Asia-Pacific Journal of Science and Technology: Volume: 26. Issue: 04. Article ID.: APST-26-04-08. Research Article

APST

Asia-Pacific Journal of Science and Technology

https://www.tci-thaijo.org/index.php/APST/index

Published by the Research and Technology Transfer Affairs Division,
Khon Kaen University, Thailand

Assessment of machine learning on sugarcane classification using Landsat-8 and Sentinel-2 satellite imagery

Teerapat Butkhot[1] and Pipat Reungsang[1,2,*]

[1]Remote Sensing and GIS Program, Department of Computer Science, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand
[2]Geo-Informatics Centre for Development of Northeast Thailand, Faculty of Science, Khon Kaen University, Khon Kaen, Thailand
[*]Corresponding author: reungsang@kku.ac.th

Received 26 March 2021
Revised 13 April 2021
Accepted 6 May 2021

Abstract

Agriculture and agricultural product development are important aspects of a country's economic development. Sugarcane is one of the key industrial crops in Thailand, Brazil, China, and India. Therefore, monitoring sugarcane growth and harvest is important for evaluating yield, optimizing logistic operations, and forecasting crop productivity. To monitor sugarcane growth more effectively and efficiently, this study aimed to classify the sugarcane cultivation regions in Chuenchom District, Maha Sarakham Province, Thailand, using Landsat-8 and Sentinel-2 satellite images. To this end, three algorithms were used for classification: support vector machine (SVM), random forest (RF), and maximum likelihood (ML). A combination of parameter sets using four bands (red, green, blue, and NIR) and two vegetation indices: normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) was set up for the classification. The overall accuracy and kappa coefficient values were computed to validate the classification results with visual interpretation of high-resolution images. Results from the study showed that RF outperformed the SVM and ML classification techniques with overall accuracy and kappa coefficient values of 75.93 and 0.616, respectively, for Landsat-8 images and 78.60 and 0.656, respectively, for Sentinel-2 images. Specifically, RF classification with red, green, blue, and NIR provided the highest accuracy for the Landsat-8 images, while RF classification with red, green, blue, and NDVI proved to be the most accurate for the Sentinel-2 images. In summary, both Landsat-8 and Sentinel-2 satellite images have great potential for sugarcane mapping using remote sensing.

Keywords: Sugarcane, Machine learning, Support vector machine, Random forest, Maximum likelihood

1. Introduction

About 110 countries produce sugar from either cane or beet. Nearly 80% of the global sugar production is via sugarcane. Recently, the top ten sugarcane producing countries, including India, Brazil, Thailand, China, the US, Mexico, Russia, Pakistan, France, and Australia accounted for nearly 70% of the global sugar output. In addition, sugarcane can also be used to produce livestock feed, fiber, and energy, particularly for biofuels (sugar-based ethanol) and co-generation of electricity (cane bagasse). Sugarcane is generally regarded as one of the most significant and efficient source of biomass for biofuel production. Given the increasing population and global economic development, there has been an increase in the need for sugarcane-related products. Similar to Brazil, China, and India, sugarcane is one of the key industrial crops in Thailand [1, 2]. As a result, monitoring sugarcane growth and harvest is important for evaluating yield, optimizing logistic operations, and forecasting crop productivity. To monitor sugarcane growth, most local industries survey manual labor to document field properties, such as the stage of growth and location of sugarcane. Such an approach is labor intensive, expensive, and efficient only at the field or farming enterprise level. Recently, with an unmanned aerial vehicle (UAV) technology, a detailed map of sugarcane plantations can be appropriately constructed from higher spatial

resolution images with sub-meter resolution to monitor crops at the field or farming enterprise level [3-6]. However, it is necessary to develop more efficient techniques at the state, regional, or country level. Some studies have reported that using remote sensing data to produce thematic maps provides effective spatial and temporal information for crop monitoring at a regional, state, or country level [7-10]. In particular, satellite images with medium spatial resolution (10-250 m) have proven to be efficient for sugarcane mapping [11]. Sentinel-1 and Sentinel-2 images have been widely used for remote sensing applications, especially for sugarcane monitoring and mapping [12, 13]. To date, numerous research experiments have been conducted on object-based image analysis (OBIA) frameworks with machine learning for crop classification techniques [14-16]. However, both Landsat-8 and Sentinel-2 satellite images have similar characteristics (in terms of spatial and spectral resolutions) but different revisiting times; few studies have focused on using Landsat-8 and Sentinel-2 to identify sugarcane using machine learning techniques [17-20]. Thus, it is worth investigating and comparing the accuracy of image classification using machine learning techniques on Landsat-8 and Sentinel-2 satellite images, especially for sugarcane cultivation.

This study focuses on exploring classification techniques for identifying sugarcane cultivation from satellite imagery by implementing pixel-based and machine learning-based (using both RF and SVM algorithms) classifications. The effectiveness of each technique was evaluated and compared using Landsat-8 and Sentinel-2 images in Chuenchom District, Maha Sarakham Province, Thailand. This paper is divided into the following sections: Section 2 describes the study area, the satellite data used, the methodology and classifier description; Sections 3 and 4 present the results with discussion and conclusion, respectively.

2. Materials and methods

2.1 Study area and data

The study area is located in Chuenchom District, which lies in the northern region of Maha Sarakham Province, Northeast Thailand (Figure 1). It extends from 16°27′9.19″ N to 16°38′36.75″ N and from 103°5′6.77″ E to 103°11′45.96″ E, covering an area of approximately 130.57 km^2. The topography is relatively flat with elevations ranging between 146 and 248 m above mean sea level. Based on the average climate statistics using the data from 1981 to 2010 [21], it was established that Maha Sarakham Province is characterized by a wet and dry climate. Its average annual rainfall is 1000-1200 mm, with an average temperature of 27.1 °C. In 2017, its land use was classified, as follows: agricultural area (82.46%), forest area (9.13%), community and building area (4.76%), waterbody area (1.93%), and miscellaneous area (1.72%) [22].

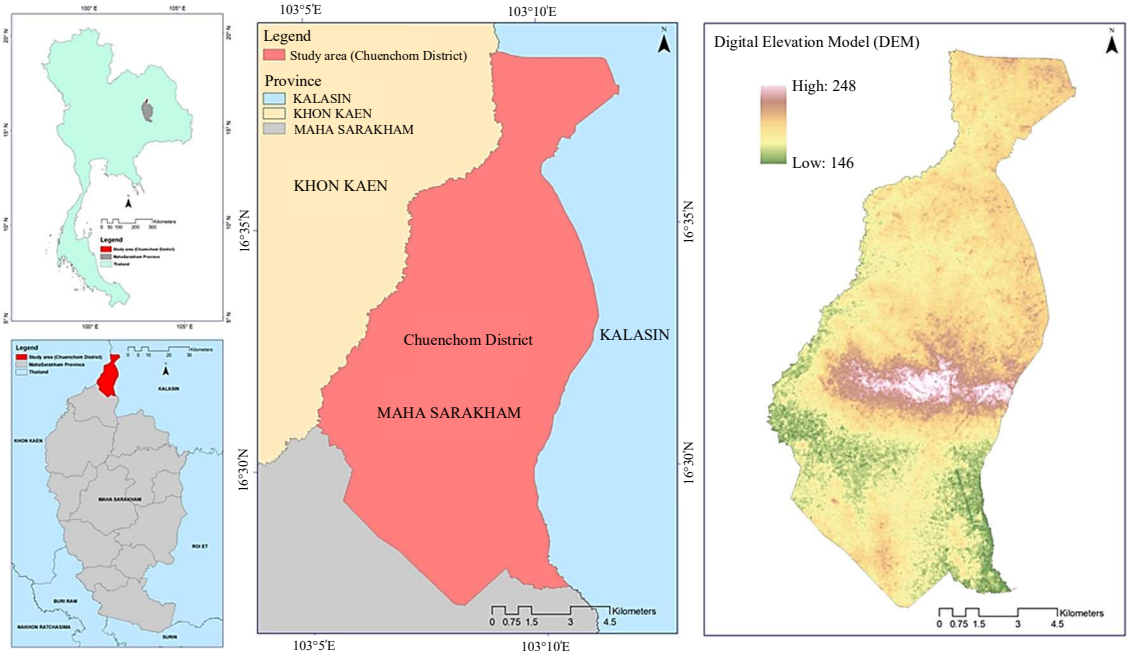

Figure 1 Location of the study area.

To identify sugarcane fields in the study area, Landsat-8 and Sentinel-2 satellite images were collected and downloaded from the U.S. Geological Survey website (https://earth explorer.usgs.gov). According to the sugarcane crop calendar in northeast Thailand, which was obtained from the Office of the Cane and Sugar

Board (OCSB), the harvesting period occurs from November to December every year. During the period before harvesting, sugarcane fields have highly condensed leaves, which provide high spectral reflectance and help differentiate the sugarcane plant from bare soil or crop residue. To minimize errors during the classification process, images were selected at a time when there was no cloud or minimal cloud cover. For these reasons, Landsat-8 satellite images (Path 128 Row 49 on December 23, 2018) along with four Sentinel-2 satellite images (48PTC, 48PUC, 48QTD, and 48QUD on December 20, 2018) were downloaded from the website via an online system. All the images had top of atmosphere (TOA) reflectance, which were rectified through radiometric and geometric corrections, including ortho-rectification and spatial registration on a Universal Transverse Mercator (UTM) grid coordinate system. Furthermore, World Geodetic System 1984 (WGS84) with grid area 48 N (zone 48 N) was applied during this pre-processing process. Once pre-processing was completed, subsets of images within the study area were selected from both Landsat-8 and Sentinel-2 before the classification process. Details of the band name, spatial resolution, and corresponding wavelengths of Landsat-8 and Sentinel-2 satellite data are shown in Tables 1 and 2, respectively. To evaluate the effectiveness of the classification techniques, referenced data were acquired by visual interpretation of Landsat-8 and Sentinel-2 images with high spatial resolution images from Google Earth to assess their accuracy.

Table 1 Details of Landsat-8 multispectral bands.

Band	Wavelength (nm)	Spatial resolution (m)
1	430-450 (Coastal Aerosol)	30
2	450-510 (Blue)	30
3	530-590 (Green)	30
4	640-670 (Red)	30
5	850-880 (NIR)	30
6; 7	1570-1650 (SWIR 1); 2110-2290 (SWIR 2)	30
8	500-680 (Panchromatic)	15
9	1,360-1,380 (Cirrus)	30
10	10,600-11,190 (TIRS 1)	100
11	11,500-12,510 (TIRS 2)	100

Table 2 Details of Sentinel-2 multispectral bands.

Band	Wavelength (nm)	Spatial resolution (m)
1	443 (Coastal Aerosol)	60
2	490 (Blue)	10
3	560 (Green)	10
4	665 (Red)	10
5; 6; 7	705; 740; 783 (Vegetation Red Edge)	20
8	842 (NIR)	10
8A	865 (Narrow NIR)	20
9	945 (Water vapour)	60
10	1375 (SWIR Cirrus)	60
11; 12	1610; 2190 (SWIR)	20

Figure 2 illustrates the proposed methodology for sugarcane classification. Initially, Landsat-8 and Sentinel-2 satellite images were downloaded and pre-processed, such that they were pan-sharpened at 15 m and 10 m,

respectively, within the study area along with the referenced images using visual interpretation. Before the classification process, groups of spectral features called "parameter sets," including original satellite image band values (red, blue, green, and NIR) and vegetation indices in spectral enhancement were prepared. Vegetation indices, including the normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) were calculated using equations (1) and (2), respectively. Although NDVI is the most popular and is widely used in remote sensing studies [23-27], the EVI was developed to emphasize the leaf surface [28, 29]. The NDVI values, which range between -1 and 1, represent the abundance and density of the vegetation. A value closer to 1 indicates an area with rich vegetation, high density, or high plant chlorophyll. Table 3 summarizes the spectral features and vegetation indices for each parameter set. To classify sugarcane, machine learning classification using random forest (RF), support vector machine (SVM) techniques, and pixel-based with maximum likelihood classification (MLC) were performed against three parameter sets, as follows: 1) red, green, blue, and NIR; 2) red, green, blue, and NDVI; and 3) red, green, blue, and EVI. Training areas included residential areas, community areas (U), water areas (W), open land areas, soil without cover (So), forestry area or forestry plantation (F), sugarcane planting area (Sg), and others such as cloud area (O). Landsat-8 satellite images had six classification categories (U, W, So, F, Sg, and O), whereas Sentinel-2 satellite images had only five classification categories (U, W, So, F, and Sg), as there was no cloud cover in the region.

2.2 Support Vector Machine (SVM)

SVM is a supervised, non-parametric statistical learning technique that can be used to solve a variety of remote sensing classification problems [30-34]. The popularity of SVM has grown in the last decade [35]. The SVM algorithm transforms training data into a higher-dimensional space and chooses the best hyperplane to distinguish between different classes or categories. The data in this algorithm are partitioned using full separation margins [36]. This machine learning algorithm employs support vectors, which are training data samples that fall on the edges of the class distribution and the kernel trick (center of the margin) [37]. The SVM model can be implemented using a variety of kernels [38], each of which has its own set of user-defined parameters [37]. The radial basis function kernel is used to implement the SVM model for multiclass classification using two parameters: a regularization parameter and a kernel bandwidth parameter. The SVM classifier requires determining the maximum number of samples to be used for defining each class. In this study, a value of 500 was set because the image data used for classification did not pre-segment the image as the default.

2.3 Random Forest (RF)

RF is a non-parametric machine learning algorithm based on the learning strategy principle that calculates the value of variables to achieve high classification accuracy [39]. It is an ensemble method that has shown excellent results in a variety of remote sensing applications [32, 40, 41]. Moreover, RF combines the responses of many classifiers to produce a final prediction and employs a replacement strategy to create new training datasets. This strategy decreases the variance and increases the classification accuracy. At each break, the RF algorithm selects a random subset of variables or predictors [42]. The performance of the classification process is determined using a majority voting scheme. The two tuning parameters for the RF algorithm include number of trees used to form an ensemble (ntree) and the number of variables/predictors used to separate the nodes (mtry). On the other hand, the best split for a node is critical for increasing classification accuracy [32, 43, 44]. The RF classification requires determining the maximum number of trees, maximum tree depth, and maximum number of samples used to define each classification type. In this study, the aforementioned parameters were set to 100, 30, and 1000, respectively. The image data used for classification have not been previously segmented.

2.4 Maximum Likelihood (ML)

ML is a supervised classification method derived from the Bayes theorem, which measures the probability that a given pixel belongs to a particular class based on the statistics for each class in each band being normally distributed. All pixels are labeled unless a probability threshold is set [45]. The class with the highest likelihood is assigned to each pixel (i.e., maximum likelihood). The pixel remains unclassified if the highest likelihood is less than the specified threshold. As mentioned at the beginning of the Section 2, six and five signature classes were selected for Landsat-8 and Sentinel-2 ML classifications, respectively. For each class, the training pixels provide values, which can be used to estimate the mean and covariances from the parameter sets. This information is used by the ML classifier to assign pixels to a particular class.

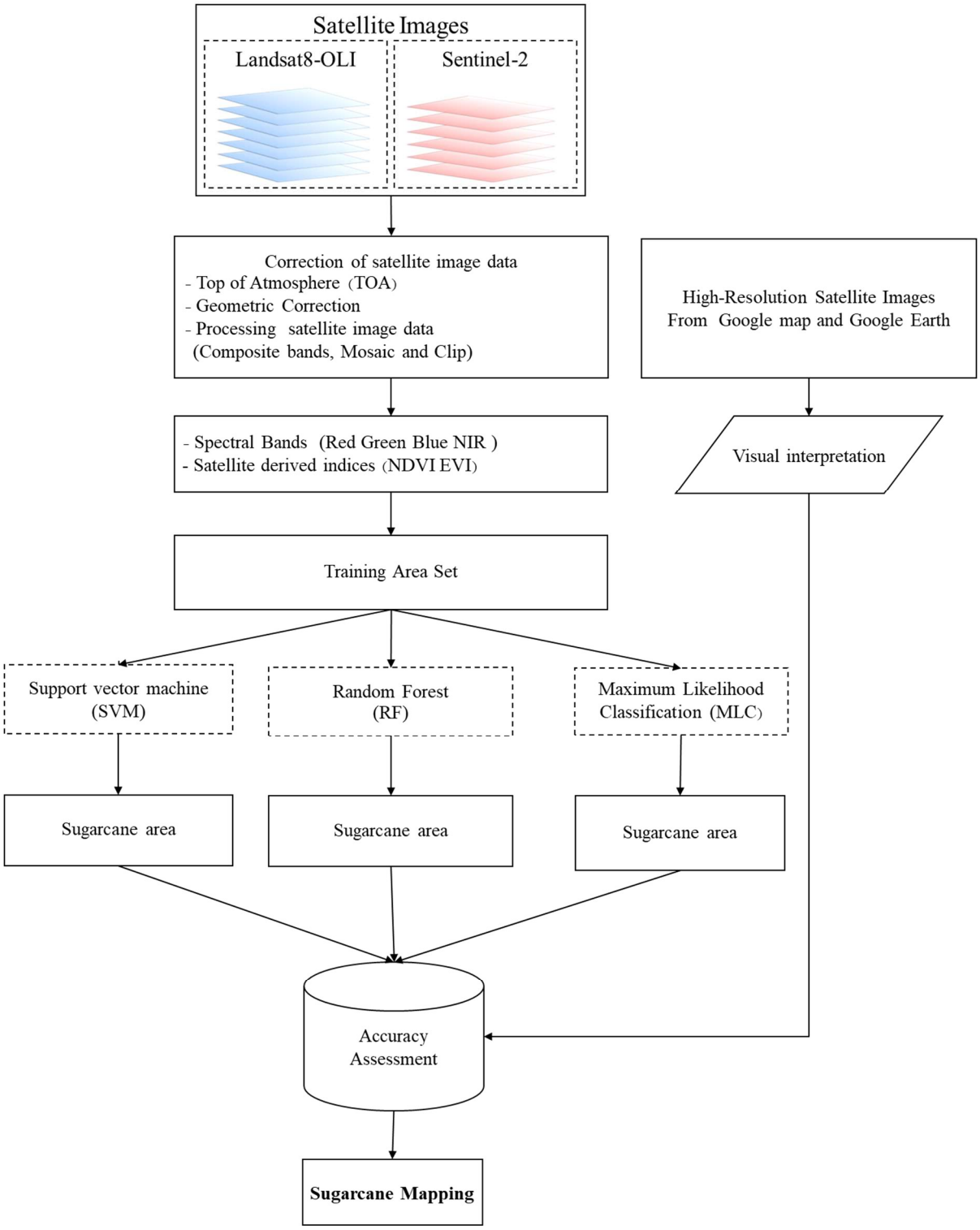

Figure 2 Framework of sugarcane classification and validation processes.

Table 3 Summary of spectral features and vegetation index for each parameter set.

Parameter Set	Bands and Index
1	Red, Green, Blue, and NIR
2	Red, Green, Blue and NDVI
3	Red, Green, Blue and EVI

2.5 Accuracy assessment

To evaluate the effectiveness of the SVM, RF, and ML techniques, sugarcane areas classified from the Landsat-8 and Sentinel-2 satellite images were overlaid and compared to the reference data. The overall accuracy assessment and kappa coefficient were calculated. The kappa coefficient [46] is an estimation of the error using KHAT statistical values to describe the degree of consistency of the two datasets using equation (3). Ranking of kappa coefficient values is shown in Table 4 [47].

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

$$EVI = G * \frac{NIR - Red}{NIR + (C1 * Red - C2 * Blue) + L}, \quad (2)$$

Where $L = 1;\ C1 = 6;\ C2 = 7.5;\ and\ G = 2.5$

$$\widehat{k} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} x_{i+} * x_{+i}}{N^2 - \sum_{i=1}^{r} x_{i+} * x_{+i}} \quad (3)$$

Where, r = sum across all rows in matrix; x_{ii} = diagonal; x_{i+}= marginal row total (row i).
x_{+i} = marginal column total (column i); N = observations.

Table 4 Degree of consistency for kappa coefficient values.

Kappa Value	Interpretation
< 0.21	Very Low Coherency
0.21 – 0.40	Low Coherency
0.41 – 0.60	Moderate Coherency
0.61 – 0.80	High Coherency
>0.80	Very High Coherency

3. Results and discussion

In this study, Landsat-8 and Sentinel-2 images were acquired during the growing season, four bands at 15 m and 10 m resolution were stacked, and the resulting images were used for sugarcane classification. The random sampling method was adopted to select training locations in both Landsat-8 and Sentinel-2 images (Figure 3). Each training location was buffered at 20 m. This was done to create training areas on different land uses, where the numbers of sampling locations on the Landsat-8 and Sentinel-2 images were 493 and 445 positions, respectively, with an approximate total area of 61.6 and 55.6 hectares, respectively (Table 5).

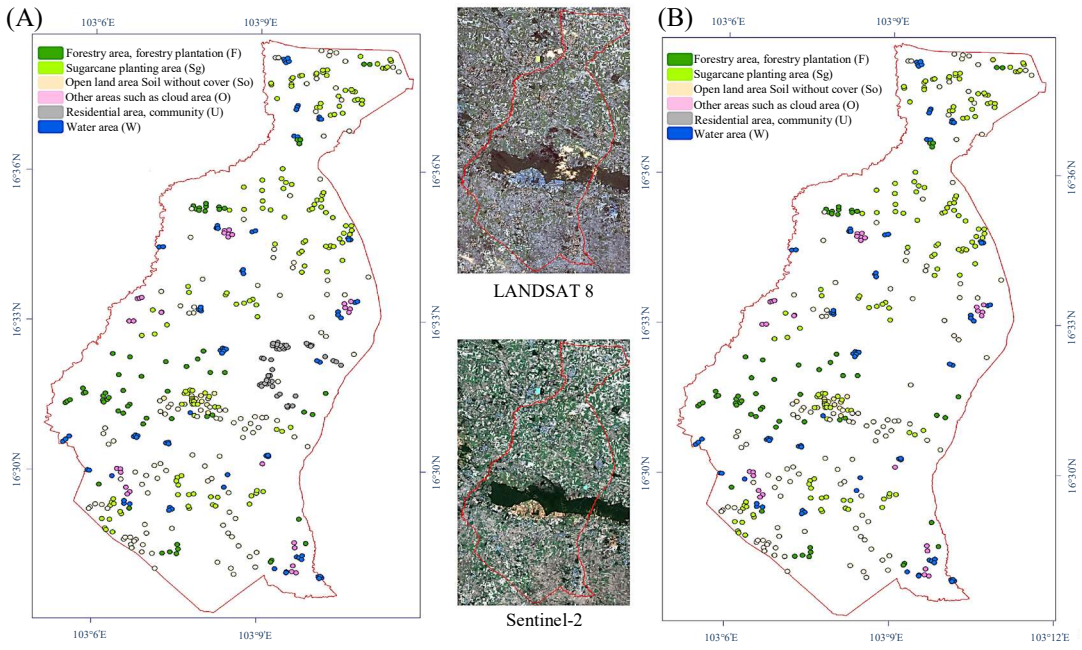

Figure 3 Training locations of Landsat-8 (A) and Sentinel-2 (B).

Table 5 Training areas in each land use categories in both Landsat-8 and Sentinel-2 images.

Land use	Landsat-8		Sentinel-2	
	Number of sampling locations	Area (km^2)	Number of sampling locations	Area (km^2)
1) Residential area, community (U)	30	0.037	30	0.037
2) Water area (W)	79	0.099	79	0.099
3) Open land area soil without cover (So)	128	0.160	128	0.160
4) Forestry area, forestry plantation (F)	62	0.077	62	0.077
5) Sugarcane planting area (Sg)	146	0.18.2	146	0.18.2
6) Others such as cloud area (O)	48	0.060	-	-
Total	493	0.616	445	0.556

In this study, land use maps were created by visual interpretation of sugarcane (Sg) and other land areas (F, So, U, W, and O) based on Landsat-8, Sentinel-2, and high-resolution satellite images obtained from Google Earth in December 2018, as shown in Figure 4 and Table 6.

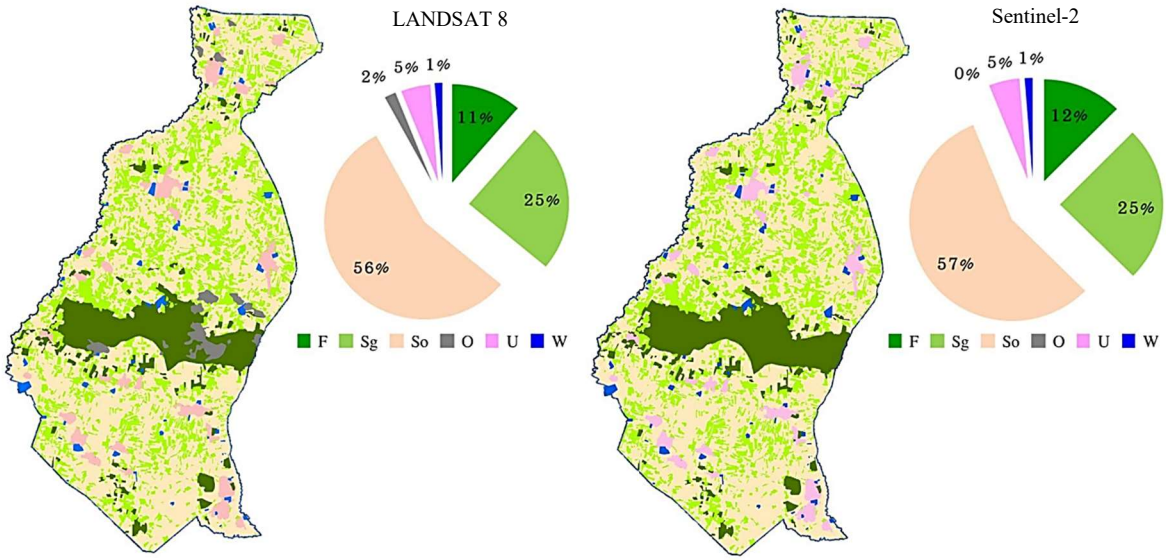

Figure 4 Map of land uses from visual interpretation.

Table 6 Summary of land use areas from visual interpretation.

Satellite		Land uses						
		Sg	F	So	U	W	O	Total
Landsat-8	Area (km^2)	32.341	14.762	73.111	6.269	1.625	2.464	130.572
	%	25	11	56	5	1	2	100
Sentinel-2	Area (km^2)	32.610	16.240	73.748	6.337	1.637	0	130.572
	%	25	12	57	5	1	0	100

To examine the performance of SVM, RF, and ML classification techniques in each parameter set in Landsat-8 and Sentinel-2 satellite images, a simple accuracy assessment was computed using the percentage of the total overlapped sugarcane areas between classification results and visual interpretation. The total overlapped area demonstrated that the classification techniques were aligned with the visual interpretation. The assessment results are listed in Table 7. The overall accuracy of Landsat-8 ranged from 77.56 to 85.86%. Furthermore, the SVM and ML provide the highest accuracy of 85.86% and 83.15%, respectively, for parameter set 1 (red, green, blue, and NIR). On the other hand, RF (77.56%) showed the lowest accuracy for parameter set 1. Overall, SVM performed better with respect to Landsat-8 images for all three parameter sets by providing higher overlapped areas, as compared to the RF and ML techniques. In Sentinel-2, SVM performed better in three parameter sets, as compared to RF and ML techniques. SVM accuracy ranged from 85.48% to 86.02%. Therefore, RF performed better than ML on Sentinel-2 images for all parameter sets, whereas ML performed better on Landsat-8 images.

Table 7 Summary of total overlapped sugarcane areas between SVM, RF, and ML from classifications and referenced data.

Satellite	Parameter Sets					
	1 (Red, Green, Blue and NIR)		2 (Red, Green, Blue and NDVI)		3 (Red, Green, Blue and EVI)	
	Area (km^2)	%	Area (km^2)	%	Area (km^2)	%
SVM						
Landsat-8	27.769	85.86	27.509	85.06	27.482	84.97
Sentinel-2	28.050	86.02	27.875	85.48	27.880	85.50
RF						
Landsat-8	25.084	77.56	25.664	79.36	25.655	79.33
Sentinel-2	27.083	83.05	26.434	81.06	26.394	80.94
ML						
Landsat-8	26.892	83.15	26.308	81.35	26.224	81.09
Sentinel-2	23.844	73.12	23.944	73.43	23.902	73.30

To assess the accuracy of SVM, RF, and ML classification on sugarcane areas, the overall accuracy and kappa coefficient were calculated. Table 8 shows the overall accuracy of SVM, RF, and ML classification on both Landsat-8 and Sentinel-2 images. Table 9 shows the kappa coefficient for Landsat-8 and Sentinel-2 images.

Table 8 Overall accuracy of SVM, RF, and ML classifications for sugarcane.

Parameter Set	Landsat-8			Sentinel-2		
	SVM	RF	ML	SVM	RF	ML
1. Red, Green, Blue, and NIR	73.45	75.93	73.71	76.52	77.13	78.01
2. Red, Green, Blue, and NDVI	73.52	74.98	73.14	78.12	78.60	76.81
3. Red, Green, Blue, and EVI	73.38	74.95	73.21	77.93	78.39	77.91

Table 9 Kappa Coefficient of SVM, RF, and ML classifications for sugarcane.

Parameter Set	Landsat-8			Sentinel-2		
	SVM	RF	ML	SVM	RF	ML
1. Red, Green, Blue, and NIR	0.587	0.616	0.579	0.635	0.637	0.643
2. Red, Green, Blue, and NDVI	0.595	0.607	0.579	0.656	0.656	0.630
3. Red, Green, Blue, and EVI	0.593	0.607	0.579	0.653	0.654	0.642

As indicated in Tables 8 and 9, RF was ideal for parameter set 1 (red, green, blue, and NIR) in Landsat-8 sugarcane classification with an overall accuracy and kappa coefficient of 75.93% and 0.616, respectively. Similarly, the performance of RF was the best for parameter set 2 (red, green, blue, and NDVI) in Sentinel-2 sugarcane classification, with an overall accuracy and kappa coefficient of 78.60% and 0.656, respectively. Overall, RF performed better than SVM and ML for both Landsat-8 and Sentinel-2 images. Classification results for Sentinel-2 provided higher values of accuracy than Landsat-8, which could be because spatial resolution is an important factor for image classification. For instance, Landsat-8 has a spatial resolution of 15 m, whereas Sentinel-2 has a spatial resolution of 10 m. Furthermore, SVM, RF, and ML classifications were performed using three parameter sets based on red, green, blue, NIR, NDVI, and EVI as classifiers, where NDVI was derived from red and NIR, and EVI was derived from red, blue, and NIR. Both NDVI and EVI are common indices used to estimate crop conditions. According to Saini and Ghosh (2018), NIR is the most important classifier for crop classification, followed by red and blue; green is not significant [48]. This explains why parameter sets 1 and 2 performed better than parameter set 3. Sugarcane areas classified by RF using Landsat-8 images with parameter set 1 and Sentinel-2 with parameter set 2 are shown in Figure 5.

Figure 5 Total classified sugarcane area by RF; Landsat-8 with parameter set 1 (A), and Sentinel-2 with parameter set 2 (B).

4. Conclusion

This study aimed to classify sugarcane areas in Chuenchom District, Maha Sarakham Province, Northeast Thailand using Landsat-8 and Sentinel-2 satellite images for December 2018 using SVM, RF, and ML classification techniques. Three parameter sets, including combinations of red, green, blue, NIR, NDVI, and EVI were considered as classifiers. Accuracy assessment in terms of overall accuracy and kappa coefficient was computed and compared for these classifiers. The results of the study illustrated that RF outperformed the SVM and ML classification techniques with overall accuracy and kappa coefficient values of 75.93 and 0.616, respectively, for Landsat-8 images and 78.60 and 0.656, respectively, for Sentinel-2 images. Specifically, RF classification with red, green, blue, and NIR provided the highest accuracy value for Landsat-8 images, while RF classification with red, green, blue, and NDVI provided the highest accuracy value for Sentinel-2 images. In summary, the results demonstrate that both Landsat-8 and Sentinel-2 satellite images have the potential to classify sugarcane cultivation; however, more accurate results can be achieved using RF classifiers.

5. Acknowledgements

We are grateful to the Remote Sensing and GIS Program at the Department of Computer Science, Faculty of Science, Khon Kaen University for supporting this project till its completion. This project was supported by the Thammasat University Research Fund under the Research University Network (RUN) initiative. We are also grateful to the reviewers and editors for their valuable comments, suggestions, and attention to detail.

6. References

[1] OECD/FAO. OECD-FAO agricultural outlook 2019-2028. Rome: OECD Publishing, Paris/Food and Agriculture Organization of the United Nations; 2019.

[2] Sawaengsak W, Prasara AJ, Gheewala SH. Assessing the socio-economic sustainability of sugarcane harvesting in Thailand. Sugar Tech. 2021;23(2):263-277.

[3] Oré G, Alcântara M, Góes J, Oliveira L, Yepes J, Teruel B, et al. (2020). Crop growth monitoring with drone-borne DInSAR. Remote Sens. 2020;12:615.

[4] Luna I, Lobo A. Mapping crop planting quality in sugarcane from UAV imagery: a pilot study in Nicaragua. Remote Sens. 2016;8:500.

[5] Som-ard J, Hossain MD, Ninsawat S, Veerachitt V. Pre-harvest sugarcane yield estimation using UAV-based RGB images and ground observation. Sugar Tech. 2018;20(6):645-657.

[6] Chea C, Saengprachatanarug K, Posom J, Wongphati M, Taira E. Sugar yield parameters and fiber prediction in sugarcane fields using a multispectral camera mounted on a small unmanned aerial system (UAS). Sugar Tech. 2020;22(4):605-621.

[7] Avci ZDU, Sunar F. Process-based image analysis for agricultural mapping: a case study in Turkgeldi region, Turkey. Adv Space Res. 2015;56:1635-1644.
[8] Foody GM. Status of land cover classification accuracy assessment. Remote Sens Environ. 2002;80:185-201.
[9] Gao F, Anderson MC, Zhang XY, Yang ZW, Alfieri JG, Kustas WP, et al. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. Remote Sens Environ. 2017;188:9-25.
[10] Korhonen L, Packalen P, Rautiainen M. Remote sensing of environment comparison of Sentinel-2 and Landsat-8 in the estimation of boreal forest canopy cover and leaf area index. Remote Sens Environ. 2017;195:259-274.
[11] Dos Santos Luciano AC, Picoli MCA, Rocha JV, Franco HCJ, Sanches GM, Leal MRLV, et al. Generalized space-time classifiers for monitoring sugarcane areas in Brazil. Remote Sens Environ. 2018;215:438-451.
[12] Kavats O, Khramov D, Sergieieva K, Vasyliev V. Monitoring of sugarcane harvest in Brazil based on optical and SAR data. Remote Sens. 2020;12:4080.
[13] Jiang H, Li D, Jing W, Xu J, Huang J, Yang J, et al. Early season mapping of sugarcane by applying machine learning algorithms to Sentinel-1A/2 time series data: a case study in Zhanjiang City, China. Remote Sens. 2019;11:861.
[14] Whyte A, Ferentinos KP, Petropoulos GP. A new synergistic approach for monitoring wetlands using Sentinels -1 and 2 data with object-based machine learning algorithms. Environ Model Softw. 2018;104:40-54.
[15] Rodriguez-galiano VF, Ghimire B, Rogan J, Chica-olmo M, Rigol-sanchez JP. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J Photogramm Remote Sens. 2012;67:93-104.
[16] Li M, Ma L, Blaschke T, Cheng L, Tiede D. A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments. Int J Appl Earth Obs Geoinf. 2016;49:87-98.
[17] Hammer RG, Sentelhas PC, Mariano JCQ. Sugarcane yield prediction through data mining and crop simulation models. Sugar Tech. 2020;22(2):216-225.
[18] Maldaner LF, Corredo LP, Canata TF, Molin JP. Predicting the sugarcane yield in real-time by harvester engine parameters and machine learning approaches. Comput Electron Agric. 2021;181:105945.
[19] Luciano ACS, Picoli MCA, Duft DG, Rocha JV, Leal MRLV, Maire G. Empirical model for forecasting sugarcane yield on a local scale in Brazil using Landsat imagery and random forest algorithm. Comput Electron Agric. 2021;184:106063.
[20] Wanga M, Liub Z, Baig MHA, Wang Y, Lib Y, Chen Y. Mapping sugarcane in complex landscapes by integrating multi-temporal Sentinel-2 images and machine learning algorithms. Land use policy. 2019;88: 104190.
[21] Agricultural Meteorology to know for Maha Sarakham [Internet]. 2014 [cited 2020 Mar 12]. Available from: http://www.arcims.tmd.go.th/DailyDATA /Agromettoknow/NE/.
[22] Landuse Northeast_GISTDA_25k [Internet]. 2017 [cited 2020 Mar 15]. Available from: https://gistdaportal.gistda.or.th/portal/home/item.html?id=6e65e3e07f3f48e6b7bf6e0161b5f61c.
[23] Lawrence RL, Ripple WJ. Comparisons among vegetation indices and bandwise regression in a highly disturbed, heterogeneous landscape. Remote Sens Environ. 1998;64:91-102.
[24] Chen D, Huang J, Jackson TJ. Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near –and short wave infrared bands. Remote Sens Environ. 2005;98:225-236.
[25] Gu Y, Brown WJ, Verdin JP, Wardlow B. A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. Geophys Res Lett. 2007;34:L06407.
[26] Bégué A, Lebourgeois V, Bappel E, Todoroff P, Pellegrino A, Baillarin F, et al. Spatio-temporal variability of sugarcane fields and recommendations for yield forecast using NDVI. Remote Sens. 2010;31:5391-5407.
[27] Fernandes JL, Ebecken NFF, Esquerdo JCDN. Sugarcane yield prediction in Brazil using NDVI time series and neural networks ensemble. Remote Sens. 2017;38:4631-4644.
[28] Huete A, Didan K, Miura T, Rodriguez EP, Gao X, Ferreira LG. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens Environ. 2002;83:195-213.
[29] Matsushita B, Yang W, Chen J, Onda Y, Qiu G. Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: a case study in high-density cypress forest. Sensors. 2007;7:2636-2651.

[30] Foody GM, Mathur A. Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification. Remote Sens Environ. 2004;93:107-117.
[31] Pal M. Random forest classifier for remote sensing classification. Int J Remote Sens. 2005;26:217-222.
[32] Sonobe R, Yamaya Y, Tani H, Wang X, Kobayashi N, Mochizuki K. Assessing the suitability of data from Sentinel-1A and 2A for crop classification. GIScience Remote Sens. 2017;54:918-938.
[33] Hawryło P, Bednarz B, Wężyk P, Szostak M. Estimating defoliation of Scots pine stands using machine learning methods and vegetation indices of Sentinel-2. Eur J Remote Sens. 2018;51:194-204.
[34] Waldner F, Canto GS, Defourny P. Automated annual cropland mapping using knowledge-based temporal features. ISPRS J Photogramm Remote Sens. 2015;110:1-13.
[35] Mountrakis G, Im J, Ogole C. Support vector machines in remote sensing: a review. ISPRS J Photogramm Remote Sens. 2011;66:247-259.
[36] Vapnik V. The nature of statistical learning theory. New York: Springer; 1995.
[37] Maxwell AE, Warner TA, Fang F, Maxwell AE, Warner TA. Implementation of machine-learning classification in remote sensing: an applied review. Int J Remote Sens. 2018;39:2784-817.
[38] Kavzoglu T, Colkesen I. A kernel function analysis for support vector machines for land cover classification. Int J Appl Earth Obs Geoinf. 2009;11(5):352-359.
[39] Breiman L. Random forests. Mach. Learn. 2001;45(1):5-32.
[40] Yin H, Dirk P, Li A, Li Z, Hostert P. Land use and land cover change in Inner Mongolia - understanding the effects of China's revegetation programs. Remote Sens Environ. 2018;204:918-930.
[41] Chan JCW, Paelinckx D. Evaluation of random forest and adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. Remote Sens Environ. 2008;112(6):2999-3011.
[42] Liaw A, Wiener M. Classification and regression by random forest. R news. 2002;2:18-22.
[43] Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. Ann Appl Stat. 2008;2:841-860.
[44] Ishwaran H, Kogalur UB. Random survival forests for R. R News. 2007;7(2):25-31.
[45] Richards JA. Remote sensing digital image analysis: an introduction. Berlin: Springer; 1999.
[46] Congalton RG. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sens Environ. 1991;37:35-46.
[47] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-174.
[48] Saini R, Ghosh SK. Crop classification on single date Sentinel-2 imagery using random forest and support vector machine. Int arch photogramm remote sens spat inf sci. 2018;XLII-5:683-688.