

Article

Tree Species Classification Using Hyperion and Sentinel-2 Data with Machine Learning in South Korea and China

Joongbin Lim ¹ , Kyoung-Min Kim ^{1,*} and Ri Jin ²

¹ Inter-Korean Forest Research Team, Division of Global Forestry, Department of Forest Policy and Economics, National Institute of Forest Science, Seoul 02455, Korea; jlim@korea.kr

² Department of Geography, Yanbian University, Yanji 133002, China; jinri0322@ybu.edu.cn

* Correspondence: greenann@korea.kr; Tel.: +82-2-961-2891

Received: 29 January 2019; Accepted: 15 March 2019; Published: 20 March 2019



Abstract: Remote sensing (RS) has been used to monitor inaccessible regions. It is considered a useful technique for deriving important environmental information from inaccessible regions, especially North Korea. In this study, we aim to develop a tree species classification model based on RS and machine learning techniques, which can be utilized for classification in North Korea. Two study sites were chosen, the Korea National Arboretum (KNA) in South Korea and Mt. Baekdu (MTB; a.k.a., Mt. Changbai in Chinese) in China, located in the border area between North Korea and China, and tree species classifications were examined in both regions. As a preliminary step in developing a classification algorithm that can be applied in North Korea, common coniferous species at both study sites, Korean pine (*Pinus koraiensis*) and Japanese larch (*Larix kaempferi*), were chosen as targets for investigation. Hyperion data have been used for tree species classification due to the abundant spectral information acquired from across more than 200 spectral bands (i.e., hyperspectral satellite data). However, it is impossible to acquire recent Hyperion data because the satellite ceased operation in 2017. Recently, Sentinel-2 satellite multispectral imagery has been used in tree species classification. Thus, it is necessary to compare these two kinds of satellite data to determine the possibility of reliably classifying species. Therefore, Hyperion and Sentinel-2 data were employed, along with machine learning techniques, such as random forests (RFs) and support vector machines (SVMs), to classify tree species. Three questions were answered, showing that: (1) RF and SVM are well established in the hyperspectral imagery for tree species classification, (2) Sentinel-2 data can be used to classify tree species with RF and SVM algorithms instead of Hyperion data, and (3) training data that were built in the KNA cannot be used for the tree classification of MTB. Random forests and SVMs showed overall accuracies of 0.60 and 0.51 and kappa values of 0.20 and 0.00, respectively. Moreover, combined training data from the KNA and MTB showed high classification accuracies in both regions; RF and SVM values exhibited accuracies of 0.99 and 0.97 and kappa values of 0.98 and 0.95, respectively.

Keywords: hyperspectral image; random forest; support vector machine; texture feature; image spectroscopy

1. Introduction

North Korea is suffering from extreme forest degradation due to food and energy shortages [1–3]. Degraded and deforested lands are vulnerable to natural disasters, such as landslides and floods, which not only cause environmental damage, but also destroy agricultural infrastructure [1–3]. This results in a vicious cycle of degradation in the forest by woodcutting to address food and fuel shortages.

Additionally, deforestation causes habitat fragmentation, which decreases the size and increases the isolation of habitat patches, causing changes in biodiversity and community structure [4,5]. Natural forests also function as carbon sinks, and can be key in reducing emissions from deforestation and forest degradation (REDD). Although the degradation of forests in North Korea has been reported as the primary national problem, it is noteworthy that 7.64 million ha of intact forests remain in North Korea [6].

Among them, 953,500 ha have been managed as protected forests [6]. North Korea has also established environmentally protected areas, of which there are nine categories: biosphere reserve, nature park, nature reserve, animal reserve, plant reserve, seabird reserve, wetland reserve, coastal resource reserve, and landscape reserve [7]. Furthermore, five tree species were reported as the dominant varieties in North Korea; these were: oak, larch, pine, deodar (cedar), and Korean pine. These species occupy 29.5% (oak), 17.5% (larch), 12.7% (pine), 8.2% (deodar), and 5.8% (Korean pine) of the total forest cover [6].

While the protected forested areas and dominant tree species of North Korea have been documented in previous studies, their spatial distributions have not yet been reported. The species-level information of forests, such as their composition and distribution, is essential for sustainable forest management [8,9]. Even though forest resource assessments are important, it has been impossible to perform forest surveys in North Korea due to its inaccessibility. In this case, remote sensing (RS) can play a key role in the surveying of forest resources in North Korea. Many research groups have attempted to determine the spatial variation of species using RS worldwide [8–20].

There are several classification methods in RS, such as spectral angle mapping [8,16,17], linear discriminant analysis [16,21], decision tree classification [22,23], artificial neural networks [24,25], convolutional neural networks [26,27], recurrent neural networks [28,29], support vector machines (SVMs) [8,18–20,30], and random forests (RFs) [13,20,31]. Among them, SVM and RF have shown very high degrees of reliability and classification accuracies in RS applications, and have been widely used in the forestry community [13,19,20]. Several groups have tried to classify plant species using RFs and SVMs [8,13,14,32,33]. In the discrimination of tropical wetland plant species in Jamaica, RFs showed accuracies of 91.8% and 84.8% for importance-ranked spectral indices and in situ reflectance spectra, respectively [33]. Additionally, RFs showed an 87.6% overall accuracy for eight savanna tree species using Compact Airborne Spectrographic Imager (CASI-1500, Itres Research Ltd., Ontario, Canada) data and waveform Light Detection and Ranging (LiDAR) data [13]. In the western Himalayas, an SVM with Hyperion data showed 82.2% accuracy and 69.62% accuracy with Landsat Thematic Mapper (TM) data. This study confirmed the potential utility of narrow spectral bands of Hyperion data in classifying tree species on hilly terrain [8]. The results of these three studies also demonstrated the classification capability of RFs and SVMs with hyperspectral data and Hyperion data. Thus, SVMs and RFs were applied in this study because a primary goal of this study was to develop a tree species classification model that can be used for the classification of trees in North Korea, where field observations for validation remain impossible.

Support vector machines with spectral features, textural features, and hyperspectral vegetation indices have shown an accuracy of 82.3% in classifying mangrove species of Qi'ao Island in China. Additionally, when the tree height information was supplemented in the classification, the accuracy was increased to 88.6% [32]. In the case of forest tree species classification using CASI data with textural features at the Liangshui National Natural Reserve area in China, an SVM showed 85.9% accuracy, and it was confirmed that combining spectral and spatial information can improve the accuracy of tree species classifications [14]. Therefore, topographic data and textural data were used in this study to increase the predictive ability of the models.

The aims of this study were to develop a tree species classification model based on RS and machine learning techniques that can be utilized for classification in North Korea. Two study sites were chosen in South Korea (the Korea National Arboretum (KNA)) and China (Mt. Baekdu (MTB), a.k.a., Mt. Changbai in Chinese) and examined to classify tree species in both regions; the latter of these sites occupies the border area between North Korea and China. As a preliminary step in developing a

classification algorithm that can be applied in North Korea, we selected two common coniferous species for investigation at both study sites: Korean pine (*Pinus koraiensis*, Siebold & Zucc.) and Japanese larch (*Larix kaempferi* (Lamb.) Carriere). It is meaningful to analyze the distribution of these two species as they are the forest species that North Korea considers economically essential [34].

Hyperion data, which are hyperspectral satellite data, have been used to classify tree species due to the abundance of spectral information acquired from more than 200 spectral bands. However, no new Hyperion data are available because the satellite has not been operated since 2017. Recently, Sentinel-2 satellite multispectral imagery has been used in tree species classification. Therefore, it is necessary to compare the two kinds of satellite data from Hyperion and Sentinel-2 to determine the possibility of species classification using the lower spectral resolution satellite. In order to achieve this goal, the following questions were posed:

1. Can Hyperion data with machine learning algorithms, such as RFs and SVMs, be adopted for tree classification?
2. Can Sentinel-2 data be used for tree species classification with RF and SVM algorithms instead of Hyperion data?
3. Can training data that were built in the KNA be used for the tree classification of MTB?

2. Materials and Methods

2.1. Study Areas

The study was conducted throughout the area of the KNA in South Korea ($37^{\circ}45' N$, $127^{\circ}10' E$) and the MTB region in China ($42^{\circ}16' N$, $127^{\circ}59' E$) (Figure 1). Due to the inaccessibility of North Korea, MTB was selected to identify the possibility of species classification in North Korea because it is a national boundary region between China and North Korea. China owns the northern part of MTB, named Mt. Changbai in Chinese, while North Korea owns the southern part of MTB, named Mt. Baekdu in Korean. Both countries have similar topographic and climate conditions in this region [35,36].

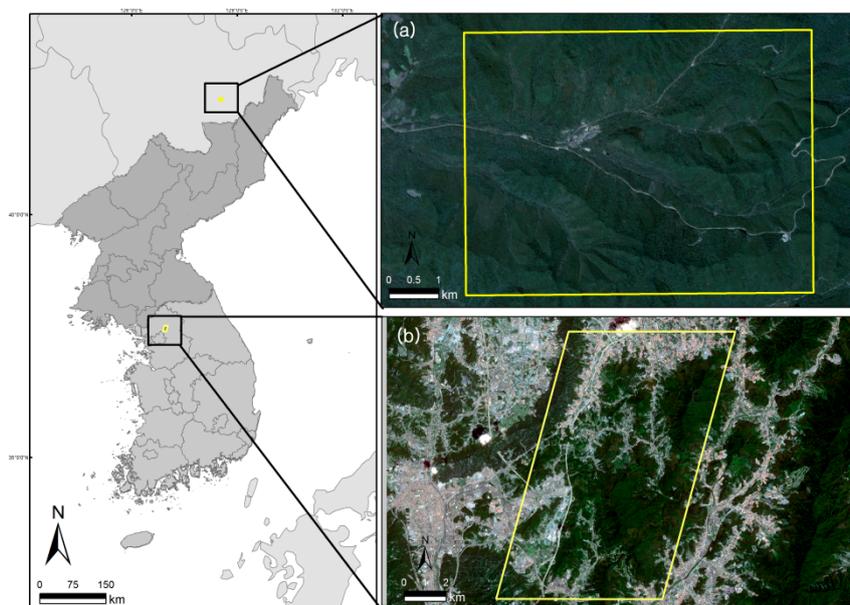


Figure 1. Study area: (a) Mt. Baekdu (MTB) region in China; (b) Korea National Arboretum (KNA) in South Korea.

The KNA is located in Gwangneung Forest. Gwangneung Forest was a royal forest, and houses the mausoleum of King Sejo of the Joseon Dynasty. Thus, it has been strictly managed to minimize human disturbance over the last 500 years. The Gwangneung Arboretum was established in 1987

in affiliation with the National Institute of Forest Science (NIFoS), controlled by the Korea Forest Service (KFS), and has been open to the public since then. It became known as the KNA on 24 May 1999. The KNA contains 1120 ha of natural forest, 100 ha of specialized gardens, a forest museum, the Korea National Herbarium, a temperate house, and the Tropical Plant Resource Center. The KNA was designated as one of the United Nations Educational, Scientific and Cultural Organization (UNESCO) biosphere reserve in June of 2010 [37].

Mt. Baekdu is the highest mountain on the Korean peninsula. The administrative area is bordered by North Korea's Yanggang Province and China's Jilin Province, with a total area of 8000 km². The climate is a typical alpine climate and experiences severe climatic changes. The average annual temperature is 6–8 °C, the maximum temperature is 18–20 °C, the January average temperature is –23 °C (lowest –47 °C), and the average daily temperature in July is 4.8 °C. The annual average humidity is 74%, and is the highest during the summer and low in winter [38]. The flora of MTB include 330 species, 47 families, and 162 genera [38]. It is called a “three-dimensional botanical garden” because it is composed of primeval forests (mixed coniferous and broadleaf forests and coniferous forests). Due to the influences of climate and topography, the temperature gradually decreases according to the altitude above sea level. The vertical distribution of plants in MTB is that of temperate broadleaf forests below 720 m above sea level, mixed forests of polar and temperate coniferous and broadleaf forests at altitudes of 720–1100 m, sub-polar coniferous forests from 1100–1700 m, polar and sub-alpine forests from 1700–2000 m, and between 2000 m and 2700 m are alpine and mossy plains. This apparent vertical distribution is unique worldwide. Mt. Baekdu was designated by North Korea as a protected vegetation area. It was registered as an International Biosphere Reserve in 1989 with an area of ~14,000 ha. In China, MTB was designated as a nature reserve in 1906 [38].

2.2. Data

In this study, EO-1 Hyperion and Sentinel-2 data were used to classify Korean pine and Japanese larch. Hyperion level 1R and 1T data and Sentinel-2 level 1C data were obtained from the official website of the United States Geological Survey (USGS) (<http://earthexplorer.usgs.gov/>) (Table 1). There was only one scene of Hyperion data in the KNA region, which was obtained on 7 September 2010. The scene has period gap with Sentinel-2 data; however, this did not influence the study results because Hyperion data were used to establish the classification algorithm procedure and to confirm the possibility of tree species classification using satellite data. Sentinel-2 data were used for tree species classification in both study regions. Additionally, PlanetScope Level 3B data were used to produce textural feature bands. A digital forest map of South Korea provided by NIFoS was used to derive training and validation data. Based on this forest map, 3000 training and 315 validation data points were randomly generated. The validation sampling number was set based on international standards [39]. Researchers at Yanbian University built training and validation data of MTB and performed the classification validation of MTB. Finally, Shuttle Radar Topography Mission (SRTM) 1-arc second digital elevation model (DEM) data, which were obtained from the USGS website, were used to generate elevation, slope, and aspect maps.

Table 1. Hyperion, Sentinel-2, and PlanetScope data information for the KNA and MTB.

	Hyperion	Sentinel-2		PlanetScope	
	KNA	KNA	MTB	KNA	MTB
Acquisition date	7 September 2010	28 April 2018 23 May 2018 2 June 2018 7 July 2018 1 August 2018 25 September 2018 30 October 2018	18 April 2018 23 May 2018 2 June 2018 22 June 2018 26 August 2016 20 September 2018 5 October 2018	12 April 2018 16 September 2017	4 April 2018 31 August 2017
Path/Row Level	116/34 1R/1T	52SCG	52TDN 1C	3B	

2.3. Methodology

The study procedure is shown in the flowchart of Figure 2. As a first step, image preprocessing, including topographic correction and atmospheric correction, was performed and then, the spectral separability between endmembers of Korean pine and Japanese larch selected from Hyperion data was investigated. The spectral similarity between endmembers of the KNA selected from Hyperion data and Sentinel-2 data was also investigated to confirm the applicability of Sentinel-2 data for classification. Hyperion bands corresponding to the Sentinel-2 wavelength were extracted and used to analyze the spectral similarity with Sentinel-2 data. The spectral similarity between endmembers of the KNA and MTB selected from multi-seasonal Sentinel-2 data (April–October) was then assessed for selecting the best seasonal Sentinel-2 image of the KNA and MTB for classification. Additionally, textural information was extracted by using the gray-level co-occurrence matrix (GLCM) [40] from PlanetScope data. Elevation, slope, and aspect maps were then generated from SRTM DEM data.

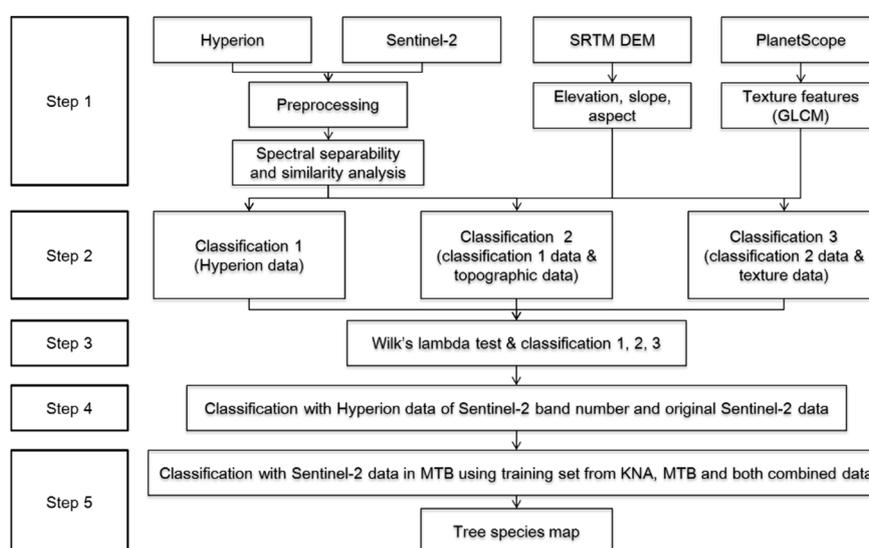


Figure 2. Flowchart of the tree species classification procedure. Abbreviations: GLCM, gray-level co-occurrence matrix.

In the second step, tree species classification using Hyperion data in the KNA was performed. During classification, additional attribute data, such as elevation, slope, aspect, and texture, were added step-by-step. Classification 1 was performed only with Hyperion data. Classification 2 used both Hyperion and topographic data (e.g., elevation, slope, and aspect). Classification 3 was accomplished with Hyperion data, topographic data, and texture data.

In the third step, Wilk's lambda test was conducted to reduce the dimensionality of Hyperion data to decrease analytical costs, and to confirm the possibility of classification using dimensionally reduced Hyperion data. The fourth step involved classifying tree species using Hyperion data with Sentinel-2 band information, as well as original Sentinel-2 data. The final step was to classify tree species on MTB using Sentinel-2 data with a training set from the KNA, MTB, and both combined. Training sets from the KNA and MTB were built using a data frame in the R statistical package ver. 3.5.1. Thus, both datasets were combined by rows using the 'rbind()' command in R.

2.4. Preprocessing

Hyperion data contained 242 bands, and many of them were found unsuitable for the analyses in this study due to the low signal-to-noise ratio. Thus, 147 bands were selected for analysis based on previous studies [8,10,41]. All bands were separately visualized, and bands that were not calibrated or had very high noise were eliminated following visual inspection. Atmospheric correction was performed using the Environment for Visualizing Images ver. 5.0 (ENVI) fast line-of-sight atmospheric

analysis of spectral hyper-cubes (FLAASH) algorithm. The Level 1R image was georeferenced using the Level 1T image. Hyperion data were projected onto the Universal Transverse Mercator (UTM) projection, World Geodetic System (WGS) 84 datum, and zone 52N.

Sentinel-2 level 1C data were orthoimages in UTM/WGS84 projection, and provided in top-of-atmosphere (TOA) reflectance. Thus, an atmospheric correction was needed to calculate the surface reflectance. In order to calculate surface reflectance, a Sen2Cor algorithm [42] was used; it was developed by the European Space Agency (ESA), which also develops, manages, and distributes Sentinel-2 data.

2.5. Texture Analysis

Textural features can help to improve the accuracy of tree species classifications [14]. The GLCM is one of most well-known texture analysis algorithms, and has been widely adopted by the RS community [43–51]. The GLCM represents the distance and angular relationship over a sub-area of an image of a specified size. It measures the spatial frequency of co-occurrences of gray pixel levels in a user-defined moving kernel to quantify texture and to form the co-occurrences of pixels in the kernel. Consideration of the window size that captures the target class should be given while calculating the GLCM texture scale. The optimal window size should be determined by the spatial resolution of the image and the size of the tree canopy [14]. In this study, the angular second moment (ASM), contrast (CON), dissimilarity (DIS), entropy (ENT), homogeneity (HOM), mean (MEAN), and variance (VARIANCE) were used. A series of GLCM texture measures were calculated according to the following equations [40]:

$$ASM = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i, j)^2; \quad (1)$$

$$CON = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} (i - j)^2 \times h_c(i, j)^2; \quad (2)$$

$$DIS = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i, j)^2 |i - j|; \quad (3)$$

$$ENT = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i, j) \times \log[h_c(i, j)]; \quad (4)$$

$$HOM = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} \frac{1}{1 + (i - j)^2} \cdot h_c(i, j); \quad (5)$$

$$MEAN = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} i \times h_c(i, j); \quad (6)$$

$$VARIANCE = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} (i - \mu)^2 h_c(i, j), \quad (7)$$

where $quant_k$ is the quantization level of band k (e.g., $2^8 = 0$ to 255) and $h_c(i, j)$ is the (i, j) th entry in one of the angular brightness value spatial-dependency matrices. Textural feature analysis was performed using the 'glcm' package in R.

2.6. Spectral Separability and Similarity Analysis

In order to investigate the spectral separability between Korean pine and Japanese larch, the Jeffries-Matusita (JM) distance was applied to the endmembers selected from Hyperion and Sentinel-2 data. Jeffries-Matusita distance values range from 0 (i.e., identical distributions) to 1.414 (i.e., complete dissimilarity), and it is generally implemented to quantify the degree of separation [8,40]. The spectral similarity between endmembers from the KNA and MTB was also assessed to investigate whether it is possible to use a training dataset of the KNA for the classification of trees in the MTB area. For the similarity test, a spectral angle mapper (SAM) algorithm, which is also commonly applied to assess spectral similarity by RS user group [8,40], was used. The SAM results range from 0 (i.e., lower similarity) to 1 (i.e., higher similarity) [52].

2.7. Classification Algorithms

2.7.1. Random Forest

The RF was developed by Leo Breiman and Adele Cutler [53]. It generates many decision trees (e.g., 500 trees) to define unclassified pixels, with its associated attributes, such as spectral reflectance, elevation, and slope, into a class. Each decision tree classifies the pixel into one class, and it votes for that class. The forest classifies the pixel to the class having the most votes from all of the trees in the forest [13,40,53–55].

Each of the individual decision trees is grown as follows [53]:

1. N samples are randomly selected with replacement from the entire training data. These N samples are used as training data in each decision tree model to generate trees. In general, approximately 70% of the total sample is extracted and used as training data in each tree model, and the remaining 30% is termed “out-of-bag” (OOB) and not used during training;
2. If there are M input variables, a number m of input variables of M is randomly selected with replacement at each node of the decision trees, and then the best node variables are determined among the m variables. The value of m is constant in an RF, and usually the square root of M , which is the total number of input variables, is used;
3. Each decision tree is created to the maximum possible size without pruning.

Each class can be weighted with a priori facts. Random forests are efficient for extensive datasets, which is an advantageous feature for analyzing sizeable RS data [40]. They also provide information on which variables are important for classification. If there were random interchange in the data for particular predictors, the variable importance would be calculated based on the degradation of the prediction [13,56]. It helps to understand which predictor(s) are driving the differences in accuracy between different classifications [57].

The variable importance can be measured using the mean decrease in Gini (MDG). The MDG measures how much a variable reduces the Gini impurity metric in a particular class [54,58]. Additionally, RFs provide OOB estimates of error rates, which are measured by counting the number of misclassifications and dividing this number by the total number of observations. Error rates can be used to choose the best fitting model [31,59,60]. In numerous studies, classifications have been performed using RFs and their superiority over other classification techniques has been demonstrated [13,59,61–64].

2.7.2. Support Vector Machine

Support vector machines were suggested by Vapnik [65]. They are advanced and useful classifiers, which can manage classification problems in hyperspectral data and have been widely used for tree species classification [14]. Support vector machines use training data to find the optimal hyperplane between classes [66,67]. At this time, it finds the optimal hyperplane between two classes that maximizes the margin between the closest training samples of the classes. The points at the boundary are called support vectors, and the middle of the margin is the optimal hyperplane separating the classes [68]. Training points located on the opposite side of the separating hyperplane have negative weights to reduce their effects. If it is not feasible to determine a linear hyperplane, a kernel function is used to convert the original data into a higher dimensional space to find a suitable hyperplane [68]. The equation is as follows [40]:

$$\min_{w,b,\zeta} \left(\frac{w^T \cdot w}{2} + C \sum_{i=1}^{\lambda} \zeta_i \right), \text{ such that} \quad (8)$$

$$y_i (w^T \cdot \Phi(x_i) + b) \geq 1 - \zeta_i, (\zeta_i \geq 0),$$

where ζ_i denotes positive slack variables used to allow some of the samples to fall on the wrong side of the hyperplane, $C \sum \zeta_i$ is a term used to penalize solutions for which ζ_i are very large, and $w^T \cdot \Phi(x_i) + b$ is a hyperplane in a higher dimensional feature space [11,40]. The basic SVM approach may be extended to optimize the nonlinear surface using the following decision function [40]:

$$f(x) = \sum_{i=1}^{\lambda} a_i y_i K(x, x_i) + b, \quad (9)$$

where a_i are nonnegative Lagrange multipliers used to search the optimal separating hyperplane, and $K(x, x_i)$ is a kernel function, which replaces the inner product $(x \cdot x_i)$ in order to solve computational problems in a higher dimensional space [40,69]. For SVM classification, the radial basis function kernel type was used with optimal gamma and cost values that were determined by a commonly used grid search approach [8,70].

Random forest and SVM classifications were performed using the ‘randomForest’ and ‘e1071’ packages in R. Accuracy assessments were performed with ‘confusionMatrix’ in the ‘caret’ package of R using validation data. The confusion matrix provides the overall accuracy, kappa statistic value, user accuracy, and producer accuracy [40]. Additionally, a receiver operating characteristic (ROC) curve was used to evaluate the predictive capability of the models. The ROC curve was produced by plotting the true positive rate against the false positive rate at various threshold settings. If a model had an area under the curve (AUC) closer to 1 and was higher than 0.5, this meant that the model had a good predictive capability [71].

2.8. Data Dimensionality Reduction

Hyperion data have 242 bands. Generally, the dimensionality is reduced before classification to decrease the processing cost and to use only the optimal bands. In addition to classification, the purpose of this study was also to investigate the applicability of Sentinel-2 data for tree species classification. The applicability of dimensionality reduction was investigated in order to determine whether dimensionally reduced Hyperion data can be used for tree species classification and if it is possible, if it is feasible to classify tree species using selected Hyperion data corresponding to Sentinel-2 band information. To reduce the dimensionality of Hyperion data, a stepwise discriminant analysis procedure based on Wilk’s lambda test statistic, which has been used for Hyperion data preprocessing by several research groups [8,15,72], was performed.

Wilk’s lambda (Λ) was given by Green and Carroll [73] as:

$$\Lambda = \frac{|W|}{|T|} = \frac{|W|}{|W + B|}. \quad (10)$$

In Equation (10), W is the within-groups sum of squares and cross-product matrix, B is the between-groups sum of squares and cross-product matrix, and T is the total sum of squares and cross-products matrix, such that:

$$T = \sum_{i=1}^g \sum_{j=1}^{n_i} (X_{ij} - \bar{X})(X_{ij} - \bar{X})^t; \quad (11)$$

$$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(X_{ij} - \bar{X}_i)^t; \quad (12)$$

$$B = \sum_{i=1}^g n_i (\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})^t, \quad (13)$$

where g is the number of groups, n_i is the number of observations in the i th group, \bar{X}_i is the mean vector of the i th group, \bar{X} is the mean vector of the all observations, and X_{ij} = the j th multivariate observation in the i th group. Wilk’s lambda ranges from 0 to 1; the values close to 0 mean that the groups are well separated, while the values close to 1 mean that the groups are poorly separated. The stepwise discriminant analysis was performed using the ‘greedy.wilks’ function of the ‘klaR’ package in R.

3. Results and Discussion

3.1. Spectral Separability and Similarity

Table 2 shows the spectral separability between Korean pine and Japanese larch selected from the Hyperion and Sentinel-2 data of the KNA and MTB. Each spectrum represents the mean spectra of each species in each dataset. The JM distance value of the KNA from Hyperion data was 0.027, and those of the KNA and MTB from Sentinel-2 data were 0.009 to 0.275, respectively. It seems from Table 2 that there was no separability between the reflectance spectra of Korean pine and Japanese larch. However, as shown in Figure 3, the reflectance spectra of Korean pine were slightly different from those of Japanese larch. According to previous studies, coniferous needles have slightly lower visible (VIS) reflectances and higher near infrared (NIR) reflectances than broad leaves [24,26]. Moreover, *Larix* sp. have similarity spectra to broad leaves [74]. Thus, it is reasonable that Korean pine exhibited a higher NIR reflectance and lower shortwave infrared (SWIR) reflectance than Japanese larch, as *Pinus* sp. have exhibited higher NIR reflectances and lower SWIR reflectances than *Larix* sp. in previous studies [74].

Table 2. Jeffries-Matusita (JM) distance values for Hyperion and Sentinel-2 image training samples of the KNA and MTB.

Data	Region	April	May	June	July	August	September	October
Hyperion	KNA						0.027	
Sentinel-2	KNA	0.051	0.072	0.074	0.086	0.009	0.057	0.091
Sentinel-2	MTB	0.275	0.107	0.094	0.075	0.093	0.120	0.025

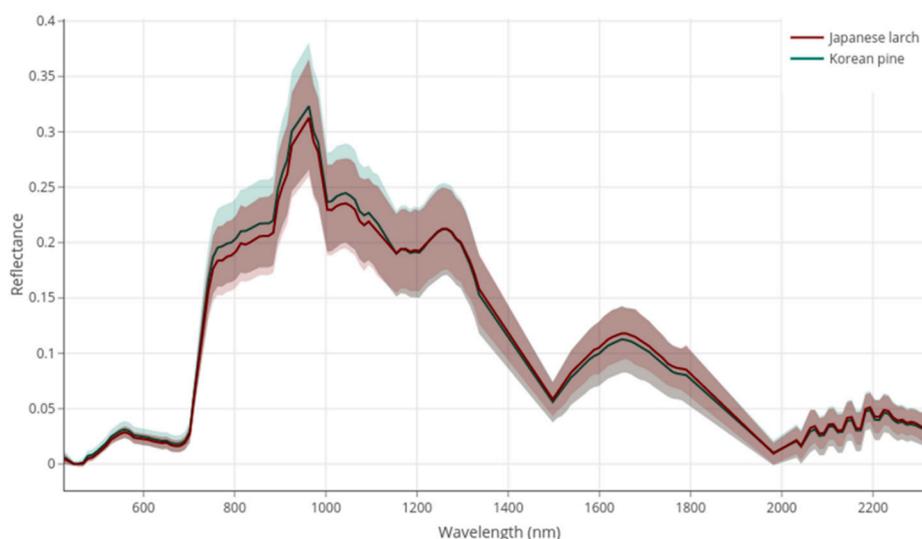


Figure 3. Visual comparison of reflectance spectra of Korean pine and Japanese larch from Hyperion images. Half-transparent regions in the figure represent standard deviations.

Table 3 shows the spectral similarity between endmember spectra derived using Sentinel-2 reflectance images of the KNA and MTB. As shown in the table, all season spectra exhibited very high SAM scores, which means that they were highly similar. Among them, spectra for May at the KNA had the highest similarity (SAM = 1) to spectra for June at MTB. Meanwhile, late May and early June are the best months for tree classification at MTB [75]; therefore, images from May at the KNA and June at MTB were chosen for classification.

Table 3. Spectral similarity values between endmember spectra of the KNA and MTB from Sentinel-2 images.

Korean Pine		MTB						
		18 April 2018	23 May 2018	2 June 2018	22 June 2018	26 August 2016	20 September 2018	5 October 2018
KNA	28 April 2018	0.966	0.995	0.994	0.991	0.992	0.996	0.988
	23 May 2018	0.945	0.999	1.000	0.999	0.999	0.997	0.980
	2 June 2018	0.948	0.999	0.999	0.998	0.998	0.997	0.981
	7 July 2018	0.919	0.956	0.955	0.953	0.953	0.954	0.943
	1 August 2018	0.942	0.998	0.999	0.999	0.998	0.996	0.978
	25 September 2018	0.937	0.999	1.000	1.000	0.999	0.996	0.975
	30 October 2018	0.944	0.999	0.999	0.999	0.998	0.997	0.980
Japanese larch		MTB						
		18 April 2018	23 May 2018	2 June 2018	22 June 2018	26 August 2016	20 September 2018	5 October 2018
KNA	28 April 2018	0.934	0.995	0.993	0.991	0.990	0.996	0.987
	23 May 2018	0.899	1.000	1.000	0.999	0.998	0.998	0.973
	2 June 2018	0.905	0.999	0.999	0.999	0.998	0.998	0.975
	7 July 2018	0.880	0.954	0.953	0.952	0.951	0.953	0.935
	1 August 2018	0.906	0.998	0.999	0.999	0.999	0.998	0.975
	25 September 2018	0.905	0.999	0.999	0.999	0.999	0.999	0.976
	30 October 2018	0.917	0.999	0.998	0.996	0.996	0.999	0.983

3.2. Classification with Hyperion Data of the Korea National Arboretum

Table 4 shows the accuracy of RF and SVM classifications using Hyperion data. In RF classification using Hyperion data, the overall accuracy of 0.82 (kappa statistic = 0.64) was achieved. The producer accuracy of Korean pine and Japanese larch in RF classifications were 0.85 and 0.80, respectively. Additionally, the user accuracy of Korean pine and Japanese larch in RF classifications were 0.78 and 0.86, respectively. The species classification performance using an SVM classifier on Hyperion data was better (overall accuracy = 85%, kappa statistic = 0.71) than RF classification. The producer accuracy of Korean pine and Japanese larch in SVM classification were 0.86 and 0.84, respectively, while the user accuracy of Korean pine and Japanese larch were 0.84 and 0.87, respectively. The producer accuracy and user accuracy increased for both species.

The distribution of species can vary depending upon geographical characteristics, and spectral reflectances can be affected by the terrain aspect. In order to consider such geographical effects, elevation, slope, and aspect maps were also utilized in this study. Table 4 shows that topographical analysis was more accurate than non-topographic analysis. The overall accuracy and kappa statistics of RF and SVM classifications were 0.83 and 0.65, and 0.86 and 0.73, respectively. Overall RFs and SVMs with topographical analyses yielded higher producer accuracies (>0.81 and >0.86, respectively) and user accuracies (>0.80 and >0.85, respectively).

Japanese larches shed their leaves in the fall and remain leafless throughout the winter period, while Korean pine is an evergreen. Thus, there is a difference in the tree canopies between Korean pine and Japanese larch, and in the timing of the first leaf buds and first autumn foliage. In order to reflect these properties, textural analysis results of both April and September images using GLCM were used as textural image bands. Table 4 shows the effects of textural image application. The overall accuracy and kappa statistics of RF and SVM classifications improved from 0.83 and 0.65, and 0.86 and 0.73, respectively, to 0.88 (RF) and 0.75 (SVM), and 0.90 (RF) and 0.81 (SVM). Overall RF and SVM gave higher producer accuracies (>0.86, >0.90) and user accuracies (>0.86, >0.90).

Table 5 shows the variable importance derived from the RF of classification 3. It shows the top 20 crucial variables within 147 bands, topographic data, and textural data. As can be seen in the table, several textural features occupy the top ranks. Topographical data (aspect) and vegetation-related bands, which are related to chlorophyll, nitrogen, and protein, were also found to be significant. The textural information and topographical factors were added stepwise in classification 2 and classification 3, and the accuracy increased as each factor was added. Additionally, the OOB estimate of the error rate decreased from 18.2% in classification 1 to 17.4% in classification 2, and to 14.1 in classification 3. Moreover, the AUC increased from 0.82, 0.85 in classification 1 to 0.83, 0.86 in classification 2, and to 0.88, 0.90 in classification 3. Therefore, topographical data and textural data were able to increase the classification accuracy of tree species.

Table 4. Species classification results using Hyperion data. Abbreviations: PA, producer accuracy; RF, random forests; SVM, support vector machine; UA, user accuracy.

Classification 1	RF			UA	SVM			UA
	Korean Pine	Japanese Larch	Total		Korean Pine	Japanese Larch	Total	
Korean pine	232	66	298	0.78	249	49	298	0.84
Japanese larch	42	257	299	0.86	39	260	299	0.87
Total	274	323	597		288	309	597	
PA	0.85	0.80			0.86	0.84		
Overall accuracy	0.82	kappa statistics	0.64	Overall accuracy	0.85	kappa statistics	0.71	
Classification 2								
Korean pine	214	54	268	0.80	229	39	268	0.85
Japanese larch	40	232	272	0.85	35	237	272	0.87
Total	254	286	540		264	276	540	
PA	0.84	0.81			0.87	0.86		
Overall accuracy	0.83	kappa statistics	0.65	Overall accuracy	0.86	kappa statistics	0.73	
Classification 3								
Korean pine	220	25	245	0.90	222	23	245	0.91
Japanese larch	36	213	249	0.86	24	225	249	0.90
Total	256	238	494		246	248	494	
PA	0.86	0.89			0.90	0.91		
Overall accuracy	0.88	kappa statistics	0.75	Overall accuracy	0.90	kappa statistics	0.81	

Table 5. Variable importance of Classification 3. Abbreviations: MDG, mean decrease in Gini.

Band	MDG	Remarks	Reference
GLCM_mean_4	186.0		
GLCM_variance_4	171.2		
GLCM_contrast_4	54.0		
GLCM_dissimilarity_4	50.7		
B6 (477.7 nm)	48.2	Chlorophyll b	[76]
GLCM_homogeneity_4	47.2		
B8 (498.0 nm)	45.8	Senescing, carotenoid, browning, soil background effects	[77]
GLCM_entropy_4	41.7		
B9 (508.2 nm)	39.0	Nitrogen	[8]
GLCM_second_moment_4	34.9		
B7 (487.9 nm)	34.5	Nitrogen	[17,78]
B10 (518.4 nm)	29.9		
B125 (1749.8 nm)	23.5	Protein	[8]
ASPECT	23.1		
B11 (528.6 nm)	22.4		
B12 (538.7 nm)	19.6		
B118 (1689.3 nm)	19.5	Lignin, starch, protein	[79]
B123 (1739.7 nm)	18.9		
B5 (467.5 nm)	18.2	Chlorophyll b	[79]
B120 (1709.5 nm)	18.2		

3.3. Wilk's Lambda Result of the Korea National Arboretum

Based on the Wilk's Lambda test, 55 bands were selected as optimal bands in the 437–2335 nm wavelength region (Table 6). Among them, 15 bands were from VIS, 15 bands from NIR, and 25 bands from SWIR regions. Usually, interspecies differences increase near longer wavelengths and are prevalent in SWIR regions [74,80]. The selected bands were related to chlorophyll in VIS, leaf and canopy structure in NIR, and water content, nitrogen, cellulose, starch, and sugar in SWIR (Table 6).

Table 6. Wilk's lambda result and importance for vegetation of selected bands.

No. of Variables	Wavelength (nm)	Wilk's Lambda	F Statistics	Remarks	Reference
1	478	0.92	501.75	Chlorophyll b	[76]
2	2143	0.81	639.25		
3	773	0.76	580.96		
4	1659	0.70	605.46	Lignin, biomass, starch, most useful to discriminate different kinds of leaves	[77]
5	722	0.68	528.05	Vegetation stress and dynamics	[77]
6	457	0.66	478.66	Chlorophyll b	[79,81]
7	681	0.64	446.96	Biomass, LAI	[77]
8	651	0.63	407.37	Chlorophyll a, b	[9]
9	498	0.62	377.42	Senescing, carotenoid, browning, soil background effects	[77]
10	1044	0.62	349.30		
11	1215	0.60	337.88	Moisture absorption	[77]
12	2133	0.60	317.32		
13	600	0.59	299.05		
14	1074	0.58	283.67		
15	1195	0.58	271.26	Water, cellulose, starch, lignin	[79]
16	488	0.58	257.28	Nitrogen	[17,78]
17	1094	0.57	244.71	Biomass, LAI	[77]
18	1175	0.57	232.93		
19	1508	0.57	222.74	Plant moisture	[77]
20	2063	0.57	213.61		
21	630	0.57	204.77	Chlorophyll b	[17,79]
22	508	0.56	196.82	Nitrogen	[8]
23	691	0.56	189.45	Biomass, LAI	[77]
24	610	0.56	182.55	Biomass, LAI	[77]
25	1023	0.56	176.17	Protein	[79]
26	468	0.56	170.27	Chlorophyll b	[79]
27	1276	0.56	164.66	Moisture absorption, starch	[9]

Table 6. Cont.

No. of Variables	Wavelength (nm)	Wilk's Lambda	F Statistics	Remarks	Reference
28	2335	0.56	159.46	Cellulose	[79]
29	2163	0.55	154.51	Lignin, sugar, protein	[9]
30	2214	0.55	149.90		
31	993	0.55	145.41		
32	1064	0.55	141.22	Plant moisture	[77]
33	963	0.55	137.25	Water, starch	[79]
34	1558	0.55	133.48		
35	1568	0.55	130.08		
36	1679	0.55	126.80	Lignin, tannin, starch, cellulose	[9]
37	923	0.55	123.64		
38	641	0.55	120.62	Chlorophyll b	[79]
39	1165	0.55	117.74		
40	1185	0.55	115.02		
41	1699	0.55	112.42	Lignin, starch, protein	[79]
42	1649	0.55	109.98	Lignin, tannin, starch, cellulose	[9]
43	864	0.55	107.61	Chlorophyll, biomass, LAI, protein	[77]
44	895	0.54	105.67		
45	803	0.54	103.62		
46	702	0.54	101.61		
47	2285	0.54	99.57		
48	2244	0.54	97.61		
49	2254	0.54	95.78		
50	671	0.54	93.98	Chlorophyll(Red 2)	[82]
51	437	0.54	92.24	Chlorophyll a	[79]
52	1326	0.54	90.55		
53	712	0.54	88.93		
54	1669	0.54	87.38	Lignin, tannin, starch, cellulose	[9]
55	1740	0.54	85.88		

Table 7 shows the accuracy of RF and SVM classifications using 55 bands of Hyperion data. The overall accuracies of RF and SVM classifications were 0.82 and 0.84, respectively, with corresponding kappa statistics of 0.65 and 0.69. Overall RF and SVM yielded high producer accuracies (>0.81 and >0.86) and user accuracies (>0.80 and >0.82). In classification 2, with 55 bands of Hyperion data, the overall accuracy and kappa statistics of RF and SVM classifications improved from 0.82 and 0.65, and 0.84 and 0.69 to 0.84 and 0.68, and 0.87 and 0.74, respectively. Overall RF and SVM yielded higher producer accuracies (>0.84 and >0.87) and user accuracies (>0.83 and >0.87). In classification 3, with 55 bands of Hyperion data, the overall accuracy and kappa statistics of classifications improved from 0.84 (RF) and 0.68 (SVM), and 0.87 (RF) and 0.74 (SVM) to 0.88 and 0.75, and 0.90 and 0.80 for RFs and SVMs, respectively.

The classified map using 55 band images matched with the 147 band image classification results as 0.96 and 0.92 for RFs and SVMs in classification 1, respectively. In classification 2, RFs and SVMs matched 0.96 and 0.93, respectively, and in classification 3, RFs and SVMs matched 0.97 and 0.94, respectively (Table 8). As a result, it was confirmed that the selection of bands can provide cost-effective results for tree species classification, and there was less of a reduction in the accuracy of RFs than SVMs.

In order to confirm the possibility of applying Sentinel-2 data for tree species classification, Hyperion bands corresponding with Sentinel-2 wavelengths were selected and applied to classification. Random forest and SVM classifications exhibited overall accuracies of 0.86 and 0.89 (kappa statistics 0.72 and 0.79), respectively (Table 7). Overall, RF and SVM yielded high producer accuracies (>0.84 and >0.89) and user accuracies (>0.84 and >0.89). As a result, the classification results showed a high degree accuracy; thus, the possibility of applying Sentinel-2 data for tree species classification was confirmed, and Sentinel-2 data were applied for classifying the tree species of the KNA and MTB.

Table 7. Species classification results of 55 bands and Sentinel-2 band combination from Hyperion data.

Classification 1	RF			UA	SVM			UA
	Korean Pine	Japanese larch	Total		Korean Pine	Japanese Larch	Total	
Korean pine	237	61	298	0.80	244	54	298	0.82
Japanese larch	44	255	299	0.85	40	259	299	0.87
Total	281	316	597		284	313	597	
PA	0.84	0.81			0.86	0.83		
Overall accuracy	0.82	kappa statistics	0.65	Overall accuracy	0.84	kappa statistics	0.69	
Classification 2								
Korean pine	223	45	268	0.83	232	36	268	0.87
Japanese larch	41	231	272	0.85	35	237	272	0.87
Total	264	276	540		267	273	540	
PA	0.84	0.84			0.87	0.87		
Overall accuracy	0.84	kappa statistics	0.68	Overall accuracy	0.87	kappa statistics	0.74	
Classification 3								
Korean pine	215	30	245	0.88	219	26	245	0.89
Japanese larch	31	218	249	0.88	24	225	249	0.90
Total	246	248	494		243	251	494	
PA	0.87	0.88			0.90	0.90		
Overall accuracy	0.88	kappa statistics	0.75	Overall accuracy	0.90	kappa statistics	0.80	
Sentinel-2 band combination from Hyperion data								
Korean pine	217	28	245	0.89	220	25	245	0.90
Japanese larch	40	209	249	0.84	28	221	249	0.89
Total	257	237	494		248	246	494	
PA	0.84	0.88			0.89	0.90		
Overall accuracy	0.86	kappa statistics	0.72	Overall accuracy	0.89	kappa statistics	0.79	

Table 8. Comparison of the results of 147 versus 55 bands using RF and SVM.

	147 versus 55 Using Random Forest	147 versus 55 Using Support Vector Machine
Classification 1	0.96	0.92
Classification 2	0.96	0.93
Classification 3	0.97	0.94

3.4. Sentinel-2 Analysis

Tree species classification was performed with Sentinel-2 data in the KNA. In RF and SVM classification using Sentinel-2 data, the overall accuracies of 0.89 and 0.86, with kappa statistics of 0.77 and 0.72, respectively, were achieved (Table 9). Overall, RFs and SVMs yielded high producer accuracies (>0.86 and >0.83, respectively) and user accuracies (>0.86 and >0.84, respectively) (Table 9). While SVMs were more accurate than RFs in classifications using Hyperion data, RFs were more accurate in classifications using Sentinel-2 data. Figure 4 shows the visual comparison of the forest map versus tree species classification maps derived from Hyperion data and Sentinel-2 data in the KNA. As shown in this figure, the shapes of the forest map, Hyperion results, and Sentinel-2 results are quite similar.

Table 9. Species classification results of the KNA using Sentinel-2 data.

Classification	RF				SVM			
	Korean Pine	Japanese Larch	Total	UA	Korean Pine	Japanese Larch	Total	UA
Korean pine	173	28	201	0.86	169	33	202	0.84
Japanese larch	16	167	183	0.91	20	162	182	0.89
Total	189	195	384		189	195	384	
PA	0.92	0.86			0.89	0.83		
Overall accuracy	0.89	kappa statistics	0.77	Overall accuracy	0.86	kappa statistics	0.72	

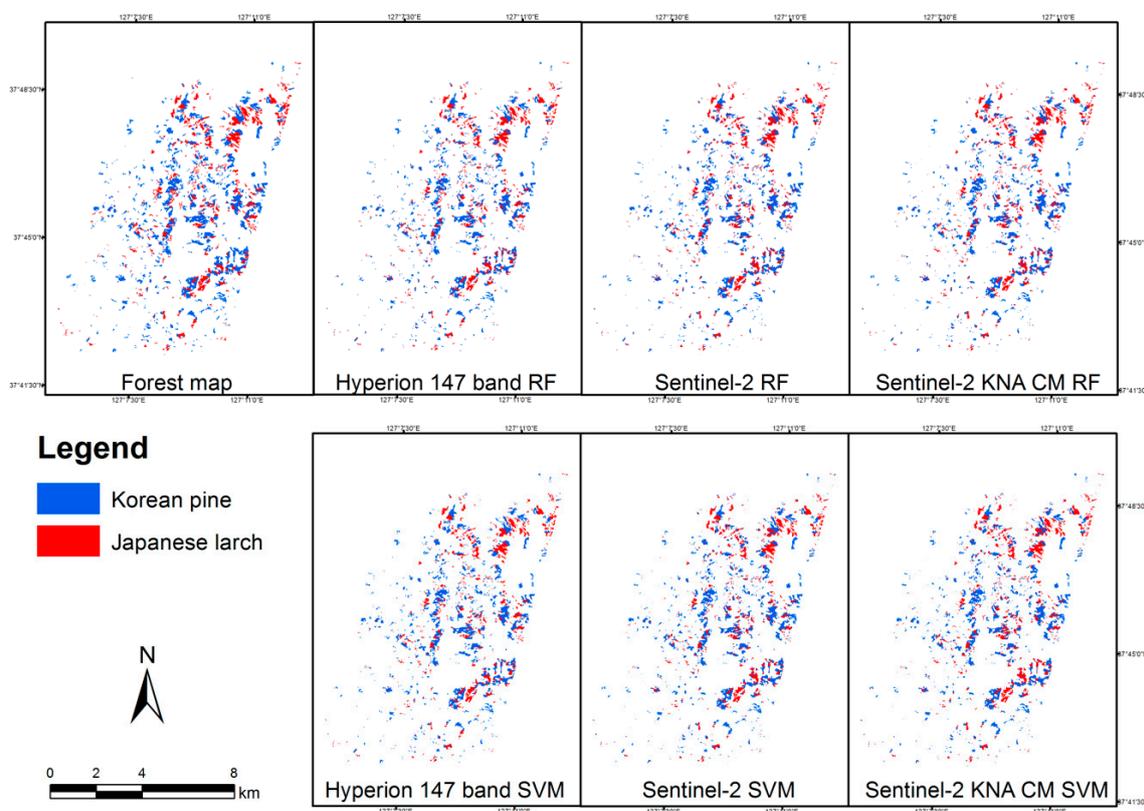


Figure 4. Species classification results of the KNA. Abbreviations: CM, combined training data model.

3.5. Classification of Mt. Baekdu

In order to investigate the possibility of using the training dataset of the KNA for tree species classification of MTB, RFs and SVMs were utilized and showed overall accuracies of 0.60 and 0.51 and kappa values of 0.20 and 0.00, respectively (Table 10). The producer accuracies of Korean pine and Japanese larch were 0.97 and 0.23, respectively, while the user accuracies of these species were 0.56 (Korean pine) and 0.87 (Japanese larch). Using the KNA data, the model classified most Japanese larch as Korean pine. It was assumed that spatial differences cause the spectral variation between the KNA and MTB, even though the same tree species are observed [12]. As shown in Figure 5, the NIR spectra of MTB are higher than those of the KNA. According to the results of previous studies (Table 6), NIR is related to leaf and canopy structures, such as vegetation stress and dynamics, the leaf area index (LAI), and starch. It was also assumed that trees in the KNA have more vegetation stress and lower LAIs than those in the MTB region. To support this hypothesis, the normalized difference vegetation index (NDVI), which is highly correlated with LAI and vegetation conditions [40], was used to compare the KNA and MTB. The mean NDVI of MTB (0.864) was higher than that of the KNA (0.836). Thus, it was assumed that the forest density or conditions caused spectral differences and classification errors. Additionally, a spectrum of Korean pine was smaller than that of Japanese larch, unlike the previously reported general patterning [74].

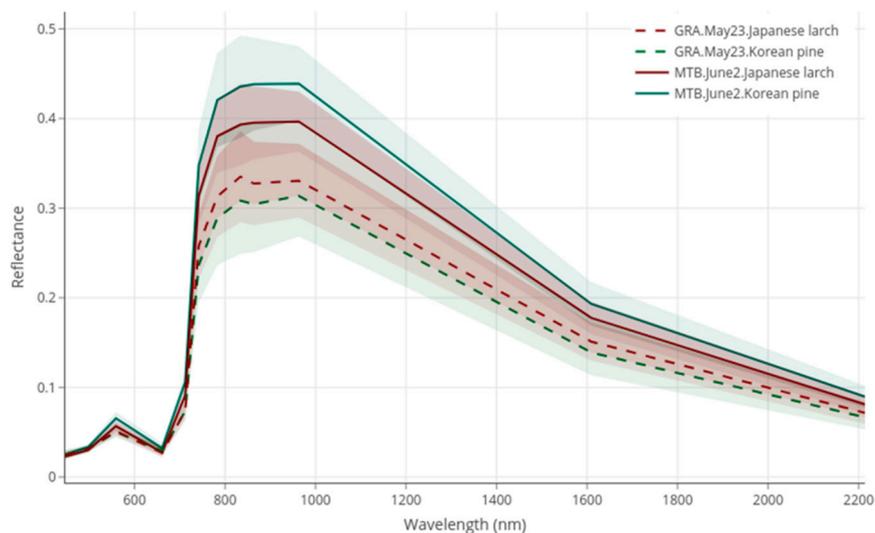


Figure 5. Spectral comparison between the KNA and MTB.

The NDVI of Korean pine in the KNA (0.823) was lower than that of Japanese larch in the KNA (0.849). In line with the comparison data between the KNA and MTB, it was assumed that Korean pines in the KNA had lower LAIs or exhibited abnormal vegetation conditions in May. The difference between the mean spectra of Korean pine and Japanese larch was small, while the variation in the two spectra overlapped in various ranges. This information could be used to classify tree species reasonably well within the same geographic location. However, it could not be adopted among different geographic locations, as has been similarly shown in previous studies [12].

In order to increase the predictive ability of the model, the training data of the KNA were combined with training data of MTB, which showed a high classification accuracy (i.e., >0.98 overall accuracy and >0.97 kappa statistics) (Table 10). The overall accuracy and kappa statistics of RF and SVM classifications were improved from 0.60 and 0.20, and 0.51 and 0.00 to 0.99 and 0.97, and 0.98 and 0.95, respectively. Overall, RFs and SVMs yielded higher producer accuracies (>0.98 and >0.97) and user accuracies (>0.98 and >0.97) (Table 10). Additionally, the combined training data model showed good performance in KNA classifications, with high overall accuracy (>0.88) and kappa statistic (>0.76).

Table 10. Classification results of MTB with Sentinel-2 data and results of the combined training data model in the MTB region and the KNA, where MTB KNA means classification of MTB species with the training data of the KNA; MTB means classification of MTB species using its own training data; MTB CM means classification of MTB species with the combined training data; KNA CM means classification of KNA species with the combined training data.

	RF				SVM			
MTB KNA	Korean Pine	Japanese Larch	Total	UA	Korean Pine	Japanese Larch	Total	UA
Korean pine	325	253	578	0.56	336	328	664	0.51
Japanese larch	11	75	86	0.87	0	0	0	0
Total	336	328	664		336	328	664	
PA	0.97	0.23			1.00	0.00		
Overall accuracy	0.60	kappa statistics	0.20	Overall accuracy	0.51	kappa statistics	0.00	
MTB								
Korean pine	335	9	344	0.97	333	8	341	0.98
Japanese larch	1	319	320	1.00	3	320	323	0.99
Total	336	328	664		336	328	664	
PA	1.00	0.97			0.99	0.98		
Overall accuracy	0.98	kappa statistics	0.97	Overall accuracy	0.98	kappa statistics	0.97	
MTB CM								
Korean pine	335	8	343	0.98	330	10	340	0.97
Japanese larch	1	320	321	1.00	6	318	324	0.98
Total	336	328	664		336	328	664	
PA	1.00	0.98			0.98	0.97		
Overall accuracy	0.98	kappa statistics	0.97	Overall accuracy	0.98	kappa statistics	0.97	
KNA CM								
Korean pine	171	28	199	0.86	169	25	194	0.87
Japanese larch	18	167	185	0.90	20	170	190	0.89
Total	189	195	384		189	195	384	
PA	0.90	0.86			0.89	0.87		
Overall accuracy	0.88	kappa statistics	0.76	Overall accuracy	0.88	kappa statistics	0.77	

Figure 6 shows the visual comparison of the forest map versus tree species classification maps derived from Sentinel-2 data in the MTB region. As shown in the figure, the shapes of the forest map and the combined training data model results are quite similar. This demonstrates that the combined training data models were accurate enough for use in predicting Korean pine and Japanese larch in both regions. Based on this result, it is assumed that Korean pine and Japanese larch in North Korea can be predicted using the developed model, as North Korea is geographically located between the KNA and MTB.

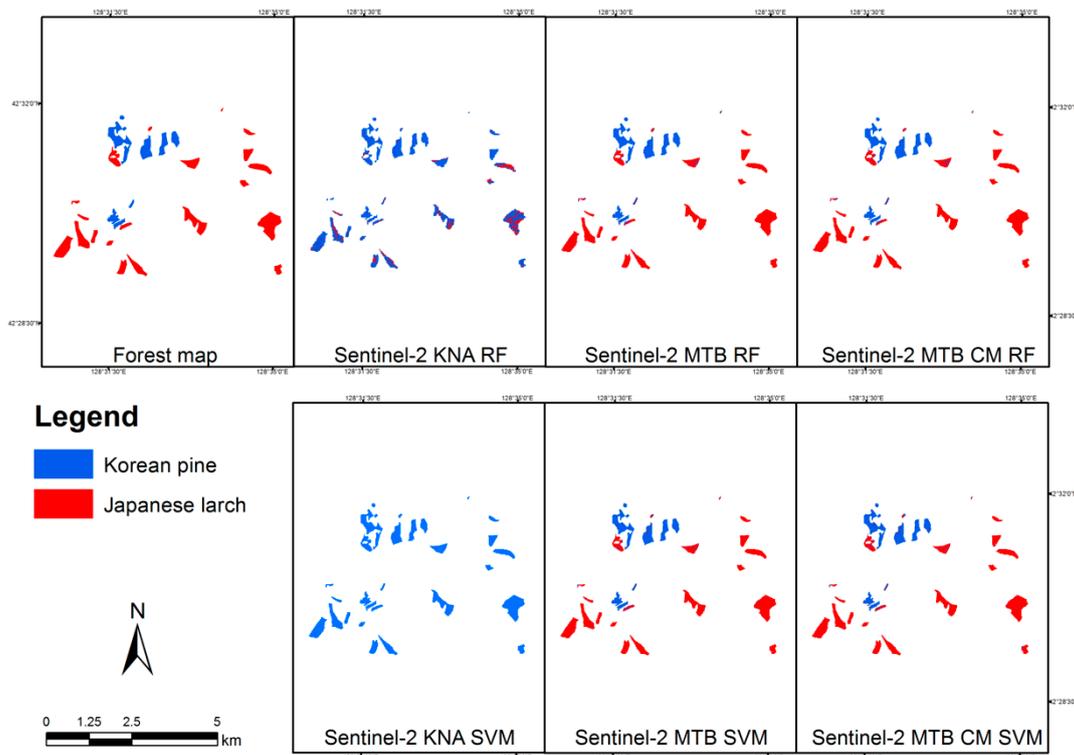


Figure 6. Species classification results of the MTB region.

4. Conclusions

In this study, a tree species classification model was developed by combining training data from South Korea with those from China in order to predict the distribution of tree species in North Korea. From the study results, three research questions were answered:

1. Hyperion data with machine learning algorithms, such as RF and SVM, can be adopted for tree classification;
2. Sentinel-2 data may be used for tree species classification with RF and SVM algorithms corresponding Hyperion data;
3. A training dataset that was built in the KNA cannot be used for tree classification of MTB. However, combined training data from the KNA and MTB showed high classification accuracies in both regions.

The results showed that the developed model had enough reliability to predict Korean pine and Japanese larch in the Mt. Baekdu region with an accuracy of >98%. The model developed in this study is able to classify tree species in inaccessible regions, such as North Korea. In order to improve the model, more tree species sampling on the Korean peninsula and across Northeast Asia should be performed in the future.

Author Contributions: Conceptualization, Joongbin Lim and Kyoung-Min Kim; methodology, Joongbin Lim and Kyoung-Min Kim; validation, Joongbin Lim, Kyoung-Min Kim and Ri Jin; formal analysis, Joongbin Lim;

investigation, Joongbin Lim; writing—original draft preparation, Joongbin Lim; writing—review and editing, Joongbin Lim, Kyoung-Min Kim and Ri Jin; visualization, Joongbin Lim; supervision, Kyoung-Min Kim; project administration, Kyoung-Min Kim.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lim, J.; Lee, K.S. Investigating flood susceptible areas in inaccessible regions using remote sensing and geographic information systems. *Environ. Monit. Assess.* **2017**, *189*, 96. [[CrossRef](#)] [[PubMed](#)]
2. Lim, J.; Lee, K.S. Flood mapping using multi-source remotely sensed data and logistic regression in the heterogeneous mountainous regions in north korea. *Remote. Sens.* **2018**, *10*, 1036. [[CrossRef](#)]
3. Lim, J. Investigation of Flood Risk Assessment in Inaccessible Regions Using Multiple Remote Sensing and Geographic Information Systems. Ph.D. Thesis, Sungkyunkwan University, Seoul, Korea, 2017.
4. Dlamini, W.M. Mapping forest and woodland loss in swaziland: 1990–2015. *Remote Sens. Appl. Soc. Environ.* **2017**, *5*, 45–53. [[CrossRef](#)]
5. Hurst, Z.M.; McCleery, R.A.; Collier, B.A.; Fletcher, R.J., Jr.; Silvy, N.J.; Taylor, P.J.; Monadjem, A. Dynamic edge effects in small mammal communities across a conservation-agricultural interface in swaziland. *PLoS ONE* **2013**, *8*, e74520. [[CrossRef](#)] [[PubMed](#)]
6. MLEP. *Democratic People's Republic of Korea Environment and Climate Change Outlook*; MLEP: Pyongyang, Korea, 2012.
7. Son, G.W. *Environmental Policies and Reality in North Korea*; Institute for Unification Education: Seoul, Korea, 2007.
8. George, R.; Padalia, H.; Kushwaha, S.P.S. Forest tree species discrimination in western himalaya using eo-1 hyperion. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *28*, 140–149. [[CrossRef](#)]
9. Sobhan, I. *Species Discrimination from a Hyperspectral Perspective*; Wageningen University: Wageningen, The Netherlands, 2007.
10. Pengra, B.W.; Johnston, C.A.; Loveland, T.R. Mapping an invasive plant, phragmites australis, in coastal wetlands using the eo-1 hyperion hyperspectral sensor. *Remote. Sens. Environ.* **2007**, *108*, 74–81. [[CrossRef](#)]
11. Su, L.; Chopping, M.J.; Rango, A.; Martonchik, J.V.; Peters, D.P.C. Support vector machines for recognition of semi-arid vegetation types using misr multi-angle imagery. *Remote. Sens. Environ.* **2007**, *107*, 299–311. [[CrossRef](#)]
12. Nidamanuri, R.R.; Zbell, B. Use of field reflectance data for crop mapping using airborne hyperspectral image. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 683–691. [[CrossRef](#)]
13. Naidoo, L.; Cho, M.A.; Mathieu, R.; Asner, G. Classification of savanna tree species, in the greater kruger national park region, by integrating hyperspectral and lidar data in a random forest data mining environment. *ISPRS J. Photogramm. Remote Sens.* **2012**, *69*, 167–179. [[CrossRef](#)]
14. Dian, Y.; Li, Z.; Pang, Y. Spectral and texture features combined for forest tree species classification with airborne hyperspectral imagery. *J. Indian Soc. Remote Sens.* **2015**, *43*, 101–107. [[CrossRef](#)]
15. Puletti, N.; Camarretta, N.; Corona, P. Evaluating eo1-hyperion capability for mapping conifer and broadleaved forests. *Eur. J. Remote Sens.* **2016**, *49*, 157–169. [[CrossRef](#)]
16. Clark, M.L.; Roberts, D.A.; Clark, D.B. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sens. Environ.* **2005**, *96*, 375–398. [[CrossRef](#)]
17. Vyas, D.; Krishnayya, N.S.R.; Manjunath, K.R.; Ray, S.S.; Panigrahy, S. Evaluation of classifiers for processing hyperion (eo-1) data of tropical vegetation. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 228–235. [[CrossRef](#)]
18. Dalponte, M.; Bruzzone, L.; Vescovo, L.; Gianelle, D. The role of spectral resolution and classifier complexity in the analysis of hyperspectral images of forest areas. *Remote Sens. Environ.* **2009**, *113*, 2345–2355. [[CrossRef](#)]
19. Baldeck, C.A.; Asner, G.P.; Martin, R.E.; Anderson, C.B.; Knapp, D.E.; Kellner, J.R.; Wright, S.J. Operational tree species mapping in a diverse tropical forest with airborne imaging spectroscopy. *PLoS ONE* **2015**, *10*, e0118403. [[CrossRef](#)]
20. Dalponte, M.; Bruzzone, L.; Gianelle, D. Tree species classification in the southern alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and lidar data. *Remote Sens. Environ.* **2012**, *123*, 258–270. [[CrossRef](#)]

21. Du, Q.; Ren, H. Real-time constrained linear discriminant analysis to target detection and classification in hyperspectral imagery. *Pattern Recognit.* **2003**, *36*, 1–12. [[CrossRef](#)]
22. Lawrence, R.; Bunn, A.; Powell, S.; Zambon, M. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sens. Environ.* **2004**, *90*, 331–336. [[CrossRef](#)]
23. Pal, M.; Mather, P.M. An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sens. Environ.* **2003**, *86*, 554–565. [[CrossRef](#)]
24. Erbek, F.S.; Özkan, C.; Taberner, M. Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. *Int. J. Remote Sens.* **2004**, *25*, 1733–1748. [[CrossRef](#)]
25. Foody, G.M. Supervised image classification by mlp and rbf neural networks with and without an exhaustively defined set of classes. *Int. J. Remote Sens.* **2004**, *25*, 3091–3104. [[CrossRef](#)]
26. Castelluccio, M.; Poggi, G.; Sansone, C.; Verdoliva, L. Land use classification in remote sensing images by convolutional neural networks. *arXiv*, 2015; arXiv:1508.00092.
27. Hu, F.; Xia, G.-S.; Hu, J.; Zhang, L. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sens.* **2015**, *7*, 14680–14707. [[CrossRef](#)]
28. Simpson, J.J.; McIntire, T.J. A recurrent neural network classifier for improved retrievals of areal extent of snow cover. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 2135–2147. [[CrossRef](#)]
29. Lyu, H.; Lu, H.; Mou, L. Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote Sens.* **2016**, *8*, 506. [[CrossRef](#)]
30. Plaza, A.; Benediktsson, J.A.; Boardman, J.W.; Brazile, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A.; et al. Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* **2009**, *113*, S110–S122. [[CrossRef](#)]
31. Chan, J.C.-W.; 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*, 2999–3011. [[CrossRef](#)]
32. Cao, J.; Leng, W.; Liu, K.; Liu, L.; He, Z.; Zhu, Y. Object-based mangrove species classification using unmanned aerial vehicle hyperspectral images and digital surface models. *Remote Sens.* **2018**, *10*, 89. [[CrossRef](#)]
33. Prospere, K.; McLaren, K.; Wilson, B. Plant species discrimination in a tropical wetland using in situ hyperspectral data. *Remote Sens.* **2014**, *6*, 8494–8523. [[CrossRef](#)]
34. Jang, Y.C.; Lee, C.H.; Kang, H.K.; Lee, S.H. Problems to be solved technologically in strengthening forest rehabilitation battle for forest restoration of whole country. *For. Sci.* **2015**, *2015*, 2–5.
35. Wikipedia. Köppen Climate Classification. Available online: https://en.wikipedia.org/wiki/K%C3%B6ppen_climate_classification (accessed on 12 March 2019).
36. Wikipedia. Humid Continental Climate. Available online: https://en.wikipedia.org/wiki/Humid_continental_climate (accessed on 12 March 2019).
37. KFS. Introduction: Korea National Arboretum. Available online: http://english.forest.go.kr/newkfsweb/html/EngHtmlPage.do?pg=/esh/org_kna/UI_KFS_0106_030110.html&mn=ENG_06_03_01 (accessed on 11 December 2018).
38. Jo, D.H. Mushrooms of mt. Baekdu (vol.1): Biota of mt. Baekdu. Available online: <https://terms.naver.com/entry.nhn?docId=3328439&cid=42526&categoryId=56828> (accessed on 11 December 2018).
39. ISO 2859-1. *Sampling Procedures for Inspection by Attributes—Part 1: Sampling Schemes Indexed by Acceptance Quality Limit (AQL) for Lot-By-Lot Inspection*; ISO: Geneva, Switzerland, 1999.
40. Jensen, J.R. *Introductory Digital Image Processing: A Remote Sensing Perspective*, 4th ed.; Pearson Series in Geographic Information Science; Pearson Education: Glenview, IL, USA, 2016.
41. Zazi, L.; Boutaleb, A.; Guettouche, M.S. Identification and mapping of clay minerals in the region of Djebel Meni (Northwestern Algeria) using hyperspectral imaging, eo-1 hyperion sensor. *Arab. J. Geosci.* **2017**, *10*. [[CrossRef](#)]
42. Louis, J.; Debaecker, V.; Pflug, B.; Main-Korn, M.; Bieniarz, J.; Mueller-Wilm, U.; Cadau, E.; Gascon, F. Sentinel-2 sen2cor: L2a processor for users. In Proceedings of the Living Planet Symposium, Prague, Czech Republic, 9–13 May 2016; p. 91.
43. Schowengerdt, R.A. *Remote Sensing: Models and Methods for Image Processing*, 3rd ed.; Academic Press: San Diego, CA, USA, 2007.

44. Peña-Barragán, J.M.; Ngugi, M.K.; Plant, R.E.; Six, J. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. *Remote Sens. Environ.* **2011**, *115*, 1301–1316. [[CrossRef](#)]
45. Culbert, P.D.; Radeloff, V.C.; St-Louis, V.; Flather, C.H.; Rittenhouse, C.D.; Albright, T.P.; Pidgeon, A.M. Modeling broad-scale patterns of avian species richness across the midwestern united states with measures of satellite image texture. *Remote Sens. Environ.* **2012**, *118*, 140–150. [[CrossRef](#)]
46. Warner, T. Kernel-based texture in remote sensing image classification. *Geogr. Compass* **2011**, *5*, 781–798. [[CrossRef](#)]
47. Luo, L.; Mountrakis, G. Integrating intermediate inputs from partially classified images within a hybrid classification framework: An impervious surface estimation example. *Remote Sens. Environ.* **2010**, *114*, 1220–1229. [[CrossRef](#)]
48. Wang, L.; Zhang, S. Incorporation of texture information in a svm method for classifying salt cedar in western china. *Remote Sens. Lett.* **2014**, *5*, 501–510. [[CrossRef](#)]
49. Jensen, J.R.; Im, J.; Hardin, P.; Jensen, R.R. Chapter 19: Image classification. In *The Sage Handbook of Remote Sensing*; Warner, T.A., Nellis, M.D., Foody, G.M., Eds.; Sage Publications: London, UK, 2009; pp. 269–281.
50. Maillard, P. Comparing texture analysis methods through classification. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 357–367. [[CrossRef](#)]
51. Hall-Beyer, M. Gcm Texture: A Tutorial v. 3.0 March 2017. Available online: <https://prism.ucalgary.ca/handle/1880/51900> (accessed on 11 December 2018).
52. De Carvalho, O.A.; Meneses, P.R. Spectral correlation mapper (scm): An improvement on the spectral angle mapper (sam). In Proceedings of the Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18, Pasadena, CA, USA, 2000; JPL Publication: Pasadena, CA, USA, 2000.
53. Breiman, L.; Cutler, A. Random Forests. Available online: https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm (accessed on 11 December 2018).
54. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
55. Hayes, M.M.; Miller, S.N.; Murphy, M.A. High-resolution landcover classification using random forest. *Remote Sens. Lett.* **2014**, *5*, 112–121. [[CrossRef](#)]
56. Prasad, A.M.; Iverson, L.R.; Liaw, A. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* **2006**, *9*, 181–199. [[CrossRef](#)]
57. Ismail, R.; Mutanga, O.; Kumar, L. Modeling the potential distribution of pine forests susceptible to siren noctilio infestations in mpumalanga, south africa. *Trans. GIS* **2010**, *14*, 709–726. [[CrossRef](#)]
58. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]
59. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* **2012**, *67*, 93–104. [[CrossRef](#)]
60. Feng, Q.; Liu, J.; Gong, J. Uav remote sensing for urban vegetation mapping using random forest and texture analysis. *Remote Sens.* **2015**, *7*, 1074–1094. [[CrossRef](#)]
61. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [[CrossRef](#)]
62. Guo, L.; Chehata, N.; Mallet, C.; Boukir, S. Relevance of airborne lidar and multispectral image data for urban scene classification using random forests. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 56–66. [[CrossRef](#)]
63. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using spot-5 hrg imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [[CrossRef](#)]
64. Rodriguez-Galiano, V.F.; Chica-Olmo, M.; Abarca-Hernandez, F.; Atkinson, P.M.; Jeganathan, C. Random forest classification of mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. *Remote Sens. Environ.* **2012**, *121*, 93–107. [[CrossRef](#)]
65. Vapnik, V. *Statistical Learning Theory*; Wiley: New York, NY, USA, 1998.
66. Foody, G.M.; Mathur, A. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a svm. *Remote Sens. Environ.* **2006**, *103*, 179–189. [[CrossRef](#)]
67. van der Linden, S.; Hostert, P. The influence of urban structures on impervious surface maps from airborne hyperspectral data. *Remote Sens. Environ.* **2009**, *113*, 2298–2305. [[CrossRef](#)]

68. Meyer, D. Support Vector Machines. Available online: <https://cran.r-project.org/web/packages/e1071/vignettes/svmdoc.pdf> (accessed on 11 December 2018).
69. Wang, J.; Neskovic, P.; Cooper, L.N. Training data selection for support vector machines. In Proceedings of the First International Conference, ICNC 2005, Changsha, China, 27–29 August 2005; Springer: Berlin/Heidelberg, Germany, 2005; pp. 554–564.
70. Mather, P.; Tso, B. *Classification Methods for Remotely Sensed Data*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2009.
71. Alice, M. How to Perform a Logistic Regression in R. Available online: <https://www.r-bloggers.com/how-to-perform-a-logistic-regression-in-r/> (accessed on 6 February 2017).
72. Ray, S.; Das, G.; Singh, J.P.; Panigrahy, S. Evaluation of Hyperspectral Indices for Lai Estimation and Discrimination of Potato Crop under Different Irrigation Treatments. *Int. J. Remote. Sens.* **2006**, *27*, 5373–5387. [[CrossRef](#)]
73. Green, P.E.; Carroll, J.D. *Analyzing Multivariate Data*; Harcourt Brace Jovanovich: Orlando, FL, USA, 1978.
74. Rautiainen, M.; Lukeš, P.; Homolová, L.; Hovi, A.; Pisek, J.; Möttus, M. Spectral properties of coniferous forests: A review of in situ and laboratory measurements. *Remote Sens.* **2018**, *10*, 207. [[CrossRef](#)]
75. Lee, B.S. Study on the selection of timing of satellite image data in forest resource survey. *Geol. Geogr.* **2013**, *1*, 44–45.
76. Curran, P.J. Remote sensing of foliar chemistry. *Remote Sens. Environ.* **1989**, *30*, 271–278. [[CrossRef](#)]
77. Thenkabail, P.S.; Enclona, E.A.; Ashton, M.S.; Legg, C.; De Dieu, M.J. Hyperion, ikonos, ali, and etm+ sensors in the study of african rainforests. *Remote Sens. Environ.* **2004**, *90*, 23–43. [[CrossRef](#)]
78. Asner, G.P. Hyperspectral remote sensing of canopy chemistry, physiology, and biodiversity in tropical rainforests. In *Hyperspectral Remote Sensing of Tropical and Sub-Tropical Forests*; CRS Press Taylor and Francis Group: Boca Raton, FL, USA; London, UK; New York, NY, USA, 2008; pp. 261–296.
79. Kumar, L.; Schmidt, K.; Dury, S.; Skidmore, A. Imaging spectrometry and vegetation science. In *Imaging Spectrometry: Basic Principles and Prospective Applications*; Meer, F.D.v.d., Jong, S.M.D., Eds.; Springer: Dordrecht, The Netherlands, 2001; pp. 111–155.
80. Hovi, A.; Raitio, P.; Rautiainen, M. A Spectral Analysis of 25 Boreal Tree Species. *Silva Fennica* **2017**, *51*. [[CrossRef](#)]
81. Curran, P.J.; Dungan, J.L.; Macler, B.A.; Plummer, S.E. The effect of a red leaf pigment on the relationship between red edge and chlorophyll concentration. *Remote Sens. Environ.* **1991**, *35*, 69–76. [[CrossRef](#)]
82. Thenkabail, P.S. Optimal hyperspectral narrowbands for discriminating agricultural crops. *Remote Sens. Rev.* **2001**, *20*, 257–291. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).