# COMPARISON OF EXPERT SYSTEM AND ARTIFICIAL NEURAL NETWORK CLASSIFICATION FOR CASSAVA AND SUGARCANE AREAS USING THEOS DATA

Wannatat TESSAWAT [a] and Suwit ONGSOMWANG [b]

, School of Remote Sensin
[a] Graduate student, School of Remote Sensing, Institute of Science, Suranaree University of Technology
[b] Assistant Professor, School of Remote Sensing, Institute of Science, Suranaree University of Technology
111 University Avenue, Muang District, Nakhon Ratchasima 30000, Thailand;
Tel: +66(0) -4422-4652; Fax. +66(0) -4422-4316
E-mail: M5010011@g.sut.ac.th

**KEY WORDS**: Expert System, Artificial Neural Network, Cassava, Sugarcane, THEOS

**Abstract:** This study aims to evaluate an optimum method for cassava and sugarcane classification using Expert Systems and Artificial Neural Network (ANN). In the study, multispectral data of THEOS was 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 in Chakkarat district, Nakhon Ratchasima province. For accuracy assessment, the overall accuracy and Kappa coefficient of Expert System for cassava and sugarcane classification varied between 76.23 and 79.51% and 63.06 and 67.69%, respectively. Meanwhile, overall accuracy and Kappa coefficient of ANN with 0.1 learning rate varied between 81.15 and 87.70% and 71.51 and 81.39%, respectively. Meanwhile, for an optimum classification method evaluation, ANN was selected as an optimum method for sugarcane and cassava classification because ANN can provide overall accuracies and Kappa coefficients higher than Expert System in all datasets. Furthermore, the combination between multispectral data and an additional data can increase accuracy of cassava and sugarcane classification using Expert System or ANN

## INTRODUCTION

Recently, Thai government encourages farmers and private companies to cultivate more alternative energy crops such as cassava, sugarcane, oil palm and maize for energy production. In northeast region, the cultivation area of cassava and sugarcane has been expanded to support ethanol industry. Nakhon Ratchasima province is the largest cassava cultivation site in Thailand. In addition, many ethanol industries in Nakhon Ratchasima province had been operated since 2008 to produce oil substitution from sugarcane, cassava, molasses and cassava waste (Office of Agricultural Economics, 2007). Thus, this study aimed to evaluate optimum method to extract cassava and sugarcane areas from remotely sensed data with Expert Systems and Artificial Neural Network (ANN) method.

## STUDY AREA

The study area is a part of Chakkarat district which is located in Nakhon Ratchasima province, northeast of Thailand. The study area is approximately 200 sq. km (Figure 1) where the topography is quite flats and slightly undulates in some areas. The cultivated areas are paddy field, field crops and orchards. The flat plain areas in the middle of district are covered by paddy field (Office of soil survey and land use planning, 2007).

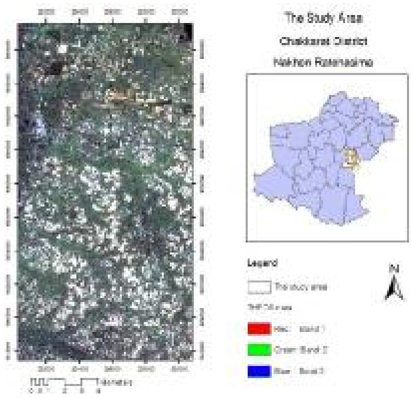


**Figure 1** The study area

## METHODS

The schematic workflow of research methodology was displayed in Figure 2. The details of each component can be summarized as following.

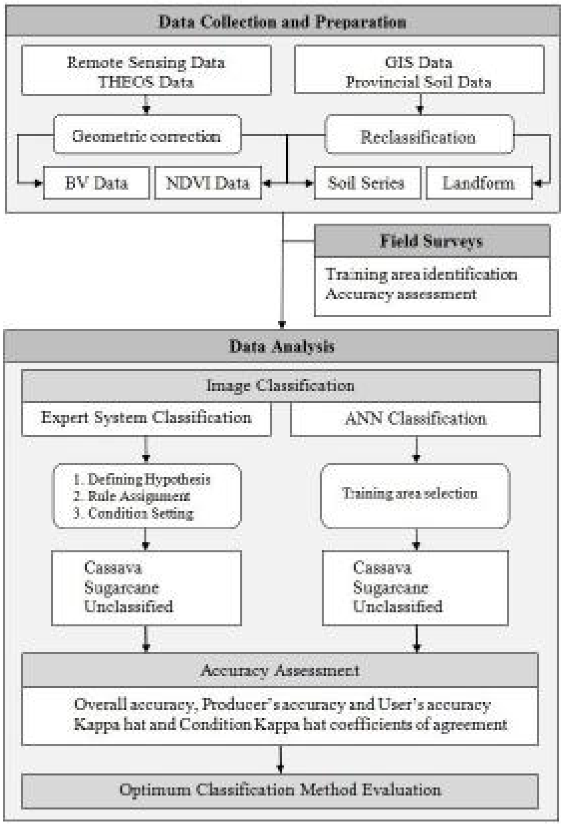


**Figure 2** The research methodology

1. Data collection and preparation

Remotely sensed data and GIS data had been collected from various sources in this study. Characteristics and sources of the collected data in the study were summarized in Table 1. In practice, remotely sensed data, THEOS data with four multispectral bands were firstly geometrically corrected by image to image rectification based on identified ground control points from ortho photographs of LDD. After that, NDVI data was created from ratio between red and near infrared bands (band 4 and 1). Whilst, GIS data were extract soil series and landform type from provincial soil data and then converted to raster format with grid size of 15 x 15 m as spatial resolution of THEOS data. Finally, 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 six predefined dataset for data analysis by image stacking included (1) Multispectral data (2) Multispectral data and NDVI (3) Multispectral data and Soil series (4) Multispectral data and Landform (5) Multispectral data, Soil series and Landform (6) Multispectral data, NDVI, Soil series and Landform.

2. Field surveys

Field surveys were conducted for training area selection and accuracy assessment, with supplementary of land use data in 2007 and sugarcane plot data in 2010.

3. Data analysis

Basically, data analysis consisted of three major steps included (1) image classification (2) accuracy assessment and (3) optimum classification method evaluation.

3.1 Image classification

In the study, six predefined datasets were separately used to classify cassava and sugarcane using Expert system and Artificial Neural Network (ANN) algorithm. In practice, cassava and sugarcane hypothesis are firstly defined for thematic classification of Expert system classification then specific conditions of each rule including multispectral data, NDVI, soil series and landform to extracted thematic classification. Herein, minimum and maximum values which were extracted from training area applied for multispectral data and NDVI rule while soil series and landform classes were extracted using overlay analysis between land use data in 2007 and soil data. At the same time, training areas were identified and selected for ANN algorithm using logistic functions with learning rate of 0.1.

3.2 Accuracy assessment

Number of sample points was firstly calculated and identified based on the multinomial distribution with stratified random sampling schema (Congalton and Green, 2008) and they then used to survey in the field as ground reference data for accuracy assessment with overall accuracy and kappa hat coefficient.

3.3 Optimum classification method evaluation

The results of accuracy assessment, overall accuracy and kappa hat coefficient were used to evaluate optimum method for cassava and sugarcane classification.

**Table 1** Details of data collection

| 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] |

[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

## RESULTS

The results of cassava and sugarcane classification of six predefined datasets using Expert system and ANN method which were classified into 3 classes (cassava, sugarcane and unclassified) were displayed in Figure 3 and Figure 4, respectively. Areas of cassava and sugarcane classification from both methods were summarized in Table 2. It was found that extracted areas of cassava and sugarcane in each datasets in each method were slightly different. However, extracted areas of cassava and sugarcane from each dataset of Expert System and ANN were not comparable and rather different.

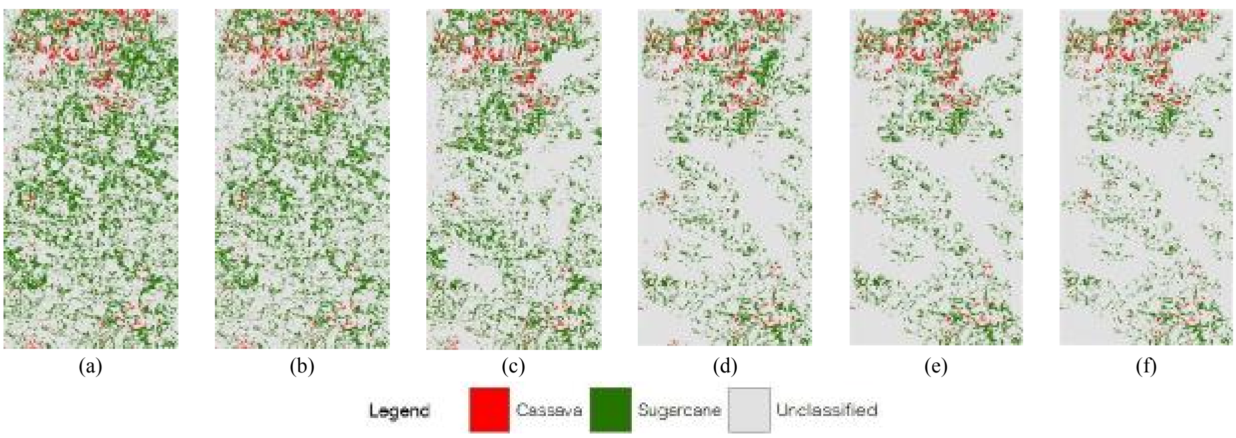


**Figure 3** Cassava and sugarcane areas from Expert Systems classification of each dataset: (a) Multispectral data (b) Multispectral data and NDVI (c) Multispectral data and Soil series (d) Multispectral data and Landform (e) Multispectral data, Soil series and Landform (f) Multispectral data, NDVI, Soil series and Landform.

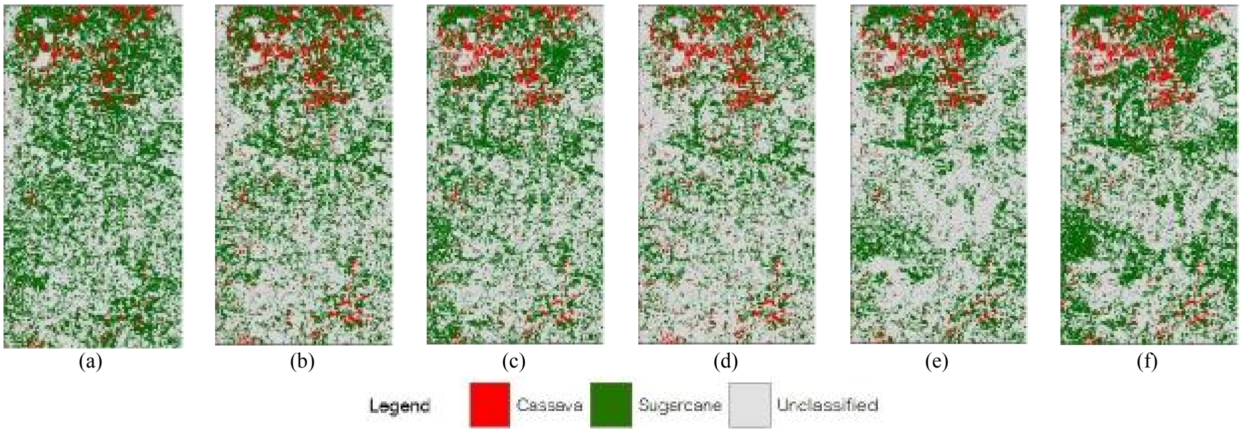


**Figure 4** Cassava and sugarcane areas from ANN classification of each dataset: (a) Multispectral data (b) Multispectral data and NDVI (c) Multispectral data and Soil series (d) Multispectral data and Landform (e) Multispectral data, Soil series and Landform (f) Multispectral data, NDVI, Soil series and Landform.

**Table 2** Areas of cassava and sugarcane classification with Expert System and ANN algorithms.

Unit: sq.km

| Dataset | Expert System | | | ANN | | |
|---|---|---|---|---|---|---|
| | **Cassava** | **Sugarcane** | **Unclassified** | **Cassava** | **Sugarcane** | **Unclassified** |
| (1) Multispectral | 5.70 | 55.95 | 151.43 | 8.04 | 88.9 | 116.15 |
| (2) Multispectral and NDVI | 5.52 | 49.15 | 158.42 | 10.48 | 62.45 | 140.15 |
| (3) Multispectral and Soil series | 5.47 | 37.75 | 169.87 | 16.37 | 107.37 | 89.35 |
| (4) Multispectral and Landform | 5.11 | 26.21 | 181.77 | 14.85 | 79.8 | 118.44 |
| (5) Multispectral, Soil series and Landform | 4.96 | 23.96 | 184.17 | 9.97 | 74.3 | 128.81 |
| (6) Multispectral, NDVI, Soil series and Landform | 4.80 | 20.67 | 187.61 | 6.57 | 94.68 | 111.84 |

In addition, accuracy assessment for cassava and sugarcane classification in each dataset of Expert System and ANN method based on error matrix between 122 classified and ground reference data using overall accuracy and Kappa coefficient were presented in Table 3. 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. Meanwhile, overall accuracy and Kappa coefficient of ANN varied between 81.15 and 87.70% and 71.51 and 81.39%, respectively. The multispectral, NDVI, soil series and landform dataset showed the highest accuracy value with overall accuracy of 87.70% and Kappa coefficient of 81.39%. At the same time multispectral dataset presented the lowest accuracy value with overall accuracy of 81.15% and Kappa coefficient of 71.51%. The overall accuracy and Kappa coefficient of multispectral, NDVI, soil series and landform dataset was higher than multispectral dataset about 6.55 and 9.88%, respectively. As these results, it was found that the combination of additional data with multispectral dataset can increase accuracy of cassava and sugarcane classification using Expert System or ANN.

For an optimum classification method evaluation, ANN was here selected as an optimum method for sugarcane and cassava classification because ANN can provide overall accuracies and Kappa coefficients higher than Expert System in all datasets. This result is similar to research work of Liu, Skidmore and Van Oosten (2002). Therefore, an optimum method for cassava and sugarcane classification is ANN. In this study, ANN can provide 87.70% of overall accuracy and 81.39% of Kappa coefficient from the best dataset of Multispectral, NDVI, Soil series and Landform data. These accuracies were higher than the best dataset of Expert System about 8.19 and 13.70%, respectively.

**Table 3** Comparison of accuracy assessment for cassava and sugarcane classification using Expert System and ANN method in six datasets.

| Dataset | Expert System | | ANN | |
|---|---|---|---|---|
| | Overall accuracy | Kappa coefficient | Overall accuracy | Kappa coefficient |
| (1) Multispectral | **76.23** | **63.06** | **81.15%** | **71.51%** |
| (2) Multispectral and NDVI | 77.05 | 64.27 | 86.07% | 78.72% |
| (3) Multispectral and Soil series | 77.87 | 65.23 | 85.25% | 77.59% |
| (4) Multispectral and Landform | **79.51** | **67.69** | 84.43% | 75.92% |
| (5) Multispectral, Soil series and Landform | 79.51 | 67.58 | 81.97% | 72.37% |
| (6) Multispectral, NDVI, Soil series and Landform | 79.51 | 67.58 | **87.70%** | **81.39%** |

## CONCLUSIONS & RECOMMENDATIONS

In conclusion, Expert System and ANN classification method can be applied for cassava and sugarcane classification using THEOS data with certain level of accuracy in each dataset. Based on accuracy assessment (overall accuracy and Kappa coefficient), an optimum method for cassava and sugarcane classification was ANN. In this study, overall accuracy and Kappa coefficient from the best predefined dataset of Multispectral, NDVI, Soil series and Landform data was 87.70 and 81.39%, respectively. In addition, the combination between multispectral data and an additional data can improve accuracy of cassava and sugarcane classification using Expert System or ANN. However, few land cover types with a single date of remotely sensed data were applied in this study (Figure 1). So, methodology from the study should be tested and verified in others areas, which have different topography or cultivation practice, to confirm conclusion of this study.

## ACKNOWLEDGEMENT

I gratefully acknowledge Geo-Informatics and Space Technology Development Agency (Public Organization) for supporting THEOS data.

## REFERENCES:

Congalton, R. G., Green, K., 2008. Assessing the accuracy of remotely sensed data : principles and practices. CRC Press, Florida, pp.

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

Office of Agricultural Economics., 2007. Licensed list of entrepreneur for ethanol plant : http://www.dede.go.th/dede/fileadmin/usr/bers/gasohol_2008/510125_Ethanol_licesees_45_mills.doc

Office of soil survey and land use planning., 2007. Land use in Chakkarat sub-district, Chakkarat district, Nahon Ratchasima province : http://www. khorat.doae.go.th.