Optimum Machine Learning Algorithm with Object-Based Image Analysis for Detecting Incompliant Land Utilization in Agricultural Land Reform Areas in Thailand

Noina. J.,[1] Ongsomwang, S.,[1*] Sritarapipat, T.[1] and Phumkokrux, S.[2]
[1]School of Mathematics and Geoinformatics, Institute of Science Suranaree University of Technology, Nakhon Ratchasima 3000, Thailand, E-mail: D6110529@g.sut.ac.th, suwit@sut.ac.th,* tanakorn.s@sut.ac.th
[2]Department, Geography, Silpakorn University, Thailand, E-mail: phumkokrux@hotmail.com
**Corresponding Author*

DOI: https://doi.org/10.52939/ijg.v21i10.4539

Abstract
The Agricultural Land Reform Office (ALRO) was established to address the country's developmental challenges by implementing land consolidation programs, which allocate land for both agricultural and residential use to farmers. Currently, it has been found that some land use under ALRO's responsibility does not comply with the Agricultural Land Reform Act, including hotels, resorts, and accommodations. The primary objectives of this study are (1) to determine the optimal machine learning algorithm among support vector machines (SVM), random forests (RF), decision trees (DT), Naïve Bayes (NB), and K nearest neighbor (KNN) for detecting incompliant land utilization in the modeling area (Wang Nam Khiao district) and (2) to validate an optimum machine learning algorithm for detecting incompliant land utilization in the test area (Pak Chong district). The research methodology, which includes eight significant steps, is implemented by applying object-based image analysis (OBIA) combined with machine learning to classify land use from Sentinel 2A imagery in the modeling area and to validate the results in the test area. The results showed that the most suitable machine learning algorithm for detecting incompliant land utilization at ALRO 4-01 plots in the modeling area was RF, as it achieved higher overall accuracy and Kappa coefficient values than SVM, DT, NB, and KNN. The derived overall accuracy and Kappa coefficient of RF were 87.45% and 79.57%, respectively. Furthermore, the selected optimal object features and algorithm from the modeling area were effectively transferred for land use classification and the detection of incompliant land use at ALRO 4-01 plots in the test area, yielding acceptable validation results. These findings can support future monitoring and enforcement of policies in ALRO-4-01 areas.

Keywords: Agricultural Land Reform Areas, Incompliant Land Utilization, Machine Learning Algorithms, Random Forests

1. Introduction

The Thai Agricultural Land Reform Act was promulgated in March 1975. There had been several critical previous attempts to limit private land ownership, for example, the famous "National Economic Plant," which was submitted to the Parliament by Pridi Phanomyong in March 1993, and the 34th Article of the Land Code of 1954, promulgated during the tenure of Field Marshall Pibun Songkharm, set the limits of landownership at fifty rai (1 rai = 0.0016 sq. km) [1]. Thailand faces numerous obstacles, including the loss of rights to agricultural land and the transformation of arable land into rented land for agricultural production. The high cost of rental land leads to a decline in production and numerous disadvantages for agriculturists. Thus, the Agricultural Land Reform Office (ALRO) was established to solve the country's developmental problems through land consolidation programs that allocate land for farming [2].

The ALRO has sought to preserve land for agricultural production, encouraged land redistribution, supported the agricultural sector, and provided funding for agricultural development to narrow the social status gap between the poorest members of society and the middle and upper classes, and to provide support to marginal occupations in areas where the number of farmers has been declining. Land rights are a significant issue for the ALRO 4-01 land certificate, as many farmers who have received it have attempted to sell it illegally. According to the ALRO law, ALRO 4-01 certificate land cannot be sold to another person and must be used solely for agricultural purposes. Additionally, farmers often fail to understand ALRO's objectives and use land unlawfully, such as by constructing hotels and other accommodations in tourist areas. In some areas, numerous people illegally occupy the ALRO 4-01 land [3].

Currently, ALRO has been addressing these issues for many decades and seeking optimal solutions to implement its regulations and enhance the operation of ALRO 4-01 land. The Mapping and Land Text Bureau, ALRO, inspects land utilization by visually interpreting high-resolution images to detect incompliant land use over numerous certified ALRO 4-01 plots. Consequently, manual inspection of incompliant land utilization requires a significant amount of time, is subjective, and is limited by human error. Additionally, high-resolution images are expensive and not always available for continuous monitoring. On the other hand, medium-resolution satellite images are more readily available and cover larger areas; however, traditional pixel-based classification methods applied to these images may not provide sufficient accuracy for detecting complex land-use changes. Recently, object-based image analysis (OBIA) has been introduced as a powerful approach that groups pixels into meaningful objects, improving classification accuracy by considering spatial and contextual information. Meanwhile, machine learning offers the potential for effective and efficient classification of remotely sensed imagery [4]. Many studies have generally found that these methods produce higher accuracy than traditional parametric classifiers, especially for complex data with a high-dimensional feature space.

Five machine learning algorithms (SVM, RF, DT, NB, and KNN) under the eCognition software and Sentinel 2A data were applied to classify land use types (urban and built-up areas, paddy fields, field crops, perennial trees and orchards, forest land, shrubland and riverine forest, water bodies, and marsh and swamp at Chaloemprakiat district, Nakhon Ratchasima province, Thailand [5]. The main objective of the study was to evaluate the most suitable machine learning algorithm for land use classification using Sentinel-2A and eCognition software, similar to the current study. This study is one of the automate cropland classifications using machine learning algorithms under OBIA in Thailand. The advantages and disadvantages of five machine learning algorithms, as reported by researchers [6][7][8] and [9], are summarized in Table 1.

Therefore, this study proposes a novel automated methodology based on object-based machine learning classification to detect incompliant land utilization in agricultural land reform areas. The specific objectives are (1) to identify an optimum machine learning algorithm for classifying land use and detecting incompliant land utilization based on the information of accuracy assessment and ALRO's rules from the modeling area, and (2) to validate an optimum machine learning algorithm for detecting incompliant land utilization in the test area.

2. Methods

2.1 Study Area

Two study areas are (1) Wang Nam Khiao District, Nakhon Ratchasima Province, as the modeling area covering an area of approximately 218 sq. km, (Figure 1(a)) and Pak Chong District, Nakhon Ratchasima Province, as the test area covering an area of approximately 109 sq. km (Figure 1(b)), are selected to classify land use type for detecting incompliant land utilization in the certified ALRO 4-01 plots. Both study areas were selected for this study because land utilization in the ALRO 4-01 plots is relatively high. 986 of 9,647 plots in 2012 and 578 of 10,499 plots in 2015 in Wang Nam Khiao District were observed as incompliant land utilization by visual interpretation. Likewise, 585 of 4,065 plots in 2013 and 485 of 5,795 plots in 2015 in Pak Chong District were observed to be incompliant land utilization [10]. According to land use data of the Land Development Department (LDD) in 2023, three dominant land use types in Wang Nam Khiao District were forest land, about 52.62%, agricultural land, about 34.18%, and urban and built-up area, about 6.45% while three dominant land use types in Pak Chong District were agricultural land, about 40.73%, forest land, about 35.45% and urban and built-up area, about 14.82% [11]. In terms of the environmental aspect, most of the forest land in both study areas is preserved as national parks, but illegal encroachment into forest land for agricultural use and for resorts and hotels is occurring in both areas. Meanwhile, tourist activity plays an essential role in both areas.

Table 1: Advantages and disadvantages of five machine learning algorithms

Algorithm	Advantages	Disadvantages	Reference
SVM	SVM is effective when the number of dimensions exceeds the number of samples. SVM works relatively well when there is a clear margin of separation between classes. SVM is more effective in high-dimensional spaces.	The SVM algorithm is not suitable for large data sets. SVM does not perform very well when the data set has more noise, i.e., the overlapping target classes. When the number of features per data point exceeds the number of training samples, the SVM will underperform.	[6]
RF	RF is less likely to overfit the data, which means it can generalize well to new data. RF can handle classification and regression problems and work well with categorical and continuous variables. RF is fast and can handle large datasets. RF can provide a measure of feature importance to support feature selection and data understanding.	RF is less prone to overfitting than a single decision tree. RF can be less interpretable than a single decision tree because it involves multiple trees. The training time of RF can be longer than other algorithms. RF requires more memory than other algorithms because it stores multiple trees.	[7]
DT	DT classifies unknown records very fast. In the presence of redundant attributes, DT works very well. DT is somewhat robust to noise when methods like overfitting are used.	In constructing a decision tree, it is not applicable to data that affects it badly. Minor changes in the data can alter the overall appearance of a decision tree. A subtree in a decision tree can be copied many times.	[8]
NB	NB is simple to implement. The conditional probabilities are easy to evaluate. NB is very fast – no iterations are needed, since the probabilities can be computed directly. NB is useful when training speed is important.	The conditional independence assumption does not always hold. In most situations, the feature shows some form of dependency. Zero probability problem: When we encounter words in the test data for a particular class not present in the training data, we might end up with zero class probabilities.	[8]
KNN	KNN is robust to noisy training data and effective when the training data is substantial. Simplicity, effectiveness, intuitiveness and competitive classification performance in many domains.	Distance-based learning is unclear about which distance to use and which attribute to use to achieve the best results. The computational cost is relatively high because it requires computing the distance between each query instance and all training samples.	[9]

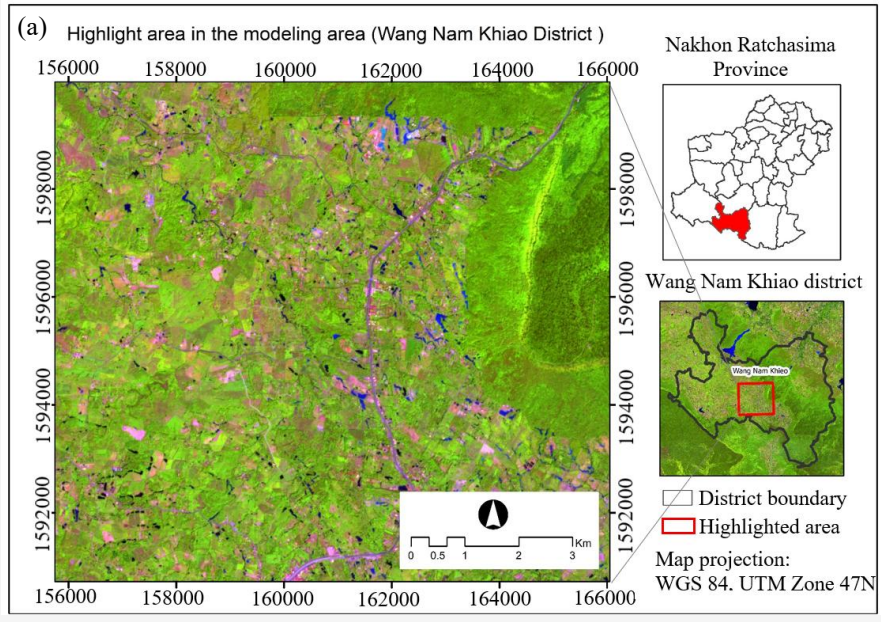


Figure 1: Study areas: (a) Modeling area (Wang Nam Khiao District) and (b) Test area (Pak Chong District) (Continue next page)

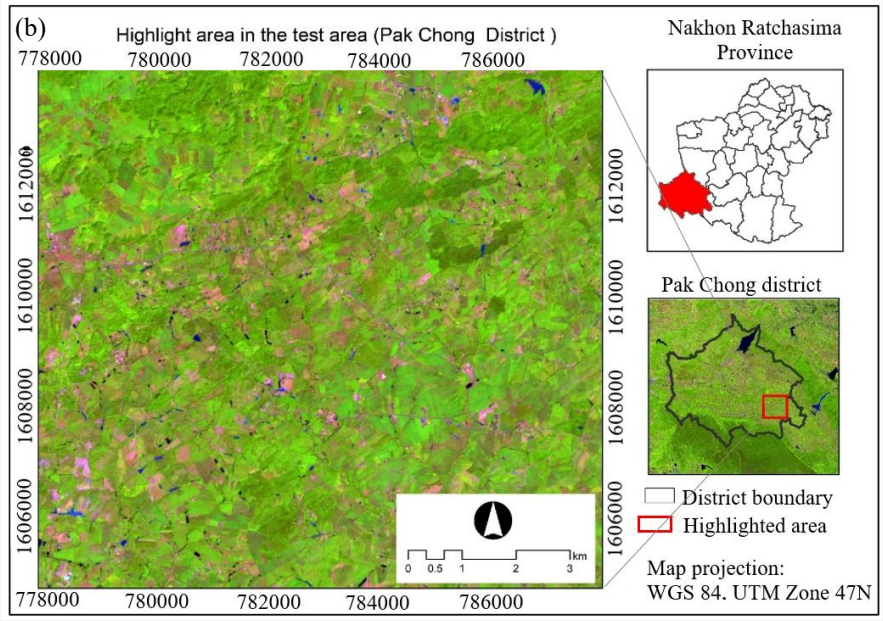


Figure 1: Study areas: (a) Modeling area (Wang Nam Khiao District) and (b) Test area (Pak Chong District) (Continue from previous page)

Table 2: Lists data collection and preparation for data analysis

Data type	Data collection	Coordinate system	Data Preparation	Reference
Remote sensing	Sentinel 2A in 2022: Acquisition date 18 October 2022,	WGS 1984 UTM Zone 47	Project transformation and spectral indices extraction	[12]
	Google Earth Image: Acquisition date 16 April 2022	WGS 1984 UTM Zone 47	Project transformation and mosaic	[13]
GIS	Registered ALRO 4-01 plots in Wang Nam Khiao district	WGS 1984 UTM Zone 47	Project transformation	[10]
	Registered ALRO 4-01 in Pak Chong district	WGS 1984 UTM Zone 47	Project transformation	[10]

2.2. Data

The collected and prepared data, which include remote sensing and GIS data, are summarized in Table 2.

2.3 Methodology

The optimal machine learning algorithm with OBIA for detecting incompliant land utilization in agricultural land reform areas is shown in Figure 2. A detailed of the methodology for detecting incompliant land utilization is provided in sections 2.3.1 to 2.3.8.

2.3.1 Study area extraction

The prepared Sentinel 2A imagery is used to extract the two study areas: Wang Nam Khiao District for the modeling area and Pak Chong District for the test area. In practice, two extended rectangular boxes, about 2 km outside the boundary of each district, were first prepared and used to define modeling and test areas.

2.3.2 Spectral indices calculation

Sentinel 2A data covering the modeling and test areas is used to calculate three spectral indices. First, the normalized difference vegetation index (NDVI), which represents vegetation feature [14], was calculated using Equation 1. The NDVI is a vital vegetation index because seasonal and inter-annual changes in vegetation growth and activity can be monitored. The ratioing can reduce many forms of multiplicative noise present in multiple bands [15]. The NDVI is easier to calculate than the Soil Adjusted Vegetation Index (SAVI), which requires a canopy background adjustment factor that accounts for differential red and near-infrared extinction within the canopy. The value of L depends on the proportional vegetation cover as well as the vegetation density [16].

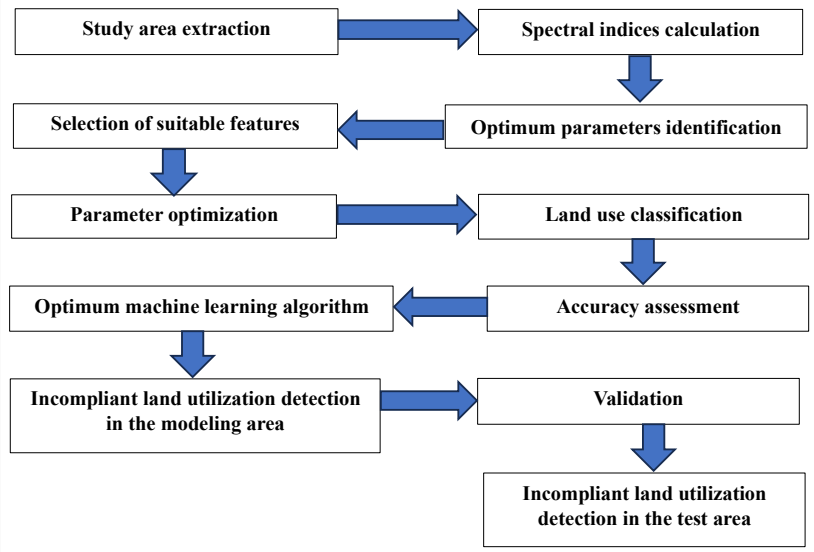


Figure 2: Incompliant land utilization detection study workflow

Second, the normalized difference built-up index (NDBI) was calculated to enhance the detection of built-up areas [17] using Equation 2. Third, the modified normalized wetness index (MNDWI), which indicates the moisture regime [18], was calculated using Equation 3.

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

Equation 1

Where *NIR* is the near infrared band, Sentinel 2A band 8, and RED is the red band, Sentinel 2A band 4.

$$NDBI = \frac{SWIR1 - NIR}{SWIR1 + NIR}$$

Equation 2

Where *SWR1* is the short-wave infrared 1 band, Sentinel 2A band 11, and NIR is the near infrared band, Sentinel 2A band 8.

$$MNDWI = \frac{GREEN - SWIR1}{GREEN + SWIR1}$$

Equation 3

Where *GREEN* is green band, Sentinel 2A band 3, and *SWR1* is short-wave infrared 1 band, Sentinel 2A band 11. After that, multispectral bands and spectral indices were combined using layer stacking for data analysis.

2.3.3 Optimum parameter identification

The systematic optimum parameter identification for multiresolution segmentation under eCognition software, including scale, shape, and compactness, was implemented in three steps based on the visual interpretation of the segmented objects as follows.

Step 1: The scale parameters will vary with values of 5, 10, 15, 20, and 25, while the weight of shape and compactness values are fixed at 0.5 and 0.5, respectively. The output in this step is the optimum scale parameter.

Step 2: The identified optimum parameter of scale was fixed, while the weights of shape varied with values of 0.1, 0.2, 0.3, 0.4, and 0.5, and the weight of compactness was fixed at 0.5. The expected output in this step is the optimum weight of the shape parameter.

Step 3: The identified optimum parameter scale and optimum weight of shape were fixed, while the weights of compactness varied with values of 0.1, 0.2, 0.3, 0.4, and 0.5. The expected output in this step is the optimum weight of the compactness parameter.

2.3.4 Selection of suitable features

Feature selection, a critical step in object-based image analysis (OBIA), was assessed using separability measurements between land use classes using the Jeffries–Matsushita (J) distance [19], with a domain value ranging from 0 to 2 (Equation 4). Higher separability distance indicates better discrimination between classes.

In this study, all pairwise J distance based on object properties (spectral values, geometric shape, and texture) among training samples of seven land use types, including (a) building areas, (b) paddy fields,

(c) field crops, (d) perennial trees and orchards, (e) forest areas, (f) water bodies and (g) rangeland were first calculated and then selected as suitable features for land use classification with the threshold value of J value equal or more than 1.75 [20].

$$J = 2(1 - e)^B$$

Equation 4

Where J is Jeffries–Matsushita distance, B is Bhattacharyya distance [21] defined in Equation 5:

$$B = \frac{1}{8}(m_1 - m_2)^2 \frac{2}{\sigma_1^2 + \sigma_2^2} + \frac{1}{2}\ln\left[\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2}\right]$$

Equation 5

Where m_1 is the mean value of class 1, m_2 is the mean value of class 2, σ_1 is the variance of class 1, and σ_2 is the variance of class 2, for the two feature distributions. If the means coincide, the first term vanishes, whereas the second term vanishes if the two feature distributions have equal variances [22].

2.3.5 Parameter optimization

The significant parameters of support vector machine (SVM), random forests (RF), and K-nearest neighbor (KNN) under the machine learning algorithms of eCognition software were optimized using Grid Search with cross-validation by Python programming before land use classification, as summary in Table 3. In the meantime, the default settings for decision trees (DT) and Naïve Bayes (NB) in the machine learning algorithm of eCognition software are used to classify land use. The optimal parameters were selected based on the training sample, achieving an overall accuracy of more than 0.80 (80%).

2.3.6 Land use classification

Land use data for the modeling areas (Wang Nam Khiao district) in 2023 were separately classified based on the suitable features from Step 2.3.4 using SVM, RF, and KNN with the optimized parameters from Step 2.3.5, as well as DT and NB with default settings. In this study, about 50 image objects of each land use are visually selected as training areas. In practice, we first visually select training areas for each land use type. Afterward, we select a classifier with optimized parameters one at a time to classify land use types based on the identified suitable features.

2.3.7 Accuracy assessment

All classified land use maps from Step 2.3.6 were assessed for thematic accuracy (overall accuracy, OA; average producer's accuracy, PA; average user's accuracy, UA; and Kappa coefficient, K) using field survey reference information from 2024 and supporting high-resolution images from Google Earth Pro from 2022. Herein, the number of random stratified sampling points was calculated using the multinomial distribution theory [23] as defined in Equation 6:

$$N = \frac{B}{4b^2}$$

Equation 6

Where b is the desired precision (e.g., 5%) for all classes and B is the upper (α/k) x 100 percentiles of the Chi-square distribution with 1 degree of freedom. In addition, the pairwise Z test in Equation 7 was applied to examine the significant difference of the derived Kappa coefficient values among the five classified land use maps based on their confusion matrices [23]:

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

Equation 7

and variance of the Kappa coefficient ($var(K)$) is calculated by Equation 8:

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

Equation 8

Table 3: Selected parameters and testing values of SVM, RF and KNN algorithm under parameter optimization

Algorithm	Parameters	Testing values under optimization
SVM with linear kernel	C value	0.01, 0.03, 0.05, 0.07, 0.09, 0.1, 0.3, 0.5, 0.7, 0.9, 1.0, 3.0, 5.0, 7.0, 9.0,10,100,1000
	Gamma value	10, 100, 1000
SVM with non-linear kernel (Radial basis function, RBF)	C value	0.01, 0.03, 0.05, 0.07, 0.09, 0.1, 0.3, 0.5, 0.7, 0.9, 1.0, 3.0, 5.0, 7.0, 9.0,10,100,1000
	Gamma value	0.1, 1, 10, 100, 1000
RF	Number of decision trees	50, 100, 150, 200
KNN	K distance	1-40

The parameters used in Equations 8 are expressed in Equations 9 to 12:

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

Equation 9

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

Equation 10

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

Equation 11

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

Equation 12

Where Z is standardized and normally distributed, k is the number of rows (land use classes) in the error matrix, n_{ii} is the number of observations in row i and column i, n_{i+} is the marginal total for row i, n_{+i} is the marginal total for column i, and n is the total number of observations. Herein, given the null hypothesis H_0: $(K_1 - K_2) = 0$, and the alternative H_1: $(K_1 - K_2) \neq 0$, H_0 is rejected if $Z \geq Z_{\alpha/2}$ [23].

2.3.8 Optimum machine learning algorithm

The optimum machine learning algorithm for detecting incompliant land utilization in agricultural land reform areas was justified based on the accuracy assessment results from Step 2.3.7. Here, the derived overall accuracy and Kappa coefficient values for five land use classification maps, using five machine learning algorithms, were compared to identify an optimum algorithm for land use classification to detect incompliant land utilization in each ALRO 4-01 plot.

2.3.9 Incompliant land utilization detection in the modeling area

The classified land use map from the optimum algorithm was used to detect incompliant land utilization in each ALRO 4-01 plot in the modeling area through overlay analysis using ESRI ArcMap software, in accordance with ALRO regulations. According to the ALRO regulation, if the percentage of building area in each plot is 10 percent or more, the land utilization in that plot is considered incompliant. Similarly, if the percentage of water body area in each plot is equal to or greater than 5 percent, the land utilization in that plot is considered incompliant land utilization. Furthermore, the extracted compliant and incompliant land utilization in each ALRO 4-01 plot was assessed in accuracy using binary change detection (Sensitivity, Specificity, Predicted Positive, Predicted Negative and Prevalence (θ) based on a binary error matrix (Table 4) [24] as defined in Equation 13:

$$Sensitivity = \frac{a}{a+c} = \frac{a}{e}$$

Equation 13

The *Sensitivity* is the proportion of cases correctly classified as having changed, as the producer's accuracy for the change class (Incompliant) [25].

$$Specificity = \frac{d}{b+d} = \frac{d}{f}$$

Equation 14

The *Specificity* in Equation 14 is the proportion of cases correctly predicted to have not changed, as producer's accuracy for the no-change class (Compliant) [25].

$$PPV = \frac{a}{a+b} = \frac{a}{g}$$

Equation 15

The *Positive predictive value (PPV)* in Equation 15 is used for the change detection map represents the user's accuracy for the change class (Incompliant) [26].

$$NPV = \frac{d}{c+d} = \frac{d}{h}$$

Equation 16

The *Negative predictive value (NPV)* in Equation 16 is used for the change detection map represents the user's accuracy for the no-change class (Compliant) [26].

$$\theta = \frac{a+c}{a+b+c+d} = \frac{e}{N}$$

Equation 17

The *Prevalence (θ)* in Equation 17 represents the probability that the ground information indicates change at a randomly selected location [24].

2.3.10 Validation

The selected optimal algorithm from Step 2.3.8 was applied to classify land use in the test area (Pak Chong District) and to assess its accuracy, as in the modeling area. The test area results were compared with those from the modeling area to evaluate the algorithm's consistency and accuracy in detecting incompliant land utilization under similar land reform conditions.

Table 4: A binary error matrix for binary change detection

		Ground reference by field survey and Google Earth image		Row total
		Incompliant plots (Change class)	Compliant plots (No change class)	
Incompliant and compliant plots based on land use classification	Incompliant plots (Change class)	a	b	g
	Compliant plots (No change class)	c	d	h
	Column total	e	f	N

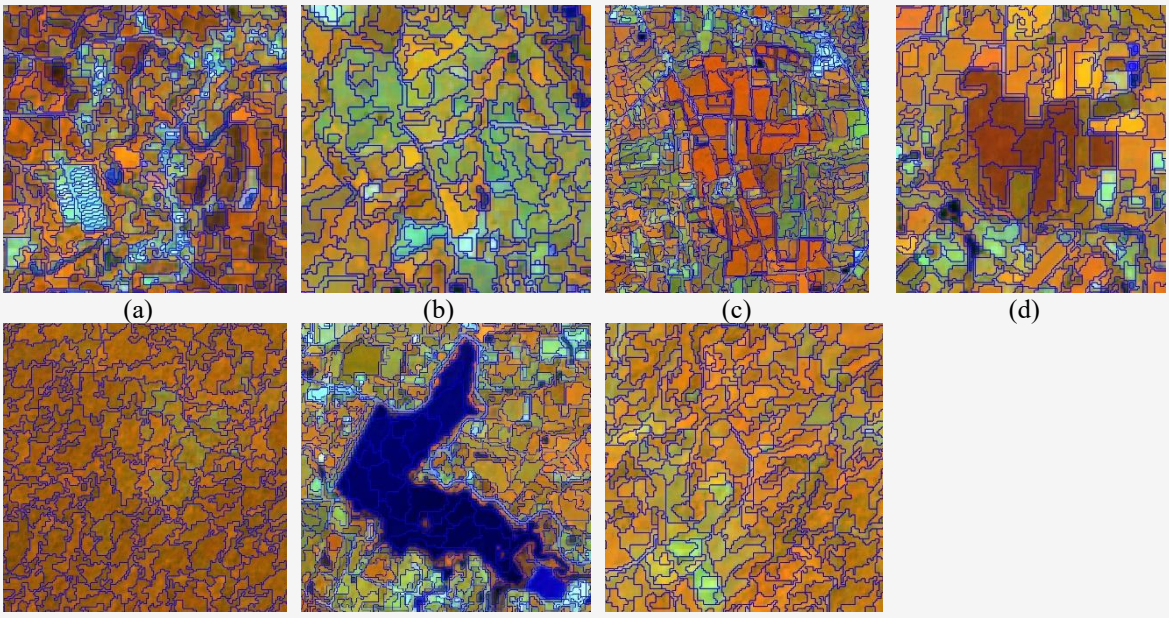


Figure 3: Image objects of land use types with optimum parameter for multiresolution segmentation: (a) building areas, (b) paddy fields, (c) field crops, (d) perennial trees and orchards, (e) forest areas, (f) water bodies and (g) rangeland

2.3.11 Incompliant land utilization detection in the test area

As in Step 2.3.9, the classified land use from Step 2.3.10 for the test area was used to detect incompliant land utilization in each ALRO 4-01 plot in accordance with ALRO regulations. Besides, the extracted compliant and incompliant land utilization in each ALRO 4-01 plot was assessed for accuracy using binary change detection.

3. Results and Discussion

3.1 Optimum Parameter Identification for Multiresolution Segmentation

The optimum parameters for multiresolution segmentation, based on scale, shape weight, and compactness weight, using Sentinel data and spectral indices under eCognition software, were 25, 0.5, and 0.5, respectively. These values were selected based on a visual interpretation of segmentation quality, as automated selection of these parameters is not available and relies on expert judgment. Figure 3 shows an example of image objects representing land use types, obtained using optimal parameters for multiresolution segmentation in eCognition software in Wang Nam Khiao District. The optimum parameter for multiresolution segmentation in the current study differs from that in previous studies [27]. They visually identified the optimal parameter for multiresolution segmentation using THEOS-1 images at 15 m resolution to detect forest changes in the Galyani Vadhana district, Chiang Mai Province. Their optimal scale parameter was 10, while the shape and compactness weights were 0.7 and 0.5, respectively. The possible cause of this phenomenon may be differences in image resolution.

3.2 Suitable Feature Selection for Land Use Classification under Object-Based Image Analysis

The results of suitable feature selection, including spectral and texture properties, for land use classification under OBIA using the J-M distance among training samples of land use types, were reported in Table 5.

Table 5: The suitable spectral and texture features for land use classification under OBIA

No	Object property	Suitable feature	Jeffries–Matushita distance
1	Spectral	Max. diff	2.00
2	features	Brightness	1.85
3		Mean of BLUE	1.99
4		Mean of GREEN	1.90
5		Mean of RED	1.97
6		Mean of NIR	2.00
7		Mean of SWIR1	2.00
8		Mean of NDVI	2.00
9		Mean of NDBI	1.98
10		Mean of MNDWI	1.99
11		Ratio of BLUE	1.99
12		Ratio of GREEN	2.00
13		Ratio of RED	1.99
14		Ratio of NIR	2.00
15		Ratio of SWIR1	1.97
16		Ratio of NDVI	1.99
17		Ratio of MNDWI	2.00
18		Standard deviation of BLUE	1.82
19		Standard deviation of GREEN	1.83
20		Standard deviation of RED	1.97
21		Standard deviation of MNDWI	1.91
22	Texture features	GLCM Mean of NDVI	1.96
23		GLCM Homogeneity of SWIR-1	1.89
24		GLCM Homogeneity of NDVI	1.81
25		GLCM Homogeneity of NDBI	1.96
26		GLCM Homogeneity of MNDWI	1.95
27		GLCM Entropy of GREEN	1.85
28		GLCM Entropy SWIR1	1.78
29		GLCM Entropy NDBI	1.94
30		GLCM Dissimilarity GREEN	1.82
31		GLCM Dissimilarity NDVI	1.95
32		GLCM Dissimilarity SWIR-1	1.84
33		GLCM Contrast of GREEN	1.79
34		GLCM Correlation NIR	1.83

Table 6: Optimum parameters of each algorithm under parameter optimization

Classifier	Parameters	Optimum value	Overall accuracy
SVM with linear kernel	C value	0.09	0.86
	Gamma value	100.00	
SVM with non-linear kernel (RBF)	C value	0.60	0.14
	Gamma value	100.00	
RF	Number of decision trees	100.00	0.88
KNN	K distance	1.00	0.81

As a result, 34 features from two object properties, spectral and texture, were identified as suitable features for classifying land use data. The highest J Values were Max diff, Mean of NIR, Mean of SWIR1, Mean of NDVI, Ratio of GREEN, Ratio of NIR, and Ratio of MNDWI value of 2.0, and the lowest J value was GLCM Entropy of SWIR1 value of 1.78. The features with J-M distances greater than 1.75 were selected as suitable for land use classification. Among these, Max. diff, several mean and ratio bands, and texture features such as GLCM Dissimilarity and Entropy performed well, with values close to or equal to 2.00, indicating excellent separability. In contrast, shape-based features such as area, border length, width, asymmetry, and shape index were excluded because their J values were below 1.75, indicating limited discriminatory power for land use classification in this context.

3.3 Parameter Optimization of Machine Learning Algorithms

The optimized parameter results for each algorithm (SVM with linear and RBF (non-linear) kernels, RF, and KNN) based on training samples with the accepted overall accuracy are summarized in Table 6. The overall accuracy values for SVM with a linear kernel, RF, and KNN are all above 80%. On the contrary, the overall accuracy of the SVM with a

non-linear kernel (RBF) is very low at 14%. The superiority of the linear kernel stems from the fact that object features create a space in which classes are already near-linearly separable. In this specific feature space, the complex RBF kernel introduces unnecessary complexity and risks overfitting, resulting in performance that is marginally worse than the simpler linear kernel. Future RBF improvement would require integrating features with higher spatial complexity (e.g., texture) to warrant a non-linear approach.

3.4 Land use Classification and Accuracy Assessment in the Modeling Area

The results of classified land use maps, using SVM with a linear kernel, RF, DT, NB, and KNN, based on the selected training areas of seven land use types as image objects (Figure 4), are displayed in Figure 5. Meanwhile, the results of the accuracy comparison are reported in Table 7. Furthermore, the results of the pairwise Z-tests for the Kappa coefficient and its variance across the five algorithms are reported in Table 8.

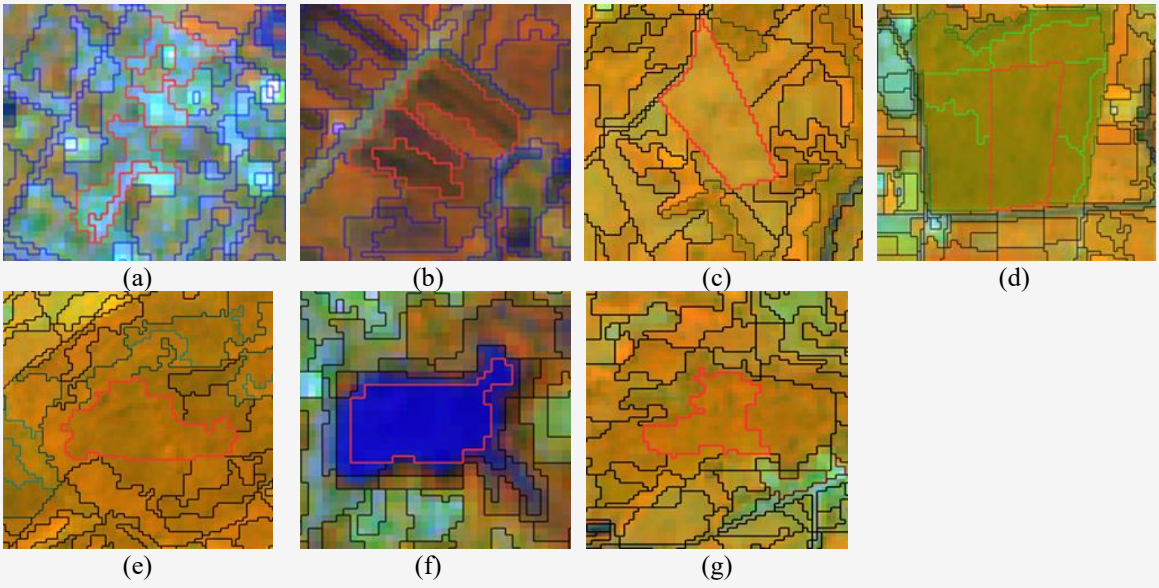


Figure 4: Example of training area of image objects for each land use type for land use classification: (a) building areas, (b) paddy fields, (c) field crops, (d) perennial trees and orchards, (e) forest areas, (f) water bodies and (g) rangeland

Table 7: Accuracy comparison of five machine learning algorithms

Algorithm	Overall accuracy (%)	Average producer' accuracy	Average user' accuracy	Kappa coefficient (%)
Random forest	87.45	83.08	79.89	79.57
SVM with linear kernel	86.47	76.41	75.15	75.52
Decision tree	83.12	76.93	75.69	73.19
Naïve Bayes	84.80	83.17	75.00	76.85
K-nearest neighbor	77.27	72.78	68.31	64.00

Table 8: Pairwise Z test of the Kappa coefficient and its variance between RF and other algorithms

Pairwise	Var(K)	Variance	Z-Statistic	Confidential level of critical value			
				80%	90%	95%	100%
RF	0.796	0.0004	0.981	1.28	1.65	1.96	2.58
NB	0.768	0.0004					
RF	0.796	0.0004	1.373	1.28*	1.65	1.96	2.58
SVM	0.755	0.0005					
RF	0.796	0.0004	2.205	1.28*	1.65*	1.96*	2.58
DT	0.732	0.0005					
RF	0.796	0.0004	5.130	1.28*	1.65*	1.96*	2.58*
KNN	0.640	0.0005					

*Note: * There is a significant difference between the Kappa coefficients*

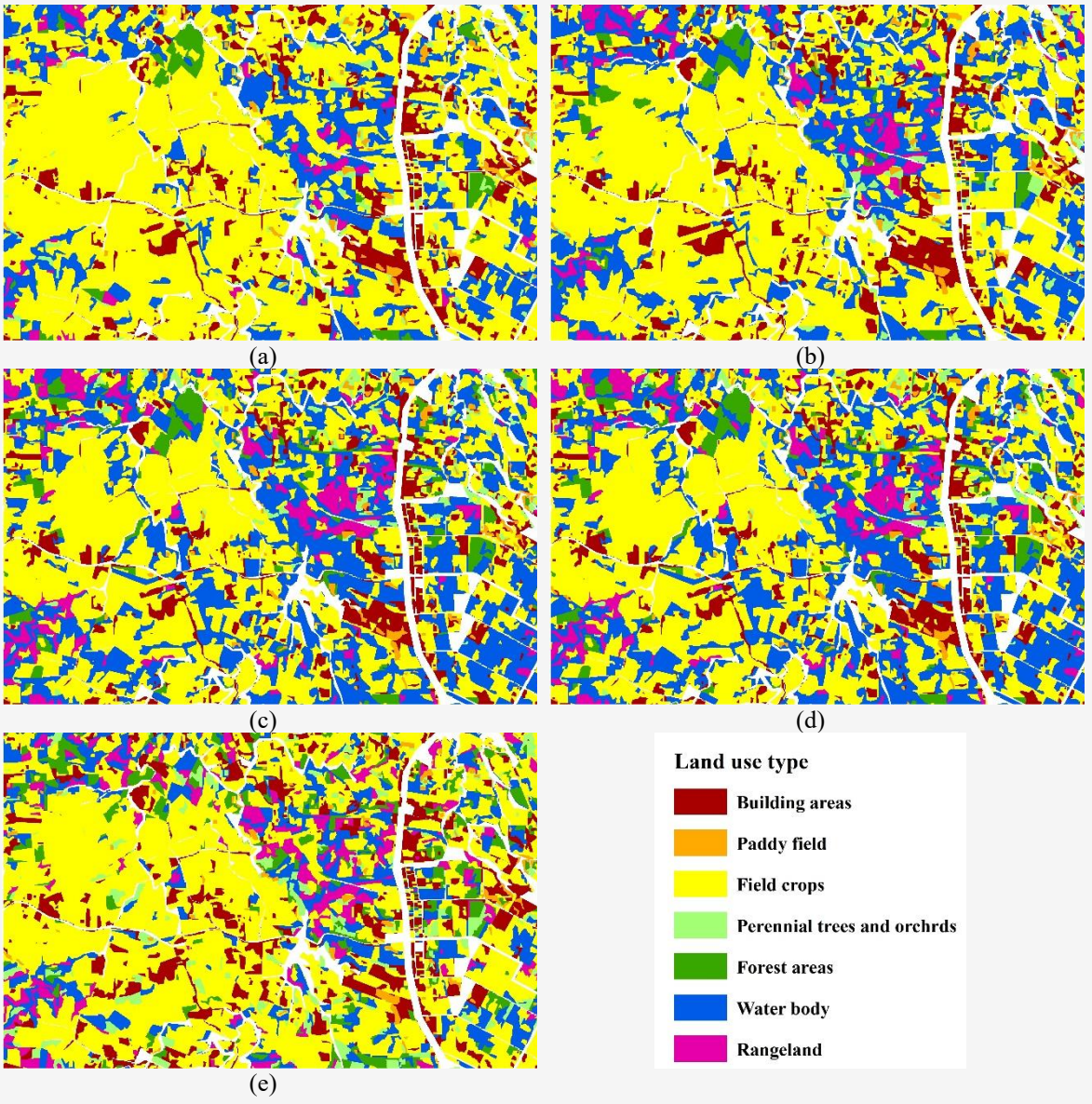


Figure 5: Classified land use data in the highlight area of Wang Nam Khiao District: (a) SVM with linear kernel, (b) RF, (c) DT, (d) NB, and (e) KNN

The pairwise Z-test is used to compare Kappa coefficients across classifiers and determine whether differences in classification accuracy are statistically significant. The result helps identify whether one algorithm performs better than another beyond what might be expected by chance. The Z statistics show that RF significantly outperforms DT and KNN at the 95% confidence level. As a result, the most suitable algorithm for high-performance land use classification in the modeling area is RF. The overall accuracy of land use maps from RF, with more than 80%, is acceptable [28]. The Kappa coefficient, which is close to 80%, represents strong accuracy between the classified and reference data [29] and [30]. Besides, the derived overall accuracy and Kappa coefficient of land use data by the RF in the current study are consistent with the previous study of Pinyoyang, who applied the RF in the EnMAP-Box software to classify urban and built-up areas, paddy fields, sugarcane, cassava, other field crops, para-rubber, perennial trees and orchards, forest land, waterbodies, rangeland, marshes and swamps, and unused land in 2019 at the Second Part of Lam Nam Chi watershed, Chaiyaphum, Thailand, from Landsat 8-OLI image. The OA and Kappa coefficient values for the land use data in 2019 were 91.37% and 88.26%, respectively [31]. Likewise, Srichaichana applied the RF in the EnMAP-Box software to

classify urban and built-up area, paddy field, rubber plantation, oil palm plantation, perennial tree and orchard, aquatic culture area, evergreen forest, mangrove forest, marsh and swamp, water body, and miscellaneous land (bare land and abandoned mine) in 2017 at Klong U-Tapao Watershed, Songkhla, Thailand from Landsat 8-OLI image. The OA and Kappa coefficient values for the land use classification were 94.32% and 87.00%, respectively [32].

Moreover, if the spatial patterns of the land use map of RF are compared with other maps of NB, SVM, DT, and KNN using spatial correlation analysis, as reported in Table 9. The correlation coefficient values between the land use map of the RF map and others show a strongly positive linear relationship, with values ranging from 0.672 to 0.899 [33]. A high correlation coefficient value suggests substantial redundancy in the information content among these land use maps [15]. This finding implies a similarity in land use patterns, as classified by these different algorithms.

3.5 Incompliant Land Utilization Detection in ALRO 4-01 Plots in the Modeling Area

The result of incompliant land utilization detection in each ALRO 4-01 plot using overlay analysis between the ALRO 4-01 plots and the percentage of building area of equal or more than 10% or water body of equal or more than 5%, based on the classified land use data of the RF algorithm over the highlight area of the modelling area is shown in Figure 6. The number of plots and the percentage of land utilization detection, categorized as either incompliant or compliant, in ALRO 4-01 plots, based on two criteria in the modeling area, are reported in Table 10. The compliance and incompliance of land utilization in each ALRO 4-01 plot, according to building and water body criteria, were verified through a ground survey and high-resolution images, as shown in Figure 7. At the same time, the results of both verified compliance and incompliance of land utilization data are applied to detect binary changes and describe the sensitivity, specificity, predicted positive, predicted negative and prevalence values, as summarized in Tables 11 and 12. The result of binary change detection based on building and water body criteria in Wang Nam Khiao District is compared in Figure 8. The sensitivity value, as producer's accuracy for the change class (Incompliant), according to building and water body criteria, was 95.63% and 90.36%, respectively. Meanwhile, the specificity, as producer's accuracy for no-change class (Compliant), according to building and water body criteria, was 82.02% and 96.38%, respectively. The prevalence value, representing the probability that ground information indicates change at a randomly selected location based on building and water body criteria, was 5.11% for buildings and 2.06% for water bodies.

Table 9: Correlation matrix and correlation coefficient of land use maps from five algorithms

Algorithm	RF	BAY	SVM	DT	KNN
RF	1.000	0.756	0.789	0.899	0.754
NB	0.756	1.000	0.713	0.729	0.672
SVM	0.789	0.713	1.000	0.743	0.735
DT	0.899	0.729	0.743	1.000	0.750
KNN	0.754	0.672	0.735	0.750	1.000

(a) (b)

Figure 6: Incompliant land utilization detection based on land use classification in the highlight area of Wang Nam Khiao District, based on two criteria: (a) Percentage of building area, (b) Percentage of water body area

(a) (b)

Figure 7: Verified incompliant land utilization detection based on field survey and Google image in the highlight area of Wang Nam Khiao District, based on two criteria:
(a) Percentage of building area, (b) Percentage of water body area

Table 10: Number and percentage of incompliant and compliant land utilization detection in ALRO 4-01 plots based on two criteria in the modeling area

Criterion	Incompliant utilization		Compliant land utilization		Total
	Number of plots	Percent	Number of plots	Percent	
Building area (= > 10 percent)	1,770	21.94	6,297	78.06	8,067
Water body area (= > 10 percent)	436	5.40	7,631	94.60	8,067

Table 11: Binary error matrix and result of binary change detection based on building criterion in Wang Nam Khiao District

Binary error matrix	Reference to incompliant land utilization (plots)	Reference to compliant land utilization (plots)	Total
Extracted incompliant land utilization (plots)	394	1,376	1,770
Extracted compliant land utilization (plots)	18	6,279	6,297
Total	**412**	**7,655**	**8,067**
Sensitivity	95.63%		
Specificity	82.02%		
Positive predictive value (PPV)	22.26%		
Negative predictive value (NPV)	99.71%		
Prevalence	5.11%		

Table 12: Binary error matrix and result of binary change detection based on water body criterion in Wang Nam Khiao District

Binary error matrix	Reference to incompliant land utilization (plots)	Reference to compliant land utilization (plots)	Total
Extracted incompliant land utilization (plots)	150	286	436
Extracted compliant land utilization (plots)	16	7,615	7,631
Total	**166**	**7,901**	**8,067**
Sensitivity	90.36%		
Specificity	96.38%		
Positive predictive value (PPV)	34.40%		
Negative predictive value (NPV)	99.79%		
Prevalence	2.06%		

As a result, binary change detection in the modeling area, based on building and water body criteria, performs well, particularly in terms of sensitivity, meaning that most incompliant land plots are correctly identified. However, the positive predictive value (PPV) for detecting incompliant land utilization in the ALRO plot based on building criteria is lower than that based on water body criteria, due to the misclassification of post-harvest agricultural fields as "Building areas." The single-date imagery (October acquisition) showed extensive bare soil and dry crop residues in these fields. Critically, the spectral signature of this bare soil closely mimics that of urban and other non-vegetated surfaces, leading the model to confuse compliant bare land with non-compliant developed areas.

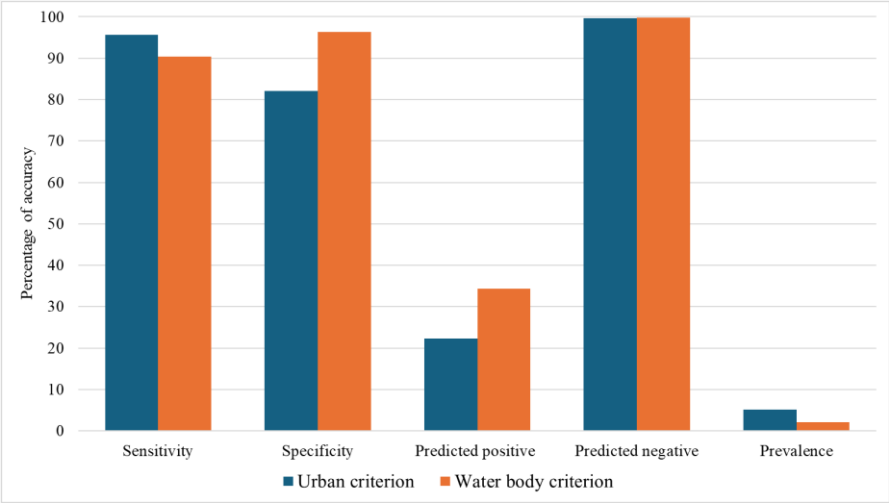


Figure 8: Comparison of results of binary change detection based on building and water body criteria in Wang Nam Khiao District

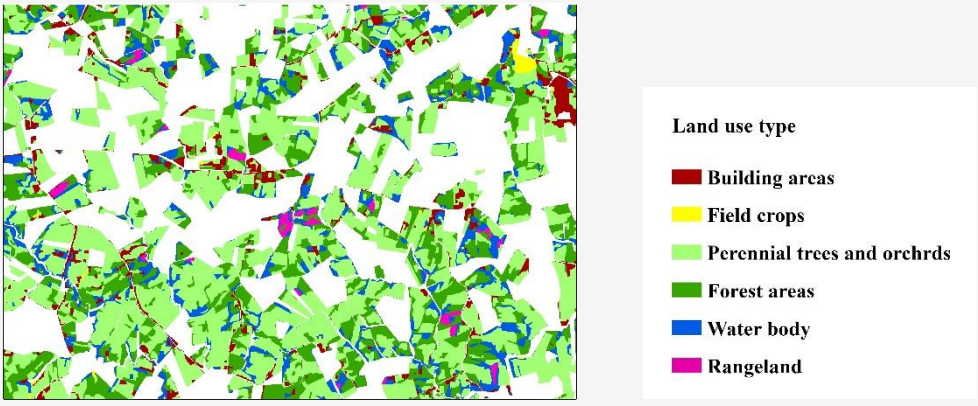


Figure 9: Classified land type data by random forests in the highlight area of Pak Chong district (test area)

3.6 Validation of an Optimum Algorithm for Detecting Incompliant Land Utilization in the Test Area

3.6.1 Land use classification and accuracy assessment

The classified land use data results for Pak Chong District, using the RF classifier as the optimal algorithm in Wang Nam Khiao, are displayed in Figure 9. It should be noted that only six land use classes, including building areas, field crops, perennial trees and orchards, forest areas, water bodies, and rangeland, were covered in the test area. As a result, the overall accuracy and Kappa coefficient of the land use map were 85.21% and 81.00%, respectively. The average PA and UA were 83.67% and 88.71%, respectively. As a result, the overall accuracy and Kappa coefficient of the classified land use map in the test area are acceptable [28][29] and [30]. Still, the PA values for building areas and water body areas were 66.67% and 85.71%, respectively, while the UA values were 96.77% and 100%, respectively. These results may be due to the spectral similarity between building areas and other land uses, such as deciduous forests or bare soil, especially in scattered building areas. The low PA value suggests that a substantial number of actual building areas were misclassified or overlooked, potentially affecting the reliability of detecting incompliant land utilization. This limitation highlights the need for higher-resolution imagery, more representative training samples, or additional object-based features (e.g., texture or shape) to improve classification performance for built-up areas in future applications.

Additionally, the results of the pairwise Z test comparing the Kappa coefficient of the land use classification with the RF algorithm in the modeling and test areas are reported in Table 13. As a result, the Kappa coefficient value of the classified land use in the test area is insignificantly different from the

Kappa coefficient value of the classified land use map in the modeling area at different confidential levels. Accordingly, the result of the classified land use map in the test area, using optimal object features and an RF algorithm, can be validated. This finding suggests that the optimal object features and algorithm in the modeling area (Wang Nam Khiao district) can be directly transferred (transferable) for land use classification in the test area (Pak Chong district).

3.6.2 Incompliant land utilization detection in ALRO 4-01 plots in the test area

The result of incompliant land utilization in each ALRO 4-01 plot, as determined by overlay analysis of all plots according to building and water body criteria, is shown in Figure 10 and reported in Table 14. The verified incompliant and compliant land utilization plots in each ALRO are displayed in Figure 11. Additionally, the results of binary change detection for both incompliant and compliant data, as determined through visual interpretation using very high-resolution images, are reported in Tables 15-16. The result of binary change detection based on building and water body criteria in Pak Chong District is compared in Figure 12. According to building and water body criteria, the sensitivity values, as producer's accuracy for change class (Incompliant), were 98.11% and 94.00%. Meanwhile, the specificity values, as producer's accuracy for no-change class (Compliant), were 76.78% and 98.49%, respectively. The prevalence value according to building and water body criteria was 9.21% and 1.24%, respectively.

Table 13: Pairwise Z test of the Kappa coefficient and its variance in the modeling and test areas

Pairwise	Var(K)	Variance	Z-Statistic	Confidential level of critical value			
				80%	90%	95%	100%
Modeling area	0.796	0.0004	0.605	1.28	1.65	1.96	2.58
Test area	0.801	0.0002					

Table 14: Number and percentage of incompliant and compliant land utilization detection in ALRO plots based on two criteria in the test area

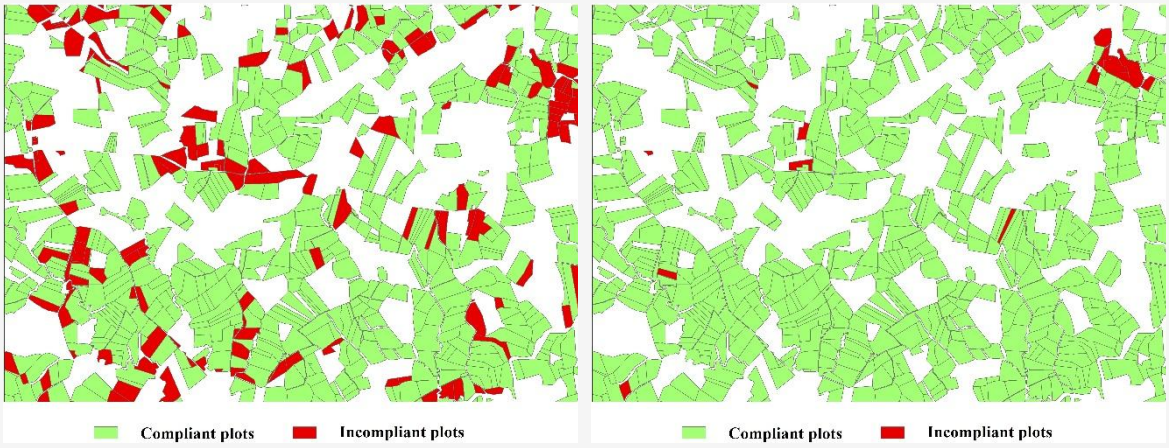
Criterion	Incompliant utilization		Compliant land utilization		Total
	Number of plots	Percent	Number of plots	Percent	
Building area (= > 10 percent)	1,213	30.1	2,814	69.87	4,027
Water body area (= > 10 percent)	107	2.65	3,920	97.34	4,027

Table 15: Binary error matrix and result of binary change detection based on building criterion in Pak Chong District

Binary error matrix	Reference to incompliant land utilization (plots)	Reference to compliant land utilization (plots)	Total
Extracted incompliant land utilization (plots)	364	849	1,213
Extracted compliant land utilization (plots)	7	2,807	2,814
Total	**371**	**3,656**	**4,027**
Sensitivity	98.11%		
Specificity	76.78%		
Positive predictive value (PPV)	33.42%		
Negative predictive value (NPV)	99.68%		
Prevalence	9.21%		

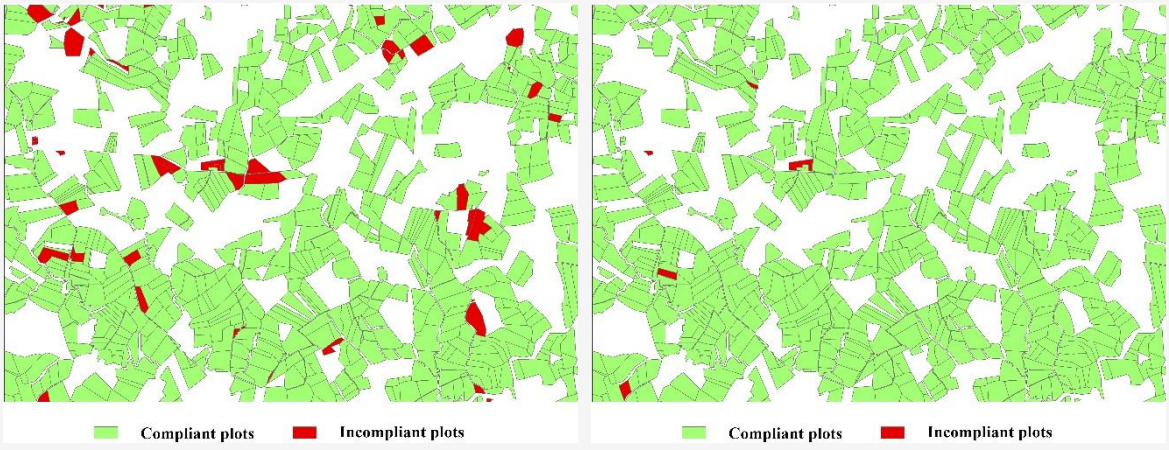
Table 16: Binary error matrix and result of binary change detection based on water body criterion in Pak Chong District

Binary error matrix	Reference to incompliant land utilization (plots)	Reference to compliant land utilization (plots)	Total
Extracted incompliant land utilization (plots)	47	60	107
Extracted compliant land utilization (plots)	3	3,917	3,920
Total	**50**	**3,977**	**4,027**
Sensitivity	94.00%		
Specificity	98.49%		
Positive predictive value (PPV)	43.93%		
Negative predictive value (NPV)	99.92%		
Prevalence	1.24%		



(a) (b)

Figure 10: Incompliant land utilization detection based on land use classification in the highlight area of Pak Chong District, based on two criteria: (a) Percentage of building area, (b) Percentage of water body area



(a) (b)

Figure 11: Verified incompliant land utilization detection based on field survey and Google image in the highlight area of Pak Chong District, based on two criteria:
(a) Percentage of building area, (b) Percentage of water body area

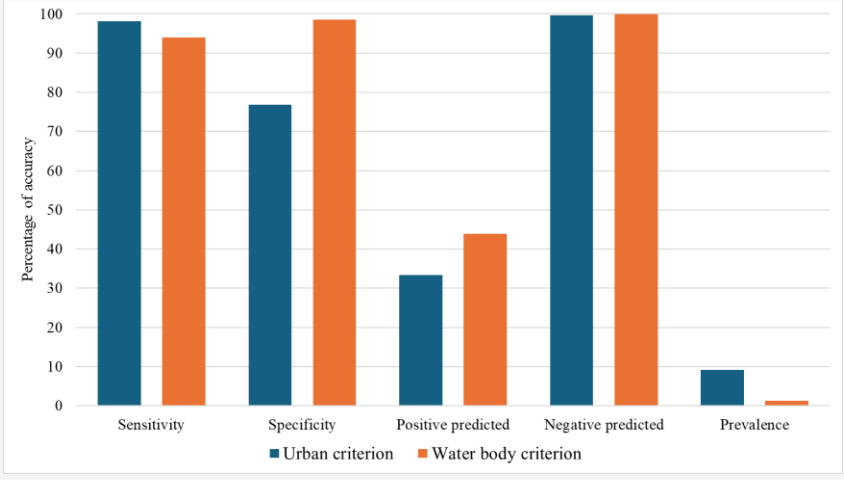


Figure 12: Comparison of results of binary change detection based on building and water body criteria in Pak Chong District

As a result, the binary error matrices indicate that the RF-based detection method performs well, particularly in terms of sensitivity, meaning most incompliant land plots are correctly identified. However, the relatively low positive predictive values (PPVs) suggest a high rate of false positives, especially in building area detection. Moreover, while visual interpretation using very high-resolution imagery ensures accuracy, it can be labor-intensive and may not scale to large areas. Nevertheless, the approach is promising for targeted monitoring in priority zones or areas with known incompliance issues.

4. Conclusions

In this paper, we have proposed a novel automated methodology based on object-based machine learning classification to detect incompliant land utilization in agricultural land reform areas. The specific objectives of the study were (1) to identify an optimum machine learning algorithm for classifying land use and detecting incompliant land utilization based on the information of accuracy assessment and ALRO's rules from the modeling area (Wang Nam Khiao District), and (2) to validate an optimum machine learning algorithm for detecting incompliant land utilization in the test area (Pak Chong District). In the study, five object-based machine learning algorithms: support vector machines (SVM), random forests (RF), decision trees (DT), Naïve Bayes (NB), and K nearest neighbor (KNN) with optimized parameters were first applied to classify land use type based on the single date of Sentinel 2A, acquired in October and assessed accuracy based on very-high resolution image of Google Earth Pro and field survey for identifying the optimum algorithm for detecting incompliant land utilization in agricultural land reform areas in the modeling area. The identified optimal machine learning algorithm for detecting incompliant land use was then validated in the test area. As a result, the most suitable algorithm for land use classification in the modeling area is the RF algorithm. The number of incompliant land utilization plots according to building and water body criteria was 1,770 (21.94%) and 436 (5.40%), respectively. After ground verification, the number of incompliant land utilization plots, based on building and water body criteria, was 394 (4.88%) and 150 (1.85%), respectively. For model validation in the test area, the number of plots with incompliant land utilization, based on building and water body criteria, was 1,213 (30.01%) and 107 (2.66%), respectively. After ground verification, the number of incompliant land utilization plots, based on building and water body criteria, was 364 (9.03%) and 47 (1.17%), respectively.

This study demonstrated the effectiveness of object-based land use classification using the RF algorithm. The proposed method not only achieved high accuracy but also demonstrated promising transferability in detecting incompliant land utilization across different districts. These findings can support future monitoring and enforcement of policies in ALRO-managed areas. Still, the misclassification is a critical temporal data challenge in this study. We recommend that future studies integrate multi-temporal image analysis to distinguish between bare soil resulting from harvesting (a transient, agricultural state) and permanent urban land use (an incompliant state), which is not possible with single-date imagery. Additionally, collecting training areas for each land use type for land use classification using object-based image analysis with RF is a critical step. The user needs to review the preliminary classification results and add more training image objects to improve classification accuracy.

Acknowledgments

The authors would like to acknowledge the Agricultural Research Development Agency (Public Organization) for the scholarship.

References

[1] Suehiro, A., (2007). Land Reform in Thailand. The Concept and Background of the Agricultural Land Reform Act of 1975. *The Developing Economies*, Vol. 19(4), 314-347. https://doi.org/10.1111/j.1746-1049.1981.tb00700.x.

[2] Pongsapich, A., (2011). Land and Agricultural Development Policies Impacting on Human Rights in Thailand. *Proceedings of Human Rights and Business: Plural Legal Approaches to Conflict Resolution*, 28 November – 1 December 2011, Institutional Strengthening and Legal Reform. Bali, Indonesia.

[3] Promsaka Na Sakolnakorn, T., Kroeksakul, P., Kaewbutee, P., Naipainit, A. and Laeheem, K., (2016). Land-Use Change under the Management of the Agricultural Land Reform Office: A Case Study in Phuket. *NIDA Development Journal*, Vol. 56(4); 121-169.

[4] Maxwell, A. E., Warner, T. A. and Fang, F., (2018). Implementation of Machine Learning Classification in Remote Sensing: An Applied Review. *International Journal of Remote Sensing*, Vol. 39(9); 2784-2817. https://doi.org/10.1080/01431161.2018.1433343

[5] Ongsomwang, S., (2023). Systematic Experiment on the Suitable Machine Learning Algorithm for Land Use Classification under Object-Based Image Analysis of eCognition Software. *Journal of Remote Sensing and GIS Association of Thailand*, Vol. 24(3); 1-40.

[6] Dhiraj, K., (2019). The Advantages and Disadvantages of SVM Algorithm. *Medium*. [Online]. Available: https://dhirajkumarblog.medium.com/top-4-advantages-and-disadvantages-of-support-vector-machine-or-svm-a3c06a2b107. [Accessed Oct. 27, 2025]

[7] Breiman, L., (2001). Random Forests. *Machine Learning*, Vol. 45; 5-32.

[8] Stein, G., Chen, B., Wu, A. S. and Hua, K. A., (2005). Decision Tree Classifier for Network Intrusion Detection with GA-Based Feature Selection. *ACMSE '05 Vol 2: Proceedings of the 43rd annual ACM Southeast Conference* - Vol. 2. 136-141. https://doi.org/10.1145/1167253.1167288.

[9] Mead, R., (201). Advantages and Disadvantages of KNN Algorithm in Machine Learning. *Medium*. [Online]. Available: https://medium.com/@madan0906/advantages-and-disadvantages-of-knn-algorithm-in-machine-learning-5bdaf72454ff. [Accessed Oct. 27, 2025]

[10] Mapping and Land Text Bureau. (2024). *ALRO 4-01 Plots in Wang Nam Khiao and Pak Chong Districts*. Mapping and Land Text Bureau, Agricultural Land Reform Office, Bangkok, Thailand.

[11] Office of Surveying and Mapping Technology. (2024). Nakhon Ratchasima Province Land Use Data in 2023. Office of Surveying and Mapping Technology, Land Development Department, Bangkok, Thailand.

[12] Modified Copernicus Sentinel data [2022]/Sentinel Hub. Available: https:/browser.dataspace.copernicus.eu. [Accessed Oct. 1, 2024].

[13] Google Earth Pro Version 7.3.6.9326, December 13, 2022. Nakhon Ratchasima, Thailand. DigitalGlobe. Available: https:/www.google.com/earth/index.html [Accessed Oct. 1, 2024]

[14] Rouse, J. W., Haas, R. H., Schell, J. A. and Deering, D. W., (1974). Monitoring Vegetation Systems in the Great Plains with ERTS. *Proceedings, 3rd Earth Resource Technology Satellite (ERTS) Symposium,* Vol. 1, 48–62.

[15] Jensen, J. R., (2015). *Introductory Digital Image Processing: A Remote Sensing Perspective*. Pearson Education, Inc., Illinois.

[16] Huete, A. R., (1988). A Soil-Adjusted Vegetation Index (SAVI). *Remote Sensing of Environment*, Vol. 25; 295–309.

[17] Zha, Y., Gao, J. and Ni, S., (2003). Use of Normalized Difference Built-up Index in Automatically Mapping Urban Areas from TM Imagery. *International Journal of Remote Sensing*, Vol. 24(3); 583-594.

[18] Xu, H., (2006). Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *International Journal of Remote Sensing*, Vol. 27(14); 3025-3033.

[19] Richards, J. A. and Jia, X., (1999). *Remote Sensing Digital Image Analysis. An Introduction*, Springer-Verlag, Berlin.

[20] Marpu, P. R., Niemeyer, I., Nussbaum, S. and Gloaguen, R., (2008). A Procedure for Automatic Object-Based Classification. *Object-based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, Thomas Blaschke, Stefan Lang and Geoffrey J. Hay (eds.). Springer-Verlag Berlin Heidelberg; 169-184.

[21] Bhattacharyya, A., (1943). On a Measure of Divergence between Two Statistical Populations Defined by Probability Distributions, *Bulletin of the Calcutta Mathematical Society*, Vol. 35; 99-109.

[22] Nussbaum, S. and Menz, G., (2008). *Object-based Image Analysis and Treaty Verification: New Approaches in Remote Sensing Applied to Nuclear Facilities in Iran.* Springer Science and Business Media., Bonn.

[23] Congalton, R. G. and Green, K., (2009). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Second Edition (2nd ed.). CRC Press., Boca Raton.

[24] Foody, G. M., (2010). Assessing the Accuracy of Land Cover Change with Imperfect Ground Reference Data. *Remote Sensing of Environment*, Vol. 114; 2271–2285. https://doi.org/10.1016/j.rse.2010.05.003.

[25] Liu, C., White, M. and Newell, G., (2009). Measuring the Accuracy of Species Distribution Models: A Review. *Proceedings 18th World IMACs/MODSIM Congress*, Cairns, Australia, 4241-4247.

[26] Liu, C., Frazier, P. and L. Kumar., (2007). Comparative Assessment of the Measures of Thematic Classification Accuracy, *Remote Sensing of Environment*, Vol. 107(4); 606–616. https://doi.org/10.1016/j.rse.2006.10.010.

[27] Rakaksorn, B., Chuchip, K. and Narangajavana, P., (2023). Development of a Semi-automatic GIS Approach for Forest Land Use Change Detection. *Journal of Remote Sensing and GIS Association of Thailand*, Vol. 24(2), 32-50.

[28] Anderson, J. R., Hardy, E. E., Roach, J. T. and Witmer, R. E., (1976). *A Land Use and Land Cover Classification System for Use with Remote Sensor Data.* Professional paper 964. U.S. Geological Survey. USGS Publications. https://doi.org/10.3133/pp964.

[29] Fitzpatrick-Lins, K., (1981). Comparison of Sampling Procedures and Data Analysis for a Land-use and Land-cover Map. *Photogrammetric Engineering and Remote Sensing*, Vol. 47(3); 343–351.

[30] Rosenfield, G. H. and Fitzpatrick-Lins, K., (1986). A Coefficient of Agreement as a Measure of Thematic Classification Accuracy. *Photogrammetric Engineering and Remote Sensing*, Vol. 52(2); 223-227.

[31] Phinyoyang, A., (2021). *Optimized Land Use and Land Cover Allocation for Flood Mitigation with Goal Programming, Mueang Chaiyaphum District, Chaiyaphum Province, Thailand,* Doctoral Dissertation. Suranaree University of Technology.

[32] Srichaichana, J., (2018). *Land Use and Land Cover Scenarios of Ecosystem Services for Optimum Water Yield and Sediment Retention in Klong U-Tapao Watershed, Songkhla, Thailand,* Doctoral Dissertation. Suranaree University of Technology.

[33] Cohen, J., (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Publishers., New Jersey.