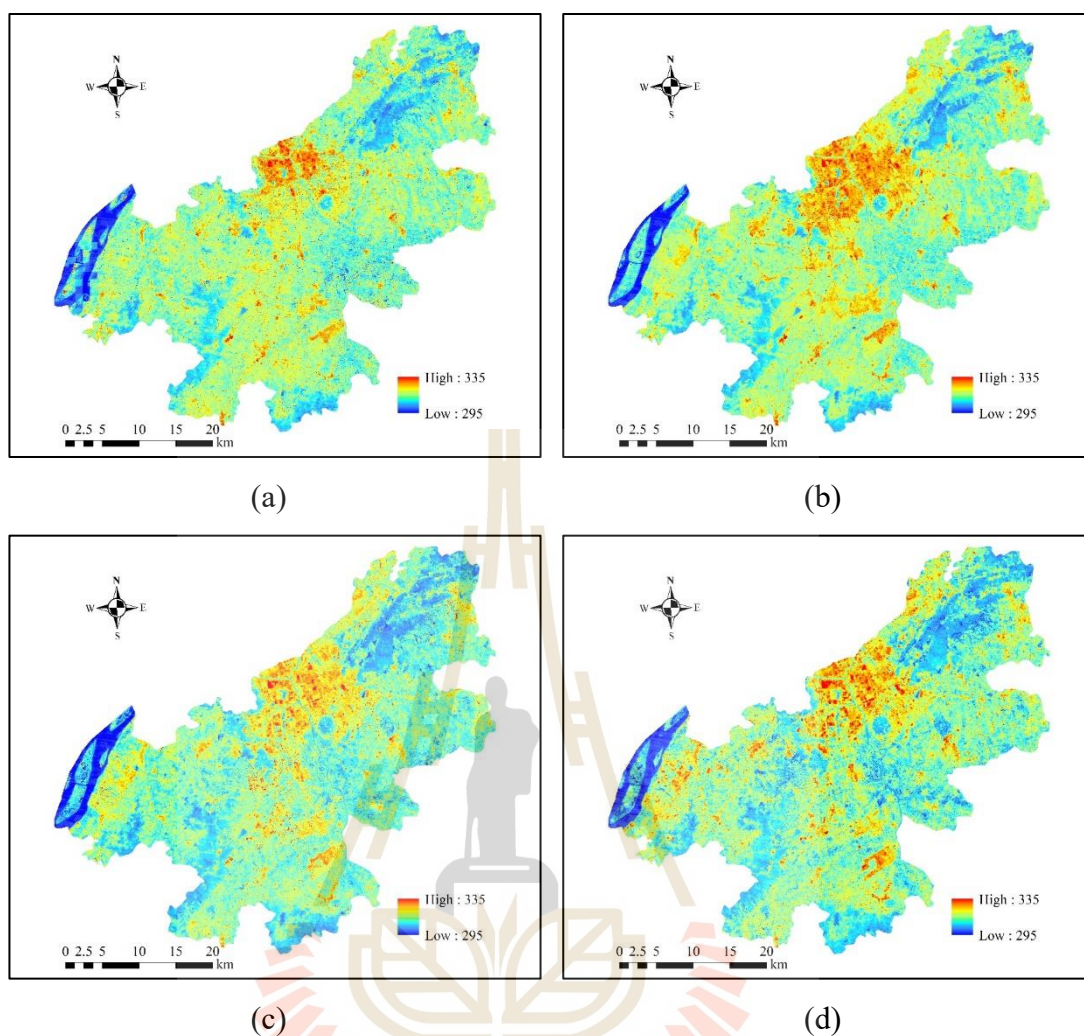


Figure 8.24 LST map of four years: (a) date 30 July 2000, (b) date 31 July 2006, (c) date 29 July 2011, (d) date 29 July 2017.

Table 8.8 LST change of the LULC change area.

Change types	LST (K)								
	2000 - 2006			2006 - 2011			2011 - 2017		
	Before	After	change	Before	After	change	Before	After	change
A to U	317.79	320.35	2.56	317.53	318.76	1.23	315.95	316.75	0.8
F to U	315.67	318.38	2.71	316.19	317.74	1.55	n. a	n. a	n. a
W to U	314.11	317.46	3.35	314.56	316.27	1.71	n. a	n. a	n. a

Note: n. a is not applied since forest land and water bodies in 2011 were not converted into urban and built-up land in 2017.

As a result, in Table 8.8, during 2000 and 2006, when agricultural land, forest land, and water bodies in 2000 were converted into urban and built-up land in 2006, the LST values increased by 2.56 K, 2.71 K, and 3.35 K, respectively. Likewise, during 2006 and 2011, when agricultural land, forest land, and water bodies in 2006 were converted into urban and built-up land in 2011, the LST values increased by 1.23 K, 1.55 K, and 1.71 K, respectively.

In the meantime, during 2011 and 2017, the LST value increased by 0.8 K when agricultural land in 2011 was converted into urban and built-up land in 2017. Additionally, forest land and water bodies in 2011 were not converted into urban and built-up land in 2017.

This finding is an expected result because when agricultural land, forest land, and water bodies are converted into urban and built-up land, LST will be increased. These phenomena indicate the impact of LULC change on LST when the green area, such as agricultural land or forest land, are changed to be urban and built-up land.

SUMMARY

Based on the harmonic term from the time domain perspective, intercept and amplitude value of LST suddenly increased, when other types of LULC were converted to urban and built land, it indicates that the average temperature and annual temperature change in urban areas are higher than in other non-urban areas. The slope value of LST decreased when other types of LULC were converted to urban and built land; it indicates that the increasing speed of LST in the LULC change area (from A to U, from F to U, from W to U,) will be reduced.

From the perspective of the spatial domain, the LST value increased, when agricultural land, forest land, and water bodies in 2000 were converted into urban and built-up land in 2006, agricultural land, forest land, and water bodies in 2006 were converted into urban and built-up land in 2011, agricultural land in 2011 was converted into urban and built-up land in 2017. These phenomena indicate the impact of LULC change on LST when the green area, such as agricultural land or forest land, are changed to be urban and built-up land.



CHAPTER IX

CONCLUSION AND RECOMMENDATION

The economic and population growth continued during the last two decades in Nanjing City, China. These phenomena induce land use and land cover change in the city. In particular, the area of urban and built-up land has continually increased over the years. As a result, it impacts the land surface temperature of this city. The ultimate goal of the study is to develop a new supervised classification method using harmonic analysis with a minimum spectral distance algorithm for multitemporal LULC mapping. Meanwhile, the specific research objectives are consisting of (1) to apply unsupervised classification algorithms for multitemporal LULC classification based on time-series Landsat datasets, (2) to develop a nearly real-time supervised LULC classification and change detection method using a time-series model and a minimum spectral distance algorithm, and (3) to reconstruct time-series LST and investigate the impact of LULC on LST between 2000 and 2017. To achieve goals and objectives, two generic unsupervised classification algorithms, namely K-Means and ISODATA, were first applied to classify multitemporal LULC maps based on the derived coefficient of harmonic analysis of time-series Landsat datasets under the MATLAB environment. Then, a new supervised classification method using harmonic analysis with a minimum spectral distance algorithm was developed and applied to classify a multitemporal LULC map under the MATLAB environment. Finally, the time-series LST was reconstructed using a harmonic analysis model and applied with the detailed LULC

data by a supervised method to examine the impact of LULC change on LST. The conclusion and recommendations of the study are presented in this chapter.

9.1 Conclusions

9.1.1 Time-series land use and land cover classification using an unsupervised method

The classification of multitemporal LULC data in 2000, 2006, 2011, and 2017 from time-series Landsat datasets with different spectral features (NIR, TOP3, TOP6, ORI6) were successfully implemented using K-Means and ISODATA algorithm. By considering the performance of the clustering algorithm for multitemporal supervised classification, it can be concluded that the accuracy of thematic maps derived by the ISODATA algorithm was slightly higher than the K-Means algorithm. Overall accuracy and Kappa hat coefficient of the classified LULC maps in four years with ORI6 spectral features varied from 77.04% to 86.48% and 63.88% to 78.53%, respectively.

Besides, the overall accuracy of the classified maps using the single spectral feature (NIR) in four different years from both algorithms was approximately 70% while the overall accuracy of the classified maps using the multi-spectral feature (TOP3, TOP6, ORI6) in four different years from both algorithms was approximately 80%. This finding indicates that the multiple spectral features could provide higher accuracy than a single spectral feature.

Furthermore, the pairwise Z-test indicates that the accuracy of the LULC classification from both algorithms was not significantly different, and the accuracy of the LULC classification using the NIR spectral feature was different from the accuracy

of LULC classification using TOP3, TOP6, and ORI6 spectral features. Additionally, whether using the K-Means algorithm or ISODATA algorithm, the accuracy of LULC classification using TOP3, TOP6, and ORI6 spectral features were insignificantly different.

9.1.2 Time-series land use and land cover classification using a supervised method

By comparing the overall accuracy of change detection according to spectral feature, LULC change detection using NIR provided the least overall accuracy (74.70%). In contrast, the combination of multispectral features provided higher accuracy, the overall accuracy of TOP3, TOP6, and ORI6 were 81.67%, 87.05%, and 88.25%, respectively.

The average overall accuracy of LULC classification in 2000, 2006, 2011, and 2017 were 86.60%, 81.53%, 84.55%, and 81.80%, respectively. Besides, the average overall accuracy with NIR, TOP3, TOP6, and ORI6 spectral features were 67.77%, 84.59%, 90.80%, and 91.31%, respectively. As a result, it infers that when the number of spectral features increases, the overall accuracy increases.

The values of the pairwise Z-test indicate that the classification accuracy of NIR was significantly different from the other spectral features. Also, the values of the pairwise Z-test shows that the classification accuracy of TOP3 was also significantly different from the other two spectral features (TOP6 and ORI6). Conversely, the pairwise Z-test indicates that the classification accuracy of TOP6 and ORI6 was insignificantly different.

9.1.3 Impact study of land use and land cover change on land surface temperature

Based on the harmonic term from the time domain perspective, intercept, and amplitude value of LST suddenly increased, when other types of LULC were converted to urban and built land. It indicates that the average temperature and annual temperature change in urban areas are higher than in other non-urban areas. The small value of the slope of LST indicates that increasing the speed of LST in the LULC change area (from A to U, from F to U, from W to U) is reduced.

From the perspective of the spatial domain, the LST value increased, when agricultural land, forest land, and water bodies were converted into urban and built-up land. These phenomena indicate the impact of LULC change on LST when the green area, such as agricultural land or forest land, are changed to be urban and built-up land.

9.2 Recommendations

The newly developed supervised classification method for multitemporal LULC mapping using harmonic analysis with a minimum spectral distance algorithm is a semi-automatic process under the MATLAB software environment, and it requires little human interference. This approach can efficiently classify multitemporal LULC maps, offering with highly accurate results based on the standard harmonic function curve for each LULC type obtained from the stable areas of each LULC type from multiple years. Moreover, this approach overcomes the limitations of image selection by considering individual pixels from the available Landsat data. For example, the scan line gaps of Landsat 7 are regarded as contaminated pixels, but available clearly observed pixels

could also be used. Also, this approach does not require the relative normalization of each image for multitemporal LULC classification and change detection because phenological and solar angle differences do not affect classification under the time-series model. Furthermore, this approach applies spatiotemporal filtering with a mode function under post-classification processing to remove unexpected errors caused by contaminated pixels (e.g., clouds and cloud shadows) or meteorological conditions (e.g., flooding and drought). More specifically, this approach can quickly generate a LULC change map for any period and can provide “from-to” change information as a post-classification comparison change detection algorithm, which is frequently applied for change detection. Multitemporal LULC data and change information can also efficiently be used for exploring the impact of LULC change on land surface temperature. However, for improving the usability of the algorithm, the recommendations are here suggested as follows.

1. This developed classification method should be examined in other areas to determine its spatial and temporal transferability.

2. In this study, the efficiency of the developed algorithm is more focused instead of the number of LULC types. Therefore, four LULC types that are persistent for an extended period are here selected to examine the capacity and efficiency of the algorithm. Future researchers who want to increase the number of LULC types are here recommended by increasing the size of the study area, such as a city or province.

3. In this study, stable areas for each LULC type between 2000 and 2017 were extracted by overlay analysis based on the classified LULC maps in 2000, 2006, 2011, and 2017 by using MLC. The error of the stable area for each LULC type is minimal since the stable areas of each LULC category are extracted from the multitemporal

LULC maps. However, LULC classification with MLC requires a great deal of time to select suitable training areas with a normal distribution. In the future, the selection of a training area of each LULC type should be implemented by using an automatic or semi-automatic process. For example, Huang et al. (2002) applied a dark object concept to automatically generate a training area for mapping forest cover changes with the advanced support vector machines algorithm.

4. To build standard reference harmonic curves for multitemporal LULC classification, almost 10,000 sample pixels were selected from stable areas of each LULC type. A large number of sample pixels can ensure the accuracy of standard reference harmonic curves, but it takes much time. In the future, it is necessary to find a suitable sample size to reduce the time for standard reference harmonic curve construction.

5. The multitemporal LULC classification is performed on the spectral values and its change of each pixel from 388 dates in this study. Conceptually, if six spectral reflectance bands of the time-series of Landsat dataset cubes (with 388 dates) are considered as six hyperspectral data cubes, spectral angle mapper (SAM) method can be examined to classify multitemporal LULC map from any single or multiple data cubes in the future.

6. Multitemporal LULC classification based on time-series Landsat data generally requires a large amount of space and time for processing and analysis. This phenomenon leads to extremely time-consuming data processing. So, cloud computing technologies can provide an alternative solution for increasing space and decreasing processing time.

7. The conceptual framework and relevant workflows of the newly developed supervised classification method using harmonic analysis with a minimum spectral distance algorithm can be used as a guideline for software developers for (semi) automatic LULC classification and mapping in the future.



REFERENCES

- Alfieri, S. M., Lorenzi, F. D., and Menenti, M. (2013). Mapping air temperature using time series analysis of LST: the SINTESI approach. **Nonlinear Processes in Geophysics**. 20(4): 513-527.
- Andrew, M. E., Wulder, M. A., and Nelson, T. A. (2014). Potential contributions of remote sensing to ecosystem service assessments. **Progress in Physical Geography**. 38(3): 328-353.
- Anyamba, A., and Eastman, J. R. (1996). Interannual variability of NDVI over Africa and its relation to El Niño/Southern Oscillation. **International Journal of Remote Sensing**. 17(13): 2533-2548.
- Arnfield, A. J. (2003). Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. **International Journal of Climatology**. 23(1): 1-26.
- Azzali, S., and Menenti, M. (2000). Mapping vegetation-soil-climate complexes in southern Africa using temporal Fourier analysis of NOAA-AVHRR NDVI data. **International Journal of Remote Sensing**. 21(5): 973-996.
- Baig, M. H. A., Zhang, L., Shuai, T., and Tong, Q. (2014). Derivation of a tasseled cap transformation based on Landsat 8 at-satellite reflectance. **Remote Sensing Letters**. 5(5): 423-431.

- Bastiaanssen, W. G. M., Menenti, M., Feddes, R. A., and Holtslag, A. A. M. (1998). A remote sensing surface energy balance algorithm for land (SEBAL). 1. Formulation. **Journal of Hydrology**. 212-213: 198-212.
- Benedetti, R., Rossini, P., and Taddei, R. (1994). Vegetation classification in the Middle Mediterranean area by satellite data. **International Journal of Remote Sensing**. 15(3): 583-596.
- Bonan, G. B., Pollard, D., and Thompson, S. L. (1992). Effects of boreal forest vegetation on global climate. **Nature**. 359: 716.
- Brown, L., Chen, J. M., Leblanc, S. G., and Cihlar, J. (2000). A Shortwave Infrared Modification to the Simple Ratio for LAI Retrieval in Boreal Forests: An Image and Model Analysis. **Remote Sensing of Environment**. 71(1): 16-25.
- Buitre, M. J. C., Zhang, H., and Lin, H. (2019). The Mangrove Forests Change and Impacts from Tropical Cyclones in the Philippines Using Time Series Satellite Imagery. **Remote Sensing**. 11(6): 688.
- Carlson, T. N., and Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. **Remote Sensing of Environment**. 62(3): 241-252.
- Chuvieco, E. (2016). **Fundamentals of Satellite Remote sensing: An Environmental Approach Second edition**. Boca Raton, Florida, USA: CRC Press/Taylor & Francis.
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess. **Journal of Official Statistics**. 6: 3-73.

- Cohen, J. (1988). **Statistical power analysis for the behavioral sciences (2nd ed.)**. Hillsdale, NJ, USA: Lawrence Erlbaum Associates.
- Cohen, W. B., Yang, Z., and Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation. **Remote Sensing of Environment**. 114(12): 2911-2924.
- Congalton, R. G., and Green, K. (2009). **Assessing the Accuracy of Remotely Sensed Data - Principles and Practices Second edition**. Boca Raton, NW, USA: CRC Press, Taylor & Francis Group.
- Cort, J. W., and Kenji, M. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. **Climate Research**. 30(1): 79-82.
- Crist, E. P., and Cicone, R. C. (1984). A Physically-Based Transformation of Thematic Mapper Data---The TM Tasseled Cap. **IEEE Transactions on Geoscience and Remote Sensing**. GE-22(3): 256-263.
- Cristóbal, J., Jiménez-Muñoz, J. C., Sobrino, J. A., Ninyerola, M., and Pons, X. (2009). Improvements in land surface temperature retrieval from the Landsat series thermal band using water vapor and air temperature. **Journal of Geophysical Research: Atmospheres**. 114(D8): D08103.
- de Jong, R., de Bruin, S., de Wit, A., Schaepman, M. E., and Dent, D. L. (2011). Analysis of monotonic greening and browning trends from global NDVI time-series. **Remote Sensing of Environment**. 115(2): 692-702.
- DeVries, B. (2015). **Monitoring tropical forest dynamics using Landsat time series and community-based data**. Wageningen, Netherlands, USA: Wageningen

University & Research, Laboratory of Geo-information Science and Remote Sensing.

Douglas, E., Niyogi, D., Frohking, S., Yeluripati, J., Pielke Sr, R., Niyogi, N., J. Vörösmarty, C., and Mohanty, U. C. (2006). Changes in moisture and energy fluxes due to agricultural land use and irrigation in the Indian Monsoon Belt.

Geophysical Research Letters. 33: L14403.

Eastman, R., and Fulk, M. (1993). Long Sequence Time Series Evaluation Using Standard Principal Components. **Photogrammetric Engineering and Remote Sensing**. 59(6): 991-996.

Feddema, J. J. (2005). The Importance of Land-Cover Change in Simulating Future Climates. **Science**. 310(5754): 1674-1678.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., et al. (2005). Global Consequences of Land Use. **Science**. 309(5734): 570-574.

Fu, P., and Weng, Q. (2016). A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with Landsat imagery. **Remote Sensing of Environment**. 175: 205-214.

Gebhardt, S., Wehrmann, T., Ruiz, M. A. M., Maeda, P., Bishop, J., Schramm, M., Kopeinig, R., Cartus, O., Kellndorfer, J., Ressl, R., Santos, L. A., and Schmidt, M. (2014). MAD-MEX: Automatic Wall-to-Wall Land Cover Monitoring for the Mexican REDD-MRV Program Using All Landsat Data. **Remote Sensing**. 6(5): 3923-3943.

- Geerken, R., Zaitchik, B., and Evans, J. P. (2005). Classifying rangeland vegetation type and coverage from NDVI time series using Fourier Filtered Cycle Similarity. **International Journal of Remote Sensing**. 26(24): 5535-5554.
- Geerken, R. A. (2009). An algorithm to classify and monitor seasonal variations in vegetation phenologies and their inter-annual change. **ISPRS Journal of Photogrammetry and Remote Sensing**. 64(4): 422-431.
- Gillanders, S. N., Coops, N. C., Wulder, M. A., Gergel, S. E., and Nelson, T. (2008). Multitemporal remote sensing of landscape dynamics and pattern change: describing natural and anthropogenic trends. **Progress in Physical Geography: Earth and Environment**. 32(5): 503-528.
- Gounaridis, D., Symeonakis, E., Chorianopoulos, I., and Koukoulas, S. (2018). Incorporating Density in Spatiotemporal Land Use/Cover Change Patterns: The Case of Attica, Greece. **Remote Sensing**. 10(7): 1034.
- Gurgel, H. C., and Ferreira, N. J. (2003). Annual and interannual variability of NDVI in Brazil and its connections with climate. **International Journal of Remote Sensing**. 24(18): 3595-3609.
- Hall-Beyer, M. (2003). Comparison of single-year and multiyear NDVI time series principal components in cold temperate biomes. **IEEE Transactions on Geoscience and Remote Sensing**. 41(11): 2568-2574.
- Hansen, J., Ruedy, R., Sato, M., and Lo, K. (2010). Global surface temperature change. **Reviews of Geophysics**. 48(4): RG4004.
- Healey, S. P., Yang, Z., Cohen, W. B., and Pierce, D. J. (2006). Application of two regression-based methods to estimate the effects of partial harvest on forest

- structure using Landsat data. **Remote Sensing of Environment**. 101(1): 115-126.
- Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., and Hobart, G. W. (2015). An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. **Remote Sensing of Environment**. 158: 220-234.
- Herold, M., Latham, J. S., Di Gregorio, A., and Schullius, C. C. (2006). Evolving standards in land cover characterization. **Journal of Land Use Science**. 1(2-4): 157-168.
- Hirosawa, Y., Marsh, S. E., and Kliman, D. H. (1996). Application of standardized principal component analysis to land-cover characterization using multitemporal AVHRR data. **Remote Sensing of Environment**. 58(3): 267-281.
- Houghton, R. A., House, J. I., Pongratz, J., van der Werf, G. R., DeFries, R. S., Hansen, M. C., Le Quéré, C., and Ramankutty, N. (2012). Carbon emissions from land use and land-cover change. **Biogeosciences**. 9(12): 5125-5142.
- Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., and Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. **Remote Sensing of Environment**. 114(1): 183-198.
- Huang, C., Wylie, B. K., Yang, L., Homer, C. G., and Zylstra, G. (2002). Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. **International Journal of Remote Sensing**. 23(8): 1741-1748.

- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., and Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote Sensing of Environment**. 83(1): 195-213.
- Huete, A., Justice, C., and Leeuwen, W. (1999). **MODIS vegetation index (MOD 13) Algorithm theoretical basis document**. <http://modis.gsfc.nasa.gov/data/atbd>.
- Immerzeel, W. W., Quiroz, R. A., and de Jong, S. M. (2005). Understanding precipitation patterns and land use interaction in Tibet using harmonic analysis of SPOT VGT-S10 NDVI time series. **International Journal of Remote Sensing**. 26(11): 2281-2296.
- J. Kauth, R., and S. Thomas, G. (1976). The Tasseled Cap -- A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by LANDSAT. **LARS Symposia**. 159: 41-51.
- Jakubauskas, M. E., Legates, D. R., and Kastens, J. H. (2002). Crop identification using harmonic analysis of time-series AVHRR NDVI data. **Computers and Electronics in Agriculture**. 37(1): 127-139.
- Jensen, J. R. (2015). **Introductory Digital Image Processing: A Remote Sensing Perspective**. NJ, USA: Prentice Hall Press.
- Jia, L., Shang, H., Hu, G., and Menenti, M. (2011). Phenological response of vegetation to upstream river flow in the Heihe River basin by time series analysis of MODIS data. **Hydrology and Earth System Sciences**. 15: 1047-1064.
- Jiménez-Muñoz, J. C., Cristobal, J., Sobrino, J. A., Soria, G., Ninyerola, M., and Pons, X. (2009). Revision of the Single-Channel Algorithm for Land Surface Temperature Retrieval From Landsat Thermal-Infrared Data. **IEEE Transactions on Geoscience and Remote Sensing**. 47(1): 339-349.

- Jiménez-Muñoz, J. C., and Sobrino, J. A. (2003). A generalized single-channel method for retrieving land surface temperature from remote sensing data. **Journal of Geophysical Research**. 108(D22): ACL 2-1.
- Jiménez-Muñoz, J. C., Sobrino, J. A., Skoković, D., Mattar, C., and Cristóbal, J. (2014). Land Surface Temperature Retrieval Methods From Landsat-8 Thermal Infrared Sensor Data. **IEEE Geoscience and Remote Sensing Letters**. 11(10): 1840-1843.
- Julien, Y., Sobrino, J. A., and Verhoef, W. (2006). Changes in land surface temperatures and NDVI values over Europe between 1982 and 1999. **Remote Sensing of Environment**. 103(1): 43-55.
- Kaufmann, R. K., and Stern, D. I. (1997). Evidence for human influence on climate from hemispheric temperature relations. **Nature**. 388: 39.
- Keenan, T. F., Gray, J., Friedl, M. A., Toomey, M., Bohrer, G., Hollinger, D. Y., Munger, J. W., O'Keefe, J., Schmid, H. P., Wing, I. S., Yang, B., and Richardson, A. D. (2014). Net carbon uptake has increased through warming-induced changes in temperate forest phenology. **Nature Climate Change**. 4: 598.
- Kennedy, R. E., Cohen, W. B., and Schroeder, T. A. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. **Remote Sensing of Environment**. 110(3): 370-386.
- Kogan, F. N. (2001). Operational space technology for global vegetation assessment. **Bulletin of the American Meteorological Society**. 82: 1949-1964.

- Lambin, E. F., Geist, H. J., and Lepers, E. (2003). Dynamics of Land-Use and Land-Cover Change in Tropical Regions. **Annual Review of Environment and Resources**. 28(1): 205-241.
- Lawrence, P. J., and Chase, T. N. (2010). Investigating the climate impacts of global land cover change in the community climate system model. **International Journal of Climatology**. 30(13): 2066-2087.
- Lawrence, P. J., Feddema, J. J., Bonan, G. B., Meehl, G. A., O'Neill, B. C., Levis, S., Lawrence, D. M., Oleson, K. W., Kluzek, E., Lindsay, K., and Thornton, P. E. (2012). Simulating the biogeochemical and biogeophysical impacts of transient land cover change and wood harvest in the Community Climate System Model (CCSM4) from 1850 to 2100. **Journal of Climate**. 25(9): 3071-3095.
- Lee, E., Chase, T. N., Rajagopalan, B., Barry, R. G., Biggs, T. W., and Lawrence, P. J. (2009). Effects of irrigation and vegetation activity on early Indian summer monsoon variability. **International Journal of Climatology**. 29(4): 573-581.
- Lee, E., He, Y., Zhou, M., and Liang, J. (2015). Potential feedback of recent vegetation changes on summer rainfall in the Sahel. **Physical Geography**. 36(6): 449-470.
- Lee, E., Sacks, W. J., Chase, T. N., and Foley, J. A. (2011). Simulated impacts of irrigation on the atmospheric circulation over Asia. **Journal of Geophysical Research: Atmospheres**. 116(D8): D08114.
- Leica. (2005). **ERDAS Field Guide**. Norcross, GA, USA: Leica Geosystems Geospatial Imaging, LLC.
- Li, Z., Tang, B., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. A. (2013). Satellite-derived land surface temperature: Current status and perspectives. **Remote Sensing of Environment**. 131: 14-37.

- Liu, D., and Cai, S. (2012). A Spatial-Temporal Modeling Approach to Reconstructing Land-Cover Change Trajectories from Multi-temporal Satellite Imagery. **Annals of the Association of American Geographers**. 102(6): 1329-1347.
- Lo, C. P., and Quattrochi, D. A. (2003). Land Use and Land Cover Change, Urban Heat Island Phenomenon, and Health Implications: A Remote Sensing Approach. **Photogrammetric Engineering and Remote Sensing**. 69(9): 1053-1063.
- Lobo, A., and Maisongrande, P. (2008). Searching for trends of change through exploratory data analysis of time series of remotely sensed images of SW Europe and NW Africa. **International Journal of Remote Sensing**. 29(17-18): 5237-5245.
- Loveland, T. R., and Defries, R. S. (2004). **Observing and Monitoring Land Use and Land Cover Change**. Washington, D.C., USA: American Geophysical Union.
- Lu, D., Li, L., Li, G., Fan, P., Ouyang, Z., and Moran, E. (2018). Examining Spatial Patterns of Urban Distribution and Impacts of Physical Conditions on Urbanization in Coastal and Inland Metropolises. **Remote Sensing**. 10(7): 1101.
- Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. Paper presented at the **5th Berkeley Symposium on Mathematical Statistics and Probability**, California.
- Mahmood, R., Pielke, R. A., Hubbard, K. G., Niyogi, D., Dirmeyer, P. A., McAlpine, C., Carleton, A. M., Hale, R., Gameda, S., Beltrán-Przekurat, A., Baker, B., McNider, R., Legates, D. R., Shepherd, M., Du, J., et al. (2014). Land cover changes and their biogeophysical effects on climate. **International Journal of Climatology**. 34(4): 929-953.

- McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. **International Journal of Remote Sensing**. 17(7): 1425-1432.
- McPherson, R. A. (2007). A review of vegetation--atmosphere interactions and their influences on mesoscale phenomena. **Progress in Physical Geography**. 31(3): 261-285.
- Menenti, M., Azzali, S., Verhoef, W., and van Swol, R. (1993). Mapping agroecological zones and time lag in vegetation growth by means of fourier analysis of time series of NDVI images. **Advances in Space Research**. 13(5): 233-237.
- Menenti, M., Jia, L., Azzali, S., Roerink, G., Gonzalez-Loyarte, M., and Leguizamon, S. (2010). **Analysis of vegetation response to climate variability using extended time series of multispectral satellite images**. Trivandrum, India: Research Signpost.
- Menenti, M., Malamiri, H. R. G., Shang, H., Alfieri, S. M., Maffei, C., and Jia, L. (2016). **Observing the response of terrestrial vegetation to climate variability across a range of time scales by time series analysis of land surface temperature**. Heidelberg, Germany: Springer Verlag.
- Mi, J., Yang, Y., Zhang, S., An, S., Hou, H., Hua, Y., and Chen, F. (2019). Tracking the Land Use/Land Cover Change in an Area with Underground Mining and Reforestation via Continuous Landsat Classification. **Remote Sensing**. 11(14): 1719.
- MOHURD. (2017). **China Urban-Rural Construction Statistical Yearbook**. Beijing, China: China Planning Press.

- Myneni, R. B. (1995). The interpretation of spectral vegetation indexes. **IEEE Transaction on Geoscience and Remote Sensing**. 33: 481-486.
- Nijbroek, R., Piikki, K., Söderström, M., Kempen, B., Turner, K. G., Hengari, S., and Mutua, J. (2018). Soil Organic Carbon Baselines for Land Degradation Neutrality: Map Accuracy and Cost Tradeoffs with Respect to Complexity in Otjozondjupa, Namibia. **Sustainability**. 10(5): 1610.
- NJMBS. (2017). **Statistical Yearbook of Nanjing**. Nanjing, China: Nanjing Municipal Bureau of Statistics.
- Pielke, R. A. (2005). Land Use and Climate Change. **Science**. 310(5754): 1625-1626.
- Qin, Z., Karnieli, A., and Berliner, P. (2001). A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. **International Journal of Remote Sensing**. 22(18): 3719-3746.
- Ready, P., and Wintz, P. (1973). Information Extraction, SNR Improvement, and Data Compression in Multispectral Imagery. **IEEE Transactions on Communications**. 21(10): 1123-1131.
- Roerink, G. J., Menenti, M., and Verhoef, W. (2000). Reconstructing cloudfree NDVI composites using Fourier analysis of time series. **International Journal of Remote Sensing**. 21(9): 1911-1917.
- Rouse, J. W., Haas, R. H., Schell, J. A., and Deering, D. W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. Paper presented at the **Proceedings of the Third ERTS Symposium**, Washington DC.
- Roy, D. P., Wulder, M. A., Loveland, T. R., C.E. W., Allen, R. G., Anderson, M. C., Helder, D., Irons, J. R., Johnson, D. M., Kennedy, R., Scambos, T. A., Schaaf,

- C. B., Schott, J. R., Sheng, Y., Vermote, E. F., et al. (2014). Landsat-8: Science and product vision for terrestrial global change research. **Remote Sensing of Environment**. 145: 154-172.
- Running, S. W., Baldocchi, D. D., Turner, D. P., Gower, S. T., Bakwin, P. S., and Hibbard, K. A. (1999). A Global Terrestrial Monitoring Network Integrating Tower Fluxes, Flask Sampling, Ecosystem Modeling and EOS Satellite Data. **Remote Sensing of Environment**. 70(1): 108-127.
- Running, S. W., Loveland, T. R., and Pierce, L. L. (1994). A vegetation classification logic-based on remote-sensing for use in global biogeochemical models. **Ambio**. 23(1): 77-81.
- Sellers, P. J., Tucker, C. J., Collatz, G. J., Los, S. O., Justice, C. O., Dazlich, D. A., and Randall, D. A. (1996). A Revised Land Surface Parameterization (SiB2) for Atmospheric GCMS. Part II: The Generation of Global Fields of Terrestrial Biophysical Parameters from Satellite Data. **Journal of Climate**. 9(4): 706-737.
- Sexton, J. O., Urban, D. L., Donohue, M. J., and Song, C. (2013). Long-term land cover dynamics by multi-temporal classification across the Landsat-5 record. **Remote Sensing of Environment**. 128: 246-258.
- Shang, H., Jia, L., and Menenti, M. (2015). Analyzing the Inundation Pattern of the Poyang Lake Floodplain by Passive Microwave Data. **Journal of Hydrometeorology**. 16(2): 652-667.
- Shiffler, R. E. (1988). Maximum Z Scores and Outliers. **The American Statistician**. 42(1): 79-80.
- Shingare, P. P., Hemane, P. M., and Dandekar, D. S. (2014). Fusion classification of multispectral and panchromatic image using improved decision tree algorithm.

Paper presented at the **2014 International Conference on Signal Propagation and Computer Technology (ICSPCT 2014)**, Rajasthan.

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

Sobrino, J. A., Jiménez-Muñoz, J. C., Soria, G., Romaguera, M., Moreno, L. G. A.-J., Plaza, A., and Martínez, P. (2008). Land Surface Emissivity Retrieval From Different VNIR and TIR Sensors. **IEEE Transactions on Geoscience and Remote Sensing**. 46(2): 316-327.

Su, Z. (2002). The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. **Hydrology and Earth System Sciences**. 6(1): 85-99.

Tang, D., Fan, H., and Zhang, Y. (2017). Review on Landsat Time Series Change Detection Methods. **Journal of Geo-information Science**. 19(8): 1069-1079.

Tortora, R. D. (1978). A Note on Sample Size Estimation for Multinomial Populations. **American Statistician**. 32: 100-102.

Townshend, J. R. G., Justice, C. O., Skole, D., Malingreau, J. P., Cihlar, J., Teillet, P., Sadowski, F., and Ruttenberg, S. (1994). The 1 km resolution global data set: needs of the International Geosphere Biosphere Programme. **International Journal of Remote Sensing**. 15(17): 3417-3441.

Tuia, D., Persello, C., and Bruzzone, L. (2016). Domain Adaptation for the Classification of Remote Sensing Data: An Overview of Recent Advances. **IEEE Geoscience and Remote Sensing Magazine**. 4(2): 41-57.

- USGS. (2017). **Landsat 8 OLI and TIRS Calibration Notices**.
<https://landsat.usgs.gov/landsat-8-18-operational-land-imager-oli-and-thermal-infrared-sensor-tirs>.
- USGS. (2019a). **Landsat 4-7 Surface Reflectance (LEDAPS) Product Guide**. USA: Department of the Interior, U.S. Geological Survey.
- USGS. (2019b). **Landsat 8 Surface Reflectance Code(LaSRC) Product Guide**. USA: Department of the Interior, U.S. Geological Survey.
- Verbesselt, J., Hyndman, R., Newnham, G., and Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. **Remote Sensing of Environment**. 114(1): 106-115.
- Verbesselt, J., Hyndman, R., Zeileis, A., and Culvenor, D. (2010). Phenological change detection while accounting for abrupt and gradual trends in satellite image time series. **Remote Sensing of Environment**. 114(12): 2970-2980.
- Verhoef, W., Menenti, M., and Azzali, S. (1996). Cover A colour composite of NOAA-AVHRR-NDVI based on time series analysis (1981-1992). **International Journal of Remote Sensing**. 17(2): 231-235.
- Vicente-Serrano, S. M., Pérez-Cabello, F., and Lasanta, T. (2008). Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of Landsat images. **Remote Sensing of Environment**. 112(10): 3916-3934.
- Vogelmann, J. E., Gallant, A. L., Shi, H., and Zhu, Z. (2016). Perspectives on monitoring gradual change across the continuity of Landsat sensors using time-series data. **Remote Sensing of Environment**. 185: 258-270.

- Voogt, J. A., and Oke, T. R. (2003). Thermal remote sensing of urban climates. **Remote Sensing of Environment**. 86(3): 370-384.
- Wang, F., Qin, Z., Song, C., Tu, L., Karnieli, A., and Zhao, S. (2015). An Improved Mono-Window Algorithm for Land Surface Temperature Retrieval from Landsat 8 Thermal Infrared Sensor Data. **Remote Sensing**. 7(4): 4268-4289.
- Weng, Q. (2009). Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. **ISPRS Journal of Photogrammetry and Remote Sensing**. 64(4): 335-344.
- Weng, Q., Fu, P., and Gao, F. (2014). Generating daily land surface temperature at Landsat resolution by fusing Landsat and MODIS data. **Remote Sensing of Environment**. 145: 55-67.
- Weng, Q., Lu, D., and Schubring, J. (2004). Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. **Remote Sensing of Environment**. 89(4): 467-483.
- White, M. A., de Beurs, K. M., Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P., O'Keefe, J., Zhang, G., Nemani, R. R., van W.J.D, L., Brown, J. F., de Wit, A., Schaepman, M., Lin, X., Dettinger, M., et al. (2009). Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982-2006. **Global Change Biology**. 15(10): 2335-2359.
- Wit, A., and Su, B. (2005). Deriving phenological indicators from SPOT-VGT data using the HANTS algorithm. Paper presented at the **2nd International VEGETATION User Conference**, Luxembourg.

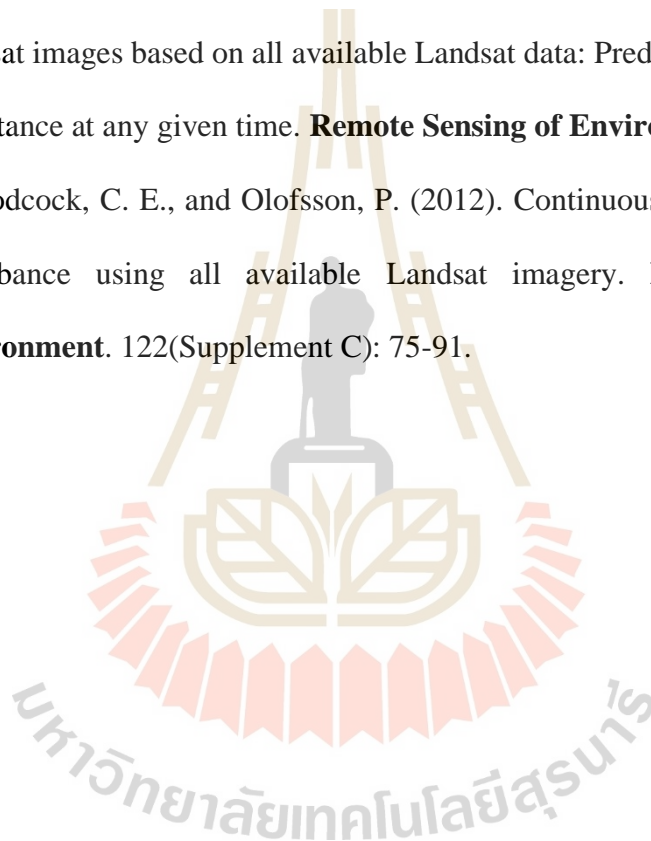
- Woodcock, C. E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S. N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P. S., Vermote, E. F., et al. (2008). Free access to Landsat imagery. **Science**. 320(5879): 1011.
- Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. **International Journal of Remote Sensing**. 27(14): 3025-3033.
- Yang, X., and Lo, C. P. (2002). Using a time series of satellite imagery to detect land use and land cover changes in the Atlanta, Georgia metropolitan area. **International Journal of Remote Sensing**. 23(9): 1775-1798.
- Zha, Y., Gao, J., and Ni, S. (2003). Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. **International Journal of Remote Sensing**. 24(3): 583-594.
- Zhou, J., Jia, L., and Menenti, M. (2015). Reconstruction of global MODIS NDVI time series: Performance of Harmonic ANalysis of Time Series (HANTS). **Remote Sensing of Environment**. 163: 217-228.
- Zhu, Z., Fu, Y., Woodcock, C. E., Olofsson, P., Vogelmann, J. E., Holden, C., Wang, M., Dai, S., and Yu, Y. (2016). Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: A case study from Guangzhou, China (2000–2014). **Remote Sensing of Environment**. 185: 243-257.
- Zhu, Z., and Woodcock, C. E. (2014a). Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for

monitoring land cover change. **Remote Sensing of Environment**. 152: 217-234.

Zhu, Z., and Woodcock, C. E. (2014b). Continuous change detection and classification of land cover using all available Landsat data. **Remote Sensing of Environment**. 144(Supplement C): 152-171.

Zhu, Z., Woodcock, C. E., Holden, C., and Yang, Z. (2015). Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. **Remote Sensing of Environment**. 162: 67-83.

Zhu, Z., Woodcock, C. E., and Olofsson, P. (2012). Continuous monitoring of forest disturbance using all available Landsat imagery. **Remote Sensing of Environment**. 122(Supplement C): 75-91.



CURRICULUM VITAE

Name: Mr. Jing Sun
Date of Birth: 11th June 1985
Place of Birth: Nanjing, Jiangsu Province, China

Education:

2007: Bachelor of Science in Geographic Information System, Nanjing University of Information Science & Technology, Nanjing, Jiangsu Province, China
2012: Master of Science in Cartography and Geographic Information System, Anhui Normal University, Wuhu, Anhui Province, China

Grant and Fellowships: SUT-Ph.D. Scholarship, Suranaree University of Technology

Position and Place of Work:

Lecturer, Department of Geographic Information Science, School of Architectural Engineering, Tongling University, Tongling, Anhui Province, China

Research Projects:

Natural Science Research Project of Colleges and Universities in Anhui Province:
Spatio-temporal Land Use and Land Cover Change Study in Hefei City Based on Landsat Time Series Data, 2019. (in Chinese).