การเปรียบเทียบความถูกต้องของการจำแนกพื้นที่ปลูกมันสำปะหลัง และอ้อยโรงงานด้วยวิธีการจำแนกระบบผู้เชี่ยวชาญและโครงข่ายประสาทเทียม จากข้อมูลดาวเทียมธีออส

นางสาววรรณทัช เทศวัฒน์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาภูมิสารสนเทศ
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2554

A COMPARATIVE ACCURACY ASSESSMENT OF EXPERT SYSTEMS AND ARTIFICIAL NEURAL NETWORK CLASSIFICATION METHODS FOR IDENTIFICATION OF CASSAVA AND SUGARCANE AREAS USING THEOS DATA

Wannatat Tessawat

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Geoinformatics

Suranaree University of Technology

Academic Year 2011

A COMPARATIVE ACCURACY ASSESSMENT OF EXPERT SYSTEMS AND ARTIFICIAL NEURAL NETWORK CLASSIFICATION METHODS FOR IDENTIFICATION OF CASSAVA AND SUGARCANE AREAS USING THEOS DATA

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's Degree.

Thesis Examining Committee

____

(Asst. Prof. Dr. Sunya Sarapirome)

Chairperson

____

(Asst. Prof. Dr. Suwit Ongsomwang)

Member (Thesis Advisor)

____

(Asst. Prof. Dr. Songkot Dasananda)

Member

____

(Assoc. Prof. Dr. Itthi Trisirisatayawong)

Member

____

(Prof. Dr. Sukit Limpijumnong)

Vice Rector for Academic Affairs

____

(Assoc. Prof. Dr. Prapun Manyum)

Dean of Institute of Science

วรรณทัช เทศวัฒน์ : การเปรียบเทียบความถูกต้องของการจำแนกพื้นที่ปลูกมันสำปะหลังและอ้อยโรงงานด้วยวิธีการจำแนกระบบผู้เชี่ยวชาญและโครงข่ายประสาทเทียมจากข้อมูลดาวเทียมธีออส (A COMPARATIVE ACCURACY ASSESSMENT OF EXPERT SYSTEMS AND ARTIFICIAL NEURAL NETWORK CLASSIFICATION METHODS FOR IDENTIFICATION OF CASSAVA AND SUGARCANE AREAS USING THEOS DATA) อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.สุวิทย์ อ๋องสมหวัง, 169 หน้า.

การจำแนกพื้นที่มันสำปะหลังและอ้อยโรงงานเป็นสารสนเทศสำคัญสำหรับอุตสาหกรรมเอทานอล ในการศึกษาครั้งนี้ มีวัตถุประสงค์หลักคือ (1) การจำแนกพื้นที่มันสำปะหลังและอ้อยโรงงานด้วยวิธีระบบผู้เชี่ยวชาญและโครงข่ายประสาทเทียม (2) เพื่อประเมินความถูกต้องของการจำแนกมันสำปะหลังและอ้อยโรงงานด้วยวิธีระบบผู้เชี่ยวชาญอ้อยและโครงข่ายประสาทเทียมและ (3) การประเมินวิธีการและชุดข้อมูลที่เหมาะสมสำหรับการจำแนกมันสำปะหลังและอ้อยโรงงาน ในการศึกษาใช้ข้อมูลแบบหลายช่วงคลื่นจากดาวเทียมธีออสเป็นข้อมูลพื้นฐานและนำมาจัดกลุ่มรวมกับข้อมูลประกอบ ได้แก่ ข้อมูลดัชนีพืชพรรณผลต่างแบบนอร์แมลไลซ์ (NDVI) ชุดดิน (soil series) และภูมิลักษณ์ (landform) เพื่อกำหนดเป็นชุดข้อมูลจำนวน 6 ชุด สำหรับใช้ในการจำแนกมันสำปะหลังและอ้อยโรงงาน จากนั้น ทำการประเมินความถูกต้องโดยอาศัยค่าความถูกต้องโดยรวมและค่าความสอดคล้องของสัมประสิทธิ์แคปปา และนำค่าสัมประสิทธิ์แคปปาที่ได้รับไปใช้ในการประเมินหาวิธีการจำแนกและชุดข้อมูลที่เหมาะสมสำหรับใช้ในการจำแนกมันสำปะหลังและอ้อยโรงงาน

ผลการศึกษา พบว่า ระบบผู้เชี่ยวชาญและโครงข่ายประสาทเทียมที่กำหนดอัตราการเรียนรู้เท่ากับ 0.1 0.2 และ 0.3 สามารถจำแนกมันสำปะหลังและอ้อยโรงงานได้ทั้ง 6 ชุดข้อมูล ซึ่งประกอบด้วย (1) ข้อมูลแบบหลายช่วงคลื่น (2) ข้อมูลแบบหลายช่วงคลื่นและดัชนีพืชพรรณผลต่างแบบนอร์แมลไลซ์ (3) ข้อมูลแบบหลายช่วงคลื่นและชุดดิน (4) ข้อมูลแบบหลายช่วงคลื่นและภูมิลักษณ์ (5) ข้อมูลแบบหลายช่วงคลื่น ชุดดิน และภูมิลักษณ์ (6) ข้อมูลแบบหลายช่วงคลื่น ดัชนีพืชพรรณผลต่างแบบนอร์แมลไลซ์ ชุดดิน และภูมิลักษณ์

ในการประเมินความถูกต้อง พบว่า ค่าความถูกต้องโดยรวมและค่าสัมประสิทธิ์แคปปาของวิธีการจำแนกโดยระบบผู้เชี่ยวชาญ มีค่าอยู่ระหว่างร้อยละ 76.23 ถึง 79.51 และร้อยละ 63.06 ถึง 67.69 ตามลำดับ และพบว่า ชุดข้อมูลแบบหลายช่วงคลื่นและภูมิลักษณ์ให้ค่าความถูกต้องสูงสุด และชุดข้อมูลแบบหลายช่วงคลื่นให้ค่าความถูกต้องต่ำสุด ในขณะเดียวกัน ค่าความถูกต้องโดยรวมและค่าสัมประสิทธิ์แคปปาของวิธีการจำแนกแบบโครงข่ายประสาทเทียมที่มีอัตราการเรียนรู้เท่ากับ

0.1 มีค่าอยู่ระหว่างร้อยละ 70.49 ถึง 78.69 และร้อยละ 62.32 ถึง 72.79 ตามลำดับ โดยพบว่า ชุดข้อมูลแบบหลายช่วงคลื่นและชุดดินให้ค่าความถูกต้องสูงสุดและชุดข้อมูลแบบหลายช่วงคลื่นให้ค่าความถูกต้องต่ำสุด นอกจากนี้ ในการทดสอบค่า Z พบว่า ค่าความถูกต้องระหว่างชุดข้อมูลแบบหลายช่วงคลื่นและชุดข้อมูลแบบหลายช่วงคลื่นและชุดดินมีความแตกต่างกันอย่างมีนัยสำคัญ ณ ระดับความเชื่อมั่นร้อยละ 80

จากผลการประเมินความถูกต้องของการจำแนก มันสำปะหลัง อ้อยโรงงาน และประเภทอื่น ๆ โดยวิธีระบบผู้เชี่ยวชาญและโครงข่ายประสาทเทียม พบว่า ค่าสัมประสิทธิ์แคปปาของวิธีโครงข่ายประสาทเทียมที่อัตราการเรียนรู้ที่ 0.1 ให้ผลลัพธ์ที่ดีที่สุด จึงเลือกให้เป็นวิธีการที่เหมาะสมสำหรับการจำแนกมันสำปะหลังและอ้อยโรงงาน ในขณะเดียวกัน ชุดข้อมูลที่เหมาะสมสูงสุดสำหรับใช้ในการจำแนกมันสำปะหลังและอ้อยโรงงานด้วยวิธีการจำแนก โดยระบบผู้เชี่ยวชาญ ได้แก่ ชุดข้อมูลแบบหลายช่วงคลื่นและภูมิลักษณ์ ในขณะที่ชุดข้อมูลที่เหมาะสมสูงสุดสำหรับการจำแนกมันสำปะหลังและอ้อยโรงงาน ด้วยวิธีการจำแนกแบบโครงข่ายประสาทเทียม ได้แก่ ชุดข้อมูลแบบหลายช่วงคลื่นและชุดดิน

จากผลการศึกษาสามารถสรุปได้ว่า ระบบผู้เชี่ยวชาญและโครงข่ายประสาทเทียมสามารถนำมาใช้จำแนกพื้นที่ปลูกมันสำปะหลังและอ้อยโรงงานจากข้อมูลดาวเทียมธีออสได้ โดยมีค่าสัมประสิทธิ์แคปปาอยู่ในระดับปานกลาง (ร้อยละ 40-80) นอกจากนี้ การจัดกลุ่มชุดข้อมูลระหว่างข้อมูลแบบหลายช่วงคลื่นของข้อมูลดาวเทียมธีออสกับข้อมูลทุติยภูมิ สามารถเพิ่มความถูกต้องในการจำแนกมันสำปะหลังและอ้อยโรงงานด้วยระบบผู้เชี่ยวชาญและระบบโครงข่ายประสาทเทียม

สาขาวิชาการรับรู้จากระยะไกล	ลายมือชื่อนักศึกษา________________________
ปีการศึกษา 2554	ลายมือชื่ออาจารย์ที่ปรึกษา__________________

WANNATAT TESSAWAT : A COMPARATIVE ACCURACY ASSESSMENT OF EXPERT SYSTEMS AND ARTIFICIAL NEURAL NETWORK CLASSIFICATION METHODS FOR IDENTIFICATION OF CASSAVA AND SUGARCANE AREAS USING THEOS DATA. THESIS ADVISOR : ASST. PROF. SUWIT ONGSOMWANG, Dr. rer. Nat. 169 PP.

EXPERT SYSTEM / ARTIFICIAL NEURAL NETWORK / CASSAVA / SUGARCANE / LAND COVER CLASSIFICATION

Classification of cassava and sugarcane areas is important information for ethanol industry. In this study, main objectives are (1) to classify cassava and sugarcane using Expert System and ANN, (2) to assess the accuracy of cassava and sugarcane classification using Expert System and ANN and (3) to evaluate an optimum method and dataset for cassava and sugarcane classification. In this study, multispectral data of THEOS were used as basic data and combined with an additional data including NDVI, soil series and landform to define 6 datasets for cassava and sugarcane extraction. Then, overall accuracy and Kappa hat coefficient of agreement were applied for accuracy assessment and the Kappa coefficients were used to identify an optimum method and dataset for cassava and sugarcane classification.

As results, Expert System and ANN with 0.1, 0.2 and 0.3 learning rate can be applied for cassava and sugarcane classification from all six datasets included (1) multispectral dataset (2) multispectral and NDVI dataset (3) multispectral and soil series dataset (4) multispectral and landform dataset (5) multispectral, soil series and landform dataset and (6) multispectral, NDVI, soil series and landform dataset.

For accuracy assessment, overall accuracy and Kappa coefficient of Expert System varied between 76.23 and 79.51% and 63.06 and 67.69%, respectively. The multispectral and landform dataset showed the highest accuracy and multispectral dataset presented the lowest accuracy. Meanwhile, overall accuracy and Kappa coefficient of ANN with 0.1 learning rate varied between 70.49 and 78.69% and 62.32 and 72.79%, respectively. The multispectral and soil series dataset showed the highest accuracy and multispectral dataset presented the lowest accuracy. Also, accuracy difference between multispectral dataset and multispectral and soil series dataset was significantly different based on Z statistic at 80% of confidence level.

Based on accuracy assessment for cassava, sugarcane and others classification of Expert System and ANN, Kappa coefficient of ANN with 0.1 learning rate, which provided the best results, was selected as an optimum method for sugarcane and cassava classification. At the same time, an optimum dataset for sugarcane and cassava classification using Expert System was multispectral and landform dataset while an optimum dataset for sugarcane and cassava classification using ANN was multispectral and soil series dataset.

In conclusion, Expert System and ANN can be used to classify cassava and sugarcane areas from THEOS data at moderate level of Kappa coefficient (40-80%). In addition, combination of THEOS multispectral and additional data can increase accuracy for cassava and sugarcane classification using Expert System and ANN.

School of Remote Sensing Student's Signature____________________

Academic Year 2011 Advisor's Signature____________________

ACKNOWLEDGEMENTS

This thesis would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First and foremost, my utmost gratitude to my advisor, Asst. Prof. Dr. Suwit Ongsomwang for his invaluable advice to complete this work. I also sincerely thank him for sparing his immensely valuable time to read this thesis. He not only provided immense knowledge but his logical ways of thinking have been of great value for me.

Besides my advisor, I am deeply grateful to Asst. Prof. Dr. Sunya Sarapirome, Asst. Prof. Dr. Songkot Dasananda, Assoc. Prof. Dr. Itthi Trisirisatayawong for serving as committee members and appreciated suggestions. I would also like to thank Dr. Dusdi Chanlikit for his helpful guidance.

I gratefully acknowledge Suranaree University of Technology for providing academic and also financial support to my research. Furthermore, I would like to express my sincere gratitude to Geo-Informatics and Space Technology Development Agency (Public Organization) for supporting THEOS data.

Special thanks go to staffs in Remote Sensing laboratory, Mr. Tinn Thirakultomorn and Mr. Winai Yaowaret for guidance and preparation of field survey map. I would also like to thank Ms. Phenpraphai Phuthong, Mr. Somporn Chobtham, Mr. Pitak Chailungka and Mr. Patiwat Littidej for kind contribution in ground checking.

I wish to thank Mr. Pichai Wongsawat and Ms. Warunee Aunphoklang for their suggestions about additional data which is helpful in this thesis. In addition, I am greatly thankful to Mr. Nutthapol Junkeaw, Mr. Anuchit Phayakkin, Mr. Rawee Rattanakom and Ms. Siriwan Ruamkaew for guiding and teaching me the essential techniques.

My warm thanks are due to my friends from the School of Remote Sensing for their kind help and moral support that made me stay with much joy and happy moments. I would also like to express my apology that I could not mention one by one personally.

Finally, a note of thanks goes to my beloved parents and my sister whose dedication, love and persistent confidence on me. I owe my loving thanks to them for being patient and encouraging me to finish this work. My graduation would not be achieved without their understanding and encouragement.

Wannatat Tessawat

CONTENTS

Page

ABSTRACT IN THAI I

ABSTRACT IN ENGLISH III

ACKNOWLEDGEMENTS V

CONTENTS VII

LIST OF TABLES XI

LIST OF FIGURES XVIII

LIST OF ABBREVIATIONS XXI

CHAPTER

I INTRODUCTION 1

1.1 Significant of the problem 1

1.2 Research objectives 2

1.3 Scope of the study 3

1.4 Study area 4

1.5 Benefit of the study 6

II LITERATURE REVIEW 7

2.1 Basic information of cassava and sugarcane 7

2.1.1 General background of cassava 7

2.1.2 General background of sugarcane 8

2.2 THEOS satellite 11

CONTENTS (Continued)

Page

2.2.1 THEOS instruments 11

2.3 Expert Systems 13

2.3.1 IMAGINE Expert Classifier 14

2.4 Artificial Neural Network (ANN) 18

2.4.1 Architecture of Neural Networks 20

2.4.2 Neural network configuration in ENVI 24

2.5 Literature review 28

2.5.1 LULC classification using Expert System 28

2.5.2 LULC classification using ANN 31

III DATA, EQUIPMENT AND METHODOLOGY 35

3.1 Data and equipment 35

3.2 Methodology 36

3.2.1 Data collection and preparation 38

3.2.2 Field survey 40

3.2.3 Data analysis 42

3.2.3.1 Image classification 42

3.2.3.2 Accuracy assessment 44

3.2.3.3 Evaluation of optimum classification methods and dataset 45

CONTENTS (Continued)

Page

IV RESULTS AND DISCUSSION 47

4.1 Data collection and preparation 47

4.1.1 Remote sensing data 47

4.1.2 GIS data 54

4.1.3 Dataset preparation 59

4.2 Sugarcane and cassava classification 61

4.2.1 Cassava and sugarcane classification using Expert System 61

4.2.2 Sugarcane and cassava classification using ANN 69

4.3 Ground truth and accuracy assessment 83

4.3.1 Calculate samples point for accuracy assessment 83

4.3.2 Accuracy assessment of Expert System 85

4.3.3 Accuracy assessment of ANN 89

4.3.3.1 ANN with learning rate 0.1 90

4.3.3.2 ANN with learning rate 0.2 96

4.3.3.3 ANN with learning rate 0.3 102

4.4 Evaluation of optimum methods for sugarcane and cassava classification 109

4.5 Evaluation of optimum dataset for sugarcane and cassava classification 111

4.5.1 Evaluation of optimum dataset for sugarcane and cassava classification using Expert system 111

CONTENTS (Continued)

Page

4.5.2 Evaluation of optimum dataset for sugarcane and cassava classification using ANN 112

V CONCLUSIONS AND RECOMMENDATIONS 116

5.1 Conclusions 116

5.1.1 Cassava and sugarcane classification using Expert System and ANN 116

5.1.2 Accuracy assessment of cassava and sugarcane classification using Expert System and ANN 117

5.1.3 Evaluation of optimum methods for sugarcane and cassava classification between Expert System and ANN 118

5.1.4 Evaluation of optimum dataset for sugarcane and cassava classification using Expert System and ANN 118

5.2 Recommendations 119

REFERENCES 120

APPENDICES 126

APPENDIX A RECTIFIED THEOS DATA 127

APPENDIX B FIELD SURVEY DATA 129

APPENDIX C ERROR MATRICES 136

APPENDIX D VARIANCE AND Z STATISTIC OF ERROR MATRICES 167

CURRICULUM VITAE 169

LIST OF TABLES

Table **Page**

2.1 Crop calendar of cassava 8

2.2 Crop calendar of sugarcane 10

2.3 THEOS panchromatic and multispectral sensor system characteristics 12

2.4 Advantages and disadvantage of Expert System 18

2.5 Advantage and disadvantage of ANN algorithm 27

3.1 Data and equipment in this study 36

3.2 Neural Network configuration for LULC classification under ENVI 45

4.1 Metadata of THEOS data 49

4.2 Area and percentage soil type in the study area 57

4.3 Combination of remotely sensed data and GIS data in each dataset for cassava and sugarcane area classification 60

4.4 Structure of hypothesis, rule and conditions for each dataset 63

4.5 Area and percentage of cassava and sugarcane classification from each dataset using Expert System 67

4.6 Area and percentage of land cover classification from each dataset using ANN with learning rate 0.1 72

4.7 Area and percentage of land cover classification form each dataset using ANN with learning rate 0.2 76

LIST OF TABLES (Continued)

Table **Page**

4.8 Area and percentage of land cover classification from each dataset using ANN with learning rate 0.3 80

4.9 Comparison of extracted area of cassava and sugarcane areas from each dataset using ANN with different learning rate 82

4.10 Summary of overall accuracy and Kappa coefficient for land cover classification by Expert Systems 86

4.11 Summary of simple PA and UA and conditional Kappa PA and UA for each category by Expert System Classification 87

4.12 Summary of overall Accuracy and Kappa coefficient for land cover classification by ANN with learning rate 0.1 94

4.13 Summary of simple PA and UA and Conditional Kappa PA and UA for each Category by ANN Classification with learning rate 0.1 95

4.14 Summary of overall accuracy and Kappa coefficient for land cover classification by ANN with learning rate 0.2 100

4.15 Summary of Simple PA and UA and conditional Kappa PA and UA for each Category by ANN Classification with learning rate 0.2 101

4.16 Summary of overall accuracy and Kappa coefficient for land cover classification by ANN with learning rate 0.3 106

4.17 Summary of simple PA and UA and conditional Kappa PA and UA for each Category by ANN Classification with learning rate 0.3 107

LIST OF TABLES (Continued)

Table **Page**

4.18 Accuracy assessment comparison of ANN classification by different learning rate 108
4.19 Accuracy assessment of ANN classification for three common categories (cassava, sugarcane and others) 110
4.20 Comparison of Kappa coefficient between Expert System and ANN classification by each dataset 111
4.21 Significant difference test by Z statistic for Expert System 112
4.22 Significant difference test by Z statistic for ANN with learning rate 0.1 113
4.23 Significant difference test by Z statistic for ANN with learning rate 0.2 114
4.24 Significant difference test by Z statistic for ANN with learning rate 0.3 115
A.1 The RMS error of image rectification... 127
B.1 Detail of sampling point.. 129
B.2 Ground photograph of each land cover class from field survey 133
C.1 Error matrices of Expert Systems classification with multispectral dataset. 136
C.2 Error matrices of Expert Systems classification with multispectral and NDVI dataset 137
C.3 Error matrices of Expert Systems classification with multispectral and soil series dataset 137
C.4 Error matrices of Expert Systems classification with multispectral and landform dataset.. 138

LIST OF TABLES (Continued)

Table **Page**

C.5 Error matrices of Expert Systems classification with multispectral, soil series and landform dataset ... 138

C.6 Error matrices of Expert Systems classification with multispectral, NDVI, soil series and landform dataset ... 139

C.7 Error matrices of ANN classification with training rate 0.1 with multispectral dataset... 140

C.8 Error matrices of ANN classification with training rate 0.1 with multispectral and NDVI dataset... 141

C.9 Error matrices of ANN classification with training rate 0.1 with multispectral and soil series dataset... 142

C.10 Error matrices of ANN classification with training rate 0.1 with multispectral and landform dataset... 143

C.11 Error matrices of ANN classification with training rate 0.1 with multispectral, soil series and landform dataset ... 144

C.12 Error matrices of ANN classification with training rate 0.1 with multispectral, NDVI, soil series and landform dataset ... 145

C.13 Error matrices of ANN classification with training rate 0.2 with multispectral dataset... 146

C.14 Error matrices of ANN classification with training rate 0.2 with multispectral and NDVI dataset... 147

LIST OF TABLES (Continued)

Table **Page**

C.15 Error matrices of ANN classification with training rate 0.2 with multispectral and soil series dataset 148

C.16 Error matrices of ANN classification with training rate 0.2 with multispectral and landform dataset 149

C.17 Error matrices of ANN classification with training rate 0.2 with multispectral, soil series and landform dataset 150

C.18 Error matrices of ANN classification with training rate 0.2 with multispectral, NDVI, soil series and landform dataset 151

C. 19 Error matrices of ANN classification with training rate 0.3 with multispectral dataset........ 152

C. 20 Error matrices of ANN classification with training rate 0.3 with multispectral and NDVI dataset........ 153

C.21 Error matrices of ANN classification with training rate 0.3 with multispectral and soil series dataset 154

C.22 Error matrices of ANN classification with training rate 0.3 with multispectral and landform dataset 155

C.23 Error matrices of ANN classification with training rate 0.3 with multispectral, soil series and landform dataset 156

C.24 Error matrices of ANN classification with training rate 0.3 with multispectral, NDVI, soil series and landform dataset 157

LIST OF TABLES (Continued)

Table **Page**

C.25 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral dataset. 158

C.26 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral and NDVI dataset 158

C.27 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral and soil series dataset. 159

C.28 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral and landform dataset............................ 159

C.29 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral, soil series and landform dataset........... 160

C.30 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral, NDVI, soil series and landform dataset 160

C.31 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral dataset. 161

C.32 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral and NDVI dataset. 161

C.33 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral and soil series dataset. 162

C.34 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral and landform dataset............................ 162

LIST OF TABLES (Continued)

Table **Page**

C.35 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral, soil series and landform dataset. 163

C.36 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral, NDVI, soil series and landform dataset. 163

C.37 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral dataset 164

C.38 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral and NDVI dataset 164

C.39 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral and soil series dataset. 165

C.40 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral and landform dataset............................ 165

C.41 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral, soil series and landform dataset. 166

C.42 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral, NDVI, soil series and landform dataset. 166

D.1 The variance and Z statistic of Expert System classification 167

D.2 The variance and Z statistic of ANN classification 167

LIST OF FIGURES

Figure		Page
1.1	The study area at Chakkarat district, Nakhon Ratchasima province	4
2.1	The Spectral reflectance curves of THEOS multispectral sensor	12
2.2	Examples of a decision tree branch	15
2.3	Split rule decision tree branch	15
2.4	A knowledge-base expert system classification of white fir on Maple Mountain, Utah State, United States of America	17
2.5	Comparison of biological and artificial neurons	19
2.6	The structure of ANN data flow in a feedforward neural network, data flow from the input layer to the hidden layer(s) and eventually to the output layer in one direction	21
2.7	The structure of data flow in a feed-forward and backpropagation neural network	24
2.8	Activation functions: (a) logistic or sigmoid and (b) hyperbolic	26
3.1	The framework of methodology	37
3.2	Sample of training areas for Expert System algorithm	41
3.3	Sample of training points for ANN algorithm	41
4.1	True color composite image of THEOS data in the study area	49
4.2	Color orthophoto of MOAC used in the study	50
4.3	Distribution of GCPs for THEOS image rectification	51

LIST OF FIGURES (Continued)

Figure **Page**

4.4 Rectified THEOS data of the study area 52

4.5 Distribution of NDVI value in the study area 53

4.6 Distribution of major land use types in 2007 in the study area 54

4.7 Distribution of soil type in the study area 57

4.8 Distribution of landform type in the study area 58

4.9 Distribution of sugarcane plot data in 2010 in the study area 69

4.10 Distribution of cassava, sugarcane and unclassified areas from each dataset using Expert System 65

4.11 Distribution of cassava, sugarcane and unclassified areas from each dataset using Expert System 66

4.12 The proportion chart of Expert Systems classification in each dataset 68

4.13 Land cover distribution from each dataset using ANN with 0.1 learning rate 70

4.14 Land cover distribution from each dataset using ANN with 0.1 learning rate 71

4.15 Proportion of land cover areas in each dataset using ANN classification with learning rate 0.1 73

4.16 Land cover distribution from each dataset using ANN with 0.2 learning rate 74

4.17 Land cover distribution from each dataset using ANN with 0.2 learning rate 75

LIST OF FIGURES (Continued)

Figure **Page**

4.18 Proportion of land cover areas in each dataset using ANN classification with learning rate 0.2 77

4.19 Land cover distribution from each dataset using ANN with 0.3 learning rate.. 78

4.20 Land cover distribution from each dataset using ANN with 0.3 learning rate.. 79

4.21 Proportion of land cover areas in each dataset using ANN classification with learning rate 0.3 81

4.22 Distribution of ground truth location for accuracy assessment 84

LIST OF ABBREVIATIONS

ANN	=	Artificial Neural Network
ASTER	=	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AVI	=	Advance Vegetation Index
BI	=	Bare Soil Index
BV	=	Brightness Value
ºC	=	Celsius
cm	=	Centimeters
CCD	=	Charge Coupled Devices
DA	=	Discriminant Analysis
DEM	=	Digital Elevation Model
ETM+	=	Enhance Thematic Mapper Plus
GCPs	=	Ground Control Points
GIS	=	Geographic Information Systems
GISTDA	=	Geo-Informatics and Space Technology Development Agency (Public Organization)
GPS	=	Global Positioning System
km	=	Kilometer
LDD	=	Land Development Department
m	=	Meter

LIST OF ABBREVIATIONS (Continued)

mm	=	Millimeter
MLC	=	Maximum Likelihood Classification
MOAC	=	Ministry of Agriculture and Cooperatives
MOU	=	Memorandum of Understanding
MS	=	Multispectral
NDVI	=	Normalized Differential Vegetation Index
NIR	=	Near Infrared
OCSB	=	Office of the Cane and Sugar Board
PA	=	Producer's Accuracy
PAN	=	Panchromatic
pH	=	Potential of Hydrogen ion
RMS error	=	The Root-Mean-Rquare error
sq. km	=	Square Kilometer
SPOT	=	Satellites Pour l'Observation de la Terre
SVMs	=	Support Vector Machines
THEOS	=	Thailand Earth Observation Satellite
TM	=	Thematic Mapper
UA	=	User's Accuracy
UTM	=	Universal Transverse Mercator
WGS	=	World Geodetic System
WI	=	Water Index
µm	=	Micrometer/ Micron

CHAPTER I

INTRODUCTION

1.1 Significant of the problem

At present, oil prices are trend to increasing continuously and the result from the upward trend effect to explore new energy sources as an alternative energy. Thailand has various crops which are potentially used as raw material to produce oil substitution. Therefore, the government encourages farmers and private companies to cultivate more alternative energy crops such as cassava, sugarcane, oil palm and maize for energy production.

In part of ethanol production, cultivation area of cassava and sugarcane has been expanded to support ethanol industry. In 2008-2009 Agricultural Land Reform Office had signed MOU with two ethanol industries for promotion on cassava cultivation to the farmers in land reform areas of 8 provinces including Nakhon Ratchasima, Buriram, Chaiyaphum, Khon Kaen, Ubon Ratchathani, Si Sa Ket, Amnat Charoen and Yasothon provinces. In addition, under special system for specific area of cassava cultivation of Department of Agricultural Extension for year 2007/2008, cassava production in Nakhon Ratchasima and Buriram provinces had been intensively managed by linkage the sources to markets and industrial sites for increasing the price (สำนักงานเกษตรจังหวัดนครราชสีมา, 2550).

Furthermore, main crops in Nakhon Ratchasima province are sugarcane, maize, cassava and rice. The most productive area of crop in this province is cassava which is the largest cultivation site in Thailand. Also, Office of Agricultural Economics had reported that there were 6 permitted ethanol industries from sugarcane, cassava, molasses and cassava waste located in Nakhon Ratchasima province in 2008 (สำนักงานเศรษฐกิจการเกษตร, 2551). However, cultivation areas of cassava and sugarcane which had been practiced for a long time as economic crops in the northeast region vary according to price incentives.

Thus, this research aimed to study how to extract cassava and sugarcane areas from remotely sensed data with Expert Systems and Artificial Neural Network (ANN) algorithm. Herein, multispectral data of Thailand Earth Observation Satellite (THEOS) was used as basic data for cassava and sugarcane classification. In addition, an optimum method and datasets for cassava and sugarcane classification were also investigated in details based on accuracy assessment.

1.2 Research objectives

The specific objectives of accuracy comparison for cassava and sugarcane areas identification using Expert Systems and ANN classification are as follows:

1.2.1 To classify cassava and sugarcane cultivation area using Expert Systems and ANN;

1.2.2 To assess the accuracy of cassava and sugarcane classification using Expert Systems and ANN;

1.2.3 To evaluate an optimum method and dataset for cassava and sugarcane classification.

1.3 Scope of the study

The study aims to evaluate the optimum method and dataset for cassava and sugarcane classification based on the accuracy of Expert Systems and ANN. The scope of the study is briefly described as following.

1.3.1 Cassava and sugarcane classification with Expert Systems was conducted using ERDAS Imagine while, cassava and sugarcane classification with and ANN was performed using ENVI.

1.3.2 Land use and land cover categories for Expert Systems is consist of cassava, sugarcane and unclassified type, while for ANN is compose of urban and built-up area, paddy field, cassava, sugarcane, forest and trees, water bodies, and bare land.

1.3.2 Six datasets which were used for cassava and sugarcane classification with Expert Systems and ANN included:

(1) Multispectral data of THEOS

(2) Multispectral data of THEOS and NDVI

(3) Multispectral data of THEOS and Soil series

(4) Multispectral data of THEOS and Landform

(5) Multispectral data of THEOS and Soil series and Landform

(6) Multispectral data of THEOS and NDVI, Soil series and Landform

1.3.4 Accuracy assessment of Expert Systems and ANN classification included overall accuracy with producer's accuracy (PA) and user's accuracy (UA) and Kappa coefficient of agreement with conditional PA and UA were calculated from the same set of field survey.

1.4 Study area

1.4.1 Location

The study area situates in Chakkarat district which is located in the east of Nakhon Ratchasima province between 220300.5 to 230800.5 E and 1647287.5 to 1667535.5 N under WGS 1984 datum of UTM coordinate with the total area of 212.625 sq. km or 132,890.625 Rai. It includes 6 sub-districts: Chakkarat, Si Lako, Hin Khon, Nong Kham, Nong Pluang, Khlong Mueang. (Figure 1.1)

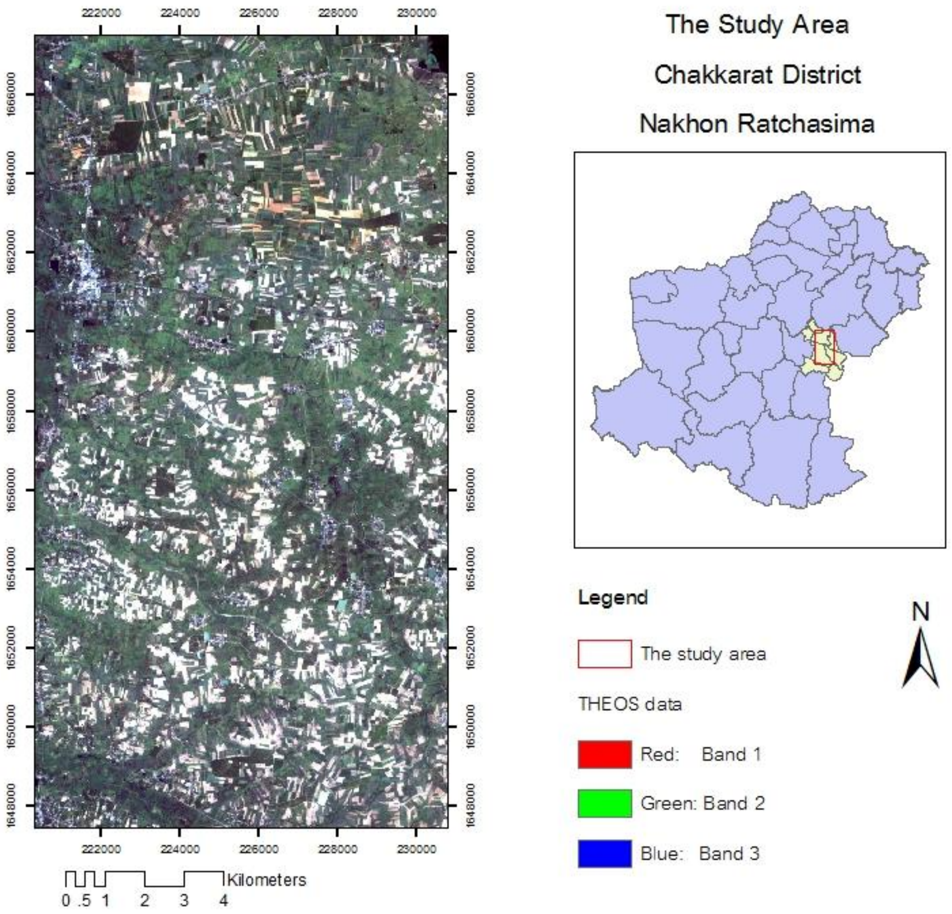


Figure 1.1 The study area at Chakkarat district, Nakhon Ratchasima province.

1.4.2 Topography

The topography is quite flats and slightly undulates in some areas. The cultivated areas are paddy field, field crops and orchard. The flat plain areas in the middle of district are covered by paddy field. Main water resources are Nong Chok and Nong Klar pond and Lam Chakkarat stream that flows from south to the north and the gullies flow through south of the area.

1.4.3 Climate

The climate of the study area is classified as a tropical monsoon climate, which is divided into 3 seasons:

- Rainy season from May to October, influenced by the southwest monsoon. The weather is damp with abundant rainfall. The maximum rainfall is in September and average raining day about 120.3 days with 221.8 mm
- Winter season is from November to February, influenced by the northeast monsoon. The wind is cold and dry and the minimum average temperature in December is 23.3 ºC.
- Summer season from March to April is hot and warm. The average maximum temperature in April is 29.8 ºC and average annual temperature is 27.1 ºC. (สำนักสำรวจดินและวางแผนการใช้ที่ดิน, 2551)

1.4.4 Land use

Based on land use data of Nakhon Ratchasima province in 2007 of Land Development Department, it shows that the most dominant land use in the study area is agricultural areas comprise of paddy field (78.35 sq. km), cassava (48.69 sq. km), and sugarcane (44.18 sq. km) (กรมพัฒนาที่ดิน, 2550).

1.4.5 Land suitability assessment

According to land suitability assessment based on FAO framework of Office of Soil Survey and Land Use Planning, Land Development Department, suitable class for cassava and sugarcane in Chakkarat district are moderately suitable and marginally suitable, respectively (สำนักสำรวจดินและวางแผนการใช้ที่ดิน, 2551).

1.5 Benefit of the study

1.5.1 The optimum method (Expert Systems or ANN) for cassava and sugarcane classification are identified. This finding will be useful to image analyst for extraction of cassava and sugarcane areas.

1.5.2 The optimum dataset for cassava and sugarcane classification with Expert Systems and ANN are identified. This discovery will be useful to planners, developers and managers in agricultural sectors for their application.

1.5.3 Results obtained from this study can be applied for monitoring and managing cassava and sugarcane cultivation in other areas.

CHAPTER II

LITERATURE REVIEW

2.1 Basic information of cassava and sugarcane

2.1.1 General background of cassava

Cassava is an important economic crop of Thailand. The plantation area in 2008 was about 6-7 million Rai. The total annual product is about 20 million tons of fresh root that can be processed into starch, pellets and other downstream industries, such as sweeteners, modified starch, and alcohol. Thailand is the largest cassava exporter of the world with value of 30,000 million baht per year. The major sources of cassava plantation areas in the country are Chaiyaphum, Kalasin, Khon Kaen, Nakhon Ratchasima, Rayong, Sa Kaeo, Chachoengsao, Chon Buri and Kanchanaburi provinces (สถาบันวิจัยพืชไร่ กรมวิชาการเกษตร, 2551).

2.1.1.1 Crop requirements of cassava

Cassava can be grown in almost any soil types but grows well in loam, sandy loam or sandy soil. Soil nutrient availability is medium and is at least 1.0 percent of organic matter. Soil depth is at least 30 cm with good drainage and good air circulation. The pH value varies between 4.5 and 8.0. The plantation areas should situate in flat plains and not be flooded with evenly distributed annual rainfall about 1,000-1,500 mm per year. The gradient of planted area should not exceed more than 5 percent (สถาบันวิจัยพืชไร่ กรมวิชาการเกษตร, 2551).

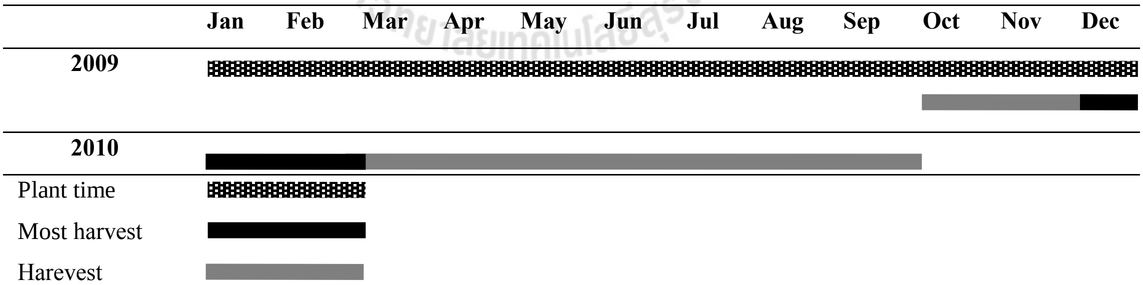
2.1.1.2 Cultivation period of cassava

The cultivation period is in the rainy season from May to June and in the end of the rainy season from October to December. Cassava can be harvested at the age of 8 months but the best practice should be in 12 months after planting. It should not be harvested during rainy season because its root will contain a low percentage of starch. (สถาบันวิจัยพืชไร่ กรมวิชาการเกษตร, 2551)

2.1.1.3 Crop calendar of cassava

Based on report of Office of Agricultural Economics in 2009 and technical document of Department of Agriculture in 2002, crop calendar of cassava can be determined as shown in Table 2.1 (กรมวิชาการเกษตร, 2545 and สำนักงานเศรษฐกิจการเกษตร, 2553)

Table 2.1 Crop calendar of cassava.



2.1.2 General background of sugarcane

Sugarcane is a main economic crop that use as raw material in sugar industry. The consumption of sugar in Thailand in 2002 was about 1.6 to 1.7 million tons per year with value of 17,000-19,000 million baht. Thailand exports to the world's sugar market more than 3 million tons with value of 20,000 to 30,000 million baht per year and it is the fourth sugar exporter of the world after Brazil, the European Union and Australia. The annual product of sugarcane is uncertainty depending on yield and plantation area in the central, northeast and east region which varies from 5.6 to 6.6 million Rai (สถาบันวิจัยพืชไร่ กรมวิชาการเกษตร, 2551).

2.1.2.1 Crop requirements of sugarcane

The plantation areas of sugarcane should be located in upland or lowland areas with no flooding. Plantation site should not higher than 1,500 meters above mean sea level with slope less than 3 percent. The optimum location of plantation site should be situated far from pollution sources and close to sugar industry plant within 60 kilometers radius.

The suitable soil for sugarcane should be loam, clay loam or sandy loam. The nutrient availability is moderate with organic matter more than 1.5 percent, phosphorus more than 10 ppm, exchangeable potassium more than 80 ppm. Soil depth should be at least 50 cm with good drainage and air circulation and pH should varies between 5.5 and 7.0 and Electrical Conductivity (EC) or salinity values should less than 4.0 decisiemens per meter.

The optimum temperature during growing stage is 30 -35 °C at day-time and 18-22 °C at night-time for maturating stage or sugarcane age about 10-11 months. Annual rainfall should be evenly distributed with amount of 1,200-1,500 mm per year during growing stage between 1 and 8 months and it should be no rain before 2 months of harvesting period.

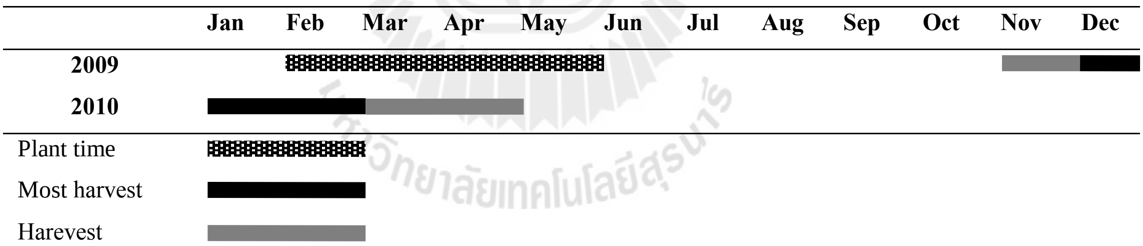
2.1.2.2 Cultivation period of sugarcane

In general, the cultivation period can be in early rainy season between February and April for irrigated area or between March and April for rainfed areas. In addition, it can also be cultivated between October and November in dry season, where soil is sandy loam area without clay or laterite layers. (กรมวิชาการเกษตร. 2545).

2.1.2.3 Crop calendar of sugarcane

Based on technical document of Department of Agriculture in 2002 and report of Office of Agricultural Economics in 2009, crop calendar of sugarcane can be determined as shown in Table 2.2 (กรมวิชาการเกษตร, 2545 and สำนักงานเศรษฐกิจการเกษตร, 2553)

Table 2.2 Crop calendar of sugarcane.



The report from Office of the Cane and Sugar Board (2010) showed that sugar industries in Nakhon Ratchasima were opened crushing in December and closed in April (สำนักงานคณะกรรมการอ้อยและน้ำตาลทราย, 2553).

Basically, crop calendar provides the phenological cycle of crop that is used to determine an optimum date for acquiring remotely sensed data. For cassava and sugarcane extraction, an optimum date should be acquired during mature stage between October and November.

2.2 THEOS Satellite

The THEOS program was developed by the Thai space agency GISTDA (Geo-Informatics and Space Technology Development Agency) with EADS Astrium as prime contractor. The satellite had a design life of five years on its operational and fuelled for minimum of 7 years in orbit (GISTDA, 2010).

2.2.1 THEOS Instruments

THEOS optical payload provides complementary images of the earth via high resolution panchromatic instrument coupled with a wide swath color instrument. The Panchromatic instrument is made exclusively of Silicium Carbide (Structure mirrors and focal plane) which ensures very high level of stability. It is designed to provide 2-meter high resolution images.

The Multispectral instrument is a 7 lens dioptric camera with 4 color filters. It is designed to provide unique large swath (90 km) images. The sensor characteristics of 2 cameras show in Table 2.3. Materials reflectance characteristic curve is shown in Figure 2.1.

The linear arrays of Charge Coupled Devices (CCD) located at the focal plane of each instruments transform the acquired and focalized radiance from ground into electronic signal. The imaging principle is the “pushbroom scanning” concept. Each line of the image is electronically scanned and successive lines are images thanks to the relation between the motion of the line of sight on the ground and the acquisition frequency.

The spectral shapes of the multispectral camera are similar to the SPOT plus a blue band. Nevertheless, the panchromatic band includes also the near infrared wavelength (GISTDA, 2010).

Table 2.3 THEOS panchromatic and multispectral sensor system characteristics.

THEOS	PAN	MS
Spectral bands and Resolution	Panchromatic 2 m	4 Multispectral 15 m
Spectral ranges	P: 0.45-0.90 µm	B0 (red) : 0.62-0.69 B1 (green) : 0.53-0.60 B2 (blue) : 0.45-0.52 B3 (NIR) : 0.77-0.90
Imaging swath	22 km	90 km
Image dynamics	8 bits among 12 bits	8 bits among 12 bits
Absolute localisation accuracy (Level 1B)	<400 m (3σ)	<400 m (3σ)
Off-nadir viewing	±50º (roll and pitch)	>117
signal to Noise Ratio	>110	

Source: GISTDA (2010)

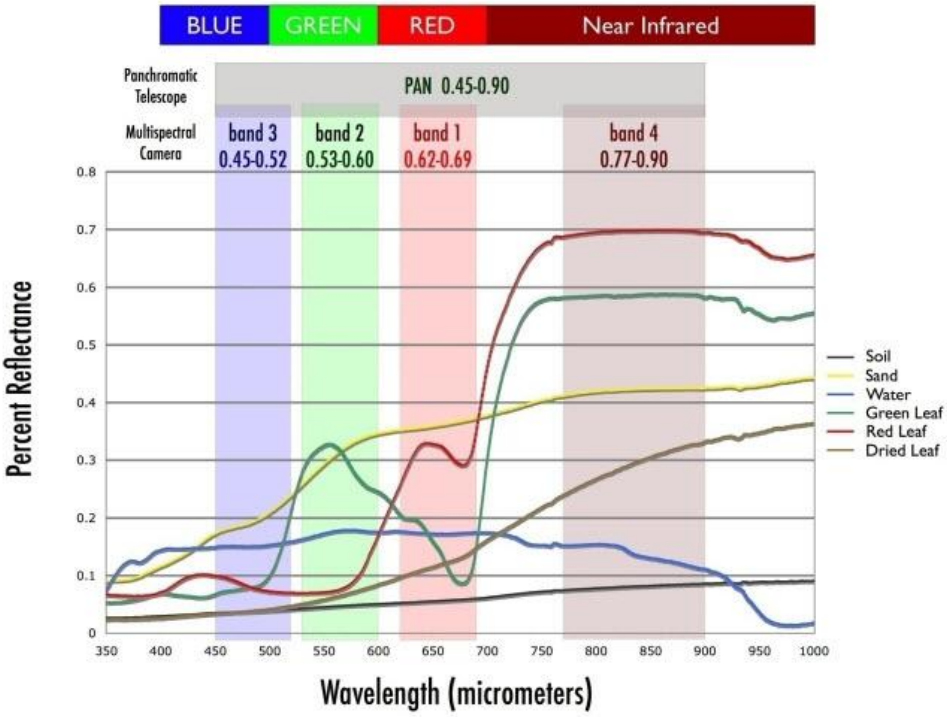


Figure 2.1 The spectral reflectance curves of THEOS multispectral sensor.

Source: GISTDA (2010)

2.3 Expert Systems

A knowledge-based expert system may be defined as “a system that uses human knowledge to solve problems that normally would require human intelligence” (PC AI, 2002).

Experts interpret remote sensing images with knowledge based on experience. However computer assisted classification utilized only very limited expert knowledge. The expert system, therefore, is a problem solving system which supports expert knowledge in a computer based system.

The following two types of knowledge are required for an expert system in remote sensing.

Knowledge about image analysis.

Procedures for image analysis can be made only with adequate knowledge about image processing and analysis. A feedback system should be introduced for checking and evaluating the objectives and the results.

Knowledge about the objects to be analyzed.

Knowledge about the objects to be recognized or classified should be introduced in addition to the ordinary classification method. The fact that forest does not exist over 3,000 meters above sea level, is one example or the type of knowledge that can be introduced.

The expert system can be integrated with a geographic information system (GIS). It is necessary to accumulate experiences and to evaluate the knowledge for an expert system to be operationally applied (Japan Association on Remote Sensing, 1993).

2.3.1 IMAGINE Expert Classifier

Expert classification can be performed using the IMAGINE Expert Classifier™. The expert classification software provides a rules-based approach to multispectral image classification, post-classification refinement, and GIS modeling. In essence, an expert classification system is a hierarchy of rules, or a decision tree, that describes the conditions under which a set of low level constituent information gets abstracted into a set of high level informational classes. The constituent information consists of user-defined variables and includes raster imageries, vector coverages, spatial models, external programs, and simple scalars.

A rule is a conditional statement, or list of conditional statements, about the variable's data values and/or attributes that determine an informational component or hypotheses. Multiple rules and hypotheses can be linked together into a hierarchy that ultimately describes a final set of target informational classes or terminal hypotheses. Confidence values associated with each condition are also combined to provide a confidence image corresponding to the final output classified image ERDAS, 2010). The IMAGINE Expert Classifier is composed of two parts: the Knowledge Engineer and the Knowledge Classifier.

2.3.1.1 Knowledge Engineer

The Knowledge Engineer provides the interface for an expert with firsthand knowledge of the data and the application to identify the variables, rules, and output classes of interest and create the hierarchical decision tree. The decision tree grows in depth when the hypothesis of one rule is referred to by a condition of another rule. The terminal hypotheses of the decision tree represent the final classes of interest. Intermediate hypotheses may also be flagged as being a class of interest. This may occurs when there is an association between classes (ERDAS, 2010).

Figure 2.2 displayed a single branch of a decision tree depicting a hypothesis, its rule and condition. However, the rule under decision tree may be split for defining the hypothesis. Herewith both conditions must still be true to fire a rule, only one rule must be true to satisfy the hypothesis as shown in Figure 2.3.

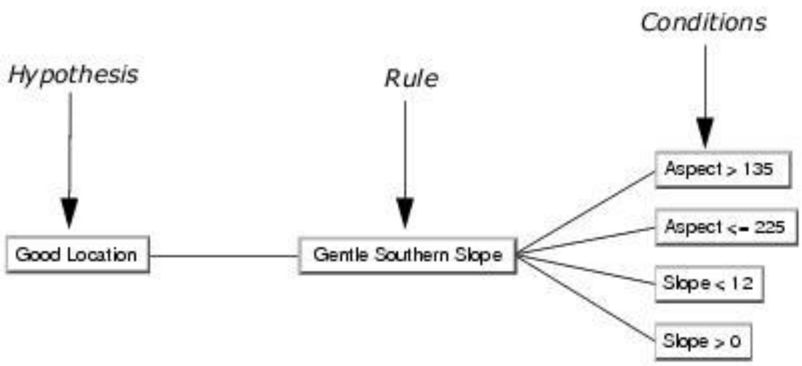


Figure 2.2 Examples of a decision tree branch.

Source: ERDAS (2010)

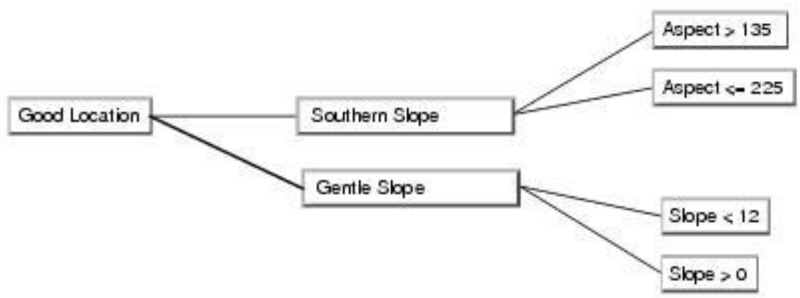


Figure 2.3 Split rule decision tree branch.

Source: ERDAS (2010)

2.3.1.2 Knowledge Classifiers

The Knowledge Classifier provides an interface for a non-expert to apply the knowledge base and create the output classification. The Knowledge Classifier is composed of two parts: an application with a user interface, and a command line executable. The user interface application allows users to input a limited set of parameters to control the use of the knowledge base. The user interface is designed as a wizard to lead users through pages of input parameters. After selecting a knowledge base, users are prompted to select classes.

After users select the input data for classification, the classification output options, output files, output area, output cell size, and output map projection, the Knowledge Classifier process can begin. An inference engine then evaluates all hypotheses at each location (calculating variable values, if required), and assigns the hypothesis with the highest confidence. The output of the Knowledge Classifier is a thematic image, and optionally, a confidence image (ERDAS, 2010).

In practice, the main working procedures of the IMAGINE Expert Classifier consist of 3 steps are as follows:

Identify the hypothesis: The expert in charge of creating the knowledge domain identifies a hypothesis (problem) to be addressed. This may be a formal hypothesis to be tested using inductive logic and confidence levels or an informal hypothesis that is in search of a logical conclusion.

Specify the expert systems rules: Heuristic rules that the expert has learned over time are the heart and soul of an expert system. If the expert's heuristic rules of thumb are indeed based on correct principle, then the expert system will most likely function properly. If the expert does not understand all the subtle nuances of the problem, has left out important variables or interaction among variables, or applied too much significance (weight) to certain variables, the expert system outcome may not be accurate. Therefore the creation of accurate, definitive rules is extremely important. Each rule provides the specific conditions to accept the hypothesis to which it belongs.

Specify the rule conditions: The expert would then specify one or more conditions that must be met for each rule (Jensen, 2005).

Example of hypothesis, rules and conditions specified in the Knowledge Engineer of IMAGINE Expert Classifier for white fir (*Abies concolor*) classification by Jensen (2005) is shown in Figure 2.4.

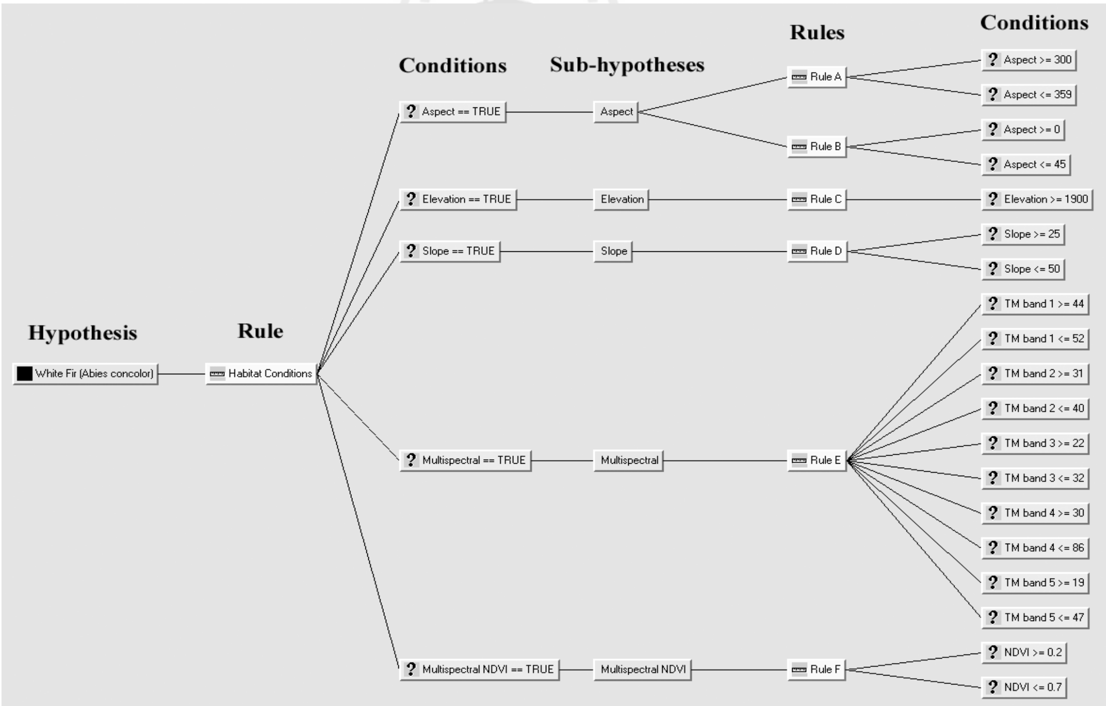


Figure 2.4 A knowledge-base expert system classification of white fir on Maple Mountain, Utah State, United States of America.

Source: Jensen (2005)

Advantage and disadvantage of Expert System was presented in Table 2.4.

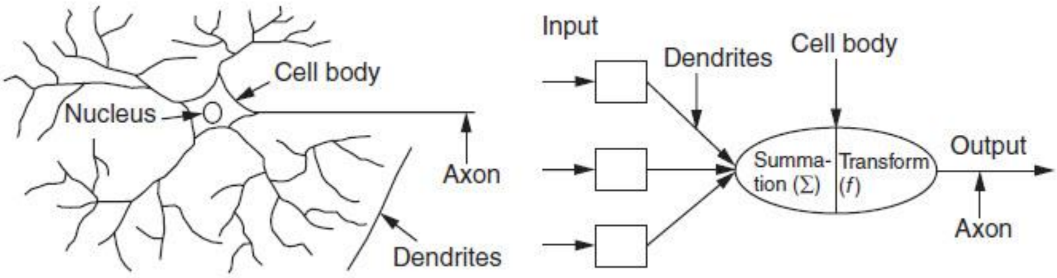
Table 2.4 Advantages and disadvantage of expert system.

Advantage
• Users can evaluate the output of the expert system and work backward to identify how a conclusion was reached. • Expert systems as nonmetric classification algorithm are being used such as decision trees, which make no assumption regarding the distribution of the data. • The decision tree can reveal nonlinear and hierarchical relationships among the input variables and use them to predict class membership. • A large body of evidence demonstrates the ability of machine-learning techniques (particularly decision trees and neural networks) to deal effectively with tasks that involve highly dimensional data.
Disadvantage
• The knowledge in a traditional expert system that must be extracted from knowledgeable experts of a domain area may be subjective and incomplete. This is because the experts may have a bias or even incorrect understanding of reality, they may not be aware of underlying rules they have used, and they may have difficulty articulating these rules. • Knowledge in an expert system is represented by logical rules made up of binary predicates. Numerical attributes have to convert to binary true/false statements, which may cause a large amount of information to be lost in the simplification process. • Most rule-based expert systems fail to generalize a predictable inference if an appropriate match with the prefect rules that must be articulated by experts cannot be obtained.

2.4 Artificial Neural Network (ANN)

ANN is one of the techniques in the area of artificial intelligence known as nature inspired. The area has been motivated by the computational mechanism of the human brain. The brain performs highly complex, nonlinear, parallel computations. The brain has a complex structure and the ability to self organize, building its own body of knowledge in what we typically refer to as experience. In the first two years of life, the human brain develops a set of connections that provides the human with a model of the world that surrounds her/ him. (Mas and Flores, 2008).

Artificial neurons are processing nodes or units that receive inputs from a number of connected nodes. All the connections between nodes are analogous to dendrites of a biological neuron. All the input signals that are fed into this artificial node are combined linearly or nonlinearly to generate an output (summation) via a transfer function (threshold) (Figure 2.5). The output is passed to other artificial neurons via another connection resembling the axon. Each processing unit functions as a simple pattern recognition machine at which the input data are evaluated against the synaptic strength and an output is produced (Gao, 2009).



(a) (b)

Figure 2.5 Comparison of biological and artificial neurons. (a) Structure of a biological neuron (b) a rendition of an artificial neuron that mimics the biological neuron.

Source: Gao (2009)

Haykin (1999) provides a general definition of an ANN as a massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects: (i) knowledge is acquired by the network from its environment through a learning process and (ii) interneuron connection strengths, known as synaptic weights, are used to store the acquired knowledge. (Mas and Flores, 2008).

ANN has many advantages over traditional computational methods. An ANN, made up of nonlinear elements, is itself nonlinear, may learn from a teacher an input–output mapping, is capable of adapting its synaptic weights to adapt to the environment, is able of dealing with incomplete information, and provides responses under uncertainty. It is worth noting that ANN are motivated or inspired by the analogy with the brain, but the motivation of creating an artificial brain lags way behind (Mas and Flores, 2008).

2.4.1 Architecture of neural networks

The three parameters essential to the architecture of an ANN model are topology, learning paradigm, and learning algorithm. Network topology refers to the manner in which all the nodes in a neural network are organized and connected, and how data and error information travel from one layer of nodes to the next. All processing units or nodes are organized into three general layers: input, hidden, and output.

Fundamentally, network topology falls into two groups: Feedforward Network and Feedback or Recurrent Network. The most common feedforward networks are exemplified by multilayer perceptrons and radial basis functions (Gao, 2009). In recurrent network, there is at least one feedback connection that corresponds to an integration operation or unit delay. Thus, the recurrent network actually represents a nonlinear dynamic system. The Hopfield model and the Boltzmann machine are Recurrent Network (Du and Swamy, 2006).

2.4.1.1 The feedforward network

In the feedforward network, the connections between neurons are in a feedforward manner. The network is usually arranged in the form of layers. There is no connection between the neurons within each layer, and no feedback between layers. A fully connected layered is a network such that every node in any layer is connected to every node in its adjacent forward layer. The inputs or outputs are processed in parallel, see Figure 2.6. When some of the connections are missing, it becomes a partially connected layered. Feedforward exhibit no dynamic properties and the networks are simply a nonlinear mapping (Du and Swamy, 2006).

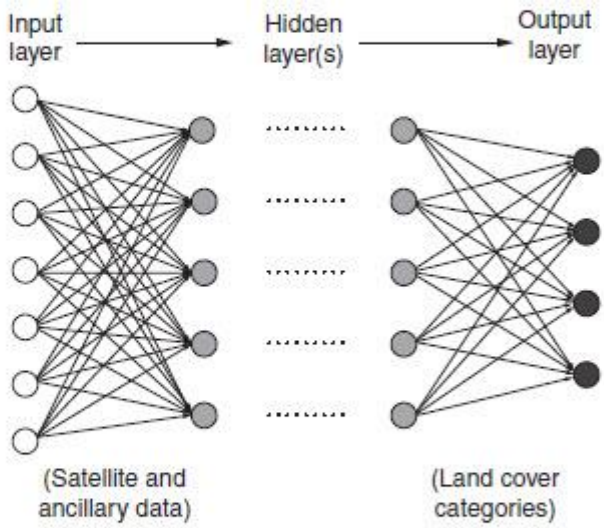


Figure 2.6 The structure of ANN data flow in a feedforward neural network, data flow from the input layer to the hidden layer(s) and eventually to the output layer in one direction. All the nodes in the same layer are fully connected to those in a layer immediately below or above. Each link is assigned a weight.

Source: Gao (2009)

2.4.1.2 Multilayer perceptron network

A multilayer perceptron network, in its most general form, can have many inputs and many outputs. In the case of prediction, there is usually one output neuron; multiple class classification requires more than one. There can be one or several hidden layers and any number of hidden neurons in each layer, in the general case where there are *n* inputs, *m* hidden neurons, and *k* outputs neurons, the intermediate stages of processing within a multilayer perceptron can be constructed as follows (Samarasinghe, 2006):

The hidden neuron input *u* and output *y* of the *j* neuron are

$$u_j = a_{0j} + \sum_{i=1}^{n} a_{ij} x_i \qquad (2.1)$$

$$y_j = f(u_j),$$

Where x_i is the i^{th} input.

a_{ij} is the weight associated with the input *i* and neuron *j*.

a_{0j} is the bias weight of hidden neuron *j* and $f(u_j)$ can be any activation function that transform u_j into a hidden neuron output y_j.

The weight sum of input v_k and the output z_k of the k^{th} output neuron can be written as

$$v_k = b_{0k} + \sum_{j=1}^{m} b_{jk} y_j \qquad (2.2)$$

$$z_k = f(v_k)$$

where *m* is the number of hidden neurons.

k is the output neurons.

b_{0k} is the bias weight of output node *k*.

b_{jk} is the weight of the connection between the j^{th} hidden neuron and the k^{th} neuron.

$f(v_k)$ is the activation function of the k^{th} output neuron, which transforms v_k into its final output.

2.4.1.3 Backpropagation algorithm

A backpropagation neural network is characterized by a feed forward topology, supervised learning, and the backpropagation learning algorithm. Data pass forward from the input layer to the output layer via the hidden layer(s). After an input is presented to the input nodes, it propagates forward in the network. An output is initially produced from this input based on randomly assigned weights. This calculated outcome is then compared with the desired output. Their discrepancy is the error signal that is subsequently propagated backward from the output nodes to the input nodes through the network (Figure 2.7). This backpropagation of errors is implemented iteratively. The synaptic strengths or weights between nodes are adjusted in each iteration to ensure the output resembles the desired outcome as closely as possible.

The backpropagation neural network model has several drawbacks in classification of multispectral remote sensing data, one of which is the need to specify and fine-tune too many parameters before the network can function optimally. Consequently, a huge amount of time is required to configure the network properly during network training (Gao, 2009).

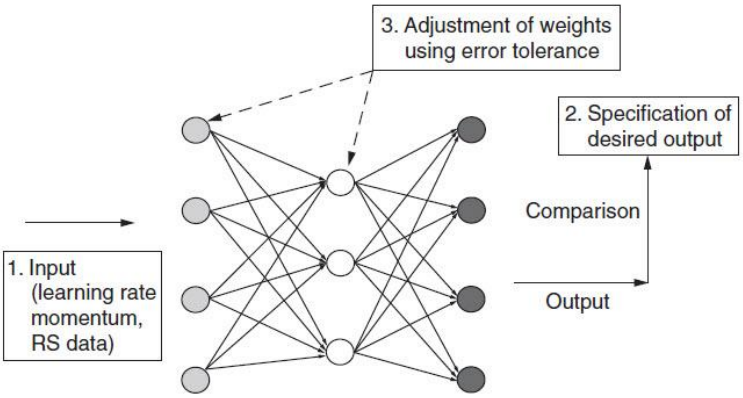


Figure 2.7 The structure of data flow in a feedforward and backpropagation neural network.

Source: Gao (2009)

2.4.2 Neural Network Configuration in ENVI

ANN-based image classification requires selection of an appropriate network model. The success of image classification depends on the proper configuration of this selected network. The potential of ANNs in image classification cannot be fully realized unless the network is optimally configured (Gao, 2009).

2.4.2.1 Training Threshold Contribution

The training threshold contribution determines the size of the contribution of the internal weight with respect to the activation level of the node that has a value from 0 to 1.0. It is used to adjust the changes to a node's internal weight. The training algorithm interactively adjusts the weights between nodes and optionally the node thresholds to minimize the error between the output layer and the desired response. Setting the Training Threshold Contribution to zero does not adjust the node's internal weights. Adjustments of the nodes internal weights could lead to better classifications but too many weights could also lead to poor generalizations.

2.4.2.2 Learning Rate

The training rate determines the magnitude of the adjustment of the weights. A higher rate will speed up the training, but will also increase the risk of oscillations or non-convergence of the training result. The Training Rate also has a value from 0 to 1.0 (ITT, 2007).

2.4.2.3 Training Momentum

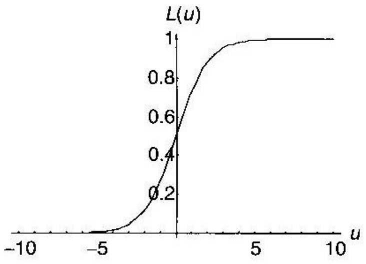
A momentum rate greater than zero allows to set a higher training rate without oscillations. A higher momentum rate trains with larger steps than a lower momentum rate. Its effect is to encourage weight changes along the current direction. Training momentum value is from 0 to 1.0 (ITT Visual Information Solutions, 2007).

2.4.2.4 Neural activation functions or Transfer functions

The neural activation functions have some important characteristics that make network vital to neural information processing. The nonlinear processing in hidden layer of multilayer perceptron works by these nonlinear activation functions.

Nonlinear means that the output of the function varies nonlinearly with the input; this aspect makes it possible for neural networks to do nonlinear mapping between inputs and outputs (Samarasinghe, 2006).

The most widely used function is *"sigmoid"*, a family of S-shaped functions that includes logistic and hyperbolic tangent functions. There are nonlinear, continuous functions that remain within some upper and lower bounds. The term “bounded” means that the output activation never reaches the large values, regardless of the input. This means that the output activation remains bounded even if the net input to the output is large (Samarasinghe, 2006).

(a) 

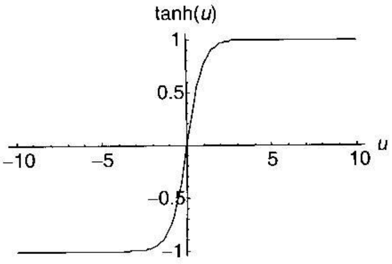
(b) 

Figure 2.8 Activation functions: (a) logistic or sigmoid and (b) hyperbolic.

Source: Samarasinghe (2006)

In Figure 2.8 (a) shows the logistic function for the rang of input u from -10 to 10, $L(u)$ denotes the output for an input u and the function has lower bound of zero and upper bound of 1. This means that the function value (or the output) range is [1,0]. At the input u = 0, the output is the midpoint (0.5), and the slope of the function, which indicates how fast the function is changing. The slope at u = 0 is 0.25 (14º). The output increases relatively quickly in the vicinity of u = 0, as input increases and approaches the upper bound much more slowly. For input below zero, the output initially decreases more rapidly, then more slowly as the lower bound is approached. The logistic function has the following mathematical formulation:

$$y = L(u) = \frac{1}{1+e^{-u}} \tag{2.3}$$

where e is the base of natural logarithm, which is a constant value of 2.71828.

Another used sigmoid function is the hyperbolic tangent function shown in Figure 2.8 (b) and given as below:

$$\tanh(u) = \frac{1+e^{-u}}{1-e^{-u}} \tag{2.4}$$

The hyperbolic tangent function has a lower bound of -1 and an upper bound of 1, making output range [-1,1] in contrast to the logistic function. Another difference is that the output at $u = 0$ is zero. The slope of the hyperbolic tangent is also higher at $u = 0$, meaning that it reaches the bounds more quickly than the logistic function. The slope at $u = 0$ here is 1.0 (i.e., 45º).

2.4.2.5 Number of Hidden Layers

For a linear classification use a value of 0 with no hidden layers, the different input regions must be linearly separable with a single hyper plane. Non-linear classifications are performed by setting the number of hidden layers to a value of 1 or greater. When the input regions are linearly inseparable and require two hyper planes to separate the classes you must have at least one hidden layer to solve the problem. Two hidden layers are used to classify input space where the different elements are neither contiguous nor connected (ITT, 2007).

Advantage and disadvantage of Expert System was presented in Table 2.5.

Table 2.5 Advantage and disadvantage of ANN algorithm.

Advantage
• A single neuron simulates the computation of a multivariate linear regression model. • A neural network makes no a priori assumptions of normal and linear data distribution due to its operation in a nonparametric fashion. • Neural networks are able to learn from existing examples adaptive, which makes the classification objective. • The nonlinear patterns are "learned" from the empirical examples instead of pre-specified" by an analysis based on prior knowledge of the datasets. • The noisy information inevitably included in the examples supplied a trained neural network with the ability to generalized, which makes neural networks robust solutions in the presence of previous unseen, incomplete, or imprecise data.

Table 2.5 (Continued).

Advantage
• A neural network can embrace data in all formats as long as the data are converted to a numeric representation. • Neural network are tolerant of noise and missing data and attempt to find the best fit for input patterns. • Neural networks continuously adjust the weights as more training data are provided in a changing environment.
Disadvantage
• Despite the excellent performance of neural networks in image classification, it is usually difficult to explain in a comprehensive fashion the process through which a given decision or output has been obtained from a neural network. The rules of image classification and interpretation learned by the neural network are buried in the weights of the neurons of the hidden layers. It is difficult to interpret these weights due to their complex nature. A neural network is often accused of being a black box. • Using neural network, an analyst might find it difficult to gain an understanding of the problem at hand because of the lack of explanatory capability to provide insight into the characteristics of the dataset. • It is difficult to incorporate human expertise to simplify, accelerate, or improve the performance of image classification; a neural network always has to learn from scratch.

2.5 Literature review

In principle, many researches on land use and land cover classification (LULC) using different methods have been carried out. There are many case studies about Expert System and ANN on LULC classification are here reviewed.

2.5.1 LULC classification using Expert System

Lawawirojwong (2002) applied the knowledge-based to develop expert classification using Landsat-7 (ETM+) imagery for land cover classification of Bang Pakong watershed. The Expert classification applied the unsupervised classification (ISODATA) which included spectral characters, GIS data (DEM and soil moisture regime) and spatial model (clump model, NDVI model, mean NDVI per zone model) to classify the conditions for land cover category identification as residential and open space area, abandoned land, mixed deciduous forest, mangrove forest and wetland, paddy field, other vegetation, and waterbodies. The percentage of accuracy for each land cover categories from Expert classification is higher than the maximum likelihood classification. Furthermore, the overall accuracy of Expert classification is about 78% and maximum likelihood classification is only 67%. Thus the accuracy of the expert classification is higher than maximum likelihood classification 11%.

Deeudomchan (2003) applied Expert systems to classify land use types of Phu Khieo Wildlife Sanctury. Land cover associated factors and some indices of Landsat TM image have been used as main factors data that comprise of altitude, slope, annual rainfall, soil and rock types, NDVI, mineral indices and the result of some arithmetic expressions of the TM bands. In addition, other supervised classifications (Parallelepiped, Minimum Distance, Mahalanobis Distance and Maximum Likelihood) were also applied for accuracy assessment comparison. The accuracy of Expert systems was about 75.42%. While accuracies of supervised classification were 70.42% (Parallelepiped), 64.61% (Minimum Distance), 69.91% (Mahalanobis Distance) and 67.86% (Maximum Likelihood). Based on accuracy assessment it was found that Expert system provided the best result for land cover classification.

Kitiphaisannon (2005) applied the Geo-Information Technology including Remote Sensing, Global Positioning System, and Geographic Information System integrated with expert system to generate a knowledge base system for age class identification of oil palm plantation. Rule base under the knowledge base for oil palm classification is generated from the relationship of multiple regressions related with Water Index (WI), Bare Soil Index (BI), Normalized Differential Vegetation Index (NDVI), and Advance Vegetation Index (AVI). These equations were used to define age class stage of oil palm plantation which can be divided into 4 classes: Young stage (1-3 years old), Intermediate stage (4-10 years old), Productive stage (11-20 years old), and Mature stage (more than 20 years old). The percentage of accuracy for each class using Expert classification was higher than the Maximum Likelihood classification. Furthermore, the overall accuracy of the Expert classification was about 63.11%, and the Maximum Likelihood classification was only about 60.33%. Thus the accuracy of the Expert classification was about 2.78% higher than the Maximum Likelihood classification.

Wentz, Nelson, Stefanov and Roy (2007) applied an expert system approach to classify land use and land cover for Delhi, India. In this study, Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data of 22 September 2003 were used. The research goals of this project were two-fold. Firstly, the research goal is to report on the extent covered by urbanization using the classified image. Thirteen different land cover categories were identified with an 85.55% overall classification accuracy based on 256 random points for validation and 50 on the ground observations. Secondly, they reported on theirs efforts to duplicate an Expert system model previously developed for Phoenix Arizona as a generalized approach for urban land use classification. Results suggested that some of the methodology could be duplicated; however there are local factors (e.g. data availability and specific land features) that required the approach to be modified.

2.5.2 LULC classification using ANN

Heinl, Walde, Tappeiner and Tappeiner (2009) investigated the performance of image classifiers for landscape-scale land cover mapping and the relevance of ancillary data for the classification success in order to assess and to quantify the importance of these components in image classification. Specifically tests were the performance of Maximum Likelihood Classification (MLC), Artificial Neural Networks (ANN) and Discriminant Analysis (DA) based on Landsat7 ETM+ spectral data in combination with topographic measures (DEM and cos(i)) and NDVI. The classifications by DA and MLC produced very similar overall accuracies for all input combinations. Accuracies were in the range of 55–60% for using only spectral data (ETM) as input variables and reached about 75% when ancillary data were included. The classifications using ANN produced higher overall accuracies for all input combinations, reaching about 75% for using only spectral data (ETM) and 85% with ancillary data. Maximum overall classification accuracy of 86.3% was achieved by using ANN and all input information, if used pixel clusters larger than 4 ha to correct the differences in the MMU of reference and image data then final maximum in this study were 89.5% (Kappa: 0.84) for MLC, 89.2% (Kappa: 0.83) for DA and 94.3% (Kappa: 0.91) for ANN classification.

While MLC and DA produced comparable results only by incorporating ancillary data into the classification process. The superiority of ANN classification was less pronounced on the level of the single land cover classes. The use of ancillary data generally increased classification accuracy and showed a similar potential for increasing classification accuracy than the selection of the classifier. Therefore, a stronger focus on the development of appropriate and optimized sets of input variables is suggested. Also the definition and selection of land cover classes has shown to be crucial and not to be simply adaptable from existing land cover class schemes.

Ashish, Mcclendon, and Hoogenboom (2009) developed an ANN-based technique for the classification of multispectral aerial images for land use in agricultural and environmental applications. The specific land-use classes included water, forest, and several types of agricultural fields. Multispectral images at a 1-m resolution were obtained for the state of Georgia, USA from a GIS data clearinghouse. These false-color images contained green, red and infrared true-color information. Three approaches were used for the preparation of the inputs to the ANN. These included histograms of the pixel intensities, textural parameters extracted from the image, and matrices of the pixels for spatial information. A probabilistic neural network was used. Seven hundred images were used for model development and 175 for independent model evaluation. The overall accuracy for the evaluation data set was 74% for the histogram approach, 71% for the spatial approach and 89% for the textural approach. The evaluation of ANNs based on various combinations of all three approaches did not show an improvement in accuracy. Also, it found that some approaches could be used selectively for certain classes. For example, the textural approach worked best for forest classes. For future studies, edge detection prior to classification, with more careful selection of each class, should be included for land use classification of multispectral images.

Dixon and Candade (2008) studied and implemented a new pattern recognition technique introduced within the framework of statistical learning theory called Support Vector Machines (SVMs), and its application to remote-sensing image classification. Standard classifiers such as ANN need a number of training samples that exponentially increase with the dimension of the input feature space. With a limited number of training samples, the classification rate thus decreases as the dimensionality increases. SVMs are independent of the dimensionality of feature space as the main idea behind this classification technique is to separate the classes with a surface that maximizes the margin between them, using boundary pixels to create the decision surface. Results from SVMs are compared with traditional Maximum Likelihood Classification (MLC) and an ANN classifier. The findings suggest that the ANN and SVM show comparable results classifiers but perform better than the traditional MLC. The ANN with 15 hidden nodes shows about 78.4% accuracy and 79.2% for SVM using the polynomial kernel and 50.6% for MLC. However, accuracy is dependent on factors such as the number of hidden nodes (in the case of ANN) and kernel parameters (in the case of SVM). The training time taken by the SVM is several magnitudes less.

Mustapha, Lim, and Mat Jafri (2010) compared the Neural Network and Maximum Likelihood approaches in land cover mapping by using high spatial resolution satellite images in Makkah city which is located in the semi-arid conditions in western of Saudi Arabia. Two algorithms were applied for classification: Maximum Likelihood classification and Neural Network classification. They were studying the performances of these methods for the purpose of land cover mapping. The experiment results indicated that the Neural Networks algorithm with 89.3% overall accuracy and 0.820 Kappa Coefficient is more reliable than the Maximum Likelihood algorithm with 80.3% and 0.672 overall accuracy and Kappa coefficient, respectively. From the results, they suggested that the increasing of the classification accuracy is due to the ability of Neural Network in handling the pixel that has many cover type (mixed pixel). The images used in this study were obtained from high spatial resolution satellite data (10m), so that much information could be extracted and neural network could overcome the speckle problem. However, complex landscapes with the mixed-pixel problem would lead to the difficulties in classification that Maximum Likelihood classifier cannot handle this kind of complex images so that many pixel cannot be classified correctly.

Suwanwerakamtorn and Supunee (2006) compared the categorization of land use with the supervised classification, maximum likelihood and Neural Network classifications. Landsat-7 ETM+ band 3, 4 and 5 images of Khon Kaen Province in the Northeast Thailand were used to identify signature set between homogenous and heterogeneous training area. The result obtained from both methods can discriminate 8 land use categories such as dry evergreen forest, deciduous forest, field crops, paddy fields, wet land, aquatic plants, water bodies and urban. The land use classification with the Maximum Likelihood method has an overall accuracy of 45.53% and Kappa Coefficient was 0.31421. At the same time land use classification with the Neural Network classification method has an overall accuracy of 43.96% and Kappa Coefficient was 0.30586. When compared with the land use map obtained from visual interpretation from the false color composite of Landsat-7 ETM+ band 4, 5 and 3 (RGB). This study concluded that land use and land cover obtained from both classifiers is not much difference. Although Maximum Likelihood classification gave better results than Neural Network but it is not significant. The reasons may come from a limitation of Landsat data resolution and the period of data acquisition. The data was recorded in dry season which almost agricultural lands were harvested. Therefore, it makes difficult to select the training areas.

CHAPTER III

DATA, EQUIPMENT AND METHODOLOGY

3.1 Data and equipment

Remotely sensed data and GIS data had been collected in this study while basic hardware and software were employed for data collection and data analysis. Data and equipment using in this study were summarized in Table 3.1.

Table 3.1 Data and equipment in this study.

Data and equipment	Data Characteristics	Source
1. Remote sensing data		
THEOS	Multispectral 4 bands (28 October 2009)	GISTDA[1]
Color orthophotos	Years 2002-2003	MOAC[2]
2. GIS data		
Land use data	2007	LDD[3]
Provincial soil data	2000	LDD
Sugarcane plot data	2010	OCSB[4]
3. Hardware/Software		
ERDAS Imaging	Software for expert system classification	Laboratory
ENVI	Software for neural network classification	Laboratory
Magellan GPS	Field survey	Personal
Vantage Point	GPS Software for managing data	Personal
Digital camera	Field survey	Personal

[1]Geo-Informatics and Space Technology Development Agency (Public Organization)
[2]Ministry of Agriculture and Cooperatives
[3]Land Development Department
[4]Office of the Cane and Sugar Board

3.2 Methodology

The methodology of the study has been divided into 3 components including data collection and preparation, field survey, and data analysis. For data analysis, there were three main steps included image classification, accuracy assessment and evaluation of an optimum classification method and dataset (Figure 3.1). The detail in each component can be described as following.

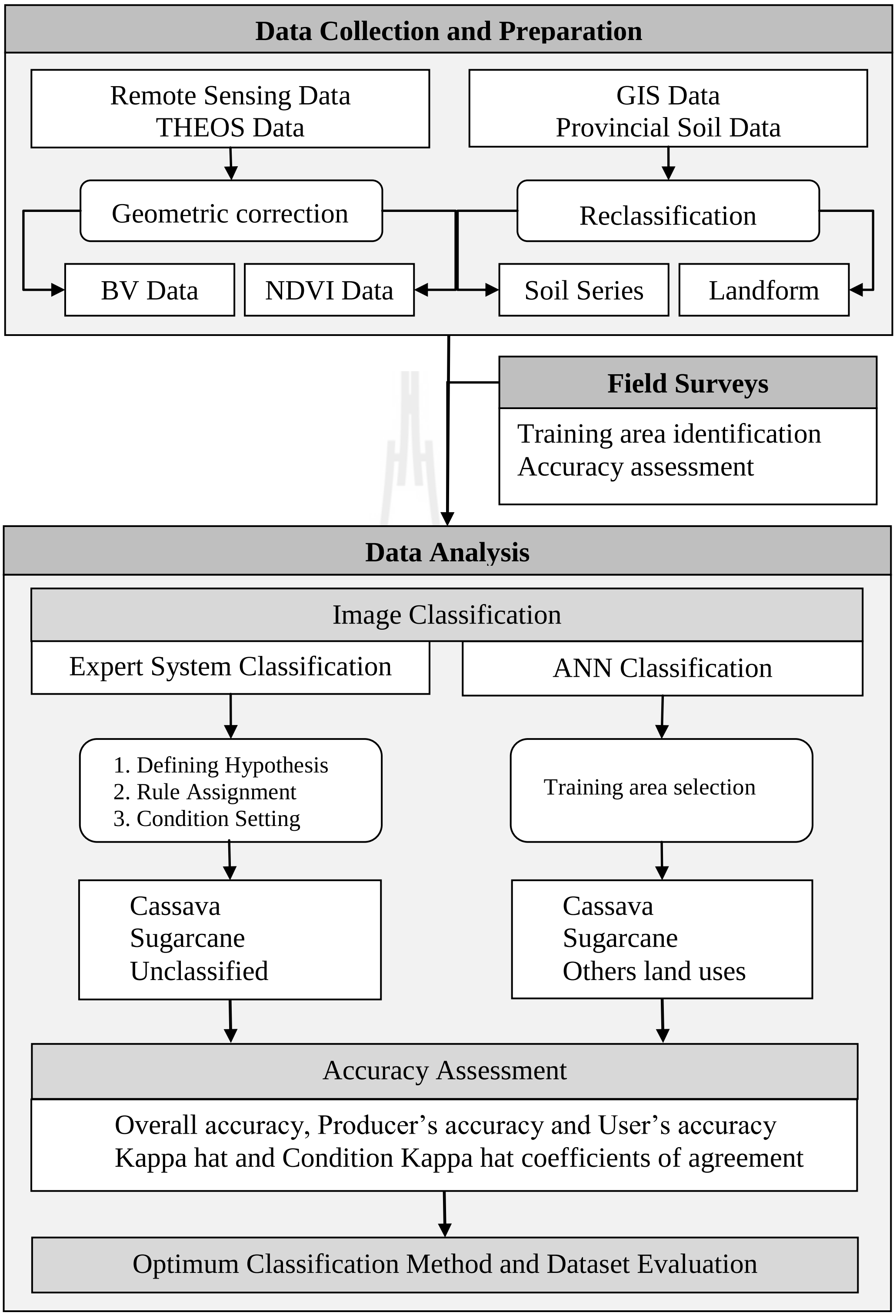


Figure 3.1 The framework of methodology.

3.2.1 Data collection and preparation

The remotely sensed and GIS data had been collected from various sources (see Table 3.1) and prepared in advance for cassava and sugarcane areas classification.

3.2.1.1 THEOS data

Acquired THEOS data with four multispectral bands were firstly geometric corrected by image to image rectification based on color orthophotos for reducing spatial distortion. In practice, GCPs were firstly collected from reference image (color orthophotos) for spatial interpolation and then data resampling were applied for intensity interpolation. In this study, second-order transformation for spatial interpolation and nearest neighbor resampling for intensity interpolation were applied with RMS errors less than 1.0 pixel (15 m).

After that, NDVI data that is the ratio between red and near infrared bands was created. For THEOS data, NDVI can be derived from ratio between band 4 and 1 as following equation:

$$NDVI = \frac{(THEOS\ Band\ 4 - THEOS\ band\ 1)}{(THEOS\ band\ 4 + THEOS\ band\ 1)} \tag{3.1}$$

3.2.1.2 Land use data

Land use data of Nakhon Ratchasima province in 2007 from LDD was firstly used to calculate the proportion of cassava and sugarcane in agricultural land in each district for study area identification. Main criteria for study area identification are based on equally proportion of cassava and sugarcane in agricultural land. Herein, Chakkarat district which provided the equally proportion of cassava and sugarcane in agricultural land was selected as study area. In addition, land use data of LDD had been used as supplementary information for field survey.

3.2.1.3 Provincial soil data

Basically, provincial soil data from LDD provide details of soil series type and theirs properties and landform. In this study, provincial soil data were used to extract soil series and landform data for cassava and sugarcane classification. In practice, provincial soil data in vector format were firstly reclassified based on their soil type and landform data and then converted to raster format with grid size of 15 x 15 m as spatial resolution of THEOS data.

3.2.1.4 Sugarcane plot data

In addition, sugarcane plot data in 2010 from OCSB was used as supplementary information for field survey and training areas selection of sugarcane.

3.2.1.5 Dataset preparation

Preprocessed and extracted data of remotely sensed data and GIS data in raster format with the same spatial resolution and coordinate system were used to create predefined dataset for data analysis by image stacking. This study has been created 6 dataset as following:

- Multispectral data
- Multispectral data + NDVI
- Multispectral data + Soil series
- Multispectral data + Landform
- Multispectral data + Soil series + Landform
- Multispectral data+ NDVI + Soil series + Landform

3.2.2 Field survey

In this study, field survey was divided into 2 steps: training area selection and accuracy assessment.

3.2.2.1 Field survey for training area selection

Field survey for training areas selection with supplementary of land use data in 2007 was conducted in May 2010. Herewith, areas which were unchanged between 2007 and 2010 had been selected as training areas for image classification. In practice, the sites and feather types of training areas for Expert system and ANN were differently selected for optimal results. Therefore Expert system considered to use polygon feather (see Figure 3.2) whilst ANN used points as training areas (see Figure 3.3). Selected training data for each classification algorithm had been applied to all dataset in this study.

3.2.2.2 Field survey data for accuracy assessment

Identified location points for accuracy assessment were visited and recorded about existing land used classes. In practice, number of ground truth points for accuracy assessment had been firstly calculated based on the multinomial distribution to determine the sample size (Congalton and Green, 2008) as:

$$N = \frac{B\Pi_i(1-\Pi_1)}{b_i^2} \tag{3.2}$$

where Π_i is the propotion of a population in the i^{th} class out of k classes that has propotion closet to 50%

b_i is the desired precision for this class

B is the upper $(\alpha/k) \times 100^{th}$ percentile of the chi square (χ^2) distribution with 1 degree of freedom

k is the number of classes

The preliminary data of population proportions in each class obtained from land use data year 2007. Then the stratified random sampling was applied for sampling point identification for accuracy assessment by using number of points that calculated based on the multinomial distribution. The most advantage of stratified random sampling is that contained all classes even the small proportion class (Jensen, 2005). Field survey for accuracy assessment was conducted in May 2011.

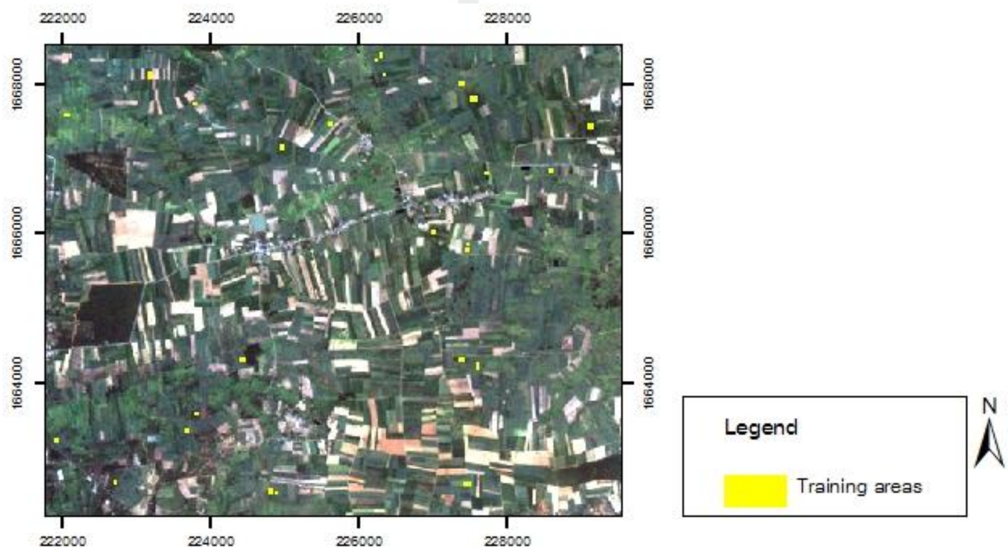


Figure 3.2 Sample of training areas for Expert System algorithm.

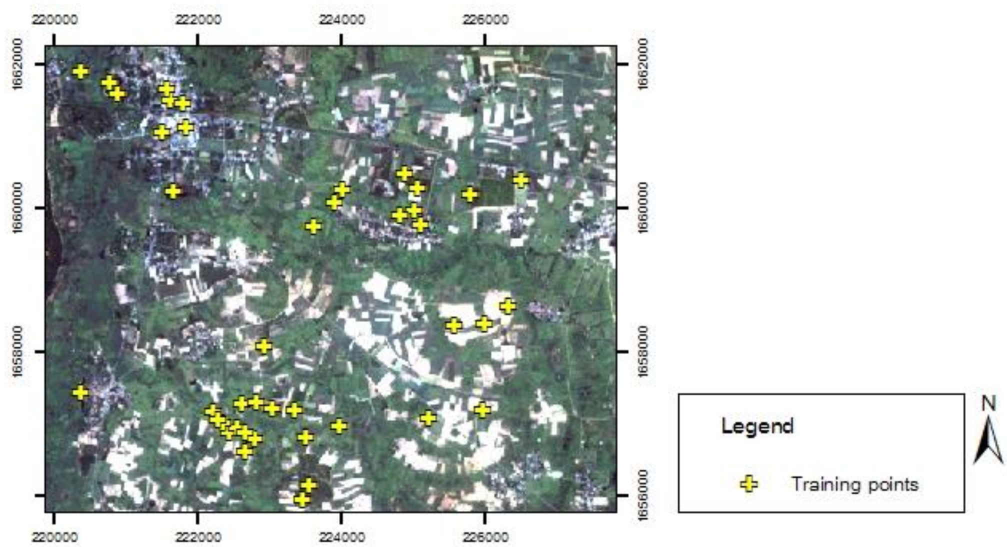


Figure 3.3 Sample of training points for ANN algorithm.

3.2.3 Data analysis

In this study, there were three main steps of data analysis included image classification, accuracy assessment and evaluation of an optimum classification method and dataset**.**

3.2.3.1 Image classification

Two main methods of image classification for sugarcane and cassava areas are Expert system and Artificial Neural Network (ANN)

(1) Cassava and sugarcane classification by Expert system

In this study, cassava and sugarcane classification by Expert system was conducted using Knowledge Engineer of ERDAS imagine. In practice, three steps are required for classification include hypothesis definition, rule assignment and condition setting. Herein, cassava and sugarcane hypothesis are firstly defined for thematic classification. Secondly, specific rule for extraction of thematic classification were assigned according to data in each dataset including multispectral data, NDVI, soil series and landform. Finally, conditions in each rule of each hypothesis were assigned as follows:

- Minimum and maximum brightness values which were extracted from training area applied for Multispectral data rule.
- Minimum and maximum NDVI values which were extracted from training area applied for NDVI rule.
- Soil series classes, which extracted using overlay analysis between land use data in 2007 and soil data, were used for Soil series rule.

- Landform classes, which extracted using overlay analysis between land use data in 2007 and soil data, were used for Landform rule.

(2) Cassava and Sugarcane classification using ANN

Cassava and sugarcane classification by ANN was conducted under ENVI program. Herein neural net function and randomly repressive configuration was applied to classify cassava and sugarcane and others land use classes. In practice, training point data of each land use class were firstly assigned and land use classes were then extracted based on neural network configuration. ANN classification allows users to use minimal training dataset and provided superior results (Hepner, 1990). In this study, major land use classes were classified include:

- Cassava
- Sugarcane
- Urban and built-up areas
- Paddy field
- Forest areas
- Water bodies
- Bare land

Furthermore, assigned standard Neural Network configuration under ENVI was set up as shown in Table 3.2.

Table 3.2 Neural Network configuration for LULC classification under ENVI.

Neural Network Parameters	Assigned value
Training Threshold Contribution	0.9
Learning Rate	0.1 - 0.3
Training Momentum	0.9
Training RMS Exit Criteria	0.1
Activate Functions	Logistic
Number of Hidden Layers	1
Number of Training Iterations	10,000

3.2.3.2 Accuracy assessment

In this study, the actual ground truth data collected from field surveys were compared with the classification maps. The results between agreement and disagreement had been shown in error matrices. Descriptive statistics include overall accuracy, producer's accuracy, user's accuracy and multivariate analytical techniques include Kappa coefficient of agreement and conditional Kappa coefficient of agreement were used for accuracy assessment (Congalton and Green, 2008) as shown in the following equations.

$$Overall\ accuracy = \frac{\sum_{i=1}^{k} n_{ii}}{n} \tag{3.3}$$

$$Producer's\ accuracy_j = \frac{n_{jj}}{n_{+j}} \tag{3.4}$$

$$User's\ accuracy_i = \frac{n_{ii}}{n_{i+}} \tag{3.5}$$

where n_{ii} is the diagonal of each column.

n_{ij} is anycell in the confuse matrix between number of samples classified into category i in the map (rows) and category j in the reference data (columns).

n is the number of catagories.

$$KHAT\ (\hat{K}) = \frac{n - \sum_{i=1}^{k} n_{ii} - \sum_{i=1}^{k} n_{i+} n_{+i}}{n^2 - \sum_{i=1}^{k} n_{i+} n_{+i}} \tag{3.6}$$

$$Conditional\ K = \frac{n n_{ii} - n_{i+} n_{+i}}{n n_{i+} - n_{i+} n_{+i}} \tag{3.7}$$

where n_{ii} is the observed agreement as previously defined.

n_{i+}, n_{+i} are the marginal values estimates the expected agreement for each category i.

n is the sample size.

3.2.3.3 Evaluation of optimum classification methods and dataset

To evaluate the optimum classification method for cassava and sugarcane classification, Kappa coefficient which is derived from diagonal and off-diagonal cell values of error matrix was applied in this study. In practice, Kappa coefficient by accuracy assessment for each dataset in each method from Expert System and ANN classification will be firstly compared. The method which provides higher accuracy will then be identified as an optimum method for cassava and sugarcane classification in each dataset.

In the meantime, an optimum dataset for cassava and sugarcane classification by Expert system and ANN will be also evaluated based on Kappa coefficient. In addition, significant different of accuracy between multispectral dataset and addition dataset will be tested using standard normal distribution or Z statistics as:

$$Z = \frac{|\widehat{K_1} - \widehat{K_2}|}{\sqrt{\widehat{var}(\widehat{K_1}) + \widehat{var}(\widehat{K_2})}} \quad (3.8)$$

where Z is normalized and standard normal distribution

$\widehat{K_1}$ is KHAT for dataset I

$\widehat{K_2}$ is KHAT for dataset II

$\widehat{var}(\widehat{K_1})$ is variance of KHAT for dataset I

$\widehat{var}(\widehat{K_2})$ is variance of KHAT for dataset II

Herewith, variance of KHAT will be calculated by:

$$\widehat{var}(\widehat{K}) = \frac{1}{n}\left\{\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2-\theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4-4{\theta_2}^2)}{(1-\theta_2)^4}\right\} \quad (3.9)$$

Where $\theta_1 = \frac{1}{n}\sum_{i=1}^{k} n_{ii}$

$\theta_2 = \frac{1}{n^2}\sum_{i=1}^{k} n_{i+}n_{+i}$

$\theta_3 = \frac{1}{n^2}\sum_{i=1}^{k} n_{ii}(n_{i+} + n_{+i})$

$\theta_4 = \frac{1}{n^3}\sum_{i=1}^{k}\sum_{j=1}^{k} n_{ij}(n_{j+} + n_{+i})^2$

In practice, given the null hypothesis H_0: $(\widehat{K_1} - \widehat{K_2}) = 0$, and the alternative H_1 : $(\widehat{K_1} - \widehat{K_2}) \neq 0$, H_0 is rejected if $Z \geq Z_{\alpha/2}$, where $\alpha/2$ is the confidence level of the two-tailed Z test and the degrees of freedom are assumed to be ∞ (infinity) (Congalton and Green, 2008).

CHAPTER IV

RESULTS AND DISCUSSION

Chapter IV explained about results of cassava and sugarcane areas classification using Expert System and ANN according to major steps of methodology including (1) Data collection and preparation, (2) Sugarcane and cassava classification, (3) Ground truth and accuracy assessment, (4) Evaluation of optimum methods for sugarcane and cassava classification and (5) Evaluation of optimum dataset for sugarcane and cassava classification.

4.1 Data collection and preparation

4.1.1 Remote sensing data

4.1.1.1 Original THEOS data

This study used THEOS data that received from GISTDA. THEOS multispectral data covered Chakkarat district, Nakhon Ratchasima province acquired on 28 October 2009. This data was consisted of 4 bands: red (0.62-0.69 μm), green (0.53-0.60 μm), blue (0.45-0.52 μm) and NIR (0.77-0.90 μm) with spatial resolution of 15 x 15 meter (Table 4.1 and Figure 4.1).

Table 4.1 Metadata of THEOS data.

Image Data Properties	**THEOS Data**
File name	SCENE T1 M 2009/10/28 03: 13:34.5 0266-03200
Satellite	THEOS
Sensor	Multispectral
Acquired Data	October 28, 2009
Format	IMAGINE image
Number of Layers:	4 (Band 1, 2, 3, 4)
Pixel Depth	Unsigned 8-bit
Compression Type	None
Projection Zone	48
Spheroid Name	WGS 1984
Datum Name	WGS 1984
Georeferenced to	UTM, Zone 48
Upper Left X	220300
Upper Left Y	1667700
Lower Right X	230800
Lower Right Y	1647450
Pixel X size (m)	15
Pixel Y size (m)	15

Source: GISTDA (2010)

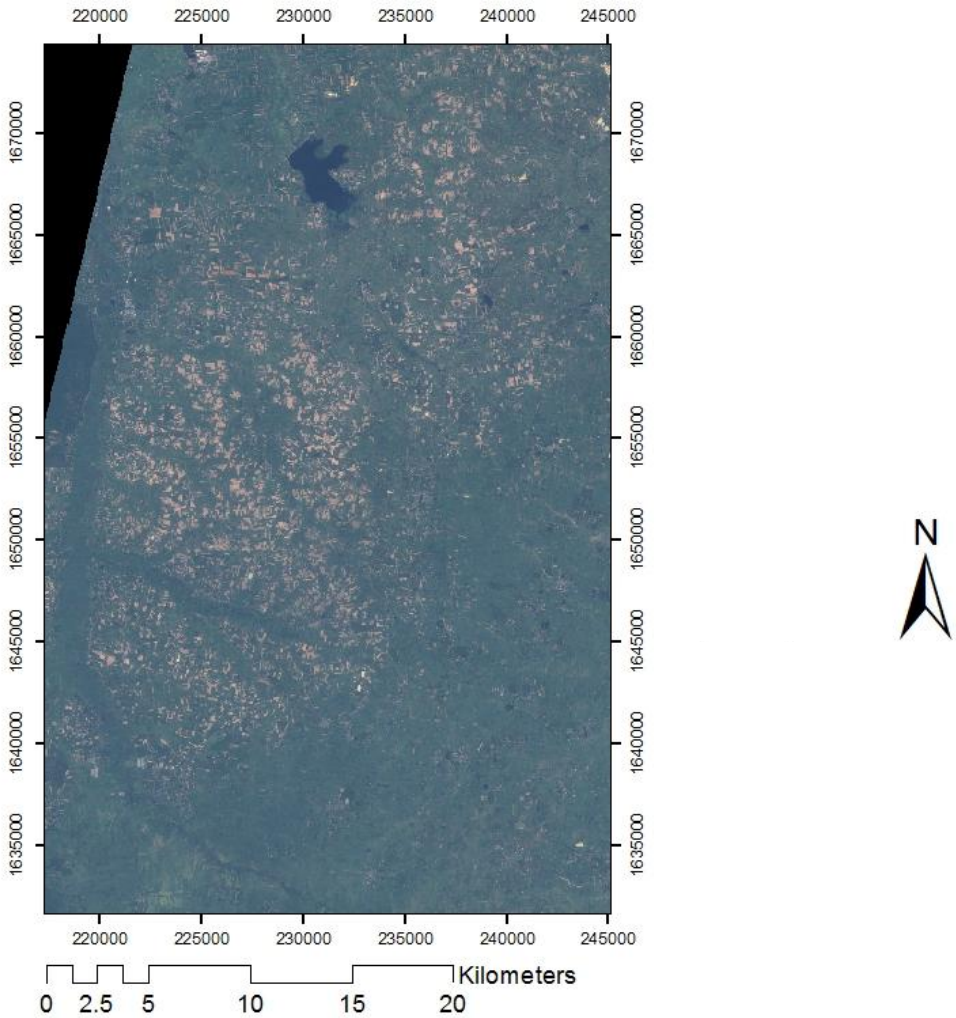


Figure 4.1 True color composite image of THEOS data in the study area.

4.1.1.2 Color orthophoto

Color orthophoto data produced in 2002 by the Ministry of Agriculture and Cooperatives had been utilized for THEOS image rectification. These data was composed of 6 scenes included 54381NE, 54381SE, 54392SE, 55384NW, 55384SW and 55393SW (Figure 4.2).

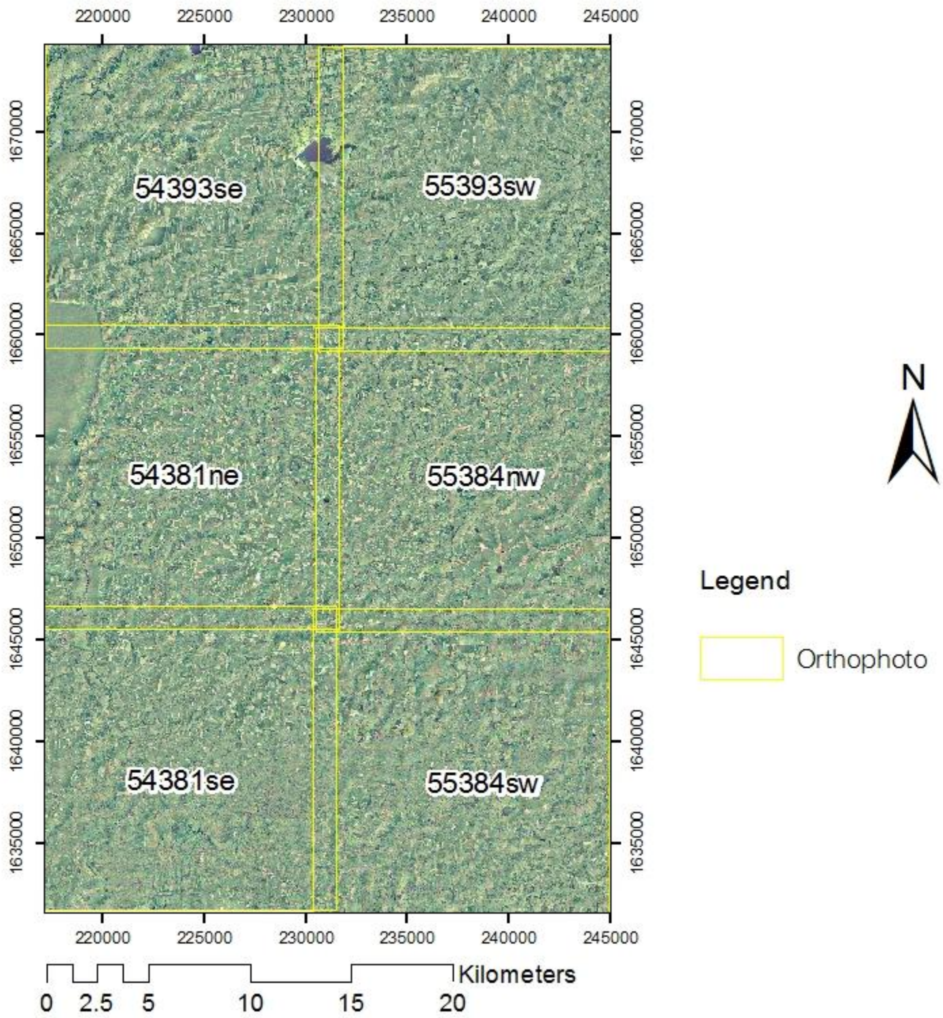


Figure 4.2 Color orthophoto of MOAC used in the study.

4.1.1.3 Rectified THEOS data

THEOS multispectral data were geometrically corrected using image to image rectification using color orthophoto as reference image. Herein, 30 GCPs (see Figure 4.3) were firstly collected from reference image (color orthophoto) for spatial interpolation with second order polynomial equation and then neighbor resampling technique was applied for intensity interpolation (Figure 4.4). The RMS error of image rectification was 0.4528 pixel or 6.792 m (See Table A.1 in Appendix A).

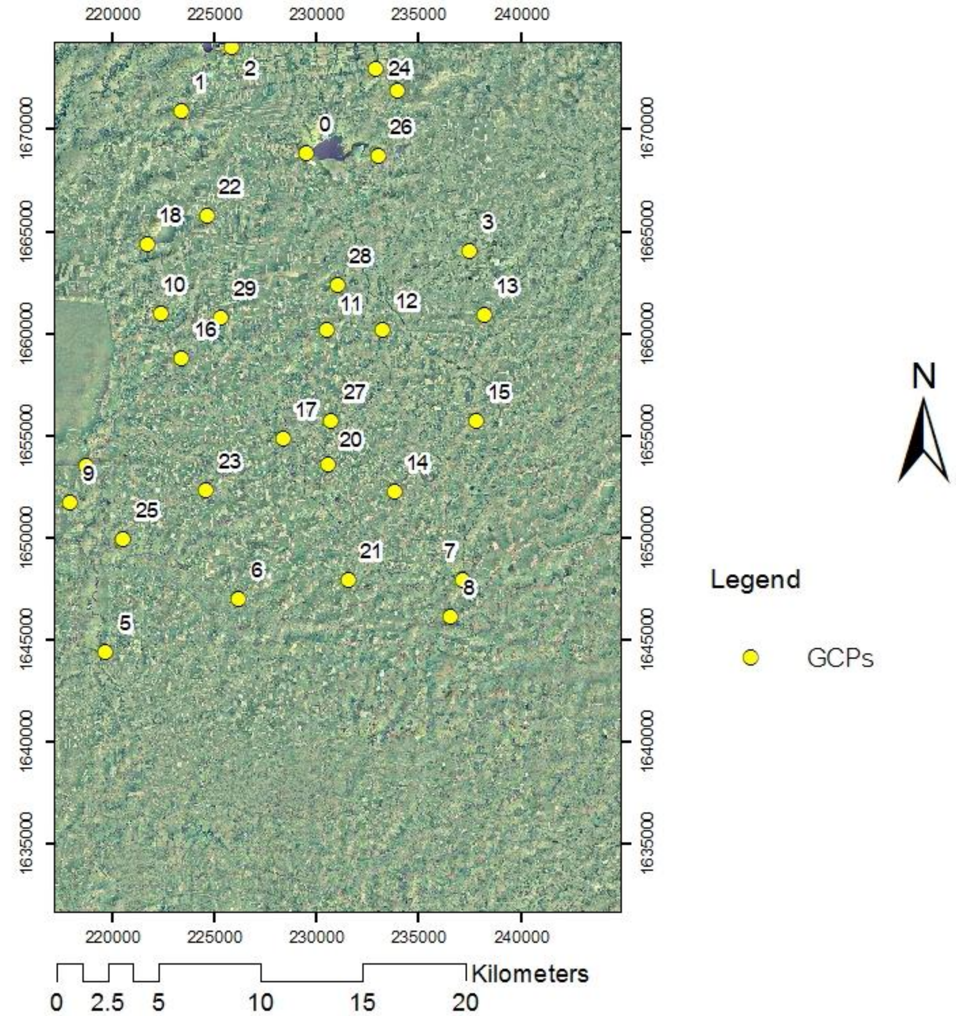


Figure 4.3 Distribution of GCPs for THEOS image rectification.

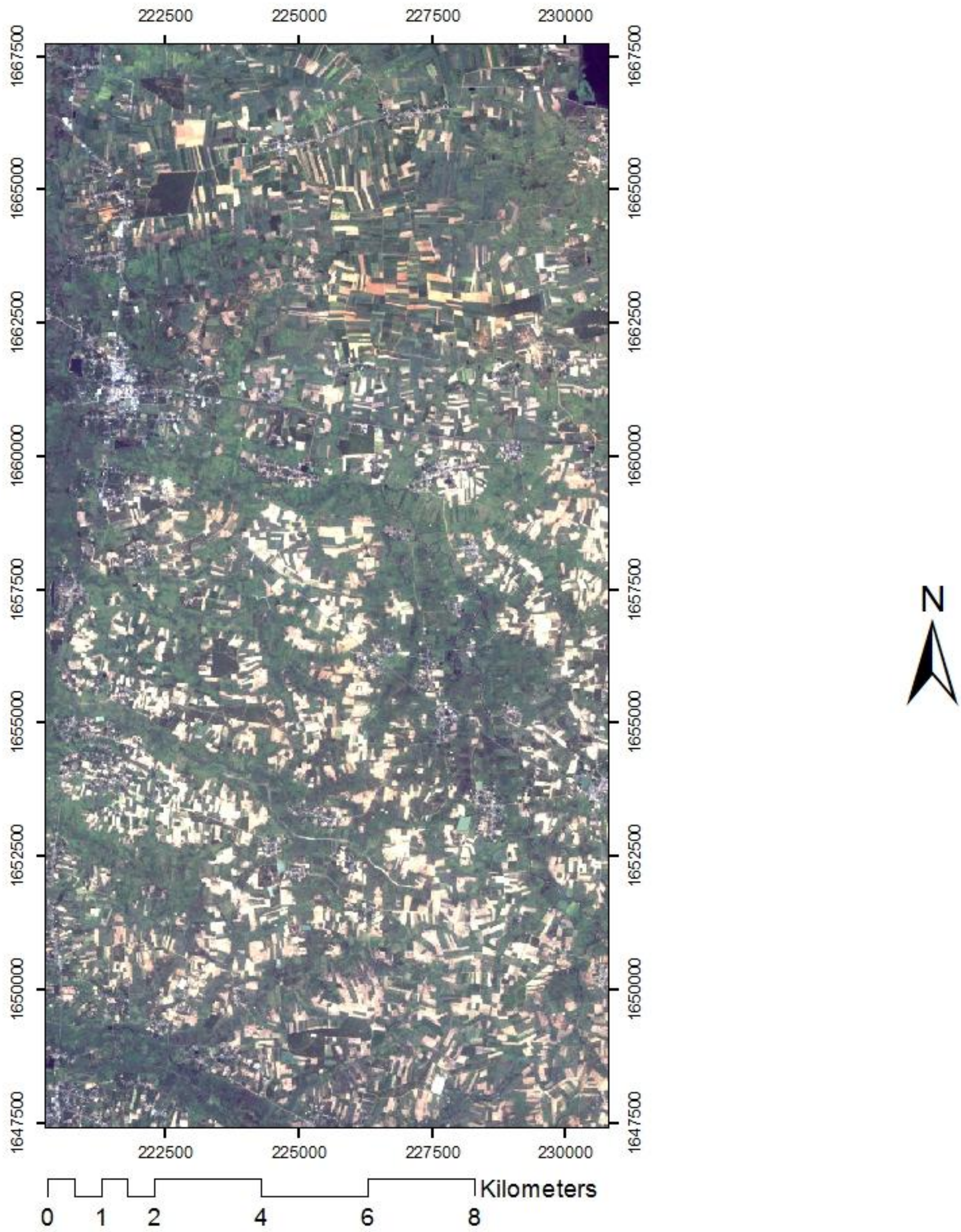


Figure 4.4 Rectified THEOS data of the study area.

4.1.1.4 NDVI

NDVI data derived from normalized different value between infrared band (band 4) and red band (band 4) of THEOS data (see Eq. 3.1). In this study, NDVI data varied between -0.51667 and 0.56364 as shown in Figure 4.5.

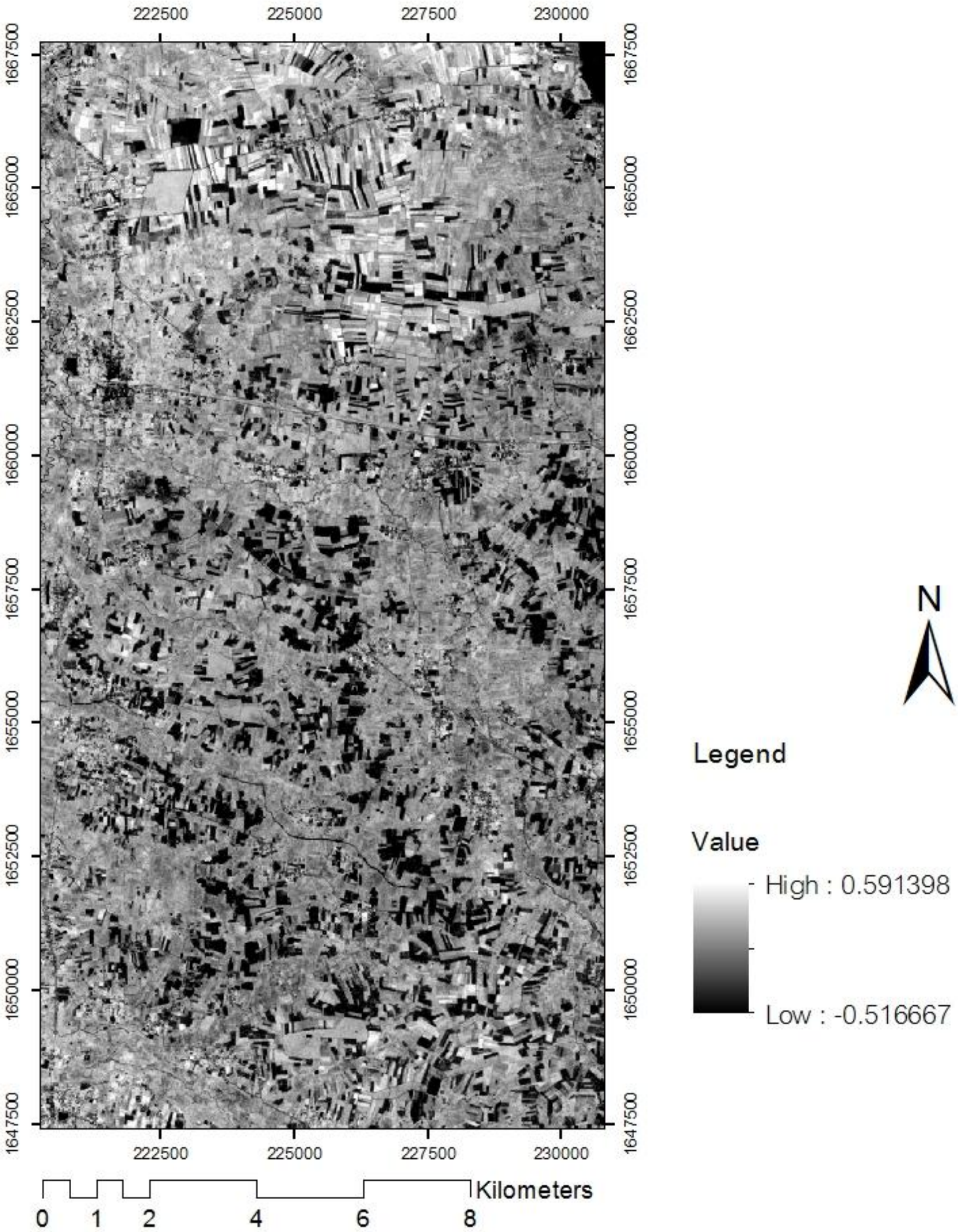


Figure 4.5 Distribution of NDVI value in the study area.

4.1.2 GIS data

4.1.2.1 Land use data

Land use data in 2007 of Nakhon Ratchasima province from LDD had been used as supplementary information for field survey and training areas selection (Figure 4.6).

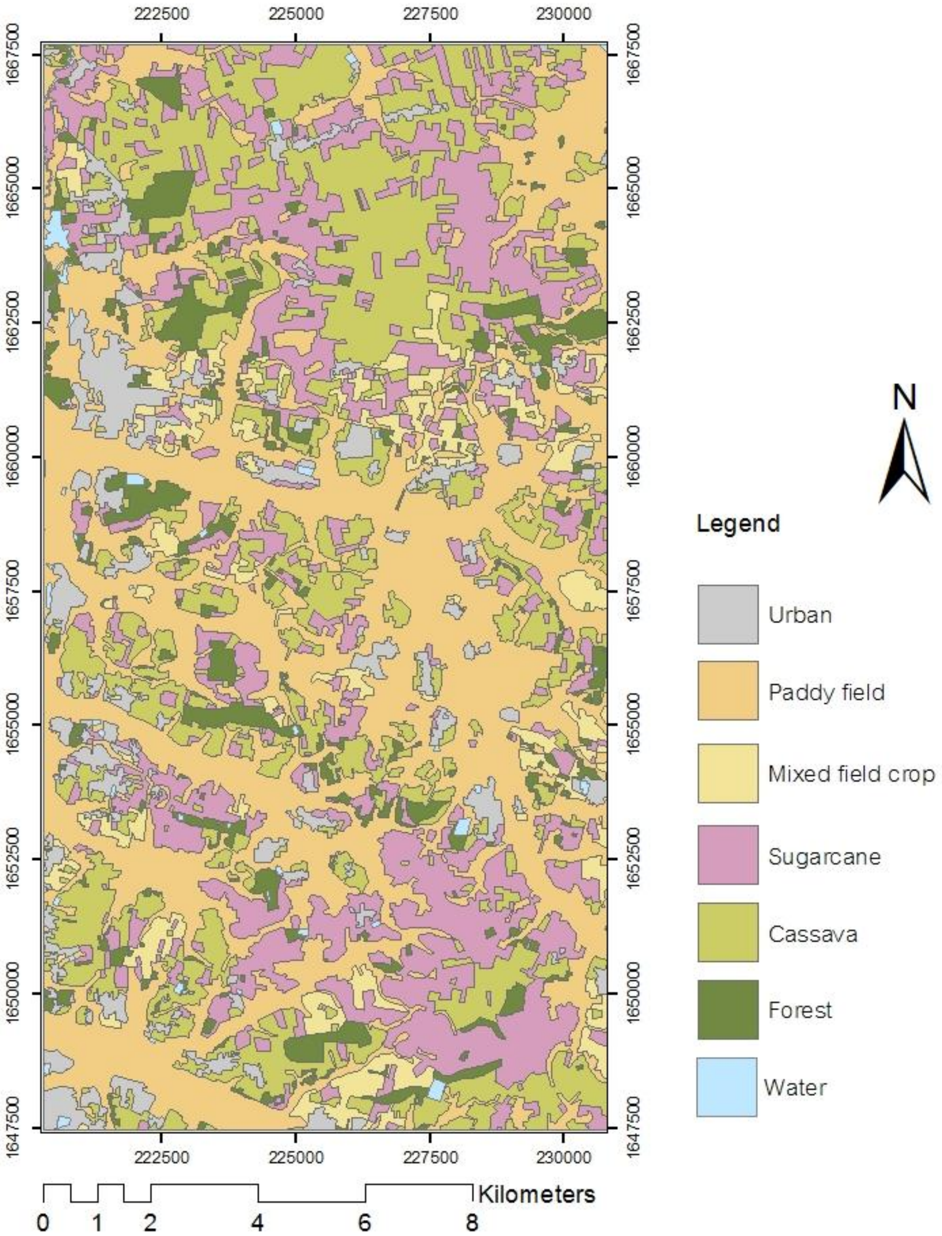


Figure 4.6 Distribution of major land use types in 2007 in the study area.

4.1.2.2 Provincial soil data

Basically, provincial soil data from LDD provide details of soil properties. In practice, soil types and landform categories of Nakhon Ratchasima province soil data were firstly extracted and then converted to raster format with the same grid size of THEOS data.

In this study, soil series that were extracted from provincial soil data were consisted of 29 classes. There were 26 soil series classes with cell value: Bli (0), By (1), Ckr (2), Cpg (3), Cpr (4), Dk (5), Ht (7), Ki (8), Kng (9), Knu (10), Ksk (11), Ksn (12), Kt (13), Ltc (14), Msk (15), Nbn (16), Ndg (17), Ng (18), Nkg (19), Nn (20), Ptk (21), Rn (22), Sda (23), St (24), Suk (25), and Ub (26). Other classes with cell value were Gully (6), Urban (27) and Water (28) as shown in Table 4.2 and Figure 4.7. At the same time, extracted landform categories with cell value from soil data included urban and built-up areas or water bodies (0), flat (1), and slightly undulates (2) as shown in Figure 4.8.

Table 4.2 Area and percentage soil type in the study area.

Soil series name	Soil code	Cell value	Area in sq. km	Percent
บัวลาย	Bli	0	12.44	4.41
บัวใหญ่	By	1	17.87	6.33
จักราช	Ckr	2	28.05	9.93
ชุมพวง	Cpg	3	29.24	10.36
จอมพระ	Cpr	4	51.36	18.19
ด่านขุนทด	Dk	5	2.67	0.94
ที่ดินร่องลึก	Gully	6	1.64	0.58
ห้วยแถลง	Ht	7	5.07	1.80
กุลาร้องไห้	Ki	8	0.38	0.13
คง	Kng	9	5.13	1.82
กระนวน	Knu	0	0.20	0.07
เขาสวนกวาง	Ksk	11	5.02	1.78
แก้งสนามนาง	Ksn	12	1.97	0.70
โคราช	Kt	13	3.92	1.39
ลำทะเมนชัย	Ltc	14	14.16	5.01
มหาสารคาม	Msk	15	3.89	1.38
หนองบุนนาก	Nbn	16	34.18	12.10
โนนแดง	Ndg	17	4.52	1.60
น้ำพอง	Ng	18	1.16	0.41
หนองกุง	Nkg	19	5.28	1.87
นครพนม	Nn	20	10.10	3.58
พระทองคำ	Ptk	21	6.67	2.36
เรณู	Rn	22	2.20	0.78
สีดา	Sda	23	0.00	0.00
สีทน	St	24	0.38	0.13
สตึก	Suk	25	9.38	3.32
อุบล	Ub	26	0.98	0.35
ที่อยู่อาศัย,วัด,โรงเรียน	Urban and built-up area	27	21.92	7.76
แหล่งน้ำ,หนอง,บึง	Water bodies	28	2.60	0.92

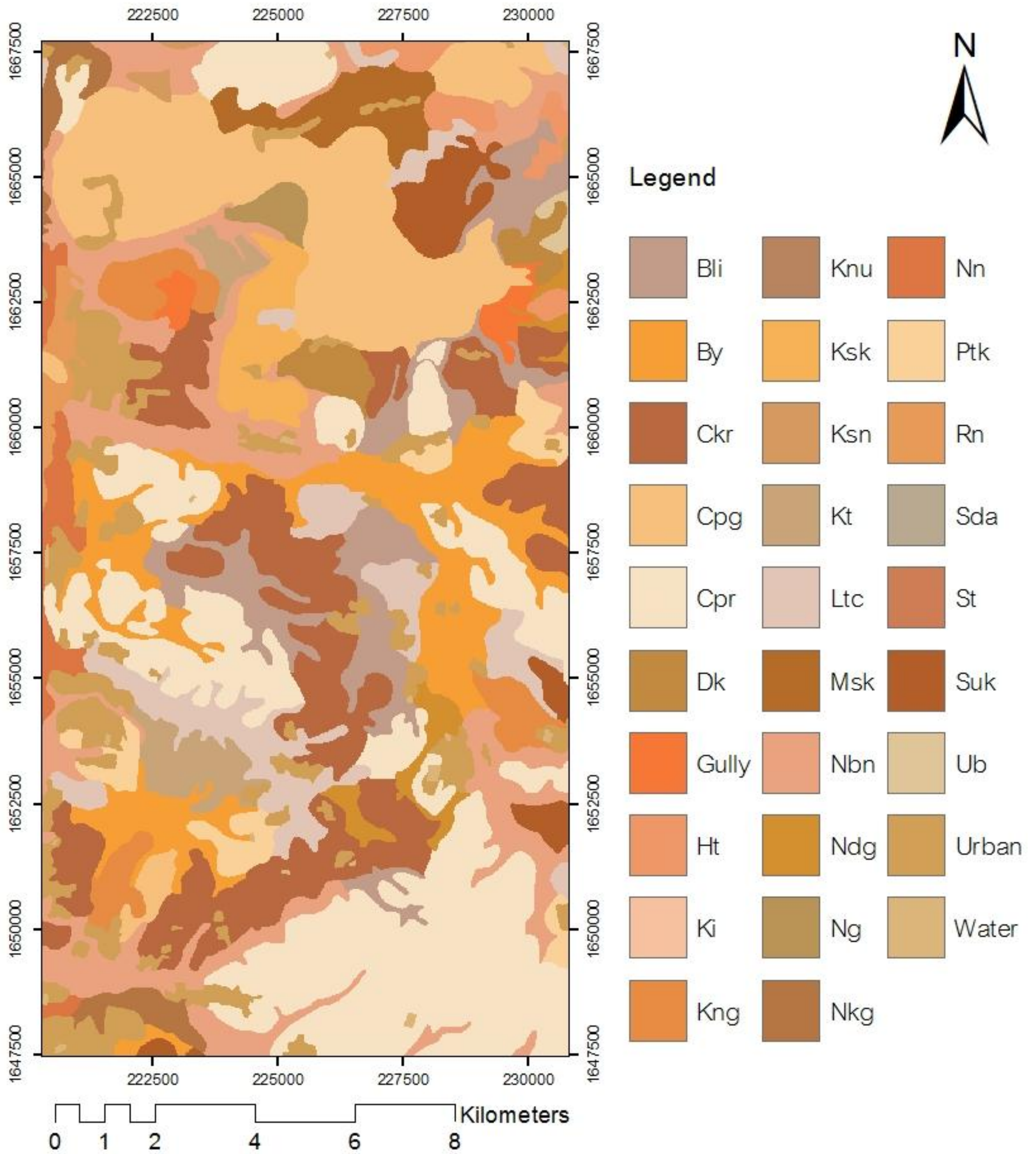


Figure 4.7 Distribution of soil type in the study area.

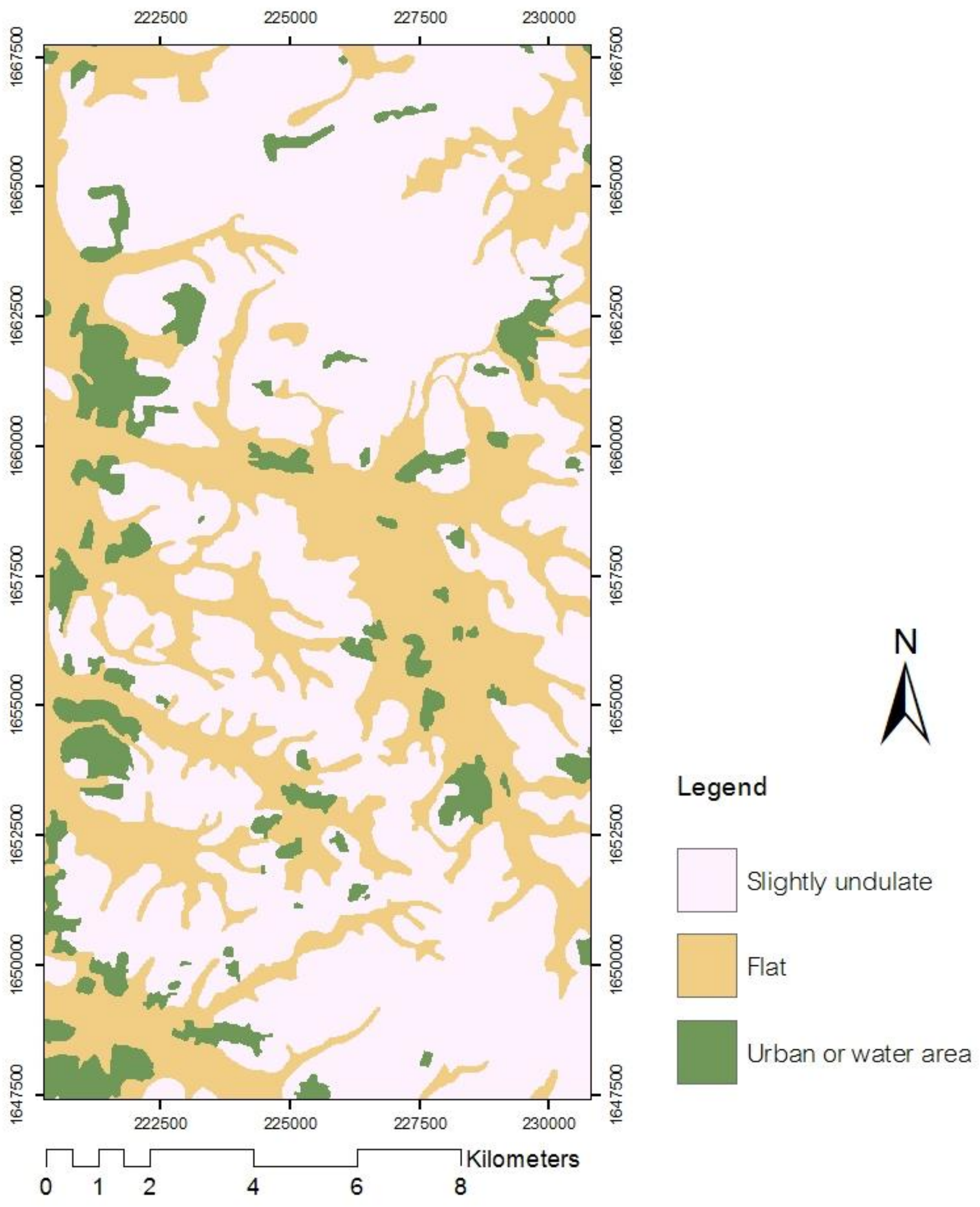


Figure 4.8 Distribution of landform type in the study area.

4.1.2.3 Sugarcane plot data

The location of sugarcane plots in year 2010 of OCSB had been used for field survey and sugarcane training areas selection. In practice, this information was simultaneously considered with land use data of LDD for training areas selection (Figure 4.9).

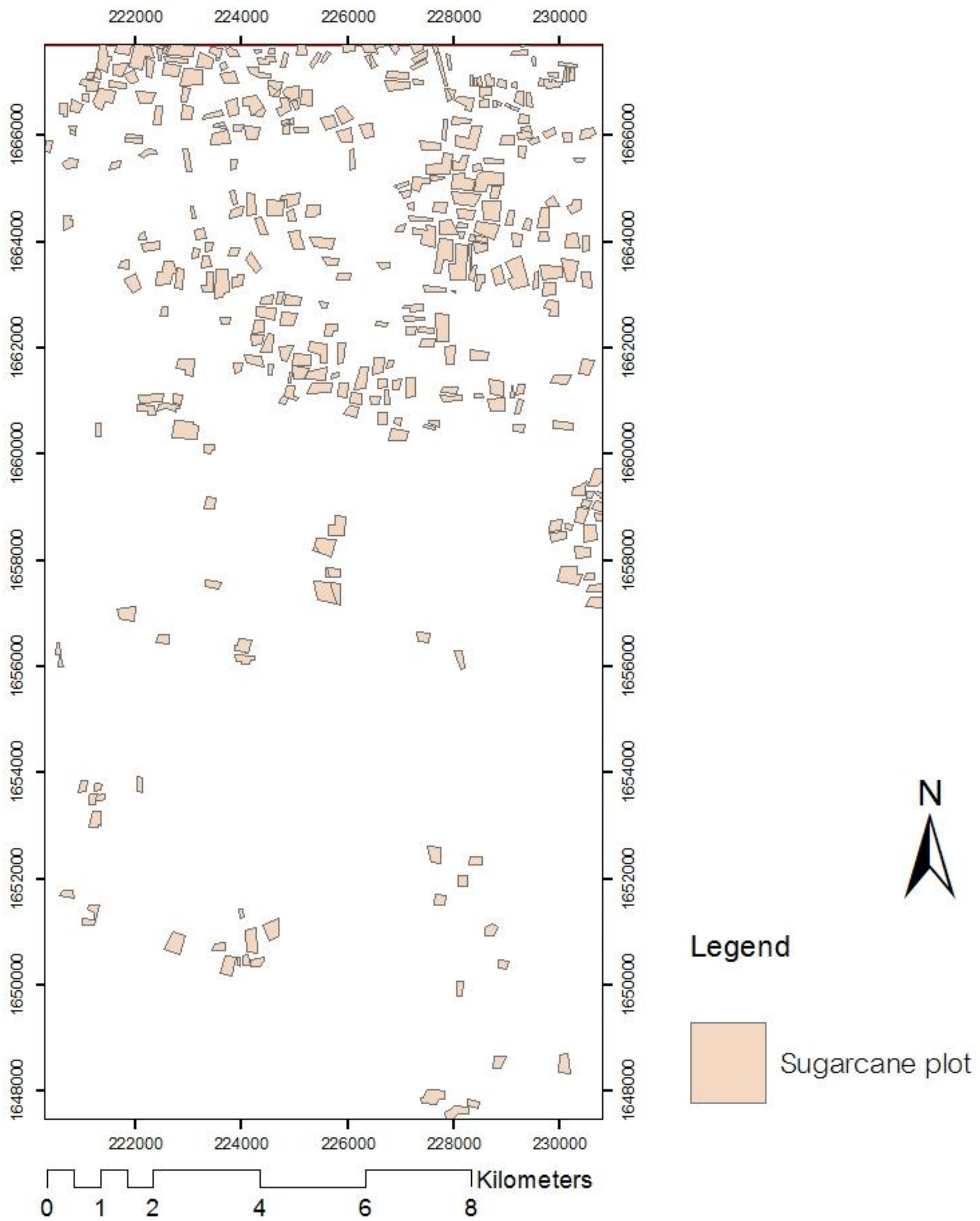
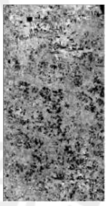
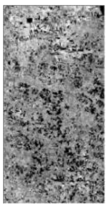


Figure 4.9 Distribution of sugarcane plot data in 2010 in the study area.

4.1.3 Dataset preparation

In this study, predefined dataset were combined between remotely sensed data and GIS data include THEOS multispectral data (4 bands), NDVI, soil series and landform categories. All data were systematically stacked in raster file with cell size of 15 x 15 m^2 and UTM WGS 1984 coordinate system. In fact, this raster file was consisting of 6 dataset for cassava and sugarcane area classification as shown in Table 4.3.

Table 4.3 Combination of remotely sensed data and GIS data in each dataset for cassava and sugarcane area classification.

Dataset	Data				
1	THEOS				
2	THEOS NDVI				
3	THEOS Soil series				
4	THEOS Landform				
5	THEOS Soil series Landform				
6	THEOS NDVI Soil series Landform				

4.2 Sugarcane and cassava classification

Under this study, two main image classification methods which included Expert System and ANN were used to classify cassava and sugarcane area with different assigned 6 datasets. In principle, both methods apply different image classification logic. Expert System does not rely on parametric or nonparametric statistics but it is nonmetric method. In contrast, ANN applies nonparametric statistic. In addition, the procedures for image classification between both methods are completely different.

In practice, two hierarchical decision trees for cassava and sugarcane areas were generated using Knowledge Engineer and classified using Knowledge Classifier under Expert System module of ERDAS Imagine. At the same time, seven land cover types included cassava, sugarcane, urban and built-up area, paddy field, forest land, water bodies and bare land were trained via Neural Net module of ENVI. After that all classified results were filtered using Majority Filtering with 3x3 windows for removing salt and pepper noise. The detail of cassava and sugarcane area classification of each classification method was here separately described in the following section.

4.2.1 Cassava and sugarcane classification using Expert System

Structure of hypothesis, rule and condition for cassava and sugarcane classification for each dataset under Knowledge Engineer was summarized as shown in Table 4.4. Three main categories in each dataset from Expert System were cassava, sugarcane and unclassified. Distribution of cassava and sugarcane areas of each dataset using Expert System classification was shown in Figure 4.10 and Figure 4.11. Area and percentage of cassava and sugarcane areas was summarized as shown in

Table 4.5. In addition, proportion of three extracted categories (cassava, sugarcane and unclassified) for each dataset was also displayed in Figure 4.12.

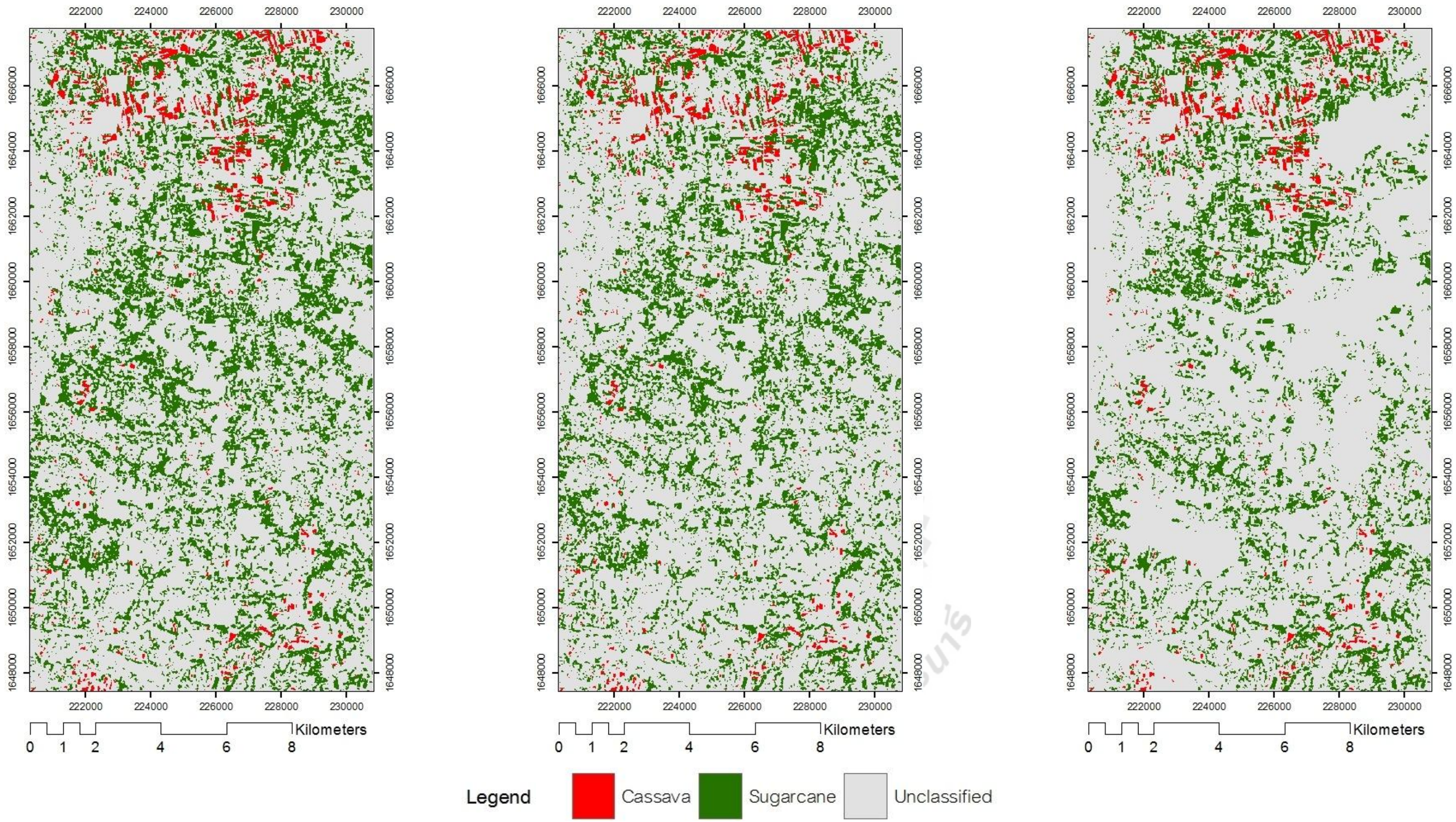
As results, it was found that area of cassava and sugarcane areas extraction using Expert System had continued to decrease when more additional data were added to multispectral dataset. Because it increases more conditions for cassava and sugarcane classification under Expert System as hierarchical decision tree classifier. As shown in Table 4.5 and Figure 4.12, multispectral dataset provides the highest cassava and sugarcane area while multispectral data, NDVI, soil series and landform dataset provides the lowest cassava and sugarcane area.

Table 4.4 Structure of hypothesis, rule and conditions for each dataset.

Predefined Dataset	Hypothesis	Rule	Condition
1. Multispectral data	Cassava	Multispectral data	Band1 61 - 82 Band2 92 - 108 Band3 111 - 124 Band4 161 - 213
	Sugarcane	Multispectral data	Band1 76 - 88 Band2 104 - 117 Band3 118 - 127 Band4 129 - 150
2. Multispectral data and NDVI	Cassava	Multispectral data	Band1 61 - 82 Band2 92 - 108 Band3 111 - 124 Band4 161 - 213
		NDVI data	0.314 - 0.529
	Sugarcane	Multispectral data	Band1 76 - 88 Band2 104 - 117 Band3 118 - 127 Band4 129 - 150
		NDVI data	0.194 - 0.316
3. Multispectral data and Soil series	Cassava	Multispectral data	Band1 61 - 82 Band2 92 - 108 Band3 111 - 124 Band4 161 - 213
		Soil series data	Soil series classes By, Ckr, Cpg, Cpr, Dk, Ht, Kng, Ksk, Kt, Ltc, Msk, Nbn, Ptk, Suk
	Sugarcane	Multispectral data	Band1 76 - 88 Band2 104 - 117 Band3 118 - 127 Band4 129 - 150
		Soil series data	Soil series classes Ckr, Cpg, Cpr, ,Dk, Ht, Kng, Ksk, Kt, Ltc, Msk, Nbn, Ng, Ptk, Suk
4.Multispectral data and Landform	Cassava	Multispectral data	Band1 61 - 82 Band2 92 - 108 Band3 111 - 124 Band4 161 - 213
		Landform data	Landform class B
	Sugarcane	Multispectral data	Band1 76 - 88 Band2 104 - 117 Band3 118 - 127 Band4 129 - 150
		Landform data	Landform class B

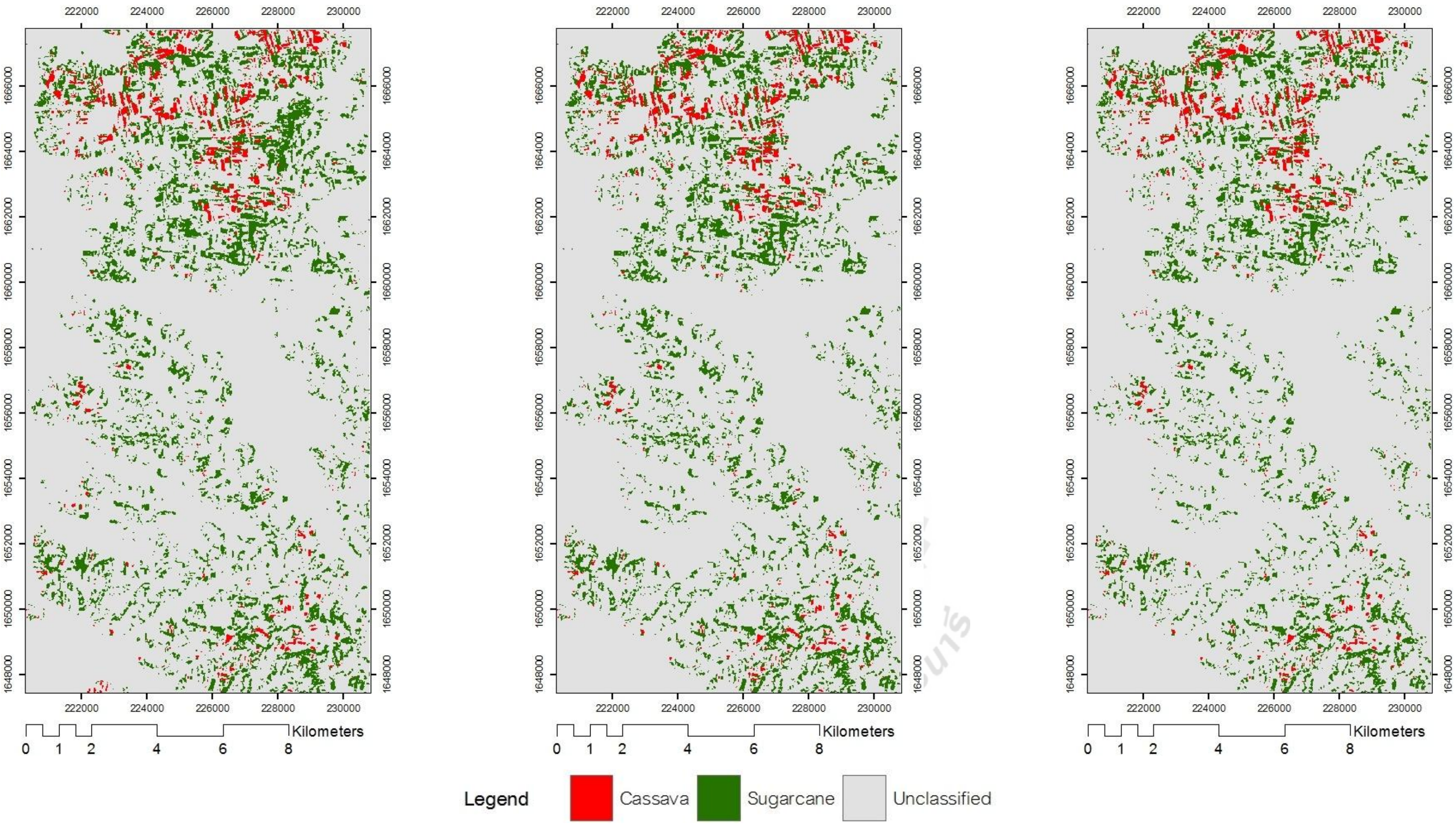
Table 4.4 (Continued).

Predefined Dataset	Hypothesis	Rule	Condition
5. Multispectral data, Soil series and Landform	Cassava	Multispectral data	Band1 61 – 82 Band2 92 - 108 Band3 111 - 124 Band4 161 – 213
		Soil series data	Soil series classes By, Ckr, Cpg, Cpr, Dk, Ht, Kng, Ksk, Kt, Ltc, Msk, Nbn, Ptk, Suk
		Landform data	Landform class B
	Sugarcane	Multispectral data	Band1 76 – 88 Band2 104 - 117 Band3 118 - 127 Band4 129 – 150
		Soil series data	Soil series classes Ckr, Cpg, Cpr, ,Dk, Ht, Kng, Ksk, Kt, Ltc, Msk, Nbn, Ng, Ptk, Suk
		Landform data	Landform class B
6. Multispectral data, NDVI, Soil series and Landform	Cassava	Multispectral data	Band1 61 – 82 Band2 92 - 108 Band3 111 - 124 Band4 161 – 213
		NDVI data	0.314 - 0.529
		Soil series data	Soil series classes By, Ckr, Cpg, Cpr, Dk, Ht, Kng, Ksk, Kt, Ltc, Msk, Nbn, Ptk, Suk
		Landform data	Landform class B
	Sugarcane	Multispectral data	Band1 76 – 88 Band2 104 - 117 Band3 118 - 127 Band4 129 – 150
		NDVI data	0.194 - 0.316
		Soil series data	Soil series classes Ckr, Cpg, Cpr, ,Dk, Ht, Kng, Ksk, Kt, Ltc, Msk, Nbn, Ng, Ptk, Suk
		Landform data	Landform class B



(a) (b) (c)

Figure 4.10 Distribution of cassava, sugarcane and unclassified areas from each dataset using Expert System: (a) multispectral dataset, (b) multispectral and NDVI dataset and (c) multispectral and soil series dataset.



(a) (b) (c)

Figure 4.11 Distribution of cassava, sugarcane and unclassified areas from each dataset using Expert System: (a) multispectral and landform dataset, (b) multispectral, soil series and landform dataset and (c) multispectral, NDVI, soil series and landform dataset.

Table 4.5 Area and percentage of cassava and sugarcane classification from each dataset using Expert System.

Dataset	Cassava		Sugarcane		Unclassified	
	Area (sq. km)	Percent	Area (sq. km)	Percent	Area (sq. km)	Percent
Multispectral data	5.70	2.68	55.95	26.26	151.43	71.07
Multispectral data and NDVI	5.52	2.59	49.15	23.06	158.42	74.34
Multispectral data and Soil series	5.47	2.57	37.75	17.71	169.87	79.72
Multispectral data and Landform	5.11	2.40	26.21	12.30	181.77	85.30
Multispectral data, Soil series and Landform	4.96	2.33	23.96	11.24	184.17	86.43
Multispectral data, NDVI, Soil series and Landform	4.80	2.25	20.67	9.70	187.61	88.04

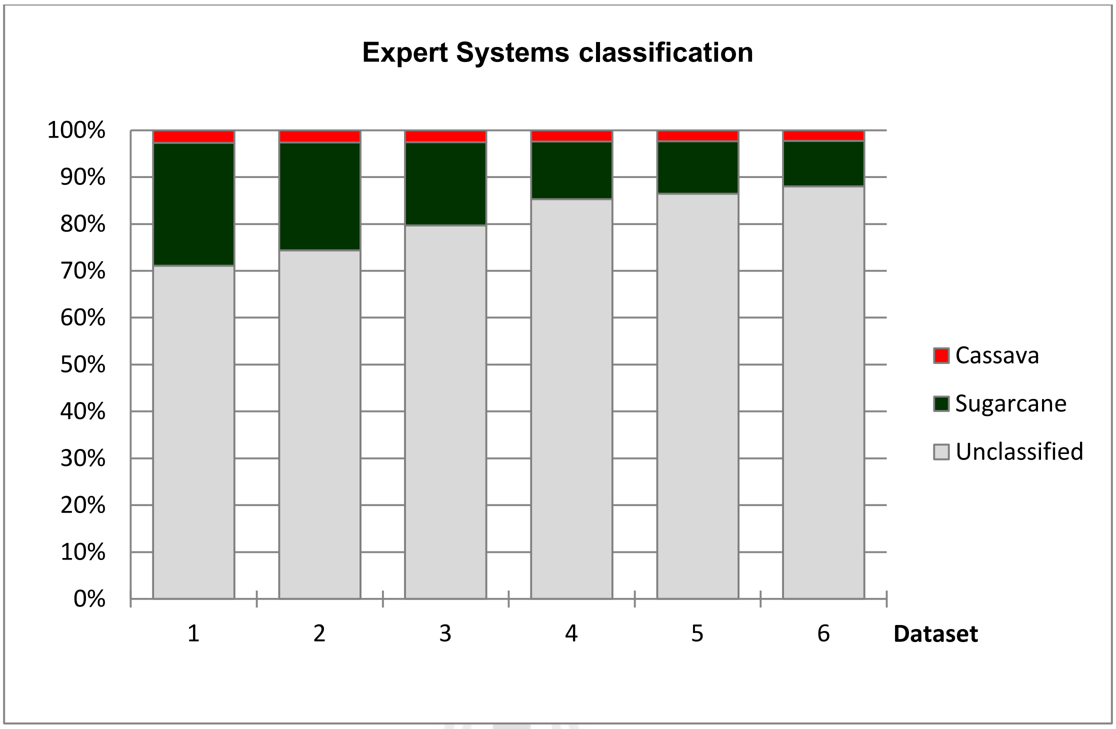


Figure 4.12 The proportion chart of Expert Systems classification in each dataset.

Remark

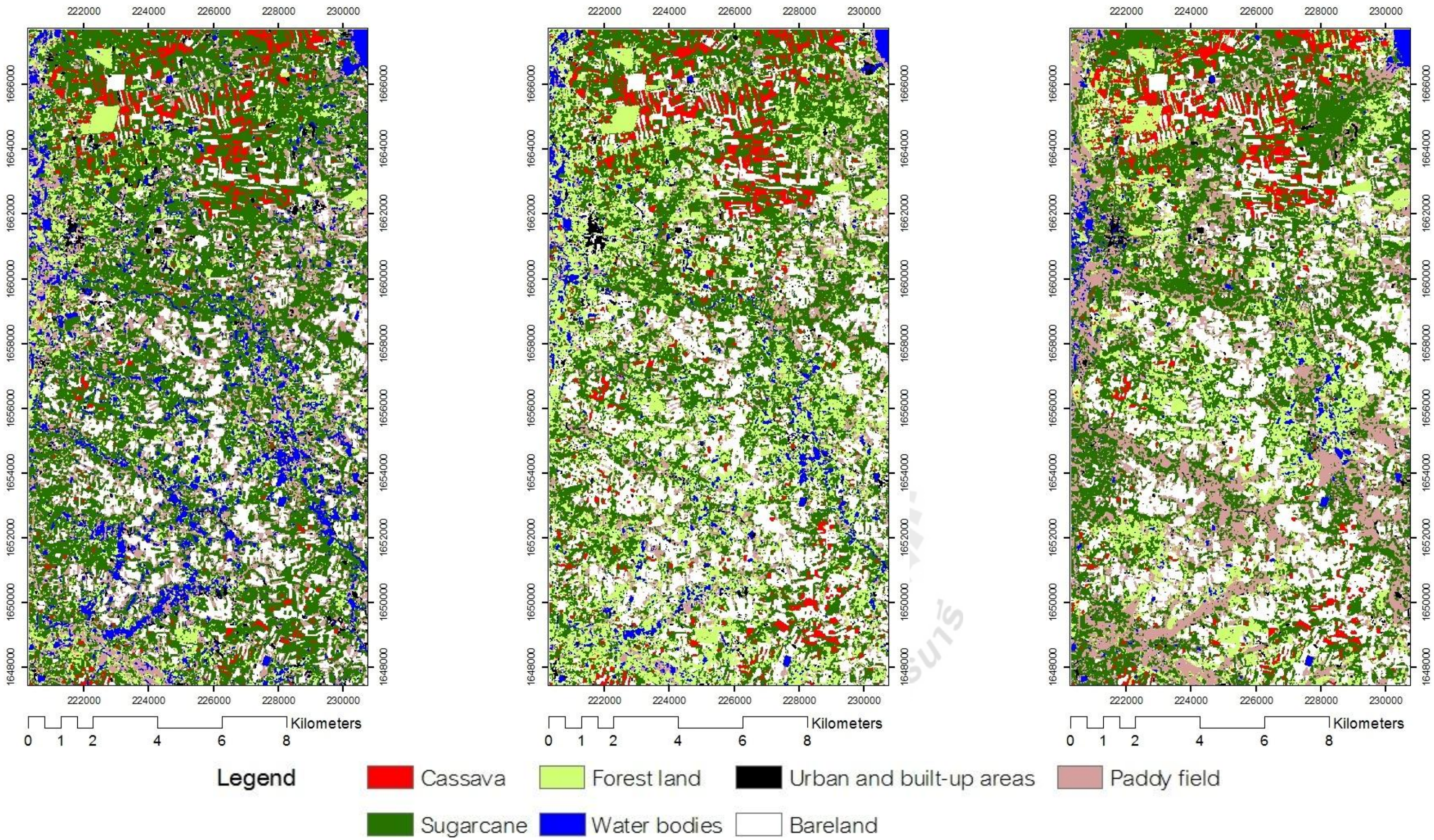
1 Multispectral data (band 1, 2, 3, 4) dataset
2 Multispectral data and NDVI dataset
3 Multispectral data and Soil series dataset
4 Multispectral data and Landform dataset
5 Multispectral data, Soil series and Landform dataset
6 Multispectral data, NDVI, Soil series and Landform dataset

4.2.2 Sugarcane and cassava classification using ANN

Neural Net of ENVI had been applied for ANN classification that was random learning rate and fixed iteration for pruning networks. A higher rate will increase the risk of oscillations or non-convergence of the training result so learning rate had been random from 0.1 to 0.3. All dataset were set 10,000 rounds for iteration which varied values of RMS error in each round. The least of RMS values was the optimum result. Normally, the results of classification will provide output classes equal to number of training data that input to network.

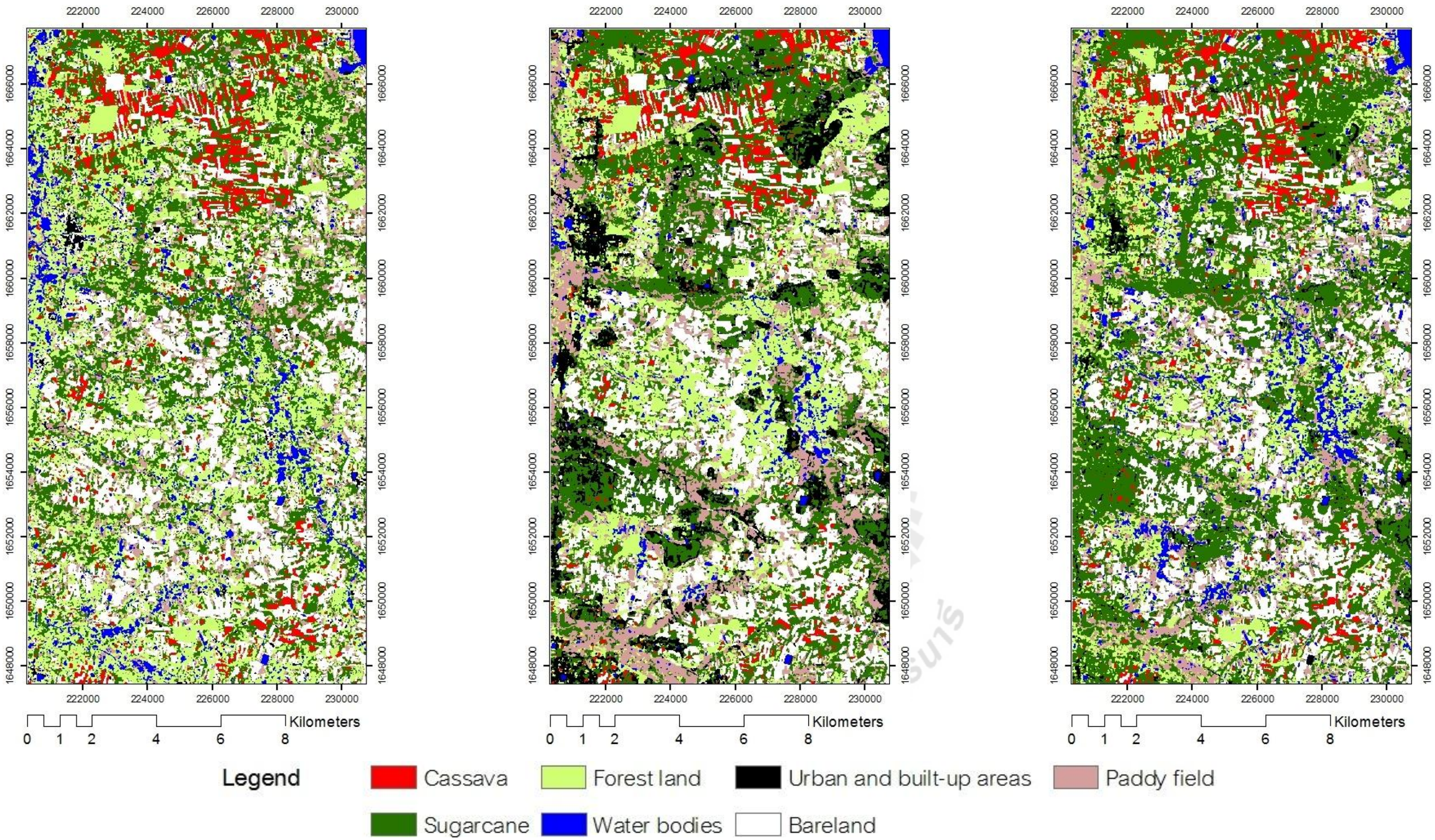
The best results of land cover classification with different learning rate (0.1, 0.2 and 0.3) were separately presented as figure and table forms. Land cover distribution using ANN with 0.1 learning rate was displayed in Figure 4.13 and Figure 4.14 while area and percentage was summarized as shown in Table 4.6 and Figure 4.15. Similarity, land cover distribution using ANN with 0.2 learning rate was displayed in Figure 4.16 and Figure 4.17 while area and percentage was summarized as shown in Table 4.7 and Figure 4. 18. Also, land cover distribution using ANN with 0.3 learning rate was displayed in Figure 4.19 to Figure 4.20 while area and percentage was summarized as shown in Table 4.8 and Figure 4.21.

.



(a) (b) (c)

Figure 4.13 Land cover distribution from each dataset using ANN with 0.1 learning rate: (a) multispectral dataset, (b) multispectral and NDVI dataset and (c) multispectral and soil series dataset.



(a) (b) (c)

Figure 4.14 Land cover distribution from each dataset using ANN with 0.1 learning rate: (a) multispectral and landform dataset, (b) multispectral, soil series and landform dataset and (c) multispectral, NDVI, soil series and landform dataset.

Table 4.6 Area and percentage of land cover classification from each dataset using ANN with learning rate 0.1.

Dataset	Land cover type (sq. km)						
	Cassava	Sugarcane	Urban and Built-up area	Paddy field	Forest land	Water body	Bare land
Multispectral data	7.52 (3.53%)	92.61 (43.46%)	4.89 (2.30%)	36.43 (17.10%)	22.25 (10.44%)	17.46 (8.20%)	31.93 (14.98%)
Multispectral data and NDVI	12.12 (5.69%)	70.96 (33.30%)	3.63 (1.70%)	19.63 (9.21%)	48.00 (22.53%)	7.84 (3.68%)	50.90 (23.89%)
Multispectral data and Soil series	10.92 (5.12%)	75.80 (35.57%)	2.35 (1.10%)	41.68 (19.56%)	29.40 (13.80%)	4.15 (1.95%)	48.79 (22.90%)
Multispectral data and Landform	13.15 (6.17%)	56.62 (26.57%)	2.88 (1.35%)	21.83 (10.24%)	60.68 (28.48%)	10.99 (5.16%)	46.94 (22.03%)
Multispectral data, Soil series and Landform	9.65 (4.53%)	65.14 (30.57%)	20.19 (9.47%)	31.86 (14.95%)	38.30 (17.97%)	5.24 (2.46%)	42.70 (20.04%)
Multispectral data, NDVI, Soil series and Landform	11.54 (5.42%)	81.18 (38.10%)	4.06 (1.91%)	26.33 (12.36%)	38.08 (17.87%)	11.50 (5.40%)	40.40 (18.96%)

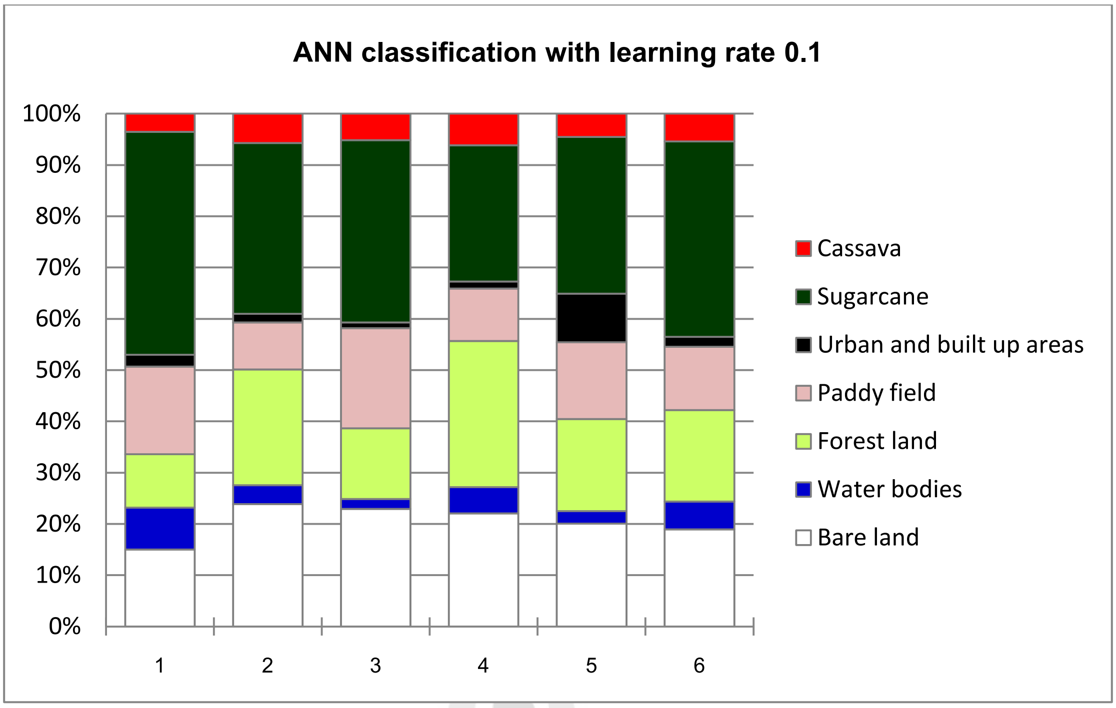


Figure 4.15 Proportion of land cover areas in each dataset using ANN classification with learning rate 0.1.

Remark

1 Multispectral data (band 1, 2, 3, 4) dataset
2 Multispectral data and NDVI dataset
3 Multispectral data and Soil series dataset
4 Multispectral data and Landform dataset
5 Multispectral data, Soil series and Landform dataset
6 Multispectral data, NDVI, Soil series and Landform dataset

As shown in Table 4.6 and Figure 4.15, multispectral data and landform dataset provides the highest cassava area while multispectral dataset provides the lowest cassava area. At the same time multispectral dataset provides the highest sugarcane area while multispectral data and landform dataset provides the lowest sugarcane area.

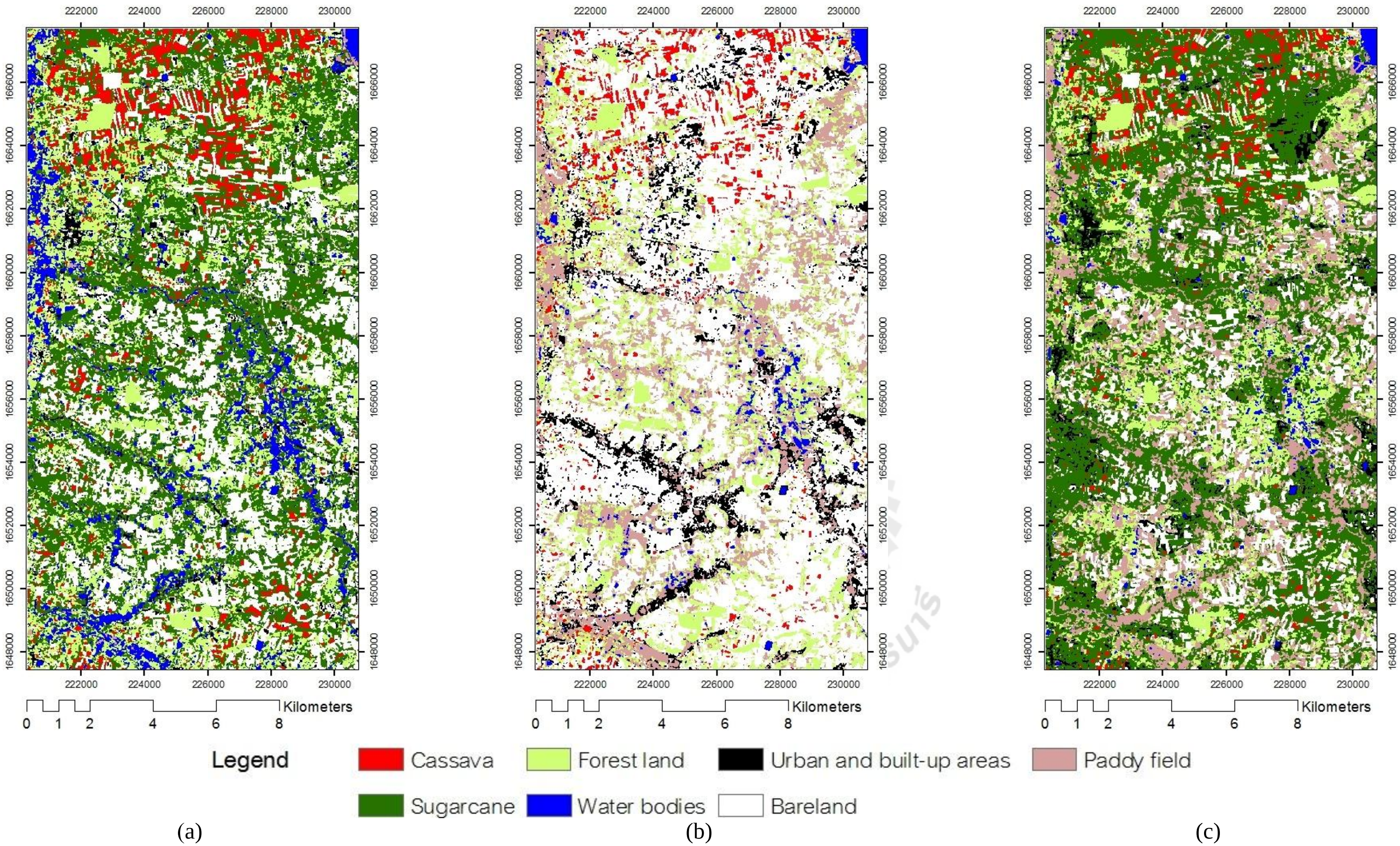
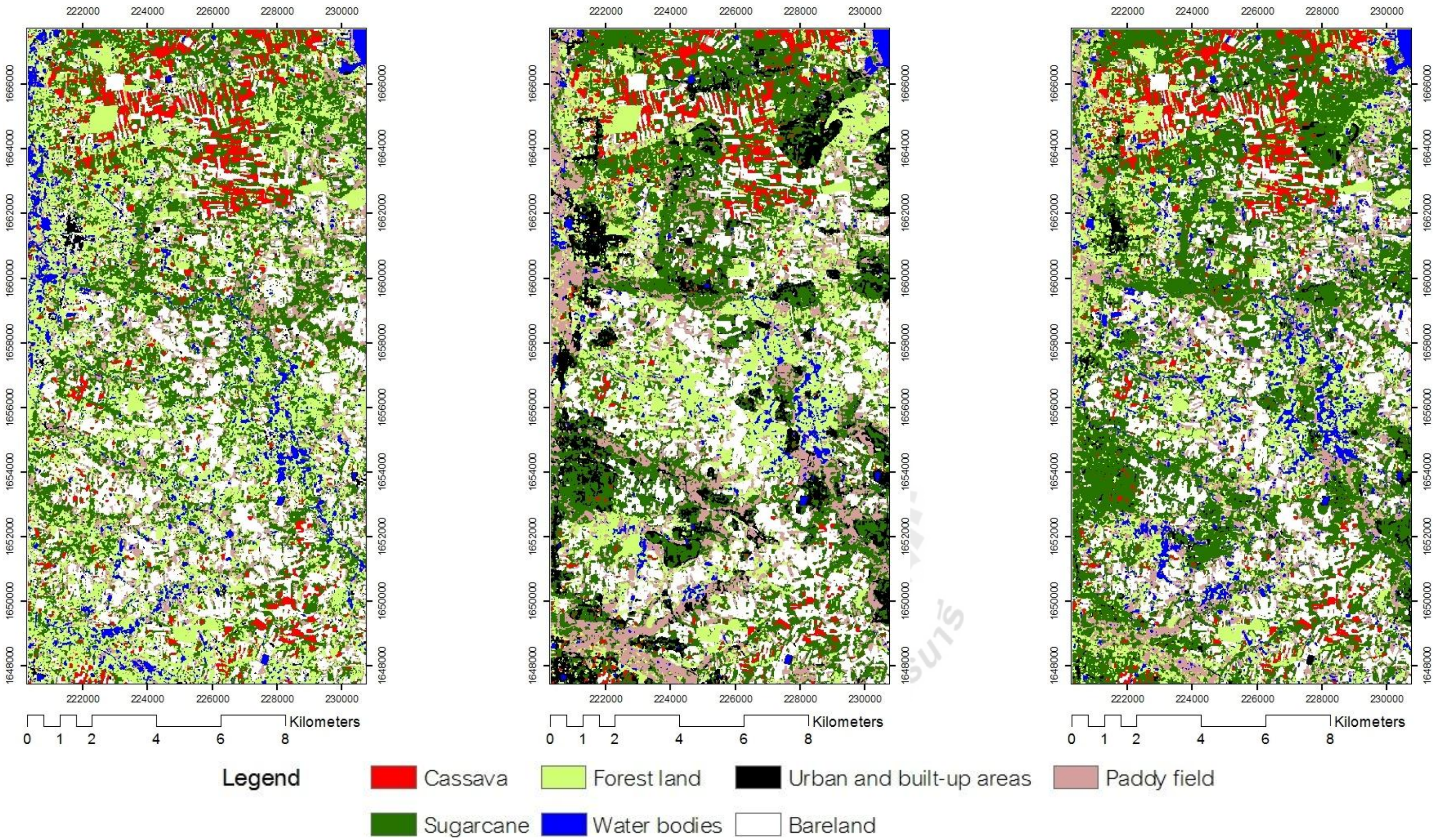


Figure 4.16 Land cover distribution from each dataset using ANN with 0.2 learning rate: (a) multispectral dataset, (b) multispectral and NDVI dataset and (c) multispectral and soil series dataset.



(a) (b) (c)

Figure 4.17 Land cover distribution from each dataset using ANN with 0.2 learning rate: (a) multispectral and landform dataset, (b) multispectral, soil series and landform dataset and (c) multispectral, NDVI, soil series and landform dataset.

Table 4.7 Area and percentage of land cover classification form each dataset using ANN with learning rate 0.2.

Dataset	Land cover type (Sq. Km)						
	Cassava	**Sugarcane**	**Urban and Built-up area**	**Paddy field**	**Forest land**	**Water body**	**Bare land**
Multispectral data	8.04 (3.77%)	88.90 (41.72%)	2.99 (1.40%)	25.46 (11.95%)	36.74 (17.24%)	5.04 (2.37%)	45.92 (21.55%)
Multispectral data and NDVI	10.48 (4.92%)	62.45 (29.31%)	4.30 (2.02%)	0.21 (0.10%)	53.85 (25.27%)	11.67 (5.48%)	70.12 (32.91%)
Multispectral data and Soil series	16.37 (7.68%)	107.37 (50.39%)	0.28 (0.13%)	14.24 (6.68%)	6.98 (3.27%)	7.92 (3.72%)	59.93 (28.13%)
Multispectral data and Landform	14.85 (6.97%)	79.80 (37.45%)	5.93 (2.78%)	1.88 (0.88%)	43.54 (20.43%)	15.29 (7.18%)	51.80 (24.31%)
Multispectral data, Soil series and Landform	9.97 (4.68%)	74.30 (34.87%)	16.57 (7.78%)	31.08 (14.59%)	38.97 (18.29%)	3.80 (1.78%)	38.39 (18.01%)
Multispectral data, NDVI, Soil series and Landform	6.57 (3.08%)	94.68 (44.43%)	6.88 (3.23%)	39.21 (18.40%)	32.73 (15.36%)	3.29 (1.54%)	29.73 (13.95%)

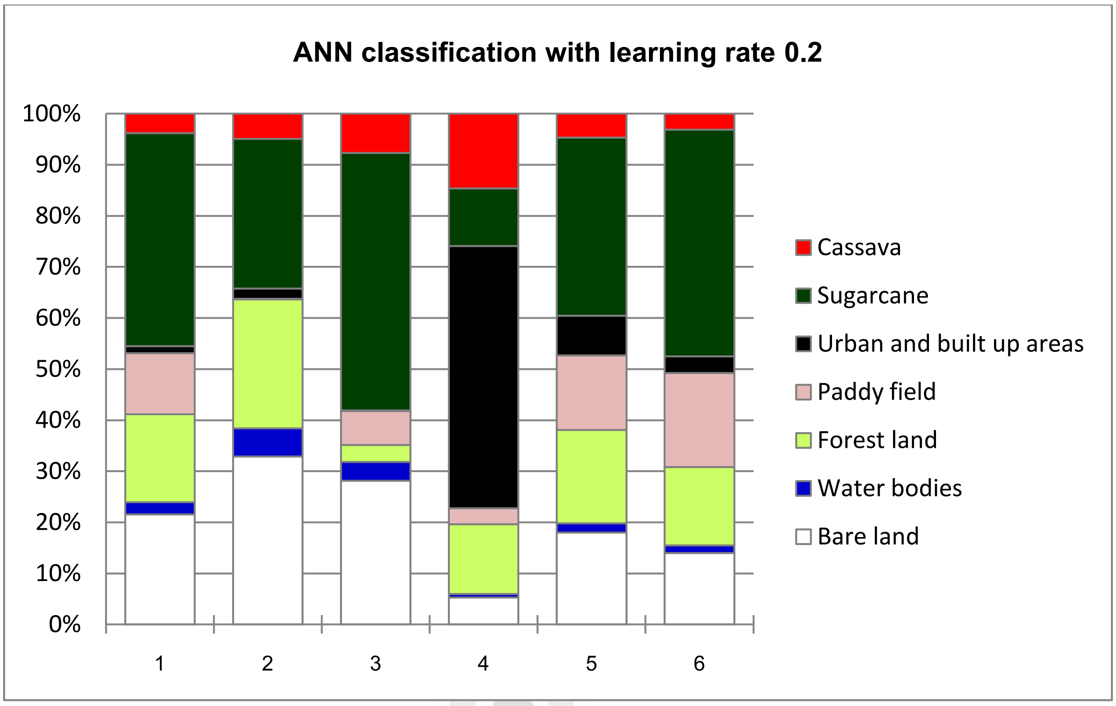


Figure 4.18 Proportion of land cover areas in each dataset using ANN classification with learning rate 0.2.

Remark

1	Multispectral data (band 1, 2, 3, 4) dataset
2	Multispectral data and NDVI dataset
3	Multispectral data and Soil series dataset
4	Multispectral data and Landform dataset
5	Multispectral data, Soil series and Landform dataset
6	Multispectral data, NDVI, Soil series and Landform dataset

As shown in Table 4.7 and Figure 4.18, multispectral data and soil series dataset provides the highest cassava area while multispectral data, NDVI, soil series and landform dataset provides the lowest cassava area. At the same multispectral data and soil series dataset provides the highest sugarcane area while multispectral data and NDVI dataset provides the lowest sugarcane area.

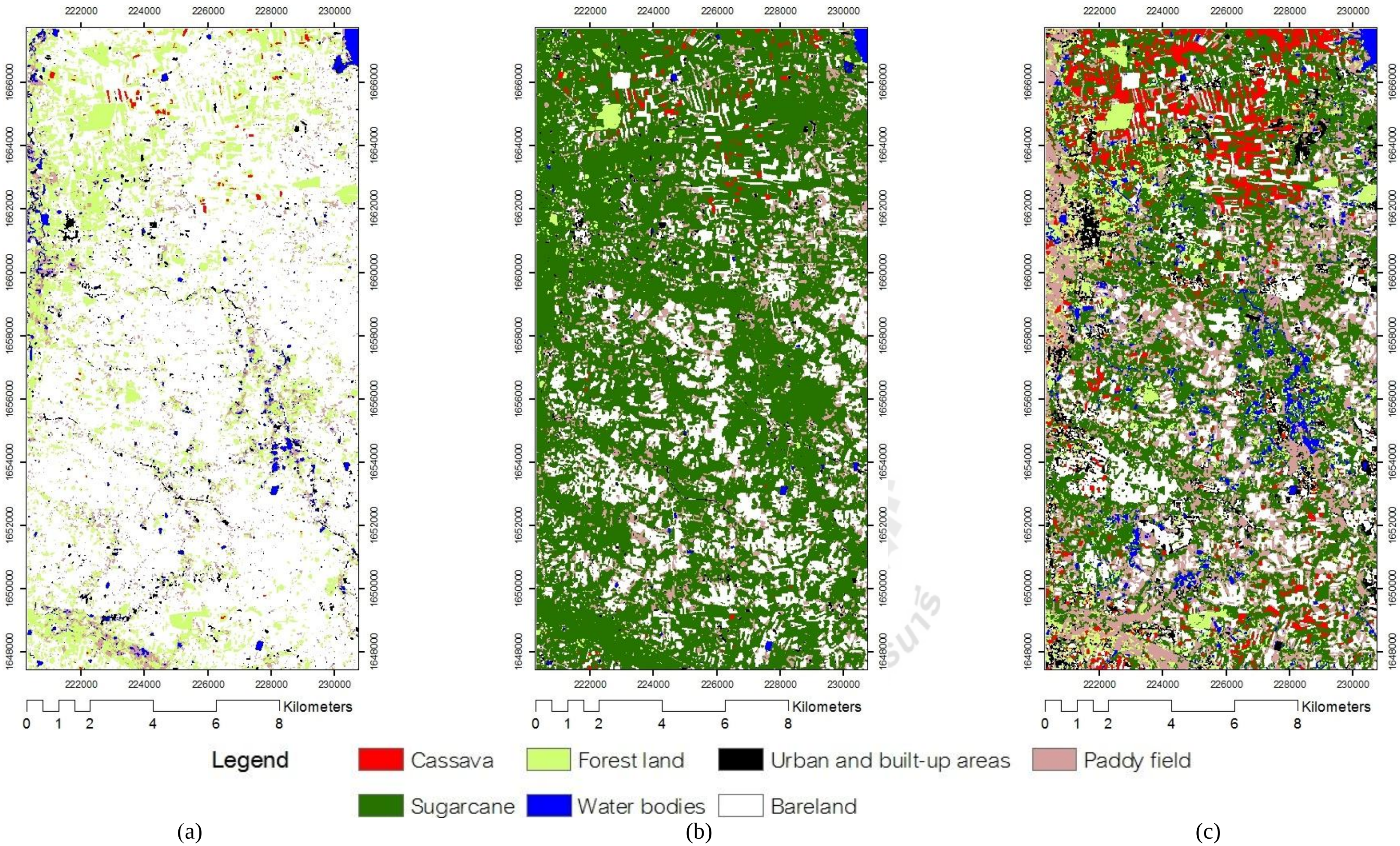
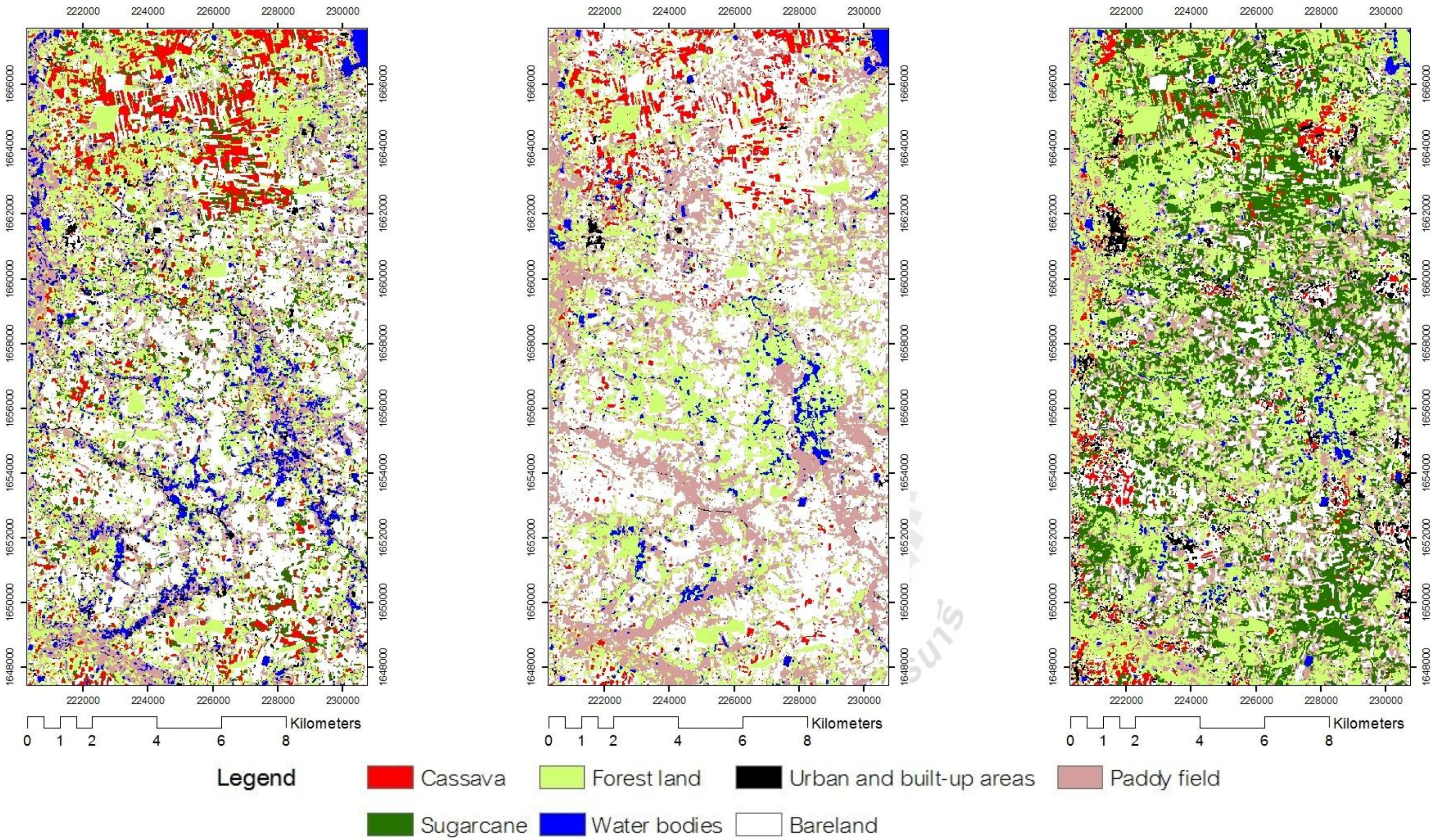


Figure 4.19 Land cover distribution from each dataset using ANN with 0.3 learning rate: (a) multispectral dataset, (b) multispectral and NDVI dataset and (c) multispectral and soil series dataset.



(a) (b) (c)

Figure 4.20 Land cover distribution from each dataset using ANN with 0.3 learning rate: (a) multispectral and landform dataset, (b) multispectral, soil series and landform dataset and (c) multispectral, NDVI, soil series and landform dataset.

Table 4.8 Area and percentage of land cover classification from each dataset using ANN with learning rate 0.3.

Dataset	Land cover type (Sq. Km)						
	Cassava	**Sugarcane**	**Urban and Built-up area**	**Paddy field**	**Forest land**	**Water body**	**Bare land**
Multispectral data	0.72 (0.34%)	99.09 (46.50%)	3.69 (1.73%)	11.16 (5.24%)	37.17 (17.44%)	3.23 (1.52%)	58.04 (27.24%)
Multispectral data and NDVI	1.56 (0.73%)	143.58 (67.38%)	2.33 (1.10%)	20.27 (9.51%)	2.09 (0.98%)	0.87 (0.41%)	42.40 (19.90%)
Multispectral data and Soil series	15.72 (7.38%)	79.19 (37.16%)	10.06 (4.72%)	44.57 (20.92%)	17.73 (8.32%)	7.57 (3.55%)	38.26 (17.96%)
Multispectral data and Landform	17.11 (8.03%)	11.86 (5.57%)	6.09 (2.86%)	35.76 (16.78%)	57.60 (27.03%)	12.32 (5.78%)	72.35 (33.95%)
Multispectral data, Soil series and Landform	10.76 (5.05%)	67.90 (31.87%)	1.36 (0.64%)	52.10 (24.45%)	49.13 (23.06%)	6.29 (2.95%)	25.54 (11.99%)
Multispectral data, NDVI, Soil series and Landform	8.97 (4.21%)	50.25 (23.58%)	7.59 (3.56%)	30.46 (14.29%)	74.44 (34.94%)	4.86 (2.28%)	36.52 (17.14%)

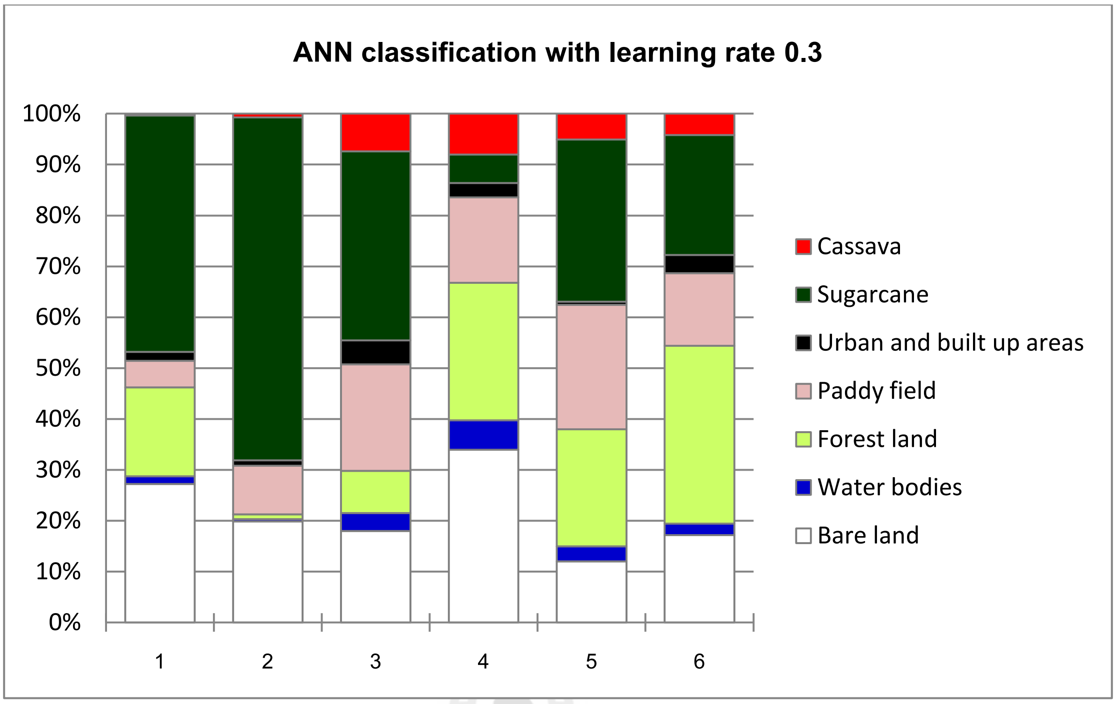


Figure 4.21 Proportion of land cover areas in each dataset using ANN classification with learning rate 0.3.

Remark

1 Multispectral data (band 1, 2, 3, 4) dataset
2 Multispectral data and NDVI dataset
3 Multispectral data and Soil series dataset
4 Multispectral data and Landform dataset
5 Multispectral data, Soil series and Landform dataset
6 Multispectral data, NDVI, Soil series and Landform dataset

As shown in Table 4.8 and Figure 4.21, multispectral data and landform dataset provides the highest cassava area while multispectral dataset provides the lowest cassava area. At the same multispectral data and NDVI dataset provides the highest sugarcane area while multispectral data and landform dataset provides the lowest sugarcane area.

Based on land cover extraction using ANN classification within specific learning rates (0.1 or 0.2 or 0.3), it was found that extracted land cover areas from each dataset are self-determining. They were no relationship among dataset. In addition, based on comparison of extracted land cover areas between three different of

ANN classification (Table 4.9), it revealed that extracted land cover areas from each dataset among learning rates are also self-determining. This uncertainty of land cover areas among various datasets or different learning rate of ANN classification might be come from training areas identification. In the study, training areas were collected from multispectral dataset and then applied for all dataset in classification

Table 4.9 Comparison of extracted area of cassava and sugarcane areas from each dataset using ANN with different learning rate.

Dataset	**ANN with learning rate of 0.1**		**ANN with learning rate of 0.2**		**ANN with learning rate of 0.3**	
	Cassava	**Sugarcane**	**Cassava**	**Sugarcane**	**Cassava**	**Sugarcane**
Multispectral data	7.52	92.61	8.04	88.90	0.72	99.09
Multispectral data and NDVI	12.12	70.96	10.48	62.45	1.56	143.58
Multispectral data and Soil series	10.92	75.80	16.37	107.37	15.72	79.19
Multispectral data and Landform	13.15	56.62	14.85	79.80	17.11	11.86
Multispectral data, Soil series and Landform	9.65	65.14	9.97	74.30	10.76	67.90
Multispectral data, NDVI, Soil series and Landform	11.54	81.18	6.57	94.68	8.97	50.25

4.3 Ground truth and accuracy assessment

4.3.1 Calculate samples point for accuracy assessment

In the study, ground truths from field survey were used for accuracy assessment based on multinomial distribution. This approach was generating an optimum sample size. Number of sample points was responsible to the desired confidence level and the desired precision (See equation 3.2).

Herein, number of sample points with 90% of confidence level and 10% of precision was 122 points were required for accuracy assessment. In practice, stratified random sampling scheme was applied to identify samples location. In fact, minimum sample point per class should be 5 points. The distribution of sampling points for accuracy assessment was displayed in Figure 4.22. Detail of sampling point and ground photograph of each land cover class from field survey was presented in Table B.1 and Table B.2, respectively in Appendix B.

The accuracies which were represented in error matrix are very effective way to represent map accuracy. In general two approaches include univariate and multivariate statistical analyses are applied for accuracy assessment using information from error matrix. Firstly, overall accuracy, producer’s accuracy (omission errors) and user’s accuracy (commission errors) are normally report for accuracy assessment. Secondly, Kappa coefficient used in accuracy assessment to statistically determine the error matrix is significantly different from another and the conditional Kappa coefficient that is agreement for an individual category. The values of conditional Kappa coefficient can range from +1 to −1. However, positive conditional Kappa coefficient values between the remotely sensed classification and the reference data are expected under three groups characterized by Landis and Koch (1977); a value greater than 0.80 (i.e., >80%) represents strong agreement, a value between 0.40 and 0.80 (i.e., 40–80%) represents moderate agreement, and a value below 0.40 (i.e., <40%) represents poor agreement (Congalton and Green, 2008).

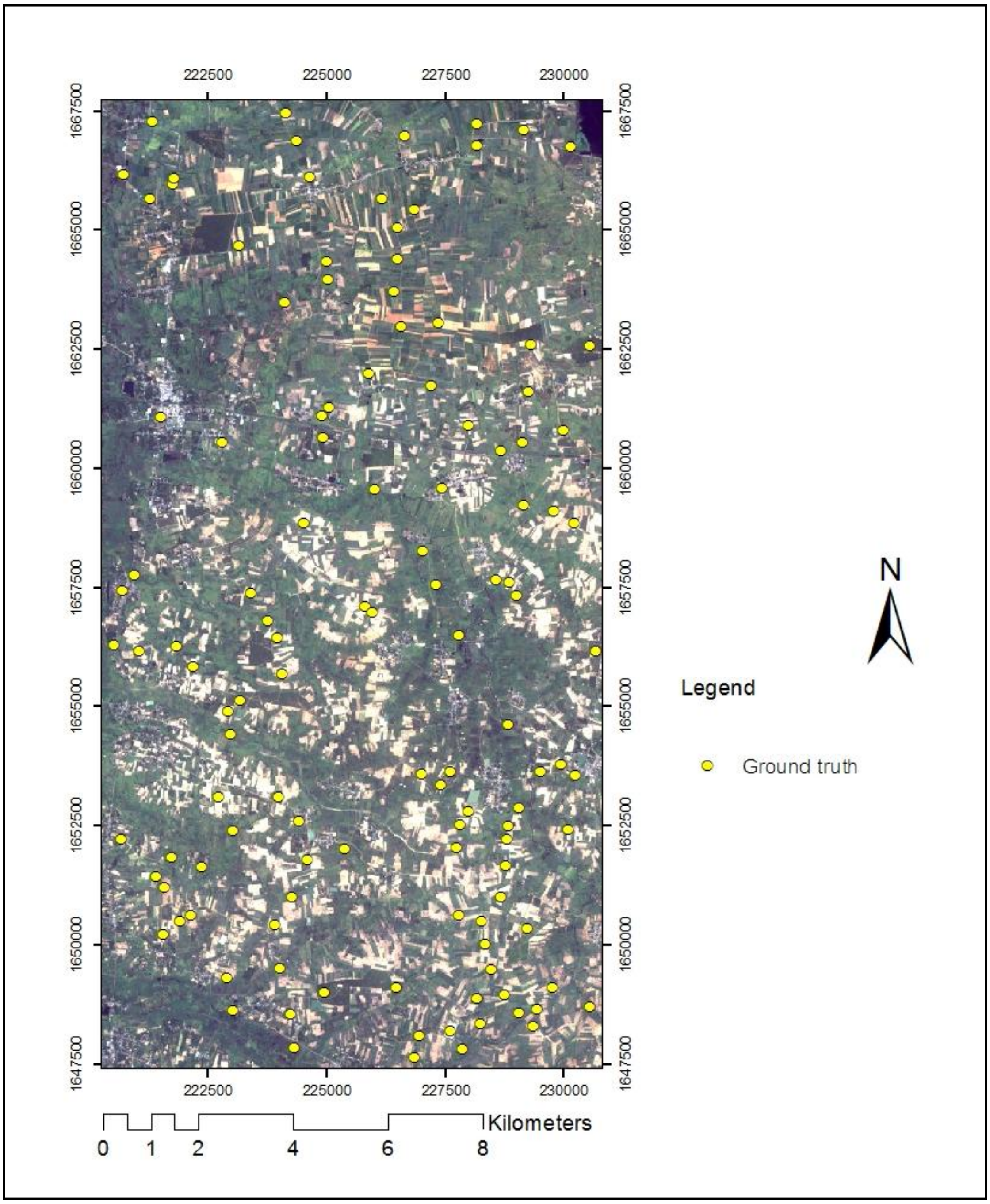


Figure 4.22 Distribution of ground truth location for accuracy assessment.

4.3.2 Accuracy assessment of Expert System

The accuracy assessments of cassava and sugarcane classification for each dataset from Expert System were presented in detail with error matrix as shown in Table C.1 to Table C.6 in Appendix C. Major results of accuracy assessment from each dataset can be summarized as shown in Table 4.10 and Table 4.11.

For multispectral dataset, overall accuracy of cassava and sugarcane classification was 76.23% while cassava and sugarcane classes provided PA and UA more than 70%. The Kappa coefficient was 63.06% which represented moderate agreement.

At the same time, for multispectral and NDVI dataset, overall accuracy of cassava and sugarcane classification was 77.05% and classes that provided PA and UA more than 70% were cassava and sugarcane. The Kappa coefficient was 64.27% which represented moderate agreement. Simultaneous, for multispectral and soil series dataset, overall accuracy of cassava and sugarcane classification was 77.87% and classes that provides PA and UA more than 70% was cassava. The Kappa coefficient was 65.23% which represented moderate agreement. Meanwhile, for multispectral and landform dataset overall accuracy of cassava and sugarcane classification was 79.51% and classes that provides PA and UA more than 70% was sugarcane. The Kappa coefficient was 67.69% which represented moderate agreement.

Meanwhile for multispectral, soil series and landform dataset and multispectral, NDVI, soil series and landform dataset overall accuracies of cassava and sugarcane classification was 79.51% and cassava and sugarcane classes provided

PA more than 60% while UA more than 80%. The Kappa coefficient was 67.58% which represented moderate agreement.

Table 4.10 Summary of overall accuracy and Kappa coefficient for land cover classification by Expert Systems.

Dataset	Description	Overall Accuracy (%)	Kappa coefficient (%)	Rank
1	Multispectral data	76.23	63.06	6
2	Multispectral data and NDVI	77.05	64.27	5
3	Multispectral data and Soil series	77.87	65.23	4
4	Multispectral data and Landform	79.51	67.69	1
5	Multispectral data, Soil series and Landform	79.51	67.58	2
6	Multispectral data, NDVI, Soil series and Landform	79.51	67.58	2

Table 4.11 Summary of simple PA and UA and conditional Kappa PA and UA for each category by Expert System Classification.

Classes	Simple PA and UA for each dataset (%)											
	1		2		3		4		5		6	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Cassava	71.43	100.00	71.43	100.00	71.43	100.00	68.57	100.00	68.57	100.00	68.57	100.00
Sugarcane	74.29	72.22	74.29	74.29	68.57	80.00	71.43	86.21	68.57	88.89	68.57	88.89
Unclassified	80.77	68.85	82.69	69.35	88.46	68.66	92.31	69.57	94.23	69.01	94.23	69.01
Classes	**Conditional Kappa coefficient for each dataset (%)**											
	1		2		3		4		5		6	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Cassava	64.06	100.00	64.06	100.00	64.06	100.00	60.87	100.00	60.87	100.00	60.87	100.00
Sugarcane	63.52	61.05	63.94	63.94	58.32	71.95	62.52	80.66	59.64	84.42	59.64	84.42
Unclassified	61.54	45.71	64.81	46.59	74.41	45.37	82.29	46.96	86.20	46.00	86.20	46.00

As shown in Table 4.10, overall accuracy of Expert system Classification using 6 dataset varied between 76.23 and 79.51% and Kappa coefficient varied between 63.06 and 67.69%. The multispectral and landform dataset showed the highest accuracy with overall accuracy of 79.51% and Kappa coefficient of 67.69%. These values provide accuracy higher than multispectral dataset about 3.28 and 4.63%, This result shows that the combination of additional data with multispectral dataset can increase accuracy of classification. This is the major advantage of Expert System for land cover classification by reducing commission error with additional condition. In the study, condition of landform Type 2 (undulating areas) that represents suitable sites for cassava and sugarcane can eliminate inclusive classes (commission errors) which usually not situated in this landform such as paddy field.

In addition, conditional Kappa coefficient of cassava and sugarcane classification from 6 dataset (Table 4.11) can be further described as follows:

(1) Cassava. PA of cassava classification of all datasets varied between 60.87 and 64.06 while UA was 100% for all dataset. In fact, datasets which provide the best PA for cassava classification were multispectral dataset, multispectral and NDVI dataset and multispectral and soil series dataset.

(2) Sugarcane. PA of sugarcane classification of all datasets varied between 58.32 and 63.94 while UA varied between 61.05 and 84.42%. In fact, dataset which provides the best PA for sugarcane classification was multispectral and NDVI dataset while datasets which provide the best UA for sugarcane classification were multispectral, soil series and landform dataset and multispectral, NDVI, soil series and landform dataset.

(3) Unclassified. This category was inattentive class that was comprised of bare land, forest, paddy field, urban and water classes. Therefore, this class was very heterogeneous in term of land cover within a class. PA of unclassified of all datasets varied between 61.54 and 86.20 while UA varied between 45.37 and 46.96%.

As results, it was found that the combination of additional data with multispectral dataset can increase accuracy of some categories. In the study, PA of sugarcane was increased by using multispectral and NDVI dataset and UA of sugarcane was also increased by using all additional datasets. Meanwhile, PA of cassava was decreased by using multispectral and landform dataset, multispectral, soil series and landform dataset and multispectral, NDVI, soil series and landform dataset but UA of cassava was no difference. UA of cassava for all datasets was 100 %.

4.3.3 Accuracy assessment of ANN

The accuracy assessments of land cover classification for each dataset from ANN with different learning rates and major results of accuracy assessment from each dataset were described and summarized as table. Herewith, the accuracy assessments ANN with learning rate 0.1, 02 and 0.3 in detail with error matrix were presented in Table C.7 to C.12, Table C.13 to C.18 and Table C.19 to C.24, respectively in Appendix C.

Major results of accuracy assessment from each dataset of ANN classification with learning rate 0.1, 0.2 and 0.3 can be summarized as shown in Table 4.12 and Table 4.13, Table 4.14 and Table 4.15, and Table 4.16 and Table 4.17, respectively. The detail of ANN accuracy assessment with different learning rate and dataset was separately described in following section.

4.3.3.1 ANN with learning rate 0.1

According to Table 4.12 and Table 4.13, overall accuracy of land cover classification for multispectral dataset was 70.49% and classes that provided PA and UA more than 70% were bare land and cassava. The Kappa coefficient was 62.32% which represented moderate agreement.

At the same time, overall accuracy of land cover classification for multispectral and NDVI dataset was 72.95% and classes that provided PA and UA more than 70% were bare land, cassava, sugarcane and urban. The Kappa coefficient was 65.99% which represented moderate agreement. Simultaneous, overall accuracy of land cover classification for multispectral and soil series dataset was 78.69% and classes that provided PA and UA more than 70% were bare land, water, cassava, sugarcane and urban. The Kappa coefficient was 72.79% which represented moderate agreement. Meanwhile, overall accuracy of land cover classification for multispectral and landform dataset was 72.13% and classes that provided PA and UA more than 70% were bare land, cassava and urban. The Kappa coefficient was 65.36% which represented moderate agreement.

Furthermore, overall accuracy of land cover classification for multispectral, soil series and landform dataset was 73.77% and classes that provided PA and UA more than 70% were bare land, water and cassava. The Kappa coefficient was 65.36% which represented moderate agreement. In addition, overall accuracy of land cover classification for multispectral, NDVI, soil series and landform dataset was 77.05% and classes that provided PA and UA more than 70% were bare land, cassava and sugarcane. The Kappa coefficient was 70.67% which represented moderate agreement.

As shown in Table 4.12, overall accuracy of ANN with learning rate 0.1 using 6 dataset varied between 70.49 and 78.69% and Kappa coefficient between 62.32 and 72.79%. The multispectral and soil series dataset showed the highest of both accuracies with overall accuracy of 78.69% and Kappa coefficient of 72.79%. These accuracies are better than multispectral dataset 8.20% and 10.47%, respectively. These findings imply that ANN classification with learning rate 0.1 can provide higher accuracy when it applied with multispectral and additional dataset more than only multispectral dataset. This agrees with previous research of Heinl, Walde, Tappeiner and Tappeiner (2009).

At the same time, conditional K for each land cover type of ANN with learning rate 0.1 using 6 dataset (Table 4.13) can be described as following.

(1) Cassava. PA of cassava classification of all datasets varied between 64.06 and 91.99% while UA varied between 91.50 and 100%. In fact, dataset which provides the best PA for cassava classification was multispectral and landform dataset while datasets which provide the best UA for cassava classification were multispectral dataset, multispectral, soil series and landform dataset and multispectral, NDVI, soil series and landform dataset.

(2) Sugarcane. PA of sugarcane classification of all datasets varied between 57.86 and 90.83% while UA varied between 44.44 and 68.34%. In fact, dataset which provides the best PA for sugarcane classification was multispectral, NDVI, soil series and landform dataset while dataset which provides the best UA for sugarcane classification was multispectral and landform dataset.

(3) Urban and built-up areas. PA of urban and built-up areas of all datasets varied between 57.93 and 79.32% while UA varied between 44.86 and 100%.

In fact, datasets which provide the best PA for urban and built-up areas classification were multispectral and NDVI dataset, multispectral and soil series dataset and multispectral and landform dataset while dataset which provides the best UA for urban and built-up areas classification was multispectral dataset.

(4) Paddy field. PA of paddy field of all datasets varied between 11.67 and 48.17% while UA varied between 32.22 and 75.60%. In fact, dataset which provides the best PA and UA for paddy field classification was multispectral, soil series and landform dataset.

(5) Forest land. PA of forest land of all datasets varied between 27.63 and 87.29% while UA varied between 28.78 and 72.77%. In fact, dataset which provides the best PA for forest land classification w as multispectral and landform dataset while dataset which provides the best UA for forest classification was multispectral dataset.

(6) Water bodies. PA of water bodies of all datasets varied between 57.57 and 100% while UA varied between 32.97 and 82.62%. In fact, datasets which provide the best PA for water-body classification were multispectral dataset, multispectral and soil series dataset, and multispectral, soil series and landform dataset while dataset which provides the best UA for urban and built-up areas classification was multispectral and soil series dataset.

(7) Bare land. PA of bare land of all datasets varied between 78.21 and 100% while UA varied between 78.21 and 90.10%. In fact, datasets which provide the best PA and UA for bare land classification were multispectral dataset, multispectral and NDVI dataset, multispectral and soil series dataset and multispectral and landform dataset.

As results about accuracy assessment of ANN with learning rate 0.1 it reveals that combination of additional data to multispectral data can increase accuracy of conditional Kappa coefficient of PA and UA in some categories. In the study, PA of cassava and forest land was increased when an additional data was added to multispectral data. At the same time PA of sugarcane was increased when multispectral, NDVI, soil series and landform dataset was used to classify land cover instead of only multispectral dataset. Similarity, three and four additional datasets increase PA for paddy field and urban and built-up areas classification, respectively. In contrast, two and three additional datasets decrease PA for bare land and water body classification.

Furthermore, UA of sugarcane and water bodies was increased when an additional data was added to multispectral data. Three additional datasets increase UA for cassava and paddy field classification. In contrast, two additional datasets decrease UA for urban and built-up areas and bare land classification. Furthermore combination of additional data to multispectral data decreased UA of forest land. (See detail in Table 4.13).

Table 4.12 Summary of overall Accuracy and Kappa coefficient for land cover classification by ANN with learning rate 0.1.

Dataset	Description	Overall Accuracy (%)	Kappa coefficient (%)	Rank
1	Multispectral data	70.49%	62.32%	6
2	Multispectral data and NDVI	72.95%	65.99%	4
3	Multispectral data and Soil series	78.69%	72.79%	1
4	Multispectral data and Landform	72.13%	65.36%	5
5	Multispectral data, Soil series and Landform	73.77%	67.00%	3
6	Multispectral data, NDVI, Soil series and Landform	77.05%	70.67%	2

Table 4.13 Summary of simple PA and UA and Conditional Kappa PA and UA for each Category by ANN Classification with learning rate 0.1.

Classes	Simple PA and UA for each dataset (%)											
	1		2		3		4		5		6	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Cassava	71.43	100.00	91.43	94.12	88.57	93.94	94.29	94.29	80.00	100.00	88.57	100.00
Sugarcane	91.43	60.38	80.00	75.68	85.71	71.43	68.57	77.42	80.00	68.29	94.29	71.74
Urban and built-up areas	60.00	100.00	80.00	100.00	80.00	100.00	80.00	100.00	80.00	57.14	60.00	50.00
Paddy field	36.36	66.67	18.18	44.44	50.00	73.33	18.18	50.00	54.55	80.00	45.45	76.92
Forest land	30.00	75.00	70.00	35.00	50.00	45.45	90.00	34.62	50.00	35.71	50.00	55.56
Water bodies	100.00	35.71	80.00	57.14	100.00	83.33	80.00	57.14	100.00	71.43	60.00	42.86
Bare land	100.00	90.91	100.00	90.91	100.00	90.91	100.00	90.91	80.00	80.00	90.00	90.00
Classes	**Conditional Kappa coefficient for each dataset (%)**											
	1		2		3		4		5		6	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Cassava	64.06	100.00	88.12	91.75	84.33	91.50	91.99	91.99	74.04	100.00	84.68	100.00
Sugarcane	84.84	44.44	71.29	65.89	78.21	59.93	57.86	68.34	69.88	55.54	90.83	60.37
Urban and built-up areas	58.99	100.00	79.32	100.00	79.32	100.00	79.32	100.00	78.78	55.31	57.93	47.86
Paddy field	29.42	59.33	11.67	32.22	42.99	67.47	12.44	39.00	48.17	75.60	38.95	71.85
Forest land	27.63	72.77	64.12	29.20	45.05	40.58	87.29	28.78	43.52	29.97	46.02	51.59
Water bodies	100.00	32.97	78.78	55.31	100.00	82.62	78.78	55.31	100.00	70.21	57.57	40.42
Bare land	100.00	90.10	100.00	90.10	100.00	90.10	100.00	90.10	78.21	78.21	89.11	89.11

4.3.3.2 ANN with learning rate 0.2

According to Table 4.14 and Table 4.15, for multispectral dataset overall accuracy of land cover classification was 68.03% and classes that provided PA and UA more than 70% were bare land, water and urban. The Kappa coefficient was 59.27% which represented moderate agreement.

At the same time, for multispectral and NDVI dataset, overall accuracy of land cover classification was 68.03% and classes that provided PA and UA more than 70% were cassava and urban while provided PA and UA 0% in paddy filed class. The Kappa coefficient was 60.09% which represented moderate agreement. Meanwhile, for multispectral and soil series dataset, overall accuracy of land cover classification was 69.67% and classes that provided PA and UA more than 70% were water and cassava. The Kappa coefficient was 60.61% which represented moderate agreement. Simultaneous, for multispectral and landform dataset overall accuracy of land cover classification was 76.23% and classes that provided PA and UA more than 70% were bare land, cassava, sugarcane and urban. The Kappa coefficient was 69.90% which represented moderate agreement.

At the same time, for multispectral, soil series and landform dataset overall accuracy of land cover classification was 70.49% and classes that provided PA and UA more than 70% were bare land and water bodies. The Kappa coefficient was 62.70% which represented moderate agreement. In addition, for multispectral, NDVI, soil series and landform dataset overall accuracy of land cover classification was 72.13% and classes that provided PA and UA more than 70% were forest and bare land. The Kappa coefficient was 64.44% which represented moderate agreement.

As shown in Table 4.14, overall accuracy of ANN with learning rate 0.2 using 6 dataset varied between 68.03and 76.23% and Kappa coefficient between 59.27 and 69.90%. The multispectral and landform dataset showed the highest of both accuracies with overall accuracy of 76.23% and Kappa coefficient of 69.90%. These accuracies are better than multispectral dataset 8.20% and 10.63%, respectively. These results reveal that ANN classification with learning rate 0.2 can also provide higher accuracy when it applied with multispectral and additional dataset more than only multispectral dataset.

At the same time, conditional K for each land cover type of ANN with learning rate 0.2 using 6 dataset (Table 4.15) can be explained as following.

(1) Cassava. PA of cassava classification of all datasets varied between 54.69 and 87.70% while UA varied between 81.05 and 100%. In fact, dataset which provides the best PA for cassava classification was multispectral and soil series dataset while datasets which provide the best UA for cassava classification were multispectral dataset, multispectral and NDVI dataset, multispectral, soil series and landform dataset and multispectral, NDVI, soil series and landform dataset.

(2) Sugarcane. PA of sugarcane classification of all datasets varied between 67.03 and 90.83% while UA varied between 39.15 and 60.37%. In fact, dataset which provides the best PA and UA for sugarcane classification was multispectral and landform dataset while dataset.

(3) Urban and built-up areas. PA of urban and built-up areas of all datasets varied between 19.34 and 79.15% while UA varied between 25.52 and 100%. In fact, datasets which provide the best PA for urban and built-up areas classification were multispectral dataset, multispectral and NDVI dataset and multispectral and landform dataset while dataset which provides the best UA for urban and built-up areas classification was multispectral and soil series dataset.

(4) Paddy field. PA of paddy field of all datasets varied between -0.83 and 48.17% while UA varied between -22.00 and 75.60%. In fact, dataset which provides the best PA for paddy field classification was multispectral, soil series and landform dataset while datasets which provide the best UA for paddy field classification were multispectral and soil series dataset and multispectral, soil series and landform dataset.

(5) Forest land. PA of forest land of all datasets varied between -3.39 and 88.60% while UA varied between -8.93 and 67.32%. In fact, dataset which provides the best PA for forest land classification was multispectral and landform dataset while dataset which provides the best UA for forest classification was multispectral, NDVI, soil series and landform dataset.

(6) Water bodies. PA of water bodies of all datasets varied between 78.21 and 100% while UA varied between 37.44 and 100%. In fact, datasets which provide the best PA for water-body classification were multispectral dataset, multispectral and soil series dataset, and multispectral, soil series and landform dataset while dataset which provides the best UA for water bodies classification was multispectral dataset.

(7) Bare land. PA of bare land of all datasets varied between 89.11 and 100% while UA varied between 55.15 and 90.10%. In fact, datasets which provide the best PA and UA for bare land classification were multispectral dataset, multispectral and NDVI dataset, multispectral and soil series dataset and multispectral and landform dataset while datasets which provide the best UA for bare land classification were multispectral dataset and multispectral and landform dataset.

As results about accuracy assessment of ANN with learning rate 0.2 it was found that combination of additional data to multispectral data can also increase accuracy of Conditional Kappa coefficient of PA and UA in some categories. In the study, four additional datasets increase PA for forest land while three additional datasets increase PA for cassava, paddy field and water bodies. In addition two additional datasets increase PA for sugarcane and bare land. In contrast, three additional datasets decrease PA for urban and built-up areas classification.

At the same time, UA of sugarcane was increased when an additional data was added to multispectral data. In addition four additional datasets increase UA for paddy field and forest land classification while one additional datasets increase UA for urban and built-up areas classification. In contrast, four additional datasets decrease UA for bare land classification and two additional datasets decrease UA for cassava classification. Furthermore combination of additional data to multispectral data decreased UA of water bodies. (See detail in Table 4.15).

Table 4.14 Summary of overall accuracy and Kappa coefficient for land cover classification by ANN with learning rate 0.2.

Dataset	Description	Overall Accuracy	Kappa coefficient	Rank
1	Multispectral data	68.03%	59.27%	6
2	Multispectral data and NDVI	68.03%	60.09%	5
3	Multispectral data and Soil series	69.67%	60.61%	4
4	Multispectral data and Landform	76.23%	69.90%	1
5	Multispectral data, Soil series and Landform	70.49%	62.70%	3
6	Multispectral data, NDVI, Soil series and Landform	72.13%	64.44%	2

Table 4.15 Summary of Simple PA and UA and conditional Kappa PA and UA for each Category by ANN Classification with learning rate 0.2.

Classes	Simple PA and UA for each dataset (%)											
	1		2		3		4		5		6	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Cassava	65.71	100.00	82.86	100.00	91.43	86.49	88.57	96.88	65.71	100.00	62.86	100.00
Sugarcane	85.71	56.60	80.00	65.12	82.86	60.42	94.29	71.74	80.00	58.33	88.57	59.62
Urban and built-up areas	80.00	80.00	80.00	80.00	20.00	100.00	80.00	80.00	40.00	28.57	60.00	60.00
Paddy field	22.73	45.45	0.00	0.00	36.36	80.00	9.09	66.67	54.55	80.00	54.55	70.59
Forest land	60.00	42.86	80.00	44.44	0.00	0.00	90.00	60.00	70.00	53.85	70.00	70.00
Water bodies	100.00	100.00	80.00	44.44	100.00	71.43	80.00	40.00	100.00	83.33	80.00	66.67
Bare land	100.00	90.91	100.00	58.82	100.00	66.67	100.00	90.91	90.00	90.00	90.00	90.00

Classes	Conditional Kappa coefficient for each dataset (%)											
	1		2		3		4		5		6	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Cassava	57.75	100.00	65.71	100.00	87.70	81.05	91.43	86.49	57.75	100.00	54.69	100.00
Sugarcane	74.74	39.15	85.71	56.60	71.74	44.49	82.86	60.42	67.03	41.57	80.08	43.37
Urban and built-up areas	79.15	79.15	80.00	80.00	19.34	100.00	20.00	100.00	36.35	25.52	58.29	58.29
Paddy field	15.07	33.45	22.73	45.45	30.68	75.60	36.36	80.00	48.17	75.60	47.19	64.12
Forest land	54.81	37.76	60.00	42.86	-3.39	-8.93	0.00	0.00	66.42	49.73	67.32	67.32
Water bodies	100.00	100.00	100.00	100.00	100.00	70.21	100.00	71.43	100.00	82.62	78.97	65.24
Bare land	100.00	90.10	100.00	90.91	100.00	63.69	100.00	66.67	89.11	89.11	89.11	89.11

4.3.3.3 ANN with learning rate 0.3

According to Table 4.16 and Table 4.17, overall accuracy of LULC classification for multispectral dataset was 58.20% and only water class that provided PA and UA more than 70%. The Kappa coefficient was 48.70% which represented moderate agreement.

At the same time, overall accuracy of land cover classification for multispectral and NDVI dataset was 55.74% and classes that provided PA and UA more than 70% were water bodies and bare land while provided PA and UA 0% in forest land. The Kappa coefficient was 41.36% which represented moderate agreement. Meanwhile, overall accuracy of land cover classification for multispectral and soil series dataset was 77.87% and classes that provided PA and UA more than 70% were cassava, sugarcane and bare land. The Kappa coefficient was 71.77% which represented moderate agreement. Simultaneous, overall accuracy of land cover classification for multispectral and landform dataset was 67.21% and only cassava provided PA and UA more than 70%. The Kappa coefficient was 60.56% which represented moderate agreement.

Furthermore, overall accuracy of land cover classification for multispectral, soil series and landform dataset was 71.31% and classes that provided PA and UA more than 70% were cassava, urban and built-up areas and bare land. The Kappa coefficient was 64.48% which represented moderate agreement. In addition, overall accuracy of land cover classification for multispectral, NDVI, soil series and landform dataset was 59.02% and only bare land provided PA and UA more than 70%. The Kappa coefficient was 49.75% this represented moderate agreement.

As shown in Table 4.16, overall accuracy of ANN with learning rate 0.3 using 6 dataset varied between 55.74 and 77.87% and Kappa coefficient between 41.36 and 71.77%. The multispectral and soil series dataset showed the highest of both accuracies with overall accuracy of 77.87% and Kappa coefficient of 71.77%. These accuracies are better than multispectral dataset 22.13% and 30.41%, respectively. These findings reveal that ANN classification with learning rate 0.3 can provide higher accuracy when it applied with almost combination of multispectral and additional dataset except multispectral and NDVI dataset which provide lower accuracies than multispectral dataset.

At the same time, conditional K for each land cover type of ANN with learning rate 0.3 using 6 dataset (Table 4.17) can be described as following.

(1) Cassava. PA of cassava classification of all datasets varied between 19.80 and 87.98% while UA varied between 84.42 and 100%. In fact, dataset which provides the best PA for cassava classification was multispectral and landform dataset while datasets which provide the best UA for cassava classification were multispectral dataset, multispectral and NDVI dataset, multispectral and soil series dataset and multispectral, soil series and landform dataset.

(2) Sugarcane. PA of sugarcane classification of all datasets varied between 23.95 and 82.35% while UA varied between 16.57 and 88.31%. In fact, dataset which provides the best PA for sugarcane classification was multispectral and soil series dataset while dataset which provides the best UA for sugarcane classification was multispectral and landform dataset.

(3) Urban and built-up areas. PA of urban and built-up areas of all datasets varied between 39.00 and 79.32% while UA varied between 42.07 and 100%.

In fact, dataset which provides the best PA for urban and built-up areas classification was multispectral, soil series and landform dataset while datasets which provide the best UA for urban and built-up areas classification were multispectral dataset, multispectral and NDVI dataset and multispectral, soil series and landform dataset.

(4) Paddy field. PA of paddy field of all datasets varied between 13.20 and 52.47% while UA varied between 47.71 and 75.60%. In fact, dataset which provides the best PA for paddy field classification was multispectral, soil series and landform dataset while datasets which provide the best UA for paddy field classification were multispectral dataset and multispectral and landform dataset.

(5) Forest land. PA of forest land of all datasets varied between -0.83 and 87.29% while UA varied between -8.93 and 68.88%. In fact, datasets which provide the best PA for forest land classification were multispectral and landform dataset and multispectral, NDVI, soil series and landform dataset while dataset which provides the best UA for forest classification was multispectral and soil series dataset.

(6) Water bodies. PA of water bodies of all datasets varied between 58.29 and 100% while UA varied between 53.66 and 100%. In fact, datasets which provide the best PA for water-body classification were multispectral dataset, multispectral and landform dataset, multispectral, soil series and landform dataset and multispectral, NDVI, soil series and landform dataset while datasets which provide the best UA for urban and built-up areas classification were multispectral dataset and multispectral and NDVI dataset.

(7) Bare land. PA of bare land of all datasets varied between 78.02 and 100% while UA varied between 42.94 and 100%. In fact, datasets which provide the best PA for bare land classification were multispectral dataset, multispectral and

NDVI dataset, multispectral and soil series dataset, multispectral and landform dataset and multispectral, soil series and landform dataset while dataset which provides the best UA for bare land classification was multispectral, soil series and landform dataset.

As results about accuracy assessment of ANN with learning rate 0.3 it was found that combination of additional data to multispectral data can also increase accuracy of conditional Kappa coefficient of PA and UA in some categories. In the study, PA of cassava was increased when an additional data was added to multispectral data. At the same time four additional datasets increase PA for urban and built-up areas classification while three additional datasets increase PA for paddy field and forest land classification. In addition one additional dataset increase PA for sugarcane classification. In contrary, two additional datasets decrease PA for water body classification and one additional datasets decreases PA for bare land classification.

Meanwhile, four additional datasets increase UA for forest land and bare land classification and two additional datasets increase UA for sugar cane classification. In contrast, four additional datasets decrease UA for paddy field and water body classification and three additional datasets decrease UA for urban and built-up area classification. Furthermore two additional datasets decrease UA for cassava classification. (See detail in Table 4.17).

In summary, when accuracy assessment of ANN classification by different learning rate are compared as shown in Table 4.18, ANN with learning rate of 0.1 provides the best results for land cover classification in almost dataset except multispectral and landform dataset. In addition, ANN with learning rate of 0.2

provides overall accuracy and Kappa coefficient higher than ANN with learning rate of 0.1. In contrast, ANN with learning rate of 0.3 provides the worst results for land cover classification in almost dataset. However, ANN with learning rate of 0.3 provides overall accuracy and Kappa coefficient higher than ANN with learning rate of 0.2 for multispectral and soil series dataset and multispectral, soil series and landform dataset. These finding suggest that ANN classification with minimal leaning rate value provides the best result but it consumes more time for training.

Table 4.16 Summary of overall accuracy and Kappa coefficient for land cover classification by ANN with learning rate 0.3.

Dataset	Description	Overall Accuracy	Kappa coefficient	Rank
1	Multispectral data	58.20%	48.70%	5
2	Multispectral data and NDVI	55.74%	41.36%	6
3	Multispectral data and Soil series	77.87%	71.77%	1
4	Multispectral data and Landform	67.21%	60.56%	3
5	Multispectral data, Soil series and Landform	71.31%	64.48%	2
6	Multispectral data, NDVI, Soil series and Landform	59.02%	49.75%	4

Table 4.17 Summary of simple PA and UA and conditional Kappa PA and UA for each Category by ANN Classification with learning rate 0.3.

Classes	Conditional Kappa coefficient for each dataset (%)											
	1		2		3		4		5		6	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Cassava	19.80	100.00	37.52	100.00	84.68	100.00	87.98	87.98	67.32	100.00	36.32	84.42
Sugarcane	80.90	48.49	75.68	16.57	82.35	60.87	23.95	88.31	44.55	46.38	60.29	44.56
Urban and built-up areas	39.00	100.00	39.00	100.00	78.78	55.31	58.64	73.93	79.32	100.00	78.41	42.07
Paddy field	30.68	75.60	13.20	47.71	41.90	56.94	48.17	75.60	52.47	71.29	13.20	47.71
Forest land	49.17	16.21	-0.83	-8.93	46.96	68.88	87.29	28.78	75.35	28.96	87.29	28.78
Water bodies	100.00	100.00	79.32	100.00	58.29	58.29	100.00	53.66	100.00	60.90	100.00	60.90
Bare land	100.00	42.94	100.00	74.86	100.00	81.85	100.00	42.94	100.00	100.00	78.02	70.29

Classes	Summary of simple PA and UA for each dataset (%)											
	1		2		3		4		5		6	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Cassava	25.71	100.00	45.71	100.00	88.57	100.00	91.43	91.43	74.29	100.00	45.71	88.89
Sugarcane	88.57	63.27	91.43	40.51	88.57	72.09	31.43	91.67	60.00	61.76	74.29	60.47
Urban and built-up areas	40.00	100.00	40.00	100.00	80.00	57.14	60.00	75.00	80.00	100.00	80.00	44.44
Paddy field	36.36	80.00	18.18	57.14	50.00	64.71	54.55	80.00	59.09	76.47	18.18	57.14
Forest land	60.00	23.08	0.00	0.00	50.00	71.43	90.00	34.62	80.00	34.78	90.00	34.62
Water bodies	100.00	100.00	80.00	100.00	60.00	60.00	100.00	55.56	100.00	62.50	100.00	62.50
Bare land	100.00	47.62	100.00	76.92	100.00	83.33	100.00	47.62	100.00	100.00	80.00	72.73

Table 4.18 Accuracy assessment comparison of ANN classification by different learning rate.

Dataset	Description	ANN Learning Rate 0.1			ANN Learning Rate 0.2			ANN Learning Rate 0.3		
		Overall Accuracy	Kappa coefficient	Rank	Overall Accuracy	Kappa coefficient	Rank	Overall Accuracy	Kappa coefficient	Rank
1	Multispectral data	70.49%	62.32%	6	68.03%	59.27%	6	58.20%	48.70%	5
2	Multispectral data and NDVI	72.95%	65.99%	4	68.03%	60.09%	5	55.74%	41.36%	6
3	Multispectral data and Soil series	78.69%	72.79%	1	69.67%	60.61%	4	77.87%	71.77%	1
4	Multispectral data and Landform	72.13%	65.36%	5	76.23%	69.90%	1	67.21%	60.56%	3
5	Multispectral data, Soil series and Landform	73.77%	67.00%	3	70.49%	62.70%	3	71.31%	64.48%	2
6	Multispectral data, NDVI, Soil series and Landform	77.05%	70.67%	2	72.13%	64.44%	2	59.02%	49.75%	4

4.4 Evaluation of an optimum method for sugarcane and cassava classification

In practice, an optimum method for sugarcane and cassava classification using Expert System and ANN method was considered from Kappa coefficient of three common categories: cassava, sugarcane and others. In theory, Kappa coefficient is discrete multivariate techniques in accuracy assessment that can be normalized or standardized error matrices by include information about the off-diagonal cell values of matrices. Hence, it is possible to directly compare between 2 methods which was the better (Jensen, 2005).

Similar to accuracy assessment results in Section 4.3, ANN accuracy assessment of three common categories (cassava, sugarcane and others) with learning rate 0.1 was the best configuration of ANN based on average value of overall accuracy and Kappa coefficient (Table 4.19). Details of three common categories accuracy assessment with error matrix for ANN classification in each learning rate were presented in Table C.25 to Table C.42 in Appendix C. So, Kappa coefficient of ANN with 0.1 learning rate and Expert Systems are here compare to identify an optimum method for cassava and sugarcane classification as shown in Table 4.20.

As results, Kappa coefficients of ANN classification with learning rate 0.1 were better than Expert System in all datasets. This result is similar to research work of Liu, Skidmore and Van Oosten (2002). Therefore, ANN classification with learning rate 01 was selected as an optimum method for sugarcane and cassava classification.

Table 4.19 Accuracy assessment of ANN classification for three common categories (cassava, sugarcane and others).

		Learning rate 0.1		Learning rate 0.2		Learning rate 0.3	
Dataset	**Description**	**Overall Accuracy**	**Kappa Coefficient**	**Overall Accuracy**	**Kappa Coefficient**	**Overall Accuracy**	**Kappa Coefficient**
1	Multispectral data	81.15%	71.51%	78.69%	67.57%	68.85%	51.25%
2	Multispectral data and NDVI	86.07%	78.72%	82.79%	73.67%	59.02%	39.93%
3	Multispectral data and Soil series	85.25%	77.59%	78.69%	68.18%	85.25%	77.59%
4	Multispectral data and Landform	84.43%	75.92%	86.89%	80.15%	77.05%	63.36%
5	Multispectral data, Soil series and Landform	81.97%	72.37%	76.23%	63.64%	76.23%	62.93%
6	Multispectral data, NDVI, Soil series and Landform	87.70%	81.39%	77.05%	65.01%	75.41%	61.78%
Average		84.43%	76.25%	80.06%	69.70%	73.64%	59.47%

Table 4.20 Comparison of Kappa coefficient between Expert System and ANN classification by each dataset.

Dataset	Description	Expert System	ANN with 0.1 learning rate
		Kappa coefficient	Kappa coefficient
1	Multispectral data	63.06%	71.51%
2	Multispectral data and NDVI	64.27%	78.72%
3	Multispectral data and Soil series	65.23%	77.59%
4	Multispectral data and Landform	67.69%	75.92%
5	Multispectral data, Soil series and Landform	67.58%	72.37%
6	Multispectral data, NDVI, Soil series and Landform	67.58%	81.39%
	Average	65.90%	76.25%

4.5 Evaluation of an optimum dataset for sugarcane and cassava classification

4.5.1 Evaluation of an optimum dataset for sugarcane and cassava classification using Expert system

In practice, an optimum dataset evaluation for sugarcane and cassava classification using Expert system is considered from Kappa coefficient. The dataset which provides the highest of Kappa coefficient is an optimum dataset for sugarcane and cassava classification.

Based on accuracy assessment in Table 4.10, multispectral and landform dataset, which was provides the highest Kappa coefficient at 67.69%, was selected as an optimum dataset for sugarcane and cassava classification using Expert system. In addition, this dataset also provided conditional Kappa coefficient of PA and UA for cassava and sugarcane more than 60%. (See detail in Table 4.11).

At the same time, pairwise testing for significance difference at 80% confidence level between two error matrices of multispectral dataset and additional dataset revealed that 5 pairwise were not significantly different as shown in Table 4.21. Detail for calculation of Z statistics was shown Table D.1 in Appendix D. This result infers that combination of multispectral data and an additional data has not significantly increased the accuracy for cassava and sugarcane classification using Expert System.

Table 4.21 Significant difference test by Z statistic for Expert System.

Pairwise	**Basic dataset**	**Additional dataset**	**Z statistics**	**Critical Value at 80% confidence level**
1	Multispectral	Multispectral and NDVI	0.1489	1.28
2	Multispectral	Multispectral and soil series	0.3013	1.28
3	Multispectral	Multispectral and Landform	0.6129	1.28
4	Multispectral	Multispectral, soil series and landform	0.6141	1.28
5	Multispectral	Multispectral, NDVI, soil series and landform	0.6141	1.28

4.5.2 Evaluation of an optimum dataset for sugarcane and cassava classification using ANN

The optimum dataset for cassava and sugarcane classification using ANN with learning rate 0.1 was multispectral and soil series dataset which provides the highest Kappa coefficient of 72.79% (see Table 4.12). In addition, this dataset provided Conditional Kappa coefficients of PA for cassava and sugarcane were 84.33 and 78.21%, respectively while conditional Kappa coefficients of UA for cassava and sugarcane were 91.50 and 59.93%, respectively. Also, it should be noted that multispectral, NDVI, soil series and landform dataset provided the best result for cassava and sugarcane classification based on conditional Kappa coefficient. The conditional Kappa coefficient of PA and UA for cassava and sugarcane were 84.68 and 90.83% and 100.00 and 60.37%, respectively. (See detail in Table 4.13).

Furthermore, pairwise testing for significance difference at 80% confidence level between two error matrices of multispectral dataset and additional dataset revealed that pairwise between multispectral dataset and multispectral and soil series dataset was significantly different as displayed in Table 4.22. (See detail in Table D.2 in, Appendix D).

Table 4.22 Significant difference test by Z statistic for ANN with learning rate 0.1.

Pairwise	**Basic dataset**	**Additional dataset**	**Z statistics**	**Critical Value at 80% confidence level**
1	Multispectral	Multispectral and NDVI	0.5157	1.28
2	Multispectral	Multispectral and soil series	**1.5002**	1.28
3	Multispectral	Multispectral and Landform	0.4279	1.28
4	Multispectral	Multispectral, soil series and landform	0.1414	1.28
5	Multispectral	Multispectral, NDVI, soil series and landform	1.1967	1.28

While, the optimum dataset for sugarcane and cassava classification using ANN with learning rate 0.2 was multispectral data and landform dataset which presented the highest Kappa coefficient of 69.90% (see Table 4.14). This dataset also provided conditional Kappa coefficient of PA and UA for cassava and sugarcane more than 60%. (See detail in Table 4.15).

Meanwhile, pairwise testing for significance difference at 80% confidence level between two error matrices of multispectral dataset and additional dataset shown that pairwise between multispectral dataset and multispectral and landform dataset was significantly different as displayed in Table 4.23. (See detail in Table D.2 in, Appendix D).

Table 4.23 Significant difference test by Z statistic for ANN with learning rate 0.2.

Pairwise	**Basic dataset**	**Additional dataset**	**Z statistics**	**Critical Value at 80% confidence level**
1	Multispectral	Multispectral and NDVI	0.1113	1.28
2	Multispectral	Multispectral and soil series	0.1785	1.28
3	Multispectral	Multispectral and Landform	**1.4891**	1.28
4	Multispectral	Multispectral, soil series and landform	0.3462	1.28
5	Multispectral	Multispectral, NDVI, soil series and landform	0.6891	1.28

At the same time, the optimum dataset for sugarcane and cassava classification using ANN with learning rate 0.3 was multispectral data and soil series dataset which provided the highest Kappa hat coefficient of 71.77% (see Table 4.16).

This dataset also provided conditional Kappa coefficient of PA and UA for cassava and sugarcane more than 60% (see Table 4.17).

Furthermore, pairwise testing for significance difference at 80% confidence level between two error matrices of multispectral dataset and additional dataset revealed that 3 pairwise were significantly different include (1) multispectral and soil series dataset, (2) multispectral and landform dataset and (3) multispectral, soil series and landform dataset as shown in Table 4.24 (See detail in Table D.2 in, Appendix D).

In summary, specific combination of multispectral data and an additional data, especially soil series and landform can increase accuracy for cassava and sugarcane classification using ANN with learning rate of 0.1 and 0.2.

Table 4.24 Significant difference test by Z statistic for ANN with learning rate 0.3.

Pairwise	**Basic dataset**	**Additional dataset**	**Z statistics**	**Critical Value at 80% confidence level**
1	Multispectral	Multispectral and NDVI	0.8790	1.28
2	Multispectral	Multispectral and soil series	**3.2495**	1.28
3	Multispectral	Multispectral and Landform	**1.6409**	1.28
4	Multispectral	Multispectral, soil series and landform	**3.1301**	1.28
5	Multispectral	Multispectral, NDVI, soil series and landform	0.1386	1.28

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

According to three specified objectives including (1) to classify cassava and sugarcane cultivation area, (2) to assess the accuracy of cassava and sugarcane classification and (3) to evaluate the optimum method and dataset for cassava and sugarcane classification from THEOS data using Expert System and ANN algorithm, major results and findings can be here summarize as following.

5.1.1 Cassava and sugarcane classification using Expert System and ANN

Expert system and ANN algorithm with learning rate of 0.1, 0.2 and 0.3 can be used to extract cassava and sugarcane areas from all specified datasets include (1) multispectral dataset, (2) multispectral and NDVI dataset, (3) multispectral and soil series dataset, (4) multispectral and landform dataset, (5) multispectral, soil series and landform dataset and (6) multispectral, NDVI, soil series and landform dataset.

In addition, classified area of cassava and sugarcane using Expert System, which was defined as a hierarchical decision-tree classifier, had continued to decrease when more additional data were added to multispectral data with specific conditions for classification. In opposite, classified area of cassava and sugarcane using ANN with different learning rate, which was freedom from normal distribution requirements with its ability to adaptively simulate complex and nonlinear pattern given proper topological structures, had no specific relationship with defined dataset for land cover classification.

5.1.2 Accuracy assessment of cassava and sugarcane classification using Expert System and ANN

Based on error matrix between classified land cover and ground information, overall accuracy and Kappa analysis were here applied to assess accuracy of classification with Expert System and ANN.

For Expert System classification, overall accuracy and Kappa coefficient varied between 76.23 and 79.51% and 63.06 and 67.69%, respectively. The multispectral and landform dataset showed the highest accuracy value with overall accuracy of 79.51% and Kappa coefficient of 67.69%. While, multispectral dataset presented the lowest accuracy value with overall accuracy of 76.23% and Kappa coefficient of 63.06%. The overall accuracy and Kappa coefficient of multispectral and landform dataset was higher than multispectral dataset about 3.28 and 4.63%, respectively. In fact, accuracy levels of all dataset were not significant different based on Z statistic at 80% of confidence level. In addition, the combination of additional data with multispectral dataset can increase accuracy of PA and UA for sugarcane classification.

Meanwhile, overall accuracy and Kappa coefficient of ANN with learning rate of 0.1, which provided the best result among three different learning rates, varied between 70.49 and 78.69% and 62.32 and 72.79%, respectively. The multispectral and soil series dataset showed the highest accuracy value with overall accuracy of 78.69% and Kappa coefficient of 72.79%. At the same time multispectral dataset presented the lowest accuracy value with overall accuracy of 70.49% and

Kappa coefficient of 62.32%. The overall accuracy and Kappa coefficient of multispectral and soil series dataset was higher than multispectral dataset about 8.20 and 10.47%, respectively. The difference of accuracy levels between multispectral dataset and multispectral and soil series dataset was significant different based on Z statistic at 80% of confidence level. In addition, the combination of additional data with multispectral dataset can increase accuracy of PA and UA for sugarcane classification. However, in the case of cassava classification only accuracy of PA was increase and accuracy of UA was constant.

5.1.3 Evaluation of an optimum methods for sugarcane and cassava classification between Expert System and ANN

ANN with learning rate of 0.1, which provided Kappa coefficient of three common categories (cassava, sugarcane and others) higher than Expert System in all dataset, was here selected as an optimum method for sugarcane and cassava classification. In addition, this method also provide simple PA and conditional kappa coefficient PA of cassava and sugarcane better than Expert system classification except in multispectral and landform dataset, which is an optimum dataset for Expert system classification.

5.1.4 Evaluation of an optimum dataset for sugarcane and cassava classification using Expert System and ANN

An optimum dataset for sugarcane and cassava classification using Expert system classification was multispectral and landform dataset. This dataset presented the highest Kappa coefficient at 67.69%. Meanwhile, an optimum dataset for sugarcane and cassava classification using ANN with learning rate of 0.1 was multispectral and soil series dataset. This dataset presented the highest Kappa coefficient at 72.79%.

Finally, it may be concluded that Expert System and ANN classification method can be applied for cassava and sugarcane areas extraction from THEOS data with certain level of accuracy in each dataset. In addition, combination between multispectral data and an additional data can improve accuracy of cassava and sugarcane classification using Expert System and ANN.

5.2 Recommendations

The possibly recommendations could be made for cassava and sugarcane classification in the future as following.

5.2.1 Based on the findings in this study, an additional data are useful and increase the accuracy of classification with Expert System and ANN, thus more relevant factors might be tested and added to multispectral dataset, such as climate and soil properties factors.

5.2.2 Methodology from the study should be tested and verified in others areas which have different topography or cultivation practice.

5.2.3 Permanent plots of cassava and sugarcane, which are the significant energy crops of the country, should be set up for phenological cycle study in field of remote sensing. These results will be applied to identify an optimum data as multi-date dataset for cassava and sugarcane classification in different growing stages.

REFERENCES

REFERENCES

กรมพัฒนาที่ดิน. (2550). ฐานข้อมูลการใช้ประโยชน์ที่ดินจังหวัดนครราชสีมา พ.ศ. 2550 [แฟ้มข้อมูล]. กรุงเทพ: กรมพัฒนาที่ดิน.

กรมวิชาการเกษตร. (2545). **เกษตรดีที่เหมาะสมสำหรับอ้อย.** กรุงเทพ: โรงพิมพ์ชุมนุมสหกรณ์การเกษตรแห่งประเทศไทย จำกัด.

สำนักงานเกษตรจังหวัดนครราชสีมา. (2550). **การใช้เทคโนโลยีการผลิตมันสำปะหลังภายใต้โครงการจัดระบบพิเศษเฉพาะพื้นที่มันสำปะหลัง ปี 2550/2551.** นครราชสีมา: โรงพิมพ์โคราชมาเก็ตติ้ง.

สำนักงานคณะกรรมการอ้อยและน้ำตาลทราย. (2553). **รายงานการผลิตน้ำตาลทรายของโรงงานน้ำตาลทั่วประเทศ ประจำปีการผลิต 2551/2552 (ฉบับปิดหีบ)** [ออนไลน์]. ได้จาก: http://www.ocsb.go.th/upload/production/ fileupload/142-8543.pdf

สำนักงานเศรษฐกิจการเกษตร. (2551). **รายชื่อผู้ประกอบการที่ได้รับการอนุญาตก่อสร้างโรงงานผลิตเอทานอล** [ออนไลน์]. ได้จาก: http://www.dede.go.th/dede /fileadmin/usr/bers/ gasohol_2008/510125_ Ethanol_licesees_45_mills.doc

สำนักงานเศรษฐกิจการเกษตร. (2553). **ข้อมูลพื้นฐานเศรษฐกิจการเกษตร ปี 2553**. กรุงเทพ: สำนักงานเศรษฐกิจการเกษตร.

สำนักสำรวจดินและวางแผนการใช้ที่ดิน. (2551). **เขตการใช้ที่ดิน ตำบลจักราช อำเภอจักราช. จังหวัดนครราชสีมา** [ออนไลน์]. ได้จาก: http://www. khorat.doae.go.th.

สถาบันวิจัยพืชไร่ กรมวิชาการเกษตร. (2551). **เอกสารแนะนำการปลูกมันสำปะหลัง.** สถาบันวิจัยพืชไร่ กรมวิชาการเกษตร. (จุลสาร).

Ashish, D., Mcclendon , R. W. and Hoogenboom, G. (2009). Land-use classification of multispectral aerial images using artificial neural networks. **International Journal of Remote Sensing**. 30(8): 1989–2004.

Bishop, Y., Fienberg, S. and Holland, P. (1975). **Discrete Multivariate Analysis: Theory and Practice.** Cambridge: MIT Press. Qouted in Congalton, R. G. and Green, K. (2008). **Assessing the accuracy of remotely sensed data: principles and practices** (2nd ed.). Florida: CRC Press.

Bossler, J. D., Jensen, J. R., McMaster, R. B. and Rizos, C. (2002). **Manual of Geospatial science and Technology**. London: Taylor & Francis. Qouted in Jensen, J. R. (2005). **Introductory digital image processing: A remote sensing perspective** (3rd ed). New Jersey: Pearson prentice hall.

Chen, L. and Lee, L. (1992) Progressive generation of control frameworks for image registration. **Photogrammetic engineering & remote sensing**. 58(9): 1303-1309. Qouted in Jensen, J. R. (2005). **Introductory digital image processing: A remote sensing perspective** (3rd ed). New Jersey: Pearson prentice hall.

Cohen, W. B.(1991). Response of Vegetation indices to changes in tree measures of leaf water stress. **Photogrammetic engineering & remote sensing**. 57(2): 195-202. Qouted in Jensen, J. R. (2005). **Introductory digital image processing: A remote sensing perspective** (3rd ed). New Jersey: Pearson prentice hall.

Congalton, R. G. and Green, K. (2008). **Assessing the accuracy of remotely sensed data : principles and practices** (2nd ed.). Florida: CRC Press.

Chuvieco, E. and Huete, A. (2010). **Fundamentals of satellite remote sensing**. Florida: CRC Press.

Deeudomchan, K. (2003). **Land use classification with expert systems method in Phu Khieo wildlife sanctuary**. M.S. thesis, Kasetsart University, Bangkok.

Dixon, B. and Candade, N. (2008). Multispectral landuse classification using neural networks and support vector machines: one or the other, or both?. **International Journal of Remote Sensing.** 29(4): 1185–1206.

Du, K.-L., and Swamy, M. N. S. (2006). **Neural Networks in a Softcomputing Framework**. Germany: Springer.

Duda, R. O., Hart, P. E. and Stork, D. G. (2001). **Pattern classification**. New York: John Wiley & Son. Qouted in Jensen, J. R. (2005). **Introductory digital image processing: A remote sensing perspective** (3rd ed). New Jersey: Pearson prentice hall.

ERDAS. (2010). **ERDAS Field Guide™**. United States of America: ERDAS, Inc.

Gao, J. (2009). **Digital analysis of remotely sensed imagery**. New York: McGraw Hill.

GISTDA. (2010). **THEOS Pilot Project under the THEOS Program**. Bangkok: Tana Press co., Ltd.

Haykin, S. (1999) **Neural Networks: A Comprehensive Foundation**. NJ: Prentice Hall. Qouted in Mas, J. F. and Flores, J. J. (2008). The application of artificial neural networks to the analysis of remotely sensed data. **International Journal of Remote Sensing**. 29(3): 617–663.

Heinl, M., Walde, J., Tappeiner, G. and Tappeiner, U. (2009). Classifiers vs. input variables - The drivers in image classification for land cover mapping. **International Journal of Applied Earth Observation and Geoinformation.** 11(6): 423–430.

Hepner, G. F. (1990). Artificial neural network classification using a minimal training set: comparison to conventional supervised classification. **Photogrammetric engineering and remote sensing**. 56(40): 469-473.

ITT Visual Information Solutions. (2007). ENVI 4.4 [Computer software]. New York.

Japan Association on Remote Sensing.(1993). **Remote sensing note**. Tokyo, Japan: Nihon Printing Co. Ltd.

Jensen, J. R. (2005). **Introductory digital image processing: A remote sensing perspective** (3rd ed). New Jersey: Pearson prentice hall.

Kitiphaisannon, J. (2005). **Expert Classification for age class identification of oil palm plantation in Krabi province**. M.S. thesis, Mahidol University, Bangkok.

Landis, J. R. and Koch, G. G. (1977). The measurement of observer agreement for categorical data. **Biometrics**. 33(1). 159–174. Qouted in Congalton, R. G. and Green, K. (2008). **Assessing the accuracy of remotely sensed data: principles and practices** (2nd ed.). Florida: CRC Press.

Lawawirojwong, S. (2002) **Expert classification for land cover mapping of Bangpakong watershed**. M.S. thesis, Mahidol University, Bangkok.

Liu, X-H., Skidmore, A.K. and Van Oosten, H. (2002). Integration of classification methods for improvement of land-cover map accuracy. **ISPRS Journal of Photogrammetry and Remote Sensing**. 56(4): 257-268.

Mas, J. F. and Flores, J. J. (2008). The application of artificial neural networks to the analysis of remotely sensed data. **International Journal of Remote Sensing.** 29(3): 617–663.

Mustapha, M. R., Lim, H. S. and Mat Jafri, M. Z. (2010). Comparison of neural network and maximum likelihood approaches in image classification. **Journal of Applied Sciences.** 10(22): 2847-2854.

PC AI. (2002). **Expert Systems** [Online]. availiable: www.PCAI.com Qouted in Jensen, J. R. (2005). **Introductory digital image processing: A remote sensing perspective** (3rd ed). New Jersey: Pearson prentice hall.

Renz, A. N. (ed.). (1999). **Remote sensing for the Earth sciences**. MD: American society for photogrammetry & remote sensing. Qouted in Jensen, J. R. (2005). **Introductory digital image processing: A remote sensing perspective** (3rd ed). New Jersey: Pearson prentice hall.

Samarasinghe, S. (2006). **Neural Network for Applied Sciences and Engineering: from fundamentals to complex pattern recognition**. Boca Raton: Auerbach Publication.

Suwanwerakamtorn, R. and Supunee, P. (2006). **A Comparison of Neural Network and Maximum Likelihood Method for Landuse/Landcover Mapping in Khon Kaen Province, Thailand** [On-line]. Available: www.a-a-r-s.org/acrs /proceeding/ ACRS2006/Papers/H-1_H3.pdf

Wentz, E. A., Nelson, D., Stefanov, W. L.and Roy, S. S. (2007). Expert system classification of urban land use/cover for Delhi, India. **International Journal of Remote Sensing.** 29(15): 4405–4427.

APPENDICES

APPENDIX A

Rectified THEOS data

Table A.1 The RMS error of image rectification.

PointID	X Input	Y Input	X Reference	Y Reference	Type	X Residual	Y Residual	RMS Error	Contrib.
GCP #1	223205.879	1670962.318	223431.786	1670915.464	Control	-0.241	0.307	0.391	0.863
GCP #2	229300.749	1668895.344	229526.275	1668846.722	Control	0.286	-0.169	0.332	0.734
GCP #3	225627.788	1674051.392	225855.420	1674003.492	Control	0.297	-0.091	0.311	0.687
GCP #4	237259.502	1664052.473	237481.076	1664005.822	Control	-0.239	-0.098	0.258	0.570
GCP #5	218549.492	1653594.465	218774.591	1653547.458	Control	-0.276	-0.006	0.276	0.609
GCP #6	219479.654	1644450.854	219709.192	1644405.214	Control	0.230	-0.381	0.444	0.982
GCP #7	225938.594	1647011.974	226166.571	1646969.157	Control	0.221	-0.110	0.247	0.544
GCP #8	236930.666	1647965.565	237154.707	1647929.103	Control	0.257	-0.233	0.347	0.766
GCP #9	236372.402	1646186.620	236597.072	1646151.535	Control	-0.115	0.164	0.200	0.441
GCP #10	217739.500	1651756.917	217965.603	1651709.946	Control	0.287	0.078	0.297	0.656
GCP #11	222149.377	1660996.853	222374.558	1660949.170	Control	0.548	-0.570	0.791	1.746
GCP #12	230329.135	1660244.514	230553.419	1660197.989	Control	0.387	-0.311	0.497	1.098

Table A.1 (Continue).

PointID	X Input	Y Input	X Reference	Y Reference	Type	X Residual	Y Residual	RMS Error	Contrib.
GCP #13	233041.090	1660186.505	233264.327	1660140.566	Control	0.092	-0.326	0.339	0.748
GCP #15	238011.723	1660956.816	238232.868	1660912.571	Control	-0.005	0.391	0.391	0.864
GCP #14	233596.269	1652300.708	233819.938	1652259.118	Control	-0.244	-0.276	0.368	0.814
GCP #16	237615.268	1655768.560	237837.032	1655726.976	Control	0.245	-0.086	0.260	0.574
GCP #17	223209.437	1658854.728	223434.617	1658808.040	Control	0.412	0.181	0.449	0.993
GCP #18	228166.190	1654925.471	228391.090	1654881.197	Control	0.043	0.594	0.595	1.314
GCP #19	221506.888	1664440.565	221731.263	1664393.757	Control	-0.343	0.312	0.464	1.024
GCP #20	233745.503	1671922.612	233970.699	1671872.280	Control	-0.033	-0.622	0.623	1.376
GCP #21	230334.236	1653620.741	230558.411	1653577.243	Control	-0.469	0.033	0.470	1.039
GCP #22	231342.574	1647977.903	231568.676	1647937.966	Control	0.005	0.095	0.095	0.209
GCP #23	224437.178	1665808.821	224661.753	1665761.013	Control	-0.421	-0.298	0.516	1.139
GCP #24	224397.022	1652384.518	224622.747	1652339.774	Control	-0.201	0.546	0.582	1.284
GCP #25	232721.981	1673022.353	232947.663	1672973.058	Control	-0.274	0.496	0.567	1.251
GCP #26	220333.052	1649969.923	220559.328	1649924.102	Control	-0.466	0.276	0.542	1.197
GCP #28	232830.229	1668761.768	233054.641	1668713.507	Control	-0.022	0.489	0.489	1.081
GCP #30	230530.981	1655752.842	230754.657	1655708.600	Control	-0.521	0.168	0.548	1.209
GCP #29	230835.143	1662437.175	231059.553	1662390.322	Control	0.566	0.023	0.566	1.250
GCP #27	225123.102	1660793.214	225347.728	1660745.639	Control	-0.005	-0.575	0.575	1.270

Control Point Error (X) = 0.3063 (Y) = 0.3335 (Total) = 0.4528

APPENDIX B

Field survey data

Table B.1 Detail of sampling point.

No.	X Projection	Y Projection	Classes
1	228358	1650014	Cassava
2	228165	1648887	Cassava
3	229248	1650358	Cassava
4	228792	1651670	Cassava
5	228801	1652221	Cassava
6	225903	1662001	Cassava
7	228172	1667224	Cassava
8	226577	1662969	Cassava
9	226503	1664394	Cassava
10	226874	1665422	Cassava
11	226182	1665671	Cassava
12	229175	1667103	Cassava
13	227357	1663051	Cassava
14	224377	1666869	Cassava
15	223163	1664677	Cassava
16	221302	1665674	Cassava
17	221795	1666086	Cassava
18	226471	1649099	Cassava
19	229173	1659231	Cassava
20	228850	1657618	Cassava
21	225819	1657105	Cassava
22	227627	1653636	Cassava
23	222744	1653096	Cassava
24	230848	1649049	Cassava
25	229065	1648583	Cassava
26	228766	1648952	Cassava

Table B.1 (Continue).

No.	X Projection	Y Projection	Classes
27	226848	1647644	Cassava
28	223029	1648619	Cassava
29	221558	1650228	Cassava
30	221419	1651443	Cassava
31	222922	1649299	Cassava
32	222940	1654895	Cassava
33	226429	1663705	Cassava
34	223414	1657391	Cassava
35	221842	1656279	Sugarcane
36	228491	1649481	Sugarcane
37	228838	1652490	Sugarcane
38	224130	1663484	Sugarcane
39	225005	1664357	Sugarcane
40	226503	1665064	Sugarcane
41	228177	1666789	Sugarcane
42	224159	1667469	Sugarcane
43	221781	1665968	Sugarcane
44	221339	1667274	Sugarcane
45	222811	1660540	Sugarcane
46	229952	1653784	Sugarcane
47	230224	1658845	Sugarcane
48	229997	1660797	Sugarcane
49	229263	1661620	Sugarcane
50	229145	1660561	Sugarcane
51	224928	1660660	Sugarcane
52	225052	1661281	Sugarcane
53	229793	1659117	Sugarcane
54	225966	1656975	Sugarcane
55	227813	1652516	Sugarcane
56	227735	1652049	Sugarcane
57	227787	1650626	Sugarcane
58	224288	1651008	Sugarcane
59	223930	1650436	Sugarcane
60	229433	1648644	Sugarcane

Table B.1 (Continue).

No.	X Projection	Y Projection	Classes
61	228248	1648359	Sugarcane
62	228695	1651016	Sugarcane
63	221936	1650490	Sugarcane
64	221604	1651217	Sugarcane
65	222389	1651636	Sugarcane
66	227207	1661741	Sugarcane
67	220688	1652217	Sugarcane
68	224088	1655689	Sugarcane
69	223980	1656439	Sugarcane
70	226650	1666971	Sugarcane
71	220729	1666164	Urban and built-up areas
72	220717	1657433	Urban and built-up areas
73	221511	1661084	Urban and built-up areas
74	227434	1659592	Urban and built-up areas
75	224439	1652594	Urban and built-up areas
76	228841	1654615	Paddy field
77	229071	1652874	Paddy field
78	230256	1653571	Paddy field
79	224911	1661105	Paddy field
80	228674	1660368	Paddy field
81	226021	1659555	Paddy field
82	229009	1657324	Paddy field
83	227796	1656507	Paddy field
84	227035	1658272	Paddy field
85	227305	1657566	Paddy field
86	220960	1657763	Paddy field
87	225400	1652019	Paddy field
88	223038	1652410	Paddy field
89	224021	1649511	Paddy field
90	230545	1648712	Paddy field
91	227863	1647805	Paddy field
92	224327	1647848	Paddy field
93	221736	1651830	Paddy field
94	222988	1654413	Paddy field

Table B.1 (Continue).

No.	X Projection	Y Projection	Classes
95	223784	1656806	Paddy field
96	222193	1655845	Paddy field
97	220536	1656287	Paddy field
98	229524	1653651	Forest land
99	230683	1656177	Forest land
100	230559	1662566	Forest land
101	227416	1653363	Forest land
102	223990	1653097	Forest land
103	227992	1652813	Forest land
104	224600	1651796	Forest land
105	226968	1648098	Forest land
106	224965	1649011	Forest land
107	223202	1655128	Forest land
108	230148	1666760	Water bodies
109	224648	1666115	Water bodies
110	228573	1657662	Water bodies
111	227622	1648193	Water bodies
112	224256	1648551	Water bodies
113	228288	1650505	Bare land
114	229323	1662610	Bare land
115	225045	1663962	Bare land
116	221068	1656160	Bare land
117	224537	1658856	Bare land
118	229374	1648288	Bare land
119	222155	1650619	Bare land
120	228003	1660900	Bare land
121	227003	1653592	Bare land
122	230103	1652425	Bare land

Table B.2 Ground photograph of each land cover class from field survey.

Land cover type	Feature
Cassava	
Sugarcane	
Urban and built-up areas	

Table B.2 (Continue).

Land cover type	Feature
Paddy field	
Forest land	
Water bodies	

Table B.2 (Continue).

Land cover type	Feature
Bare land	

APPENDIX C

Error matrices

Table C.1 Error matrices of Expert Systems classification with multispectral dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	25	0	0	25	71.43%	100.00%	64.06%	100.00%
Sugarcane	0	26	10	36	74.29%	72.22%	63.52%	61.05%
Unclassified	10	9	42	61	80.77%	68.85%	61.54%	45.71%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 76.23%

Overall Kappa coefficient = 63.06%

Table C.2 Error matrices of Expert Systems classification with multispectral and NDVI dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	**Unclassified**	**Cassava**	**Sugarcane**	**Row Total**			**PA**	**UA**
Cassava	25	0	0	25	71.43%	100.00%	64.06%	100.00%
Sugarcane	0	26	9	35	74.29%	74.29%	63.94%	63.94%
Unclassified	10	9	43	62	82.69%	69.35%	64.81%	46.59%
Column Total	35	35	52	122	-	-	-	-
Overall Classification Accuracy = 77.05%								
Overall Kappa coefficient = 64.27%								

Table C.3 Error matrices of Expert Systems classification with multispectral and soil series dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	**Unclassified**	**Cassava**	**Sugarcane**	**Row Total**			**PA**	**UA**
Cassava	25	0	0	25	71.43%	100.00%	64.06%	100.00%
Sugarcane	0	24	6	30	68.57%	80.00%	58.32%	71.95%
Unclassified	10	11	46	67	88.46%	68.66%	74.41%	45.37%
Column Total	35	35	52	122	-	-	-	-
Overall Classification Accuracy = 77.87%								
Overall Kappa coefficient = 65.23%								

Table C.4 Error matrices of Expert Systems classification with multispectral and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	24	0	0	24	68.57%	100.00%	60.87%	100.00%
Sugarcane	0	25	4	29	71.43%	86.21%	62.52%	80.66%
Unclassified	11	10	48	69	92.31%	69.57%	82.29%	46.96%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 79.51%

Overall Kappa coefficient = 67.69%

Table C.5 Error matrices of Expert Systems classification with multispectral, soil series and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	24	0	0	24	68.57%	100.00%	60.87%	100.00%
Sugarcane	0	24	3	27	68.57%	88.89%	59.64%	84.42%
Unclassified	11	11	49	71	94.23%	69.01%	86.20%	46.00%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 79.51%

Overall Kappa coefficient = 67.58%

Table C.6 Error matrices of Expert Systems classification with multispectral, NDVI, soil series and landform dataset.

Classified Data	**Ground reference**				**PA**	**UA**	**Conditional**	
Class Name	**Unclassified**	**Cassava**	**Sugarcane**	**Row Total**			**PA**	**UA**
Cassava	24	0	0	24	68.57%	100.00%	60.87%	100.00%
Sugarcane	0	24	3	27	68.57%	88.89%	59.64%	84.42%
Unclassified	11	11	49	71	94.23%	69.01%	86.20%	46.00%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 79.51%

Overall Kappa coefficient = 67.58%

Table C.7 Error matrices of ANN classification with training rate 0.1 with multispectral dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	**Cassava**	**Sugarcane**	**Urban and built-up areas**	**Paddy field**	**Forest land**	**Water bodies**	**Bare land**	**Row Total**			**PA**	**UA**
Cassava	25	0	0	0	0	0	0	25	71.43%	100.00%	64.06%	100.00%
Sugarcane	10	32	0	4	7	0	0	53	91.43%	60.38%	84.84%	44.44%
Urban and built-up areas	0	0	3	0	0	0	0	3	60.00%	100.00%	58.99%	100.00%
Paddy field	0	3	1	8	0	0	0	12	36.36%	66.67%	29.42%	59.33%
Forest land	0	0	0	1	3	0	0	4	30.00%	75.00%	27.63%	72.77%
Water bodies	0	0	0	9	0	5	0	14	100.00%	35.71%	100.00%	32.97%
Bare land	0	0	1	0	0	0	10	11	100.00%	90.91%	100.00%	90.10%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-
Overall Classification Accuracy = 79.51%												
Overall Kappa coefficient = 67.58%												

Table C.8 Error matrices of ANN classification with training rate 0.1 with multispectral and NDVI dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	**Cassava**	**Sugarcane**	**Urban and built-up areas**	**Paddy field**	**Forest land**	**Water bodies**	**Bare land**	**Row Total**			**PA**	**UA**
Cassava	32	2	0	0	0	0	0	34	91.43%	94.12%	88.12%	91.75%
Sugarcane	2	28	0	4	3	0	0	37	80.00%	75.68%	71.29%	65.89%
Urban and built-up areas	0	0	4	0	0	0	0	4	80.00%	100.00%	79.32%	100.00%
Paddy field	1	3	0	4	0	1	0	9	18.18%	44.44%	11.67%	32.22%
Forest land	0	2	0	11	7	0	0	20	70.00%	35.00%	64.12%	29.20%
Water bodies	0	0	0	3	0	4	0	7	80.00%	57.14%	78.78%	55.31%
Bare land	0	0	1	0	0	0	10	11	100.00%	90.91%	100.00%	90.10%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-
Overall Classification Accuracy = 72.95% Overall Kappa coefficient = 65.99%												

Table C.9 Error matrices of ANN classification with training rate 0.1 with multispectral and soil series dataset.

Classified Data	**Ground reference**								**PA**	**UA**	**Conditional**	
Class Name	**Cassava**	**Sugarcane**	**Urban and built-up areas**	**Paddy field**	**Forest land**	**Water bodies**	**Bare land**	**Row Total**			**PA**	**UA**
Cassava	31	2	0	0	0	0	0	33	88.57%	93.94%	84.33%	91.50%
Sugarcane	2	30	1	4	5	0	0	42	85.71%	71.43%	78.21%	59.93%
Urban and built-up areas	0	0	4	0	0	0	0	4	80.00%	100.00%	79.32%	100.00%
Paddy field	1	3	0	11	0	0	0	15	50.00%	73.33%	42.99%	67.47%
Forest land	0	0	0	6	5	0	0	11	50.00%	45.45%	45.05%	40.58%
Water bodies	0	0	0	1	0	5	0	6	100.00%	83.33%	100.00%	82.62%
Bare land	1	0	0	0	0	0	10	11	100.00%	90.91%	100.00%	90.10%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 78.69%

Overall Kappa coefficient = 72.79%

Table C.10 Error matrices of ANN classification with training rate 0.1 with multispectral and landform dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	33	2	0	0	0	0	0	35	94.29%	94.29%	91.99%	91.99%
Sugarcane	2	24	0	4	1	0	0	31	68.57%	77.42%	57.86%	68.34%
Urban and built-up areas	0	0	4	0	0	0	0	4	80.00%	100.00%	79.32%	100.00%
Paddy field	0	3	0	4	0	1	0	8	18.18%	50.00%	12.44%	39.00%
Forest land	0	6	0	11	9	0	0	26	90.00%	34.62%	87.29%	28.78%
Water bodies	0	0	0	3	0	4	0	7	80.00%	57.14%	78.78%	55.31%
Bare land	0	0	1	0	0	0	10	11	100.00%	90.91%	100.00%	90.10%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 72.13%

Overall Kappa coefficient = 65.36%

Table C.11 Error matrices of ANN classification with training rate 0.1 with multispectral, soil series and landform dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	**Cassava**	**Sugarcane**	**Urban and built-up areas**	**Paddy field**	**Forest land**	**Water bodies**	**Bare land**	**Row Total**			**PA**	**UA**
Cassava	28	0	0	0	0	0	0	28	80.00%	100.00%	74.04%	100.00%
Sugarcane	5	28	0	2	4	0	2	41	80.00%	68.29%	69.88%	55.54%
Urban and built-up areas	0	1	4	1	1	0	0	7	80.00%	57.14%	78.78%	55.31%
Paddy field	0	3	0	12	0	0	0	15	54.55%	80.00%	48.17%	75.60%
Forest land	1	3	0	5	5	0	0	14	50.00%	35.71%	43.52%	29.97%
Water bodies	0	0	0	2	0	5	0	7	100.00%	71.43%	100.00%	70.21%
Bare land	1	0	1	0	0	0	8	10	80.00%	80.00%	78.21%	78.21%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 73.77%

Overall Kappa coefficient = 67.00%

Table C.12 Error matrices of ANN classification with training rate 0.1 with multispectral, NDVI, soil series and landform dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	31	0	0	0	0	0	0	31	88.57%	100.00%	84.68%	100.00%
Sugarcane	3	33	1	4	5	0	0	46	94.29%	71.74%	90.83%	60.37%
Urban and built-up areas	0	0	3	0	0	2	1	6	60.00%	50.00%	57.93%	47.86%
Paddy field	1	2	0	10	0	0	0	13	45.45%	76.92%	38.95%	71.85%
Forest land	0	0	0	4	5	0	0	9	50.00%	55.56%	46.02%	51.59%
Water bodies	0	0	0	4	0	3	0	7	60.00%	42.86%	57.57%	40.42%
Bare land	0	0	1	0	0	0	9	10	90.00%	90.00%	89.11%	89.11%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 77.05%

Overall Kappa coefficient = 70.67%

Table C.13 Error matrices of ANN classification with training rate 0.2 with multispectral dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	23	0	0	0	0	0	0	23	65.71%	100.00%	57.75%	100.00%
Sugarcane	11	30	0	8	4	0	0	53	85.71%	56.60%	74.74%	39.15%
Urban and built-up areas	0	0	4	1	0	0	0	5	80.00%	80.00%	79.15%	79.15%
Paddy field	1	5	0	5	0	0	0	11	22.73%	45.45%	15.07%	33.45%
Forest land	0	0	0	8	6	0	0	14	60.00%	42.86%	54.81%	37.76%
Water bodies	0	0	0	0	0	5	0	5	100.00%	100.00%	100.00%	100.00%
Bare land	0	0	1	0	0	0	10	11	100.00%	90.91%	100.00%	90.10%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-
Overall Classification Accuracy = 68.03% Overall Kappa coefficient = 59.27%												

Table C.14 Error matrices of ANN classification with training rate 0.2 with multispectral and NDVI dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	29	0	0	0	0	0	0	29	82.86%	100.00%	77.51%	100.00%
Sugarcane	5	28	0	9	1	0	0	43	80.00%	65.12%	69.11%	51.08%
Urban and built-up areas	0	0	4	1	0	0	0	5	80.00%	80.00%	79.15%	79.15%
Paddy field	0	0	0	0	0	1	0	1	0.00%	0.00%	-0.83%	-22.00%
Forest land	0	5	0	5	8	0	0	18	80.00%	44.44%	76.54%	39.48%
Water bodies	0	0	0	5	0	4	0	9	80.00%	44.44%	78.41%	42.07%
Bare land	1	2	1	2	1	0	10	17	100.00%	58.82%	100.00%	55.15%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-
Overall Classification Accuracy = 68.03% Overall Kappa coefficient = 60.09%												

Table C.15 Error matrices of ANN classification with training rate 0.2 with multispectral and soil series dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	32	4	0	1	0	0	0	37	91.43%	86.49%	87.70%	81.05%
Sugarcane	2	29	0	7	10	0	0	48	82.86%	60.42%	71.74%	44.49%
Urban and built-up areas	0	0	1	0	0	0	0	1	20.00%	100.00%	19.34%	100.00%
Paddy field	0	2	0	8	0	0	0	10	36.36%	80.00%	30.68%	75.60%
Forest land	0	0	0	4	0	0	0	4	0.00%	0.00%	-3.39%	-8.93%
Water bodies	0	0	0	2	0	5	0	7	100.00%	71.43%	100.00%	70.21%
Bare land	1	0	4	0	0	0	10	15	100.00%	66.67%	100.00%	63.69%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 69.67%

Overall Kappa coefficient = 60.61%

Table C.16 Error matrices of ANN classification with training rate 0.2 with multispectral and landform dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	31	0	0	1	0	0	0	32	88.57%	96.88%	84.51%	95.62%
Sugarcane	4	33	0	8	1	0	0	46	94.29%	71.74%	90.83%	60.37%
Urban and built-up areas	0	0	4	1	0	0	0	5	80.00%	80.00%	79.15%	79.15%
Paddy field	0	0	0	2	0	1	0	3	9.09%	66.67%	6.80%	59.33%
Forest land	0	2	0	4	9	0	0	15	90.00%	60.00%	88.60%	56.43%
Water bodies	0	0	0	6	0	4	0	10	80.00%	40.00%	78.21%	37.44%
Bare land	0	0	1	0	0	0	10	11	100.00%	90.91%	100.00%	90.10%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 76.23%

Overall Kappa coefficient = 69.90%

Table C.17 Error matrices of ANN classification with training rate 0.2 with multispectral, soil series and landform dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	23	0	0	0	0	0	0	23	65.71%	100.00%	57.75%	100.00%
Sugarcane	11	28	2	3	3	0	1	48	80.00%	58.33%	67.03%	41.57%
Urban and built-up areas	0	1	2	4	0	0	0	7	40.00%	28.57%	36.35%	25.52%
Paddy field	1	2	0	12	0	0	0	15	54.55%	80.00%	48.17%	75.60%
Forest land	0	4	0	2	7	0	0	13	70.00%	53.85%	66.42%	49.73%
Water bodies	0	0	0	1	0	5	0	6	100.00%	83.33%	100.00%	82.62%
Bare land	0	0	1	0	0	0	9	10	90.00%	90.00%	89.11%	89.11%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-
Overall Classification Accuracy = 70.49% Overall Kappa coefficient = 62.70%												

Table C.18 Error matrices of ANN classification with training rate 0.2 with multispectral, NDVI, soil series and landform dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	22	0	0	0	0	0	0	22	62.86%	100.00%	54.69%	100.00%
Sugarcane	12	31	1	5	3	0	0	52	88.57%	59.62%	80.08%	43.37%
Urban and built-up areas	0	0	3	0	0	1	1	5	60.00%	60.00%	58.29%	58.29%
Paddy field	1	4	0	12	0	0	0	17	54.55%	70.59%	47.19%	64.12%
Forest land	0	0	0	3	7	0	0	10	70.00%	70.00%	67.32%	67.32%
Water bodies	0	0	1	1	0	4	0	6	80.00%	66.67%	78.97%	65.24%
Bare land	0	0	0	1	0	0	9	10	90.00%	90.00%	89.11%	89.11%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 72.13%

Overall Kappa coefficient = 64.44%

Table C. 19 Error matrices of ANN classification with training rate 0.3 with multispectral dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	9	0	0	0	0	0	0	9	25.71%	100.00%	19.80%	100.00%
Sugarcane	9	31	0	5	4	0	0	49	88.57%	63.27%	80.90%	48.49%
Urban and built-up areas	0	0	2	0	0	0	0	2	40.00%	100.00%	39.00%	100.00%
Paddy field	0	2	0	8	0	0	0	10	36.36%	80.00%	30.68%	75.60%
Forest land	17	1	0	2	6	0	0	26	60.00%	23.08%	49.17%	16.21%
Water bodies	0	0	0	0	0	5	0	5	100.00%	100.00%	100.00%	100.00%
Bare land	0	1	3	7	0	0	10	21	100.00%	47.62%	100.00%	42.94%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 58.20%

Overall Kappa coefficient = 48.70%

Table C. 20 Error matrices of ANN classification with training rate 0.3 with multispectral and NDVI dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	16	0	0	0	0	0	0	16	45.71%	100.00%	37.52%	100.00%
Sugarcane	19	32	0	18	10	0	0	79	91.43%	40.51%	75.68%	16.57%
Urban and built-up areas	0	0	2	0	0	0	0	2	40.00%	100.00%	39.00%	100.00%
Paddy field	0	3	0	4	0	0	0	7	18.18%	57.14%	13.20%	47.71%
Forest land	0	0	0	0	0	1	0	1	0.00%	0.00%	-0.83%	-8.93%
Water bodies	0	0	0	0	0	4	0	4	80.00%	100.00%	79.32%	100.00%
Bare land	0	0	3	0	0	0	10	13	100.00%	76.92%	100.00%	74.86%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-
Overall Classification Accuracy = 55.74% Overall Kappa coefficient = 41.36%												

Table C.21 Error matrices of ANN classification with training rate 0.3 with multispectral and soil series dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	31	0	0	0	0	0	0	31	88.57%	100.00%	84.68%	100.00%
Sugarcane	2	31	0	6	4	0	0	43	88.57%	72.09%	82.35%	60.87%
Urban and built-up areas	0	0	4	0	1	2	0	7	80.00%	57.14%	78.78%	55.31%
Paddy field	2	3	1	11	0	0	0	17	50.00%	64.71%	41.90%	56.94%
Forest land	0	0	0	2	5	0	0	7	50.00%	71.43%	46.96%	68.88%
Water bodies	0	0	0	2	0	3	0	5	60.00%	60.00%	58.29%	58.29%
Bare land	0	1	0	1	0	0	10	12	100.00%	83.33%	100.00%	81.85%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-
Overall Classification Accuracy = 77.87%												
Overall Kappa coefficient = 71.77%												

Table C.22 Error matrices of ANN classification with training rate 0.3 with multispectral and landform dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	32	3	0	0	0	0	0	35	91.43%	91.43%	87.98%	87.98%
Sugarcane	1	11	0	0	0	0	0	12	31.43%	91.67%	23.95%	88.31%
Urban and built-up areas	0	0	3	1	0	0	0	4	60.00%	75.00%	58.64%	73.93%
Paddy field	1	2	0	12	0	0	0	15	54.55%	80.00%	48.17%	75.60%
Forest land	0	13	0	4	9	0	0	26	90.00%	34.62%	87.29%	28.78%
Water bodies	0	0	0	4	0	5	0	9	100.00%	55.56%	100.00%	53.66%
Bare land	1	6	2	1	1	0	10	21	100.00%	47.62%	100.00%	42.94%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 67.21%

Overall Kappa coefficient = 60.56%

Table C.23 Error matrices of ANN classification with training rate 0.3 with multispectral, soil series and landform dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	**Cassava**	**Sugarcane**	**Urban and built-up areas**	**Paddy field**	**Forest land**	**Water bodies**	**Bare land**	**Row Total**			**PA**	**UA**
Cassava	26	0	0	0	0	0	0	26	74.29%	100.00%	67.32%	100.00%
Sugarcane	8	21	1	2	2	0	0	34	60.00%	61.76%	44.55%	46.38%
Urban and built-up areas	0	0	4	0	0	0	0	4	80.00%	100.00%	79.32%	100.00%
Paddy field	0	4	0	13	0	0	0	17	59.09%	76.47%	52.47%	71.29%
Forest land	1	10	0	4	8	0	0	23	80.00%	34.78%	75.35%	28.96%
Water bodies	0	0	0	3	0	5	0	8	100.00%	62.50%	100.00%	60.90%
Bare land	0	0	0	0	0	0	10	10	100.00%	100.00%	100.00%	100.00%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 71.31%

Overall Kappa coefficient = 64.48%

Table C.24 Error matrices of ANN classification with training rate 0.3 with multispectral, NDVI, soil series and landform dataset.

Classified Data	Ground reference								PA	UA	Conditional	
Class Name	Cassava	Sugarcane	Urban and built-up areas	Paddy field	Forest land	Water bodies	Bare land	Row Total			PA	UA
Cassava	16	2	0	0	0	0	0	18	45.71%	88.89%	36.32%	84.42%
Sugarcane	15	26	0	2	0	0	0	43	74.29%	60.47%	60.29%	44.56%
Urban and built-up areas	0	2	4	1	0	0	2	9	80.00%	44.44%	78.41%	42.07%
Paddy field	1	2	0	4	0	0	0	7	18.18%	57.14%	13.20%	47.71%
Forest land	3	3	0	11	9	0	0	26	90.00%	34.62%	87.29%	28.78%
Water bodies	0	0	1	2	0	5	0	8	100.00%	62.50%	100.00%	60.90%
Bare land	0	0	0	2	1	0	8	11	80.00%	72.73%	78.02%	70.29%
Column Total	35	35	5	22	10	5	10	122	-	-	-	-

Overall Classification Accuracy = 59.02%

Overall Kappa coefficient = 49.75%

Table C.25 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	25	0	0	25	71.43%	100.00%	64.06%	100.00%
Sugarcane	10	32	10	52	91.43%	61.54%	85.06%	46.07%
Unclassified	0	3	42	45	80.77%	93.33%	69.53%	88.38%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 81.15%

Overall Kappa coefficient = 71.51%

Table C.26 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral and NDVI dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	32	2	0	34	91.43%	94.12%	88.12%	91.75%
Sugarcane	2	28	7	37	80.00%	75.68%	71.29%	65.89%
Unclassified	1	5	45	51	86.54%	88.24%	76.87%	79.50%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 86.07%

Overall Kappa coefficient = 78.72%

Table C.27 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral and soil series dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	31	2	0	33	88.57%	93.94%	84.33%	91.50%
Sugarcane	2	30	9	41	85.71%	73.17%	78.48%	62.38%
Unclassified	2	3	43	48	82.69%	89.58%	71.47%	81.85%
Column Total	35	35	52	122	-	-	-	-
Overall Classification Accuracy = 85.25% Overall Kappa coefficient = 77.59%								

Table C.28 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	33	2	0	35	94.29%	94.29%	91.99%	91.99%
Sugarcane	1	23	5	29	65.71%	79.31%	55.02%	70.99%
Unclassified	1	10	47	58	90.38%	81.03%	81.67%	66.95%
Column Total	35	35	52	122	-	-	-	-
Overall Classification Accuracy = 84.43% Overall Kappa coefficient = 75.92%								

Table C.29 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral, soil series and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	28	0	0	28	80.00%	100.00%	74.04%	100.00%
Sugarcane	5	28	8	41	80.00%	68.29%	69.88%	55.54%
Unclassified	2	7	44	53	84.62%	83.02%	72.80%	70.40%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 81.97%

Overall Kappa coefficient = 72.37%

Table C.30 Error matrices of ANN classification of three common categories with training rate 0.1 with multispectral, NDVI, soil series and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	31	0	0	31	88.57%	100.00%	84.68%	100.00%
Sugarcane	3	33	9	45	94.29%	73.33%	90.95%	62.61%
Unclassified	1	2	43	46	82.69%	93.48%	72.22%	88.63%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 87.70%

Overall Kappa coefficient = 81.39%

Table C.31 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	23	0	0	23	65.71%	100.00%	57.75%	100.00%
Sugarcane	11	30	9	50	85.71%	60.00%	75.79%	43.91%
Unclassified	1	5	43	49	82.69%	87.76%	71.07%	78.66%
Column Total	35	35	52	122	-	-	-	-
Overall Classification Accuracy = 78.69%								
Overall Kappa coefficient = 67.57%								

Table C.32 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral and NDVI dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	29	0	0	29	82.86%	100.00%	77.51%	100.00%
Sugarcane	5	28	8	41	80.00%	68.29%	69.88%	55.54%
Unclassified	1	7	44	52	84.62%	84.62%	73.19%	73.19%
Column Total	35	35	52	122	-	-	-	-
Overall Classification Accuracy = 82.79%								
Overall Kappa coefficient = 73.67%								

Table C.33 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral and soil series dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	32	4	1	37	91.43%	86.49%	87.70%	81.05%
Sugarcane	2	29	16	47	82.86%	61.70%	72.11%	46.29%
Unclassified	1	2	35	38	67.31%	92.11%	52.52%	86.24%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 78.69%

Overall Kappa coefficient = 68.18%

Table C.34 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	31	0	1	32	88.57%	96.88%	84.51%	95.62%
Sugarcane	4	32	8	44	91.43%	72.73%	86.59%	61.76%
Unclassified	0	3	43	46	82.69%	93.48%	72.22%	88.63%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 86.89%

Overall Kappa coefficient = 80.15%

Table C.35 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral, soil series and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	23	0	0	23	65.71%	100.00%	57.75%	100.00%
Sugarcane	11	27	9	47	77.14%	57.45%	62.82%	40.33%
Unclassified	1	8	43	52	82.69%	82.69%	69.84%	69.84%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 76.23%

Overall Kappa coefficient = 63.64%

Table C.36 Error matrices of ANN classification of three common categories with training rate 0.2 with multispectral, NDVI, soil series and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	22	0	0	22	62.86%	100.00%	54.69%	100.00%
Sugarcane	12	29	9	50	82.86%	58.00%	70.95%	41.10%
Unclassified	1	6	43	50	82.69%	86.00%	70.67%	75.60%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 77.05%

Overall Kappa coefficient = 65.01%

Table C.37 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	9	0	0	9	25.71%	100.00%	19.80%	100.00%
Sugarcane	9	31	8	48	88.57%	64.58%	81.16%	50.34%
Unclassified	17	4	44	65	84.62%	67.69%	67.07%	43.69%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 68.85%

Overall Kappa coefficient = 51.25%

Table C.38 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral and NDVI dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	16	0	0	16	45.71%	100.00%	37.52%	100.00%
Sugarcane	19	32	28	79	91.43%	40.51%	75.68%	16.57%
Unclassified	0	3	24	27	46.15%	88.89%	30.85%	80.63%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 59.02%

Overall Kappa coefficient = 39.93%

Table C.39 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral and soil series dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	31	0	0	31	88.57%	100.00%	84.68%	100.00%
Sugarcane	2	31	10	43	88.57%	72.09%	82.35%	60.87%
Unclassified	2	4	42	48	80.77%	87.50%	68.30%	78.21%
Column Total	35	35	52	122	-	-	-	-
Overall Classification Accuracy = 85.25% Overall Kappa coefficient = 77.59%								

Table C.40 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	32	3	0	35	91.43%	91.43%	87.98%	87.98%
Sugarcane	1	10	0	11	28.57%	90.91%	21.49%	87.25%
Unclassified	2	22	52	76	100.00%	68.42%	100.00%	44.96%
Column Total	35	35	52	122	-	-	-	-
Overall Classification Accuracy = 77.05% Overall Kappa coefficient = 63.36%								

Table C.41 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral, soil series and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	26	0	0	26	74.29%	100.00%	67.32%	100.00%
Sugarcane	8	20	5	33	57.14%	60.61%	41.25%	44.76%
Unclassified	1	15	47	63	90.38%	74.60%	80.12%	55.74%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 76.23%

Overall Kappa coefficient = 62.93%

Table C.42 Error matrices of ANN classification of three common categories with training rate 0.3 with multispectral, NDVI, soil series and landform dataset.

Classified Data	Ground reference				PA	UA	Conditional	
Class Name	Unclassified	Cassava	Sugarcane	Row Total			PA	UA
Cassava	16	2	0	18	45.71%	88.89%	36.32%	84.42%
Sugarcane	15	26	2	43	74.29%	60.47%	60.29%	44.56%
Unclassified	4	7	50	61	96.15%	81.97%	92.31%	68.57%
Column Total	35	35	52	122	-	-	-	-

Overall Classification Accuracy = 75.41%

Overall Kappa coefficient = 61.78%

APPENDIX D

Variance and Z statistic of error matrices

Table D.1 The variance and Z statistic of Expert System classification.

Dataset	KHAT	Variance
Multispectral and NDVI dataset	0.6306	0.0037
Multispectral and soil series dataset	0.6433	0.0036
Multispectral and Landform dataset	0.6561	0.0034
Multispectral, soil series and landform dataset	0.6815	0.0032
Multispectral, NDVI, soil series and landform dataset	0.6815	0.0032
Multispectral and NDVI dataset	0.6815	0.0032

Table D.2 The variance and Z statistic of ANN classification.

Dataset	KHAT	Variance
ANN with learning rate 0.1		
Multispectral and NDVI dataset	0.6232	0.0027
Multispectral and soil series dataset	0.6599	0.0024
Multispectral and Landform dataset	0.7279	0.0022
Multispectral, soil series and landform dataset	0.6536	0.0024
Multispectral, NDVI, soil series and landform dataset	0.6700	0.0024
Multispectral and NDVI dataset	0.7067	0.0022

Table D.2 (Continue).

Dataset	KHAT	Variance
ANN with learning rate 0.2		
Multispectral and NDVI dataset	0.5927	0.0029
Multispectral and soil series dataset	0.6009	0.0025
Multispectral and Landform dataset	0.6061	0.0027
Multispectral, soil series and landform dataset	0.6990	0.0022
Multispectral, NDVI, soil series and landform dataset	0.6270	0.0027
Multispectral and NDVI dataset	0.6444	0.0027
ANN with learning rate 0.3		
Multispectral and NDVI dataset	0.4870	0.0028
Multispectral and soil series dataset	0.4136	0.0041
Multispectral and Landform dataset	0.7177	0.0022
Multispectral, soil series and landform dataset	0.6056	0.0024
Multispectral, NDVI, soil series and landform dataset	0.6448	0.0026
Multispectral and NDVI dataset	0.4975	0.0029

CURRICULUM VITAE

Name Miss WANNATAT TESSAWAT

Date of Birth April 28, 1985

Place of Birth Ubon Ratchathani

Education 2002 - 2006 Bachelor of Information Science (Information Studies), Suranaree University of Technology, Nakhon Ratchasima, Thailand.

Publications

Ongsomwang, S. and Tessawat, W. (2010). Cassava and Sugarcane Area Classification from THEOS Data using Expert System. **Geoinfotech 2010.** Impact Convention Center, Muang Thong Thani, Nonthaburi, Thailand.

Grants and Fellowships Potential scholarship, Suranaree University of Technology