Article

Mapping Rubber Tree Stand Age Using Pléiades Satellite Imagery: A Case Study in Thalang District, Phuket, Thailand

Werapong Koedsin[1,a,]* **and Alfredo Huete**[2,b]

1 Tropical Plants Biology Research Unit, Faculty of Technology and Environment, Prince of Songkla University, Phuket Campus, Kathu, Phuket, Thailand
2 Plant Functional Biology and Climate Change Cluster (C3), University of Technology, Sydney, NSW 2007, Australia
E-mail: [a]werapong.g@phuket.psu.ac.th (Corresponding author), [b]alfredo.huete@uts.edu.au

Abstract. The rubber stand age information is an important variable for determining the distribution of carbon pools and fluxes in rubber plantation ecosystems as well as for production management. This study demonstrates the capability of high spatial resolution satellite imagery as Pléiades Satellite Imagery with the help of feature selection (i.e., Sequential Forward Floating Selection) to improve the accuracy of rubber stand age mapping at a part of Thalang district, Phuket, Thailand. The 238 sample plots were used to classification and accuracy assessment. This study found that the Pléiades imagery with the help of Sequential Forward Floating Selection can successfully classify rubber stands age between less than 7 years, 7-15 years and more than 15 years, respectively. The total testing accuracy was improved from 94.07% to 94.92% and 96.61% after applying the Principal Component and the Sequential Forward Floating Selection algorithms, respectively. Since the methodology proposed in this study can accurately classify 3 classes of rubber stand age, it is anticipated that this methodology can be used as a guideline for rubber tree stand age mapping in other study areas.

Keywords: Rubber tree, *Hevea brasiliensis*, remote sensing, age classification, feature selection.
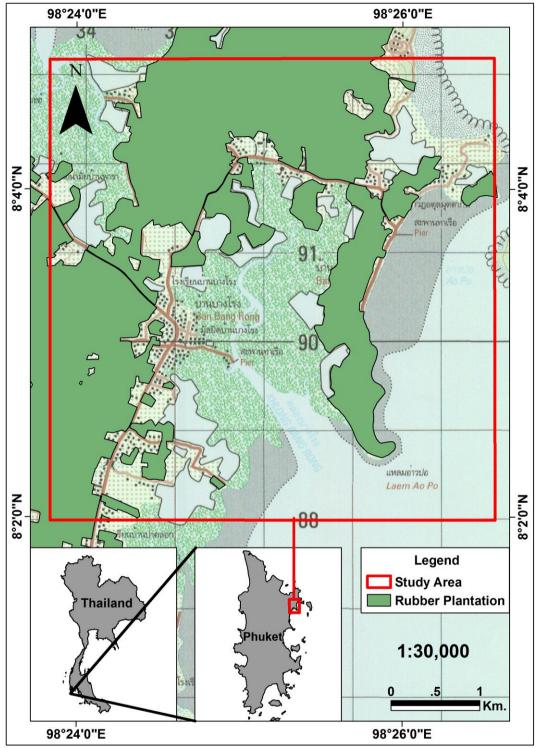
1. Introduction

Rubber tree (*Hevea brasiliensis* Muell.) plays an important role as a source of natural rubber and wood products [1–3]. Over the last several decades rubber plantations have expanded by more than a million hectares in mainland areas of Thailand, China, Laos, Vietnam, Cambodia and Myanmar [4–7]. Southeast Asia accounts for 97% of the global natural rubber supply [7, 8]. Rapid expansion of rubber plantations has occurred in this area due to the increasing demand for rubber products. This large expansion of rubber plantations plays a crucial role in changing regional environments that significantly affect human well-being and ecosystem services (e.g. climate, carbon stocks, and biodiversity) [9, 10]. Accurate information of rubber plantation parameters such as area, age, volume, and spatial distribution, are therefore, needed for effective production management or ecosystem modelling [2, 11]. The stand age of plants, is one important variable for determining the distribution of carbon pools and fluxes in forest ecosystems as well as for production management [12–15]. Unfortunately, the acquisition of rubber plantation parameters by field survey is difficult over vast areas as it is time-consuming, expensive, and labor-intensive [11]. However, such difficulties have been significantly improved by earth observation remote sensing techniques. It is obvious that remote sensing instruments are now operationally used for mapping and monitoring rubber tree at the broad level [1, 2, 6, 8-9, 11, 16–20]. A few of related studies claim the advantage of exploiting such remote sensing data to characterize the rubber stand age with regression and classification [2, 6, 11, 19]. The previous work [2] found that near-infrared and middle-infrared bands of TM data could provide the best correlation with the rubber stand age in Malaysia. Moreover, [14] found that linear combination of NDVI, ETM4/ETM3 and the brightness component of the tasseled cap transformation [21] were better predictors of stand age than any other combination of transformed ETM+ bands. Recently, [19] used multispectral bands, NDVI and Principal Component transformation as input of rubber stand age classification in Northeast Thailand. It is evident that the spectral bands as well as the vegetation indices are useful to use as input of rubber stand age classification. However, the use of several bands and similarity of vegetation indices subject to co-linearity (i.e., redundant spectral information) also imposes the risk of over fitting when the classification is performed [22]. Fortunately, such problems can be handled by techniques of feature selection. In the remote sensing literature, feature selection is one of the popular approaches to reduce spectral dimension [22–24].
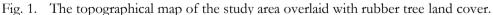
Consequently, this study investigates further into the potential of remote sensing for rubber stand age classification. The aim of this work is to prove for the first time whether the capability of high spatial resolution satellite imagery with the help of feature selection (i.e., Sequential Forward Floating Selection (SFFS)) can improve the accuracy of rubber stand age mapping in south of Thailand (i.e., Paklok subdistrict, Thalang district, Phuket province). The satellite data used is the Pléiades multispectral image. The final classification results were statistically tested against independent test data sets.

2. Methodology

2.1. Study Sites

The study site is at a part of the Paklok sub district, Thalang district, Phuket province, in the south of Thailand. It is located between 98° 23' 50.4306" E to 8° 4' 47.6754"N (Fig. 1).The average elevation in the region is 21 meters above mean sea level. The climate of the study area is tropical with a mean annual temperature of 27.6^0C and a mean annual rainfall of 2240 mm. The dry period occurs during April to November, and the rest of the year is dominated by the monsoons (rainy season).It is the favorable climate for rubber tree. The rubber tree (*Hevea brasiliensis* species) is a main crop in this area: nearly 900 ha were planted with rubber trees.

Fig. 1.   The topographical map of the study area overlaid with rubber tree land cover.

2.2.   Image Pre-Processing and Field Data Collection

The Pléiades image was acquired on March 17, 2013, within the rubber tapping season. The image has 4 bands and 2 meter spatial resolution (see the characteristics of the satellite in Table 1). The image was then re-projected into a UTM zone 47N and was geometrically corrected by using 21 ground control points (GCPs) gathered throughout the image utilizing features such as road intersections, bridges, and other easily identifiable landscape features. Aerial orthoimages acquired in 2002 of Phuket area, obtained from the Land Development Department, Ministry of Agriculture and Cooperatives, Thailand were used as reference

maps. Nearest neighbor resampling was used in the geometrical transformations to minimize statistical properties changes of the datasets [25].The final positional accuracy of the image after resampling is less than the size of one pixel. The image was then applied radio metric correction to normalize satellite images for factors such as sensor degradation, Earth–Sun distance variation, incidence angle, view angle, and time of data gathering. The process involved converting Digital Number (DN) into radiance using calibration coefficients, consequently, the simple algorithm to correct the atmospheric noise as QUick Atmospheric Correction (QUAC) was used to correct the atmospheric noise.

Table 1.   The characteristic of Pléiades Satellite Imagery [26].

Resolution	Panchromatic 50 cm Multispectral 2.0 m
Nominal swath width	20 km at Nadir
Bands	Pan: 450-830 nm Blue: 430-550 nm Green: 500-620 nm Red: 590-710 nm Near IR: 740-940 nm

The field data collection was conducted during the rubber tapping season between March and April, 2013. The land-use map from Land Development Department, Ministry of Agriculture and Cooperatives was used for the identification of the rubber planted area. A stratified random sampling method was used for locating the rubber tree sampling plots. The stratification was done by clustering the study area into 6 clusters using a K-Mean method. The date of planting from Phuket Office of the rubber replanting aid fund, Ministry of Agriculture and Cooperatives and information from the landowner were used for documenting the age of each stand. The relative positioning technique was used to record the position in the Universal Transverse Mercator (UTM) coordinate system of each stand. Then, the stand age of each sampling station was classified into three age intervals (i.e., less than 7 years, 7- 15 years and more than 15 years). There were 237 sampling stations in total throughout the study area (see Fig. 2). The stations were randomly divided into two groups for the purpose of image classification and validation (Table 2).
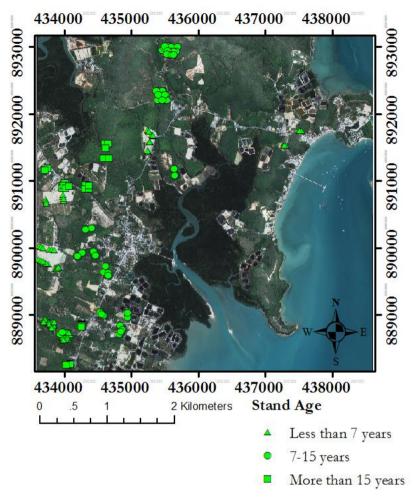
Fig. 2.   The distribution of sample plots throughout the study area.

Table 2.   The number of training and testing samples per stand age.

Rubber stand age (years)	Training samples	Testing samples
Less than 7	69	69
7 to 15	28	27
More than 15	22	22

2.3.   Vegetation Indices

In this study, three commonly used vegetation indices were used as the input of the feature selection technique. The Normalized Difference Vegetation Index (NDVI) [27], Enhanced Vegetation Index (EVI) [28] and Soil-adjusted Vegetation Index (SAVI) [29] were computed as the Eq. (1)–(3).

2.3.1.   Normalized Difference Vegetation Index

$$NDVI = \frac{NIR - \mathrm{Re}\,d}{NIR + \mathrm{Re}\,d} \tag{1}$$

where Red and NIR stand for the spectral reflectance measurements acquired in the red and near-infrared band, respectively.

2.3.2. Enhanced Vegetation Index

$$EVI = 2.5 * \frac{(NIR - \text{Re}\,d)}{(NIR + C_1 * \text{Re}\,d - C_2 * Blue + L)} \qquad (2)$$

where Blue, Red and NIR stand for the spectral reflectance measurements acquired in red blue and near-infrared bands, respectively. And C1, C2, and L are coefficients to correct for atmospheric condition. (For standard values are L=1, C1=6, and C2=7.5).

2.3.3. Soil-adjusted Vegetation Index

$$SAVI = \frac{(NIR - \text{Re}\,d)}{(NIR + \text{Re}\,d + L)} * (1 + L) \qquad (3)$$

where NIR and Red stand for the spectral reflectance measurements acquired in the visible (red) and near-infrared regions, respectively and L is the soil brightness correction factor (L=0.5 is the default value used).

2.4. Sequential Forward Floating Search and Classification

In this study, the Sequential Forward Floating Search (SFFS) was used as the feature selection protocol. The all multispectral bands, three selected vegetation indices (i.e., NDVI, EVI and SAVI) were used as the input features of SFFS procedure. The SFFS method is a suboptimal search algorithm that collects the spectral features that have the highest objective values until the number of features reaches the pre-defined critical number selected by the user. The three steps for the SFFS algorithm are detailed as follows. The first step (inclusion) recognizes the most significant (i.e., highest objective function value) feature amongst the unselected features with respect to the current set of selected features. The second step (conditional exclusion) identifies the least significant feature in the current feature set and removes it except it is the feature just added in step 1. Finally, is a continuation of the feature exclusion process that is based on the significance of the least significant feature in the current feature set as compared to previous feature sets of the same cardinality [30]. In this study the Mahalanobis distance was used as the objective function [31].

The best features combination from SFFS result was used as the input band for Maximum Likelihood classification. The multispectral band data (i.e., RGB and Near Infrared Band (NIR)) were analyzed using Principal Component Analysis (PCA) and the results compared to the SFFS in order to see if there was any bias in the final classification results. The summarized of the classification process were showed as Fig. 3.
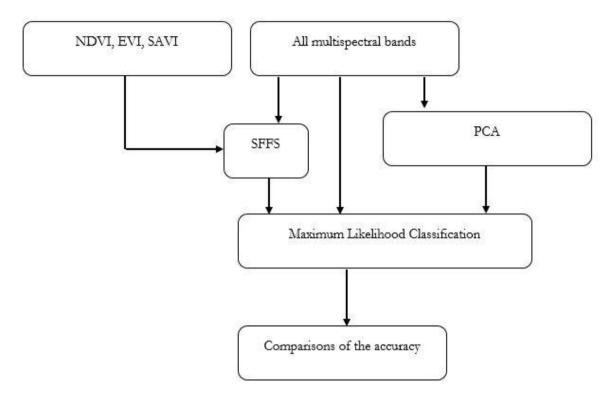
Fig. 3.   The flowchart of classification and comparisons.

3.   Results and Discussion

For searching the appropriate features combination, the all multispectral bands and vegetation indices (i.e., NDVI, EVI and SAVI) were used as input features for SFFS methods. This study used features sizes that were varied from 1 to 6. The number of features with the objective function values are shown in Fig. 4. It was found that the use of two features (i.e., NIR and EVI) gained the highest Mahalanobis distance of 11.10. In addition, it was found that each features size was selected using NIR and EVI data except in the case of one feature which was based only on SAVI (See Fig. 4). This result agrees with the previous work of [2, 11] who reported that the near and middle infrared bands of Landsat TM data could provide the best correlation with rubber stand age and volume of rubber tree.

The features selected using SFFS were used as input bands for Maximum Likelihood classification. For the purpose of comparison, all spectral bands and PCA with 2 components were used for classification. All the classification methods used the same training and testing datasets to avoid bias. The classification results are shown in Table 3. The total testing accuracy was improved from 94.07% to 94.92% and 96.61% after applying the PCA and the SFFS algorithm, respectively. The rubber stand age distribution from the best classification method (i.e., SFFS) is shown in Fig. 5. The final classification image somewhat appeared more than one stand age in same rubber plot, it is caused by used the pixel-based sampling as stratified random sampling method. The confusion matrix is shown in Table 4 indicates the degree of incorrect classification based of the evidence of field ground-truth observation. Confusion mostly occurred between 7-15 year old stands and more than 15 year old stands. Fortunately, the confusion between classes was decreased after application of the SFFS algorithm. In particular, the confusion between 7-15 year old and more than 15 year old classes were improved (compare the highlighted area and user's accuracy in Table 4). The improved accuracy might have resulted from the selected features used in the SFFS analysis were strongly correlated with the stand age of the trees [2, 11, 14]. The previous works [2] and [11] reported that near-infrared and middle-infrared bands of TM data could provide the best correlation with rubber stand age. Likewise, [14] found that linear combination of NDVI, ETM4/ETM3 and the brightness component of the tasseled cap transformation were better predictors of stand age than any other combination of transformed ETM+ bands. The present study has shown that NIR information and combination of transformed bands (such as vegetation indices, NDVI, EVI and SAVI) can be used to accurately estimate

the stand age of rubber trees. Although, the study area is small but we hope the results and methodology in this study can be used as guideline information for future study in other area in South of Thailand and Malaysia.
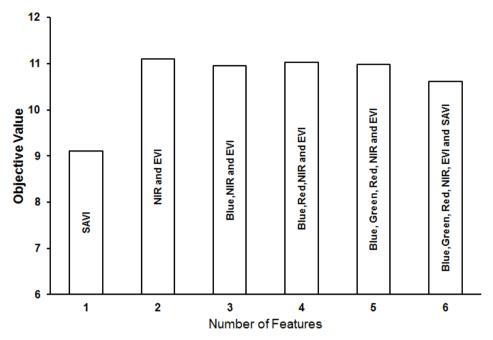

Fig. 4.   A comparison between the class separability of six different number-of-features varied from 1 to 6 features selected by the SFFS algorithm.

Table 3.   Results of Accuracy Assessment. Where RGB is Red, Green and Blue Bands, OA is Overall Accuracy and Kappa is the Kappa coefficient.

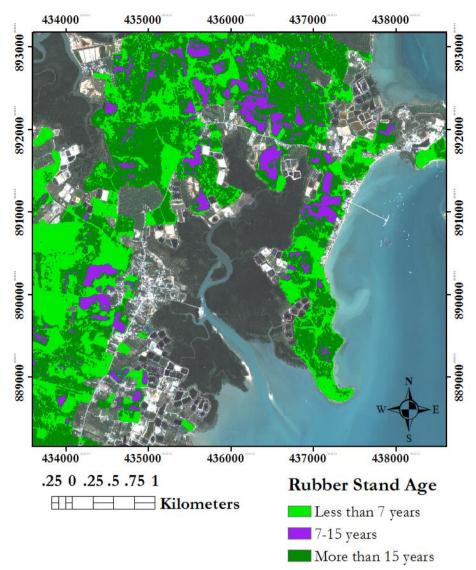| Methods	OA	Kappa
All spectral bands (RGB+NIR)	94.07	0.8970
PCA (2 components)	94.92	0.9112
SFFS (NIR + EVI)	96.61	0.9407

Fig. 5.   The rubber stand age distribution of the study area when using Maximum Likelihood classification aided by SFFS feature selection.

Table 4.   (a) The confusion matrix, producer's and user's accuracy of the all spectral bands (Overall Accuracy = 94.07%), (b) The confusion matrix, producer's and user's accuracy of the transformed band by PCA methods (Overall Accuracy = 94.92%), and (c) The confusion matrix, producer's and user's accuracy of the bands combination selected by the SFFS search algorithm (Overall Accuracy = 96.61%).

(a) All Spectral bands			
Class (years)	Less than 7	7-15	More than 15	Total
Less than 7	68	0	0	68
7-15	0	22	1	23
More than 15	1	5	21	27
Total	69	27	22	118

Class (years)	Producer's accuracy %	User's accuracy %
Less than 7	98.55	100.00

7-15	81.48	95.65
More than 15	95.45	77.78

(b) PCA methods

Class (years)	Less than 7	7-15	More than 15	Total
Less than 7	69	0	0	69
7-15	0	22	1	23
More than 15	0	5	21	26
Total	69	27	22	**118**

Class (years)	Producer's accuracy	User's accuracy
Less than 7	100.00	100.00
7-15	81.48	95.65
More than 15	95.45	80.77

(c) SFFS methods

Class (years)	Less than 7	7-15	More than 15	Total
Less than 7	69	0	0	69
7-15	0	24	1	25
More than 15	0	3	21	24
Total	69	27	22	**118**

Class (years)	Producer's accuracy	User's accuracy
Less than 7	100.00	100.00
7-15	81.89	96.00
More than 15	95.45	87.50

4. Conclusion

This study investigated the potential of remote sensing for rubber stand age classification. The all of Pléiades multispectral bands and three selected vegetation indices were used as the input bands of SFFS for selection the best feature used in classification process. The mapping accuracy of all multispectral bands and PCA were compared with the SFFS result in order to see if there was any bias in the final classification results.

The results demonstrated that high spatial resolution satellite imagery with the help of SFFS can successfully classify rubber stands age. The overall accuracy was improved from 94.07% to 94.92% and 96.61% after applying the PCA and the SFFS algorithms, respectively. In addition, this study has shown that NIR band and combination of transformed bands as EVI can be used to accurately estimate the stand age of rubber trees. Since the methodology proposed in this study can accurately classify 3 classes of rubber stand age, it is anticipated that this methodology can be used as a guideline for rubber tree stand age mapping in other study areas. This is important management information. The next stage of this work will be compare the pixel based with object-based technique to mapping the rubber tree stand age.

Acknowledgments

References

[1] M. N. Suratman, G. Q. Bull, D. G. Leckie, V. LeMay, and P. L. Marshall, "Modelling attributes of Rubberwood (*Hevea brasiliensis*) stands using spectral radiance recorded by Landsat Thematic Mapper in Malaysia," in *Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International*, vol. 4, pp. 2087–2090.

[2] M. N. Suratman, G. Q. Bull, D. G. Leckie, V. M. Lemay, P. L. Marshall, and M. R. Mispan, "Prediction models for estimating the area, volume, and age of rubber (*Hevea brasiliensis*) plantations in Malaysia using Landsat TM data," *Int. For. Rev.*, vol. 6, no. 1, pp. 12–12, 2004.

[3] L. J. L. W. X. Guishui and A. Fen, "Status of rubber-wood processing and utilization in Hainan Province," *Chin. J. Trop. Agric.*, vol. 3, p. 67–70, 2010.

[4] A. D. Ziegler, J. M. Fox, and J. Xu, "The rubber juggernaut," *Science*, vol. 324, no. 5930, pp. 1024–1025, 2009.

[5] C. C. Mann, "Addicted to rubber," Sci., vol. 325, no. 5940, pp. 564–566, 2009.

[6] Z. Li and J. M. Fox, "Rubber tree distribution mapping in Northeast Thailand," *Int. J. Geosci.*, vol. 2, no. 4, 2011.

[7] J. Fox and J.-C. Castella, "Expansion of rubber (*Hevea brasiliensis*) in Mainland Southeast Asia: What are the prospects for smallholders?," *J. Peasant Stud.*, vol. 40, no. 1, pp. 155–170, 2013.

[8] Z. Li and J. M. Fox, "Mapping rubber tree growth in mainland Southeast Asia using time-series MODIS 250 m NDVI and statistical data," *Appl. Geogr.*, vol. 32, no. 2, pp. 420–432, 2012.

[9] J. Dong, X. Xiao, B. Chen, N. Torbick, C. Jin, G. Zhang, and C. Biradar, "Mapping deciduous rubber plantations through integration of PALSAR and multi-temporal Landsat imagery," *Remote Sens. Environ.*, vol. 134, pp. 392–402, 2013.

[10] H. Li, T. M. Aide, Y. Ma, W. Liu, and M. Cao, "Demand for rubber is causing the loss of high diversity rain forest in SW China," *Biodivers. Conserv.*, vol. 16, no. 6, pp. 1731–1745, 2007.

[11] B. Chen, J. Cao, J. Wang, Z. Wu, Z. Tao, J. Chen, C. Yang, and G. Xie, "Estimation of rubber stand age in typhoon and chilling injury afflicted area with Landsat TM data: A case study in Hainan Island, China," *For. Ecol. Manag.*, vol. 274, pp. 222–230, 2012.

[12] M. A. Wulder, R. S. Skakun, W. A. Kurz, and J. C. White, "Estimating time since forest harvest using segmented Landsat ETM+ imagery," *Remote Sens. Environ.*, vol. 93, no. 1, pp. 179–187, 2004.

[13] Q. Zhang, G. Pavlic, W. Chen, R. Latifovic, R. Fraser, and J. Cihlar, "Deriving stand age distribution in boreal forests using SPOT VEGETATION and NOAA AVHRR imagery," *Remote Sens. Environ.*, vol. 91, no. 3, pp. 405–418, 2004.

[14] R. Sivanpillai, C. T. Smith, R. Srinivasan, M. G. Messina, and X. B. Wu, "Estimation of managed loblolly pine stand age and density with Landsat ETM+ data," *For. Ecol. Manag.*, vol. 223, no. 1, pp. 247–254, 2006.

[15] E. B. Racine, N. C. Coops, B. St-Onge, and J. Bégin, "Estimating forest stand age from LiDAR-derived predictors and nearest neighbor imputation," *For. Sci.*, vol. 60, no. 1, pp. 128–136, Feb. 2014.

[16] A. Ekadinata, A. Widayati, and G. Vincent, "Rubber agroforest identification using object-based classification in Bungo District, Jambi, Indonesia," in *25th Asian Conference on Remote Sensing,* Chiang Mai, Thailand, 2004, pp. 22–26.

[17] A. Ekadinata and G. Vincent, "Rubber agroforests in a changing landscape: analysis of land use/cover trajectories in Bungo district, Indonesia," *For. Trees Livelihoods*, vol. 20, no. 1, pp. 3–14, 2011.

[18] J. Dong, X. Xiao, S. Sheldon, C. Biradar, and G. Xie, "Mapping tropical forests and rubber plantations in complex landscapes by integrating PALSAR and MODIS imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 74, pp. 20–33, 2012.

[19] W. Putklang, S. Maneechot, and C. Mongkolsawat, "Assessing Thaichote satellite data in support of mapping rubber tree plantation in Northeast Thailand," in the *33rd Asian Conference on Remote Sensing 2012, ACRS 2012*, vol. 3, pp. 2269–2275.

[20] C. Mongkolsawat and W. Putklang, "Rubber tree expansion in forest reserve and paddy field across the greater Mekong sub-region, northeast Thailand based on remotely sensed imagery," in the *33rd Asian Conference on Remote Sensing 2012, ACRS 2012*, vol. 1, pp. 214–219.

[21] R. J. Kauth and G. S. Thomas, "The tasselled cap–a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat," in *LARS Symposia*, 1976, p. 159.

[22] G. B. Sreekala and S. K. Subodh, "Hyperspectral data mining," in *Hyperspectral Remote Sensing of Vegetation*. P. S. Thenkabail, G. J. Lyon, and A. Huete, Eds. Boca Raton, FL, USA: CRC Press, 2011.

[23] C. Vaiphasa, A. K. Skidmore, W. F. de Boer, and T. A. Vaiphasa, "Hyperspectral band selector for plant species discrimination," *ISPRS J. Photogramm.*, vol. 62, pp. 225–235, 2007.

[24] L. Gomez-Chova, J. Calpe, G. Camps-Valls, J. Martin, E. Soria, J. Vila, L. Alonso-Chorda, and J. Moreno, "Feature selection of hyperspectral data through local correlation and SFFS for crop classification," *IEEE Int. Geosci. Remote Sens. Symp. Proc.*, vol. 1, pp. 555–557, 2003.

[25] G. M. Foody, D. S. Boyd, and M. E. J. Cutler, "Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions," *Remote Sens. Environ.*, vol. 85, no. 4, pp. 463–474, 2003.

[26] Astrium GEO-Information Services, "Spectral modeling," in *Pléiades Imagery User Guide*. France, 2012, pp. 93.

[27] J. W. Rouse Jr, R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS," *NASA Spec. Publ.*, vol. 351, p. 309, 1974.

[28] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote Sens. Environ.*, vol. 83, no. 1, pp. 195–213, 2002.

[29] A. Huete, "A soil-adjusted vegetation index (SAVI)," *Remote Sens. Environ.*, vol. 25, no. 3, pp. 295–309, 1988.

[30] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern Recognit. Lett.*, vol. 15, no. 11, pp. 1119–1125, 1994.

[31] P. Somol and P. Pudil, "Feature selection toolbox," *Pattern Recognit.*, vol. 35, no. 12, pp. 2749–2759, 2002.