Taylor & Francis
Taylor & Francis Group

Journal of the Chinese Institute of Engineers

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tcie20

Effectiveness of the red-edge band of RapidEye in land cover classification

Tzu-Ying Chen, Hui-Hsin Chen, Tee-Ann Teo & Peter Tian-Yuan Shih

To cite this article: Tzu-Ying Chen, Hui-Hsin Chen, Tee-Ann Teo & Peter Tian-Yuan Shih (2023) Effectiveness of the red-edge band of RapidEye in land cover classification, Journal of the Chinese Institute of Engineers, 46:1, 21-30, DOI: 10.1080/02533839.2022.2141339

To link to this article: https://doi.org/10.1080/02533839.2022.2141339

JOURNAL OF THE CHINESE INSTITUTE OF ENGINEERS
2023, VOL. 46, NO. 1, 21–30
https://doi.org/10.1080/02533839.2022.2141339

Taylor & Francis
Taylor & Francis Group

CIVIL ENGINEERING

Effectiveness of the red-edge band of RapidEye in land cover classification

Tzu-Ying Chen, Hui-Hsin Chen, Tee-Ann Teo and Peter Tian-Yuan Shih

Department of Civil Engineering, National Yang Ming Chiao Tung University, No 1001 Daxue Rd East Dist, Hsinchu City 300093, Taiwan

ABSTRACT

This study examined the effectiveness of the red-edge band using RapidEye satellite images for land cover classification. The analysis comprises three schemes for evaluating the effectiveness of the red-edge band: principal component analysis (PCA), vegetation index, and supervised image classification. The factor loadings computed by means of PCA were applied to analyze the importance of each band in the training samples. The analysis results of the factor loadings indicated that the red-edge band performed better than the visible band in the vegetation region. When rice paddy and peanuts were classified using the NDVI_RE, the improvement in accuracy was approximately 7%. Further, the accuracy of rice paddy classification using CMFI_RE was improved by approximately 6%. It can thus be inferred that the red-edge band made a certain contribution to vegetation classification. In land cover classification using reflectance, the accuracy of the support vector machine (SVM) was higher than that of the maximum likelihood classifier (MLC), the iterative self-organizing data analysis technique, and the K-means algorithm. When the red-edge band was included, the overall accuracy improved from 1% to 3%. The results of our experiments indicated that the red-edge band contributed marginally to land cover classification.

ARTICLE HISTORY
Received 23 March 2022
Revised 17 September 2022
Accepted 11 October 2022

CO EDITOR-IN-CHIEF
Ou, Yu-Chen

ASSOCIATE EDITOR
Ou, Yu-Chen

KEYWORDS
Red-edge, image classification, PCA, RapidEye

1. Introduction

In recent years, the spectral resolution of several commercial high-resolution satellites (e.g. RapidEye, WorldView-2, WorldView-3) has improved owing to the incorporation of an additional red-edge band. The red-edge band lies between the red and near-infrared (NIR) bands

Corresponding author e-mail: Tee-Ann Teo, tateo@nycu.edu.tw

and can be used to obtain a great amount of spectral information for land cover classification. Numerous studies have examined the effectiveness of different vegetation indices (VIs) obtained from multispectral bands. For example, Lyon et al. (1998) used seven VIs to investigate change in land cover. The authors found that the normalized difference vegetation index (NDVI) was the best indicator of vegetation change than other indices. Agapiou, Hadjimitsis, and Alexakis (2012) compiled a list of 71 broadband and narrowband VIs to evaluate the benefits of using VIs for crop classification. Although broadband NIR can be used to effectively classify vegetated and non-vegetated areas, it might overlook minor spectral changes. Their experimental results showed that the narrowband VIs increased the difference between archeological and non-archeological areas up to 20%.

Several studies have demonstrated that the red-edge band improves classification accuracy, especially in terms of distinguishing between different plants or crops. Schuster, Förster, and Kleinschmit (2012) designed three combinations – red-edge, NDVI, and NDVI_RedEdge – to evaluate the benefits of including the red-edge band when acquiring RapidEye satellite image. Both support vector machines (SVMs) and the maximum likelihood classifier (MLC) were used to further strengthen the explanatory power of the classification results. The experimental results indicated that the additional red-edge band marginally increased the overall accuracy (OA) by 2%. Moreover, the red-edge band improved the OA for the individual categories in distinguishing different plants. Similarly, Adelabu, Mutanga, and Adam (2014) evaluated the effect of the red-edge band on the RapidEye satellite images used to classify insect defoliation levels. Both SVM and random forest (RF) classifiers were used to perform three accuracy assessments. When the red-edge band was included, the OA of SVM and RF increased by approximately 19% and 21%, respectively. Kim and Yeom (2014) compared the effectiveness of the red-edge and texture features. The results indicated that the broadband red-edge band was slightly superior in terms of its texture features when classifying crop conditions in relatively homogeneous rice paddy environments. Li et al. (2017) performed a spectral analysis by using VIs for land cover classification in an arid region to discuss the effectiveness of the red-edge band. The experimental results demonstrated that the red-edge band contributed to a marginal improvement in classification accuracy. Apart from traditional machine learning, Saini and Ghosh (2019) employed extreme gradient boosting (XGBoost) to analyze the effect of the red-edge band on land cover classification by using RapidEye multispectral imagery. They proved that the results obtained using all classifiers with the red-edge band were effective and reported that XGBoost can prove beneficial in solving more complex classification problems. Instead of using the red-edge band for land cover classification, Meng et al. (2021)

CONTACT Tee-Ann Teo tateo@nycu.edu.tw Department of Civil Engineering, National Yang Ming Chiao Tung University, No 1001 Daxue Rd East Dist, Hsinchu City 300093, Taiwan

used the SCOPE model to quantitatively estimate the physicochemical parameters of vegetation. Their experimental results indicated that the red-edge band also improved the retrieval of vegetation parameters.

Principal component analysis (PCA) can be used in conjunction with data transformation to identify important components. PCA can also be employed to reduce the dimensionality of data (Saegusa, Sakano, and Hashimoto 2004). Munyati (2004) applied PCA to reduce the dimensionality of a dataset and compute the eigenvalues and factor loadings of the principal components of each original band. The first and second principal components usually have higher explanations, and the factor loadings of each of the bands can be computed to determine the contributions of individual bands.

The majority of previous studies have focused on examining the effectiveness of the red-edge band in image classification applications. However, the capability of the classifier can impact the effectiveness of the red-edge band. Instead of using image classification, we analyzed the separability of the training samples to reduce the effect of the image classifier. We also provided a different perspective from factor loadings, as factor loadings is rarely discussed on the red-edge band. Therefore, the contribution of this study is to discuss the effectiveness of red-edge from the perspective of factor loadings and different classifiers. We not only analyzed the results of vegetation areas but also compared the effects of non-vegetation areas with and without the red-edge band.

This study aimed at evaluating the effectiveness of the RapidEye red-edge band. In the first part of this study, we adopted the Jeffries–Matusita (JM) distance to discuss separability, and used PCA to discuss the explanations provided by different bands. The effectiveness indicated the improvement of spectral separation in JM and contribution to the major principal component in PCA. In the second part, the classification results were compared using different VIs from the red-edge and NIR bands. The effectiveness referred to the improvement of classification accuracy using red-edge-derived VIs. In the third part, land cover classification was performed using a supervised classifier. The effectiveness indicated the improvement of classification accuracy when red-edge was available.

2. Material and methods

2.1. Study area and test images

The study area was located in the middle western part of Taiwan (Figure 1), between longitude 120°E and 121°E and latitude 23.6°N and 24°N. The test images were RapidEye multispectral satellite images with the red-edge band. The RapidEye image captured on 16 October 2015, shows fields with crops, such as rice paddy and peanut. A RapidEye basic (Level 1B) image comprises the blue, green, red, red-edge, and NIR bands, with wavelengths ranging from 410–510 nm, 520–590 nm, 630–685 nm, 690–730 nm, and 760–850 nm, respectively. The spatial resolution was 6.5 m, and the radiometric resolution was 12 bits. We adopted a rational function model (Teo and Huang 2013) and ground control points to perform image orthorectification. For radiometric correction, ATCOR in PCI Geomatica was used to convert the digital number of an image to a reflectance value.

Figure 1. Test image on 10/16/2015.

2.2. Ground truth data of crops

The ground truth data for evaluation were based on a farmland cover map, produced by the Taiwan Agricultural Research Institute on the basis of field investigations. The attributes of the farmland cover map were the crop type, crop height, growth stage of the crop, and survey date. We determined the crop types in the training samples by using the ground truth and the aforementioned attributes. The field investigation date and RapidEye imaging date were within a month of each other. The main crop types included rice paddy and peanuts.

2.3. Training samples selection

According to the characteristics of the study area, we divided the land cover into four categories: water body, bare earth, rice paddy, and peanut. The rice paddy and peanut regions were selected directly from the farmland cover map, while the training samples of the water body and bare earth were digitized directly from the satellite images. The selection criteria of the training and testing samples and their numbers are described as follows: (1) the peanut regions were obtained from the farmland cover map with the crop height attribute of > 20 cm; (2) the rice paddy areas were covered with non-harvest rice paddies in the farmland cover map with the crop height attribute of > 40 cm; (3) the bare earth was covered by sand from river beds and fallow lands; and (4) the water body included seas, lakes, and rivers, constituting 23% of the whole training samples. To avoid the class imbalance problem, the proportion of each class was similar, and the proportions of the peanut regions, rice paddy, bare earth, and water body were 23%, 21%, 21%, and 34%, respectively. Table 1 summarizes the number of training and test data sets. Figure 2 presents an

Table 1. Number of training and testing datasets.

Class	Train Unit:polygons	Test Unit:polygons	Train Unit:pixels	Test Unit:pixels	Train+Test Unit:pixels	Train Unit:%	Test Unit:%
Peanuts	271	466	5946	10,224	16,170	37%	63%
Rice paddies	258	415	5660	9104	14,764	38%	62%
Bare earth	49	75	5884	9006	14,890	40%	60%
Water	10	37	5093	18,844	23,937	21%	79%
Sum	588	993	22,583	47,179	69,762		

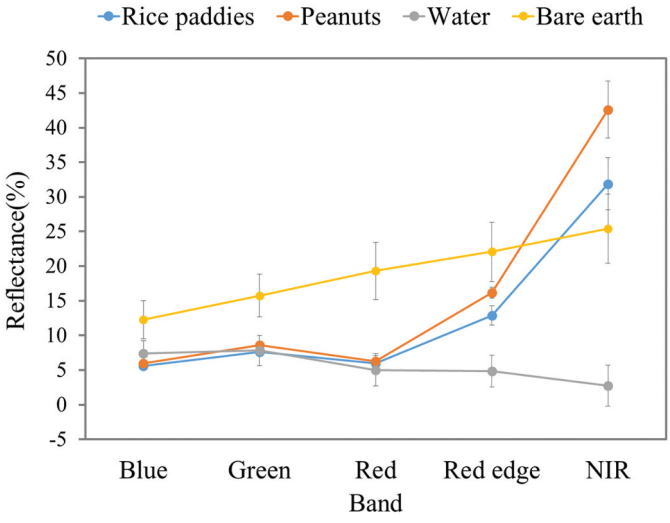

Figure 2. Mean reflectance values of different types of land cover in each band.

analysis of the reflectance differences between these four classes in different spectral bands. It was easy to separate bare earth from the other classes. The rice paddy and peanuts could not be separated in the visible bands; however, they could be distinguished in the red-edge and NIR bands.

2.4. Evaluating the effectiveness of the red-edge band

In this study, we analyzed the effectiveness of the red-edge band for land cover classification. In the first part, the JM distance was used to discuss separability, and PCA was used to discuss the explanations of different bands. Both JM and PCA were applied to the training dataset. In the second part, the classification results obtained from the red-edge and NIR bands by using the different VIs were be compared by test dataset. In this part, the effectiveness of the red-edge band with different VIs was assessed by test dataset. In the third part, land cover classification was performed by using the supervised and unsupervised classifiers. The SVM, MLC, iterative self-organizing data analysis technique (ISODATA), and K-means were applied to classify the four types of land cover.

After the second and third part classification, several accuracy indices, such as the OA and Kappa coefficients (κ), were applied for accuracy analysis. The OA was computed by dividing the total number of correct pixels by the total number of pixels in the error matrix (Equation 1). The Kappa (κ) (Equation 2) was a measure of the difference between the actual agreement between the reference data and an automated classifier and the chance agreement between the reference data and a random classifier. The variance of Kappa could be computed by using Equation 3 through the delta method (Congalton and Green 1999).

$$OA = \left(\frac{1}{N}\sum_{i=1}^{n} N_{ii}\right) \times 100\%, \tag{1}$$

$$\kappa = \frac{N\sum_{i=1}^{n} N_{ii} - \sum_{i=1}^{n}(N_{i+} * N_{+i})}{N^2 - \sum_{i=1}^{n}(N_{i+} * N_{+i})}, \tag{2}$$

$$\sigma_\kappa^2 = \frac{1}{N}\left[\frac{P_0(1-P_0)}{(1-P_1)^2} + \frac{2(1-P_0)(2P_0P_1 - P_2)}{(1-P_1)^3} + \frac{(1-P_0)^2(P_3 - 4P_1^2)}{(1-P_1)^4}\right], \tag{3}$$

where n = number of rows in the confusion matrix; N = total number of observations included in the confusion matrix; N_{ii} = number of correct observations in the diagonal direction; N_{i+} = total number of observations in the direction of row i; and N_{+i} = total number of observations in the direction of column i. $P_0 = \frac{1}{N}\sum_{i=1}^{n} N_{ii}$; $P_1 = \frac{1}{N^2}\sum_{i=1}^{n}(N_{i+} * N_{+i})$; $P_2 = \frac{1}{N^2}\sum_{i=1}^{n} N_{ii} * (N_{i+} + N_{+i})$; and $P_3 = \frac{1}{N^3}\sum_{i=1}^{n} N_{ii} * (N_{i+} + N_{+i})^2$.

In this study, the OA and Kappa coefficients were included in the accuracy assessment. Moreover, a Kappa analysis and a pairwise Z-test were conducted to determine whether the two classifications were significantly different (Weih and Riggan 2010). The Z-test was used to test for statistical significance. We performed this test to evaluate whether the classifications, with and without the red-edge band, were significantly different from each other. The critical Z-score values in the 95% confidence level were −1.96 and +1.96 standard deviations. If the value of Z-score was larger or less than 1.96, the null hypothesis was rejected and the alternative hypothesis was accepted. Thus, the p-value associated with a 95% confidence level was 0.05. The pairwise Z-scores and probabilities (p-values, α) were calculated for each combination of the two classifications (Equation 4). This study applied the Z-test by testing a hypothesis (p-values $\alpha = 0.05$ and $Z_{0.05} = 1.96$): If the p-value was < 0.05, the null hypothesis was rejected, meaning that the two classifications were considered to be statistically significantly different.

$$Z = \frac{|\kappa_1 - \kappa_2|}{\sqrt{\sigma_{\kappa_1}^2 + \sigma_{\kappa_2}^2}}, \tag{4}$$

where $\kappa_1 and \kappa_2$ are the Kappa coefficients for the two classifications, and $\sigma_{\kappa 1}$ and $\sigma_{\kappa 2}$ are the standard deviations of the two aforementioned Kappa coefficients for the two classifications.

2.4.1. Analysis of training samples

2.4.1.1. Jeffries–Matusita (JM) distance. In this study, the JM distance, like the divergence, was used to measure the statistical separability between two classes in terms of their spectral reflectance (Equation 5) (Swain and Davis 1978).

$$\mathrm{JM} = J_{\mathrm{ij}} = \left[2(1 - e^{-\alpha})\right]^{\frac{1}{2}}, \tag{5}$$

$$\alpha = \frac{1}{8}(\boldsymbol{U_i} - \boldsymbol{U_j})^T \left(\frac{\sum_i + \sum_j}{2}\right)^{-1} (\boldsymbol{U_i} - \boldsymbol{U_j}) + \frac{1}{2} log_e \left[\frac{\frac{1}{2}\left|\sum_i + \sum_j\right|}{\sqrt{\left|\sum_i\right| \times \left|\sum_j\right|}}\right], \tag{6}$$

where (*i, j*) = two classes; $\boldsymbol{U_i}$ = mean value of class *i*; $\boldsymbol{U_j}$ = mean value of class *j*; $\boldsymbol{\Sigma_i}$ = covariance of class *i*; $\boldsymbol{\Sigma_j}$ = covariance of class *j*.

The JM distance ranged from 0 to $\sqrt{2}$ (≈1.414). A JM distance of 1.414 indicated that the two classes were well separated. Usually, a JM distance greater than $\sqrt{1.9}$ (≈1.386) was considered acceptable for distinguishing between two classes (Shafri, Anuar, and Saripan 2009). If the JM distance between two classes was less than 1, the two classes were said to be spectrally overlapped and should be merged into a single class.

2.4.1.2. Principal component analysis (PCA). In PCA, an image was simply transformed by linear combinations of weighted multispectral images. The PCA transformation coefficients were statistical quantities called eigenvectors. PCA only changed the data representation and did not change the data content, so the total variation before and after PCA transformation remains the same (Equation 7) (Jensen 1996).

$$\sum_{i=1}^{n} Var(x_i) = \sum_{i=1}^{n} Var(PC_i), \tag{7}$$

whereHT x_i= variable, PC_i= principal component, n = number of variables, and $Var()$ = variance.

Explanation was the percentage of the total variance in the data explained by each component (Equation 8). A component with a high value of explanation was more informative than other components. It was also considered more important than other components. It was possible to determine how each band loads or was associated with each principal component by computing its factor loading (C_{ji}) (Equation 9) (Hair et al. 1998), which ranges from −1 to 1. The factor loadings (C_{ji}) represented the contribution of each band (i.e. blue, green, red, red-edge, and NIR) to each principal component (PC_i). They represented the degree of correlation between each band and the principal components. If the factor loading of a band was greater than 0.5, the band makes a large contribution to the corresponding principal component (Hair et al. 1998).

$$Explanation(\%) = \frac{Var(PC_i)}{\sum_{i=1}^{n} Var(PC_i)} \times 100\%, \tag{8}$$

$$C_{ji} = \boldsymbol{eigenvectors} \times \sqrt{\lambda_j}. \tag{9}$$

where $\lambda_j = j^{th}$ eigenvalue.

This study used main component ($\Sigma(x_{j+})^2$) to analyze the contribution of each band. The first two major principal components (PC_1 and PC_2) were selected and ${PC_1}^2 + {PC_2}^2$ was main component from PC_1 and PC_2. The larger the main component, the more important it was with each band. In this study, the factor loading was calculated under two different conditions. Under the first condition, the entire image was used to calculate the factor loadings. This condition was used to compare the contributions of different bands to the major principal component. Under the second condition, different types of land covers, such as vegetation, water, and bare earth, were used to calculate the factor training areas. This condition was used to compare the contributions of different bands in different land covers.

2.4.2. Land cover classification using different vegetation indices

The VIs can be divided into broadband and narrowband indices. RapidEye images are broadband images. Nevertheless, Agapiou, Hadjimitsis, and Alexakis (2012) and Baluja et al. (2012) used RapaidEye to calculate the VIs for narrowband images. Thus, this study used the VIs designed from broadband and narrowband images, including the NDVI, cropping management factor index (CMFI), optimized soil-adjusted vegetation index (OSAVI), transformed chlorophyll absorption ratio index (TCARI), chlorophyll absorption ratio index (CARI), and modified chlorophyll absorption ratio index (MCARI).

In this study, two land cover categories were considered for decision tree classification. The first category represented the classification of vegetation and non-vegetation areas and was used to compare the effectiveness of the red-edge band for vegetation classification by using different VIs. The second category included four classes: water body, bare earth, rice paddy, and peanut. The thresholds for decision tree classification were extracted from the training data.

A VI was a type of index obtained by transforming multispectral bands to obtain new information. It could also reveal the health status of vegetation. For example, the NDVI was the most effective and commonly used index for examining the vegetation condition. Spectral models and indices were also being developed to improve vegetation sensitivity by accounting for atmospheric and soil effects. In this section, the effectiveness of the red-edge band will be analyzed by using various VIs (Table 6). Only the VI was used for classification, and a simple decision tree method was used to classify different land covers.

2.4.3. Land classification using reflectance values

This section analyses accuracy of image classification when red-edge band was available. This study assessed whether there were any significant differences between the classification results obtained with and without the red-edge band. The supervised classification procedure classified the unknown points by using the spectral information of the training samples. Gong, Shen, and

Table 2. JM distances computed with and without consideration of the red-edge band.

	4 bands without red-edge			
	Peanuts	Rice paddies	Bare earth	Water
Peanuts	-	1.144	1.374	1.410
Rice paddies		-	1.413	1.414
Bare earth			-	1.400
Water				-
	5 bands with red-edge			
	Peanuts	Rice paddies	Bare earth	Water
Peanuts	-	1.163	1.376	1.412
Rice paddies		-	1.414	1.414
Bare earth			-	1.402
Water				-
	Improvement			
	Peanuts	Rice paddies	Bare earth	Water
Peanuts	-	0.019	0.002	0.002
Rice paddies		-	0.001	0.000
Bare earth			-	0.002
Water				-

Lu (2019) mentioned that a large number of classifiers have been developed and that the MLC and SVM have emerged as the two representative classifiers in recent decades. The MLC employs probability density functions for classifying an unidentified pixel by computing the probability of the pixel value belonging to each category (Jensen 1996). The SVM aims to find a margin in the feature space to separate different classes, with the selected hyperplane being the one that maximizes the margins. If a hyperplane is farthest from its support vector, it is called the best classification hyperplane. The common unsupervised classifications employed are the ISODATA (Jensen 1996) and K-means (Jensen 1996) clustering method. The ISODATA classification method involves calculating the Euclidean distance from a point to the center of any category. If the distance to the center of the category is the shortest, it is classified into this category, with the separation between the categories being the largest. K-means divides data into the number of groups specified by the user during classification and finds the category with the shortest Euclidean distance from each point to the center of each category. All these four classifiers are used to examine the effectiveness of the red-edge band.

3. Results and discussions

3.1. Analysis of training samples

3.1.1. Jeffries–Matusita distance for the training data sets

In this study, the JM distance was calculated to examine the separability of the training samples between peanut, rice paddy, bare earth, and water. Table 2 compared the JM distances calculated with and without the red-edge band. The JM distance of 1.386 was used as a criterion to discriminate between two classes. Regardless of whether it was calculated with or without considering the red-edge band, the JM distances of peanut/rice paddy and peanut/bare earth in the crop fields were less than 1.386, implying that it was difficult to separate these two classes. Although the JM distance of peanut/rice paddy and peanut/bare earth was less than 1.386, the additional red-edge band increased the JM distance of peanut/rice paddy from 1.144 to 1.163 and that of peanut/bare earth from 1.374 to 1.376. The improvement rate of peanut/rice paddy was 0.019, which was relatively higher than the other in-between classes. When an additional red-edge band was available, the other in-between classes also showed marginal improvement.

3.1.2. Principal component analysis of training set

Figure 3 shows the factor loadings obtained using entire images. The explanations of the first and second principal components were 84% and 13%, respectively. Therefore, the first and second principal components accounted for over 97% of the explanation; these two components were more explanatory than the other components. We will only discuss the first two principal components because the last three principal components exhibited lower levels of variation. The factor loadings of the blue, green, red, and red-edge bands were greater than 0.5 for the first principal component. Only the red-edge and NIR bands had factor loadings greater than 0.5 for the second principal component. These results indicated that all the bands contributed to the first two principal components. Moreover, the contributions of the red-edge band to both the principal components were more descriptive.

To obtain a better understanding of the case with different types of categorized land covers, we extracted the vegetation, bare earth, and water regions to calculate the factor loadings. Table 3 shows the factor loadings in the vegetation regions, and the last column presents the summation of the first two principal components. The summation of the first two principal components of the NIR band (i.e. 0.990) was the most explanatory and that of the red-edge band (i.e. 0.981) was the second explanatory for the vegetation region. Tables 4 and 5 show the factor loadings of the water and bare earth regions. The red-edge band had the lowest factor loading for the bare earth region. In other words, the red-edge band made a small contribution to the bare earth region.

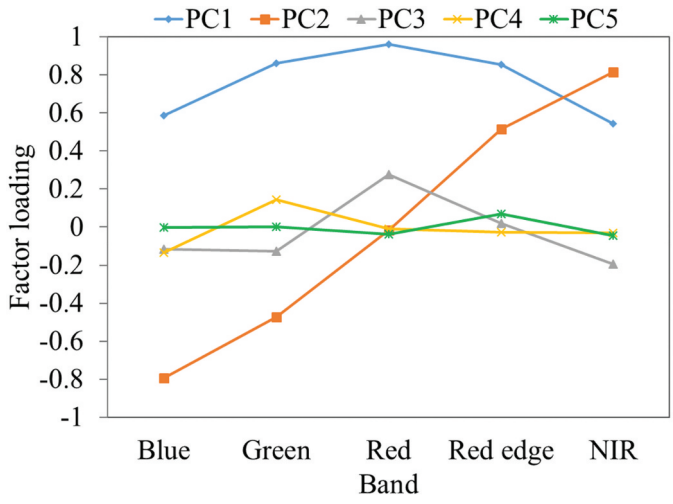

Figure 3. Factor loadings of entire area.

In addition, the red-edge band made the second most contribution to the vegetation region and made the least contribution to the bare earth region. However, it made the most contribution to the water region. Across all the study areas, the explanatory variation of each band exceeded 90%; therefore, each band contributed meaningfully to the principal components. In the vegetation area, each band made an important contribution to the first principal component, and the factor loadings of the red-edge and NIR bands revealed that these bands offered more information for identifying vegetation. In the water body, the red-edge band revealed the highest factor loading among the five bands; thus, the red-edge band made a weighty contribution in the water region.

3.2. Land cover classification using different vegetation indices

If the OA was higher than 80% and Kappa was higher than 0.6, the classification result was considered good (Landis and Koch 1977; Fitzgerald and Lees 1994). The results of the two categories were summarized in Table 7. When using different VIs, the classification accuracies, from high to low, were in the following order: CMFI > MCARI > TCARI > OSAVI > NDVI > CARI. Because most of the VIs showed OA values higher than 90%, they yielded good classification results, except the CARI (the blue band was included in the CARI). To compare the accuracies obtained using the NIR and red-edge bands, after changing the NDVI, CMFI, and OSAVI from the NIR band to the red-edge band, only the accuracy of the NDVI improved. There was a significant difference between the results when the red-edge band was considered and ignored for the CMFI and OSAVI in the Z-test; however, there was no significant difference for the NDVI. It can thus be inferred that the NIR band was more effective for two-category classification than the red-edge band. In terms of the VIs obtained using the red-edge band, the NDVI_RE and CMFI_RE were superior to the OSAVI_RE. Therefore, this study suggests that the NDVI or CMFI be used as a VI in the red-edge band. To summarize, different VIs were used to divide land cover into vegetation and non-vegetation areas, and most VIs showed good classification accuracies. The top-five VIs were CMFI, CMFI_RE MCARI, TCARI, and NDVI_RE.

Although the results reported by Adelabu, Mutanga, and Adam (2014) revealed a 20% improvement with the use of the red-edge band for classifying insect defoliation, this study found that the OA difference between the classification of

Table 3. Factor loadings in vegetation region.

	PC_1	PC_2	PC_3	PC_4	PC_5	$PC_1^2+ PC_2^2$
Blue	0.940	−0.299	−0.072	0.140	0.038	0.974
Green	0.978	−0.095	0.182	0.020	−0.040	0.965
Red	0.913	−0.378	−0.079	−0.114	−0.058	0.977
Red-edge	0.958	0.251	0.010	−0.076	0.117	0.981
NIR	0.787	0.608	−0.060	0.032	−0.071	0.990

Table 4. Factor loadings in bare earth region.

	PC_1	PC_2	PC_3	PC_4	PC_5	$PC_1^2+ PC_2^2$
Blue	0.990	−0.106	−0.079	−0.050	−0.029	0.990
Green	0.995	−0.029	−0.056	0.073	−0.002	0.992
Red	0.995	−0.083	0.007	−0.024	0.051	0.997
Red edge	0.974	0.186	0.131	0.001	−0.020	0.983
NIR	0.035	0.999	−0.034	−0.005	0.005	0.999

Table 5. Factor loadings in water region.

	PC_1	PC_2	PC_3	PC_4	PC_5	$PC_1^2+ PC_2^2$
Blue	0.585	−0.792	−0.115	−0.134	−0.001	0.969
Green	0.860	−0.472	−0.128	0.145	0.001	0.963
Red	0.960	−0.015	0.277	−0.008	−0.036	0.922
Red edge	0.854	0.515	0.021	−0.026	0.068	0.994
NIR	0.545	0.814	−0.195	−0.030	−0.044	0.959

Table 6. Equations of all vegetation indices.

Index	Equation	Reference
NDVI	$\frac{\rho_{NIR}-\rho_{Red}}{\rho_{NIR}+\rho_{Red}}$	(Rouse et al. 1974)
NDVI_RE	$\frac{\rho_{RE}-\rho_{Red}}{\rho_{RE}+\rho_{Red}}$	(Rouse et al. 1974)
CMFI	$\frac{\rho_{Red}}{\rho_{NIR}+\rho_{Red}}=\frac{1}{2}*\left(1-\frac{\rho_{NIR}-\rho_{Red}}{\rho_{NIR}+\rho_{Red}}\right)$	(Wan et al. 2017)
CMFI_RE	$\frac{\rho_{Red}}{\rho_{RE}+\rho_{Red}}=\frac{1}{2}*\left(1-\frac{\rho_{RE}-\rho_{Red}}{\rho_{RE}+\rho_{Red}}\right)$	
OSAVI	$\frac{1.16*(\rho_{800}-\rho_{670})}{\rho_{800}+\rho_{670}+0.16}$	(Rondeaux, Steven, and Baret 1996)
OSAVI_RE	$\frac{1.16*(\rho_{800}-\rho_{700})}{\rho_{700}+\rho_{670}+0.16}$	(Rondeaux, Steven, and Baret 1996)
TCARI	$3CA(\rho_{700}-\rho_{670})-0.2\times(\rho_{700}-\rho_{550})5\frac{\rho_{700}}{\rho_{670}})$	(Haboudane et al. 2002)
CARI	$\frac{(\rho_{700}-\rho_{500})*\rho_{670}+\rho_{670}+(\rho_{550}-((\rho_{700}-\rho_{500})*\rho_{670})*\rho_{550}))}{\sqrt{((\rho_{700}-\rho_{500})\times 00I2^2}}$	(Broge and Leblanc 2001)
MCARI	$(\rho_{700}-\rho_{670})-0.2\times(\rho_{700}-\rho_{550})\times 5\frac{\rho_{700}}{\rho_{670}})$	(Daughtry et al. 2000)

_RE: new index with red-edge band instead of NIR band
ρ_{Red},ρ_{RE}, and ρ_{NIR}: Red band, red-edge band, and NIR band
ρ_{500}: Blue band **ρ_{550}**: Greed band **ρ_{670}**: Red band
ρ_{700}: Red-edge band **ρ_{800}**: NIR band

Note. Compiled from Agapiou et al. (2012) and Baluja et al. (2012).

vegetation and non-vegetation areas was only approximately 1%. Most of the VIs attained accuracy levels of 95%. Kim and Yeom (2014) also showed a slight OA improvement in paddy rice classification. The classification accuracy apparently depends on the investigated land-use class.

The results of the accuracy analysis conducted with four classes are summarized in Table 8. The classification accuracy was in the following descending order: CMFI, TCARI, MCARI, NDVI, OSAVI, and CARI. The TCARI and MCARI were narrowband VIs that use the red-edge band. The benefits of the red-edge band for vegetation classification are that these two VIs are superior to traditional NDVI. The NDVI_RE, CMFI_RE, and OSAVI_RE were used as VIs when the band was changed from NIR to red-edge. The α values of all of these indices were smaller than 0.05, indicating a significant difference between the results obtained using the NIR and red-edge bands. When rice paddy and peanuts were classified using the NDVI_RE, the classification accuracy was improved by almost 7%. Moreover, the accuracy of the rice paddy with the CMFI was improved by approximately 6%. The accuracies of the NDVI_RE and CMFI_RE were superior to those of the OSAVI_RE. It can thus be inferred that the red-edge-derived VI made a certain contribution to the vegetation.

3.3. Land classification using reflectance values

In this study, two land cover categories were considered for supervised and unsupervised classification. The first category included water, bare earth, and vegetation. Vegetation was further

Table 7. Classification assessment of all VIs with two categories.

VI	OA (%)	KAPPA	σ_{kappa}	Z-TEST
NDVI	95.2383	0.9049	0.4252	α > 0.05
NDVI_RE	95.6523	0.9132	0.4074	
CMFI	97.9120	0.9582	0.2861	α < 0.05
CMFI_RE	96.7925	0.9358	0.3529	
OSAVI	95.2384	0.9049	0.4252	α < 0.05
OSAVI_RE	84.1507	0.6803	0.7367	
TCARI	96.6249	0.9325	0.3610	-
CARI	59.1537	0.3530	0.8693	-
MCARI	96.6249	0.9325	0.3610	-

Table 8. Classification assessment of all VIs by using four categories.

	Individual accuracy							
Coefficient Index	Water	Bare earth	Rice paddies	Peanuts	OA	Kappa	σ_{kappa}	Ztest
NDVI	100	85.89	50.47	63.61	70.76	0.66	0.0074	α < 0.05
NDVI_RE	99.63	69.94	62.43	70.98	68.80	0.62	0.0059	
CMFI	99.99	81.44	55.21	73.52	74.55	0.54	0.0056	α < 0.05
CMFI_RE	99.95	67.47	61.21	68.26	68.65	0.62	0.0057	
OSAVI	100	82.04	39.42	54.72	70.04	0.62	0.0055	α < 0.05
OSAVI_RE	100	77.58	1.65	51.35	60.25	0.46	0.0054	
TCARI	99.94	69.54	61.03	82.48	72.84	0.59	0.0058	-
CARI	99.09	96.89	44.63	34.76	62.99	0.64	0.0056	-
MCARI	99.84	70.32	59.96	87.78	71.01	0.60	0.0059	-

Table 9. Result of three categories with and without the red-edge band.

Classifier	No. of bands	Water	Bare earth	Vegetation	OA	Kappa	σ_{kappa}	Z-test
MLC	4	100.00	94.61	99.86	98.38	0.97	0.0014	$\alpha < 0.05$
	5	100.00	96.20	99.97	98.90	0.98	0.0011	
SVM	4	100.00	98.35	99.34	99.23	0.99	0.0000	$\alpha > 0.05$
	5	100.00	98.35	99.34	99.23	0.99	0.0010	
ISODATA	4	100.00	69.27	93.94	86.50	0.79	0.0037	$\alpha < 0.05$
	5	100.00	60.80	89.04	80.96	0.71	0.0042	
Kmeans	4	100.00	83.30	99.63	94.39	0.91	0.0025	$\alpha > 0.05$
	5	100.00	84.31	98.94	94.54	0.92	0.0025	

Table 10. Result obtained with four categories, with and without the red-edge band.

Classifier	No. of bands	Water	Bare earth	Rice paddies	Peanut	OA	Kappa	σ_{kappa}	Z-test
MLC	4	100.00	70.42	63.47	86.72	75.10	0.67	0.0057	$\alpha < 0.05$
	5	100.00	71.59	66.50	87.27	77.00	0.69	0.0057	
SVM	4	100.00	82.64	62.71	84.29	80.06	0.73	0.0055	$\alpha < 0.05$
	5	100.00	70.42	63.47	86.72	81.77	0.76	0.0055	
ISODATA	4	100.00	68.42	3.64	58.13	66.05	0.54	0.0053	$\alpha < 0.05$
	5	100.00	64.29	2.08	60.94	63.17	0.50	0.0054	
Kmeans	4	100.00	68.39	3.64	58.14	66.05	0.54	0.0053	$\alpha < 0.05$
	5	100.00	66.72	2.12	59.38	63.36	0.50	0.0054	

divided into peanuts and rice paddy in the second category. A total of four classes were considered: water body, bare earth, rice paddy, and peanut. In this section, the benefits of using the red-edge band and other four bands for land cover classification will be compared. Each classifier was to perform classification twice: once without the red-edge band (a total of four bands) and once with the red-edge band (a total of five bands). The accuracy assessment included the OA and Kappa coefficients. In addition, a pairwise Z-test was conducted to determine whether the two classifications were significantly different.

Regarding the results of the three categories, the OA values of most classifiers were higher than 94% and the Kappa values reached 0.9, except for ISODATA. Most classifiers yielded good classification results and could be used to classify these three categories (Table 9). A comparison of the results obtained, with and without using the red-edge band, revealed that the red-edge band marginally improved the classification results, except for ISODATA. Because the SVM classified these three classes satisfactorily in the hyperplane, there was no significant difference between the set of four bands and five bands in SVM. Further, the p-value of SVM was greater than 0.05. The improvement rate of the MLC in bare earth was approximately 2%. Each p-value was smaller than 0.05, meaning there was a significant difference between the results obtained with and without using the red-edge band. According to the accuracy analysis results, the use of the red-edge band in MLC and SVM contributed to an approximately 2% improvement in classification accuracy for rice paddies and peanuts.

In the results obtained for the four categories (Table 10), vegetation was divided into peanuts and rice paddy. By using the red-edge band as a classification feature, for both supervised and unsupervised classifications, the OA and Kappa improved. Each p-value was smaller than 0.05, meaning there was a significant difference between the results obtained with and without using the red-edge band. However, the accuracies for unsupervised classification for rice paddies and peanuts were less than 60%. According to the accuracy analysis results, the use of the red-edge band in MLC and SVM contributed to an approximately 2% improvement in classification accuracy for rice paddies and peanuts. A comparison between the MLC and SVM, regardless of the red-edge band usage, revealed that SVM was superior to MLC by approximately 5%. In summary, the red-edge band improved the classification of rice paddies and peanuts by approximately 1–3% with supervised classifications. The effectiveness of the red-edge band was related to the classification targets. Owing to the different classification targets, the improvement in accuracy achieved by this study by using the red-edge band (i.e. 2%) was different from that achieved by Adelabu, Mutanga, and Adam (2014) (i.e. 20%) but similar to the result obtained by Saini and Ghosh (2019) (i.e. 3%). The accuracies of classification were related to land-use classes. The OA of the supervised classifier improved when classifying rice paddy and peanuts. Moreover, the Z-test indicated a significant difference between the results obtained using four and five bands. These findings imply that the red-edge band slightly improved the accuracy and influenced the classification process.

4. Conclusions and future works

The objective of this study was to evaluate the benefits of using the red-edge band of RapidEye for land cover classification. To this end, the effectiveness of the red-edge band was examined by using training samples, VIs, and land cover classification. We employed training samples to evaluate the effectiveness of the red-edge band by factor loading, which in the context of the red-edge band, is rarely considered in the literature. While RapidEye is the first satellite to implement the red-edge band, many satellite sensors can include the red-edge band to increase the number of multi-spectral bands. The major findings of this study are that there was only a slightly improvement in accuracy when the red-edge band was considered for particular land covers. In contrast to the previous studies on this topic, our study provides a conservative view of the broadband red-edge band and offers a different perspective on red-edge classification accuracy. Following experimentation, we

identified an issue with selecting the broadband red-edge by providing solid proof. This study demonstrated that the effectiveness of red-edge classification accuracy apparently depends on the investigated land-use class. The conclusions and suggestions are summarized as follows:

(1) During the evaluation of factor loadings by means of PCA, the contribution of the red-edge band to the categorization of vegetation was considered more important than that of the visible bands. The contribution of the red-edge band was similar to that of the NIR band.
(2) The classification of VIs by using the red-edge band can effectively separate vegetation and non-vegetation regions. The OA was found to be higher than 90%. These results indicated that the CMFI, NDVI, and MCARI were relatively high-precision VIs. By using the red-edge band instead of NIR in the NDVI and CMFI, the OA of rice paddy classification can be improved by 6–7%.
(3) Based on the accuracy analysis results of land cover classification, SVM was considered more accurate than MLC, ISODATA, and K-means. The red-edge band improved the accuracy of peanut classification by approximately 2%. Regardless of the red-edge band usage, the performance of SVM was approximately 5% better than that of the MLC.

This study focused on the availability of the broadband red-edge from RedEdge. Results can be better interpreted by using the narrow-band red-edge compared to the broadband red-edge. Therefore, our future research will aim at analyzing the four narrow red-edge bands from Sentinel-2. Moreover, there have only been a few studies that have applied spectral simulation for RapidEye and Sentinel-2 spectral datasets. Our future work will also analyze the effectiveness of the red-edge band on real and spectral simulation datasets.

Acknowledgments

The authors would like to thank Taiwan Agricultural Research Institute for providing farmland cover map.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported in part by the Ministry of Interior, Taiwan under grant number [111CCL030C].

ORCID

Tee-Ann Teo http://orcid.org/0000-0002-7625-5662

References

Adelabu, S., O. Mutanga, and E. Adam. 2014. "Evaluating the Impact of Red-Edge Band from Rapideye Image for Classifying Insect Defoliation Levels." *ISPRS Journal of Photogrammetry and Remote Sensing* 95: 34–41. doi:10.1016/j.isprsjprs.2014.05.013.

Agapiou, A., D.G. Hadjimitsis, and D.D. Alexakis. 2012. "Evaluation of Broadband and Narrowband Vegetation Indices for the Identification of Archaeological Crop Marks." *Remote Sensing* 4 (12): 3892–3919. doi:10.3390/rs4123892.

Baluja, J., M.P. Diago, P. Balda, R. Zorer, F. Meggio, F. Morales, and J. Tardaguila. 2012. "Assessment of Vineyard Water Status Variability by Thermal and Multispectral Imagery Using an Unmanned Aerial Vehicle (UAV)." *Irrigation Science* 30 (6): 511–522. doi:10.1007/s00271-012-0382-9.

Broge, N.H., and E. Leblanc. 2001. "Comparing Prediction Power and Stability of Broadband and Hyperspectral Vegetation Indices for Estimation of Green Leaf Area Index and Canopy Chlorophyll Density." *Remote Sensing of Environment* 76 (2): 156–172. doi:10.1016/S0034-4257(00)00197-8.

Congalton, R., and K. Green. 1999. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. Boca Raton, Florida: CRC/ Lewis Publishers.

Daughtry, C.S.T., C.L. Walthall, M.S. Kim, E.B. De Colstoun, and J.E. Mc Murtrey. 2000. "Estimating Corn Leaf Chlorophyll Concentration from Leaf and Canopy Reflectance." *Remote Sensing of Environment* 74 (2): 229–239. doi:10.1016/S0034-4257(00)00113-9.

Fitzgerald, R. W., and B. G. Lees. 1994. "Assessing the Classification Accuracy of Multisource Remote Sensing Data." *Remote Sensing of Environment* 47 (3): 362–368. doi:10.1016/0034-4257(94)90103-1.

Gong, X., L. Shen, and T. Lu. 2019. "Refining Training Samples Using Median Absolute Deviation for Supervised Classification of Remote Sensing Images." *Journal of the Indian Society of Remote Sensing* 47 (4): 647–659. doi:10.1007/s12524-018-0887-7.

Haboudane, D., J.R. Miller, N. Tremblay, P.J. Zarco-Tejada, and L. Dextraze. 2002. "Integrated Narrow-Band Vegetation Indices for Prediction of Crop Chlorophyll Content for Application to Precision Agriculture." *Remote Sensing of Environment* 81 (2): 416–426. doi:10.1016/S0034-4257(02)00018-4.

Hair, J.F., W.C. Black, B.J. Babin, R.E. Anderson, and R.L. Tatham. 1998. *Multivariate Data Analysis*. New Jersey: Prentice Hall Incorporation.

Jensen, J.R. 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective*. New Jersey: Prentice Hall Press.

Kim, H.O., and J.M. Yeom. 2014. "Effect of Red-Edge and Texture Features for Object-Based Paddy Rice Crop Classification Using Rapideye Multispectral Satellite Image Data." *International Journal of Remote Sensing* 35 (19): 7046–7068. doi:10.1080/01431161.2014.965285.

Landis, J. R., and G. G. Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics* 33 (1): 159–174. doi:10.2307/2529310.

Li, X.J., G. Chen, J.Y. Liu, W.T. Chen, X.W. Cheng, and Y.W. Liao. 2017. "Effects of Rapideye Imagery's Red-Edge Band and Vegetation Indices on Land Cover Classification in an Arid Region." *Chinese Geographical Science* 27 (5): 827–835. doi:10.1007/s11769-017-0894-6.

Lyon, J.G., D. Yuan, R.S. Lunetta, and C.D. Elvidge. 1998. "A Change Detection Experiment Using Vegetation Indices." *Photogrammetric Engineering and Remote Sensing* 64 (2): 143–150. doi:10.13140/R.G.2.2.32449.25449.

Meng, L., Y. Huang, N. Zhu, Z. Chen, and X. Li. 2021. "Mapping Properties of Vegetation in a Tidal Salt Marsh from Multispectral Satellite Imagery Using the SCOPE Model." *International Journal of Remote Sensing* 42 (2): 422–444. doi:10.1080/01431161.2020.1809737.

Munyati, C. 2004. "Use of Principal Component Analysis (PCA) of Remote Sensing Images in Wetland Change Detection on the Kafue Flats, Zambia." *Geocarto International* 19 (3): 11–22. doi:10.1080/10106040408542313.

Rondeaux, G., M. Steven, and F. Baret. 1996. "Optimization of Soil-Adjusted Vegetation Indices." *Remote Sensing of Environment* 55 (2): 95–107. doi:10.1016/0034-4257(95)00186-7.

Rouse, J.W., R.H. Haas, D.W. Deering, and J.A. Schell, 1974. "Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation." Final Report, RSC 1978–4, Remote Sensing Center, Texas A&M Univ., College Station. pp. 371.

Saegusa, R., H. Sakano, and S. Hashimoto. 2004. "Nonlinear Principal Component Analysis to Preserve the Order of Principal Components." *Neurocomputing* 61: 57–70. doi:10.1016/j.neucom.2004.03.004.

Saini, R., and S.K. Ghosh. 2019. "Analyzing the Impact of Red-Edge Band on Land Use Land Cover Classification Using Multispectral Rapideye Imagery and Machine Learning Techniques." *Journal of Applied Remote Sensing* 13 (4): 044511. doi:10.1117/1.JRS.13.044511.

Schuster, C., M. Förster, and B. Kleinschmit. 2012. "Testing the Red Edge Channel for Improving Land-Use Classifications Based on

High-Resolution Multispectral Satellite Data." *International Journal of Remote Sensing* 33 (17): 5583–5599. doi:10.1080/01431161.2012.666812.

Shafri, H.Z.M., M.I. Anuar, and M.I. Saripan. 2009. "Modified Vegetation Indices for Ganoderma Disease Detection in Oil Palm from Field Spectroradiometer Data." *Journal of Applied Remote Sensing* 3 (1): 033556. doi:10.1117/1.3257626.

Swain, P. H., and S. M. Davis. 1978. *Remote Sensing : The Quantitative Approach*. New York: McGraw-Hill International Book.

Teo, T. A., and S. H. Huang. 2013. "Automatic co-registration of Optical Satellite Images and Airborne LiDAR Data Using Relative and Absolute Orientations." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6 (5): 2229–2237. doi:10.1109/JSTARS.2012.2237543.

Wan, S., S. H. Chang, C. T. Peng, and Y. K. Chen. 2017. "A Novel Study of Artificial Bee Colony with Clustering Technique on Paddy Rice Image Classifications." *Arabian Journal of Geosciences* 10 (9): 1–13. doi:10.1007/s12517-017-2992-2.

Weih, R.C., and N.D. Riggan. 2010. "Object-Based Classification VS. Pixel-Based Classification: Comparative Importance of Multi-Resolution Imagery." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 38 (4): C7.