forests

Article

Uncovering the Potential of Multi-Temporally Integrated Satellite Imagery for Accurate Tree Species Classification

Sungeun Cha [1,†], Joongbin Lim [1,†], Kyoungmin Kim [1], Jongsoo Yim [1] and Woo-Kyun Lee [2,*]

[1] Forest ICT Research Center, National Institute of Forest Scinece, Seoul 02455, Republic of Korea
[2] Department of Environmental Science & Ecological Engineering, Korea University, Seoul 02841, Republic of Korea
* Correspondence: leewk@korea.ac.kr; Tel.: +82-2-3290-3016
† These authors equally contributed to the paper.

Abstract: In this study, prior to the launch of compact advanced satellite 500 (CAS500-4), which is an agriculture and forestry satellite, nine major tree species were classified using multi-temporally integrated imageries based on a random forest model using RapidEye and Sentinel-2. Six scenarios were devised considering the composition of the input dataset, and a random forest model was used to evaluate the accuracy of the different input datasets for each scenario. The highest accuracy, with accuracy values of 84.5% (kappa value: 0.825), was achieved by using RapidEye and Sentinel-2 spectral wavelengths along with gray-level co-occurrence matrix (GLCM) statistics (Scenario IV). In the variable importance analysis, the short-wave infrared (SWIR) band of Sentinel-2 and the GLCM statistics of RapidEye were found to be sequentially higher. This study proposes an optimal input dataset for tree species classification using the variance error range of GLCM statistics to establish an optimal range for window size calculation methodology. We also demonstrate the effectiveness of multi-temporally integrated satellite imageries in improving the accuracy of the random forest model, achieving an approximate improvement of 20.5%. The findings of this study suggest that combining the advantages of different satellite platforms and statistical methods can lead to significant improvements in tree species classification accuracy, which can contribute to better forest resource assessments and management strategies in the face of climate change.

Keywords: tree species classification; multi-temporally integrated imageries; compact advanced satellite 500 (CAS500-4); random forest (RF)

check for updates

Citation: Cha, S.; Lim, J.; Kim, K.; Yim, J.; Lee, W.-K. Uncovering the Potential of Multi-Temporally Integrated Satellite Imagery for Accurate Tree Species Classification. *Forests* **2023**, *14*, 746. https://doi.org/10.3390/f14040746

Academic Editor: Chunying Ren

Received: 13 March 2023
Revised: 29 March 2023
Accepted: 4 April 2023
Published: 5 April 2023

1. Introduction

Forest management is a critical field that is significantly impacted by climate change. In the Republic of Korea, tree species serve as a fundamental component of forest resource information, including basic forest statistics and annual forestry statistics. Reliable information on tree species is essential for various types of forest investigations, and the accuracy of forest type maps is heavily dependent on the availability of such information. However, the interpretation of forest type maps based on aerial photographs is subjective and can vary based on the skill of the individual interpreter [1]. This underscores the need to introduce remote sensing techniques in forest surveys, as they can significantly reduce the actualization period and improve the sustainability, efficiency, and wide-area detectability of forest surveys. By utilizing remote sensing in forest surveys, it is possible to obtain reliable and consistent information on tree species that is critical for effective forest management.

An effective methodology for tree species classification that matches the actual species present in the Republic of Korea's forests requires a combination of high-resolution satellite imageries and ground data. The ground data should include information on forest type, species, structure, and topography, which can be used to train and validate the classification model. Additionally, the integration of multi-temporal images can analyze the seasonal and spatial variations of the forest and improve the overall accuracy of the classification. Previous studies on tree species classification based on satellite imagery have primarily used spectral characteristics [2,3]. However, the method is mainly used to distinguish forest types, such as broad-leaved or needle-leaved forests, because it is challenging to classify species with similar spectral characteristics using multi-spectral bands of satellite imageries. In addition, the limited availability of domestic satellite imagery in the Republic of Korea's forest sector has led to the utilization of overseas satellite images such as Sentinel, Landsat, and MODIS. To address this, the government plans to develop a next-generation, medium-sized satellite to meet public demand. CAS500-4, with a spatial resolution of 5 m and five spectral bands, was developed as a result of a feasibility study and is set to launch in 2025. Its main objective is to monitor and evaluate forest resources on the Korean Peninsula. Furthermore, the satellite has a short time resolution of 1–3 days, which allows for wide-area monitoring in the Republic of Korea. Applying a classification model learned from a single-period image classification for nationwide tree species classification has limitations due to the risk of overfitting [4]. Thus, it is necessary to use multi-temporal imageries because the spectral characteristic values vary even in the same tree species depending on the time or region of satellite imagery acquisition [5–10]. The lack of red-edge and/or SWIR bands makes the data insufficient, particularly for ecosystem monitoring and management. Currently, there are no domestic satellites that include WVP or SWIR bands; therefore, satellites such as Landsat and Sentinel need to be used for analyzing forest resources. Since RapidEye and Sentinel satellite imageries differ in spatial and spectral resolutions from those of CAS500-4, some errors may occur during integration of heterogeneous imageries. Nevertheless, studies using integrated images for classification by combining specific bands have been conducted, and it is clear that the classification accuracy of tree species has improved [11,12].

The forest-type map is produced by digitizing aerial photographs and confirming forest-type division and attribute information through field surveys. Although it is difficult to visually interpret the morphological characteristics of these crowns using satellite imageries, the difference in wavelength according to the characteristics of the crowns has been estimated to be calculable in many studies [9,13–18]. These results confirm the importance of texture information in tree species classification analysis. The growth patterns of trees are highly dependent on environmental conditions such as precipitation and atmospheric temperature. Multi-temporal vegetation indices provide intrinsic phenomenological characteristics for each tree species [19,20]. Forest types can be characterized by seasonal (or climatological) changes in vegetation indices [21,22]. The research on the classification of the Republic of Korea mainly focuses on forest type classification, such as coniferous, broadleaf, and mixed forests, rather than classification by tree species. In particular, there is still a limitation that studies on the classification of non-native species in the Republic of Korea have been conducted using low-resolution time series images [20]. Therefore, it is necessary to develop an effective tree species classification methodology that uses high-resolution images to match the species actually present in the forests of the Republic of Korea.

This study investigates the effectiveness of multi-temporal satellite imagery integration for ecosystem monitoring and management through forest resource analysis. The study emphasizes the necessity of using multi-temporal imageries to capture the variability in spectral characteristic values of tree species. The study also highlights the insufficiency of domestic satellites in terms of including red-edge and/or SWIR bands for ecosystem monitoring and management, and proposes the use of Sentinel-2 for analyzing forest resources. Overall, the study aims to enhance ecosystem monitoring and management through tree species classification using the integration of multi-temporal satellite imageries.

2. Materials and Methods

The research flow for tree species classification is shown in Figure 1. It involved pre-processing of multi-temporal images from RapidEye and Sentinel-2, preparation of input datasets comprising wavelengths, gray-level co-occurrence matrix (GLCM) of NIR band, and vegetation indices, and conducting Jeffries–Matusita Distance (JMD) analysis to identify separation of variables. Random forest models were run using different input data combinations in Scenarios I, II, III, IV, V, and VI. These scenarios consisted of different combinations of input materials, with Scenarios I–III using RapidEye images and Scenarios IV–VI integrating Sentinel-2 images with wavelengths not present in RapidEye. We aimed to determine which combination of input data was most effective for tree species classification by using various combinations of input data as shown below (* RE: RapidEye; S2: Sentinel-2):

- Scenario I—wavelengths of RE;
- Scenario II—wavelengths of RE, GLCM statistics;
- Scenario III—wavelengths of RE, GLCM statistics, vegetation indices from RE;
- Scenario IV—wavelengths of RE and S2, GLCM statistics;
- Scenario V—wavelengths of RE and S2, GLCM statistics, vegetation indices from RE;
- Scenario VI—wavelengths of RE and S2, GLCM statistics, vegetation indices from RE and S2.

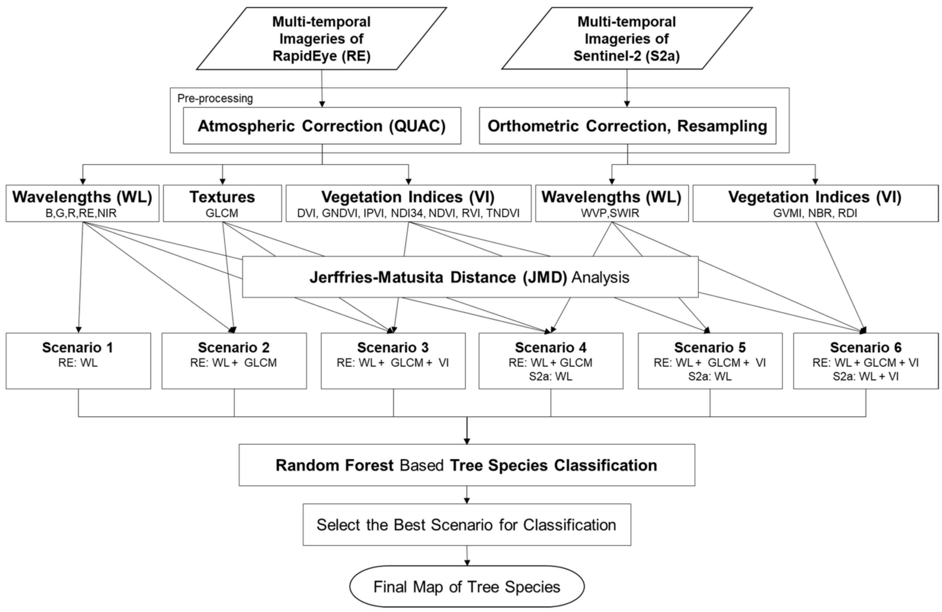

Figure 1. Research flow for the study.

2.1. Study Sites

This study was conducted in the area of the Korea National Arboretum (KNA) located in the Gwangneung Forest between Namyangju and Pocheon City in Gyeonggi Province, Republic of Korea (37°42′ N–37°49′ N, 127°5′ E–127°13′ E) (Figure 2). The Gwangneung Forest was a royal forest surrounding King Sejo's tomb during the Joseon Dynasty and has been strictly maintained to minimize human disturbance for over five hundred years [23]. Established in 1987, the KNA has 15 specialized botanical gardens, forest museums, herbariums, and seed banks. It has been open to the public since 1999 and was designated as a UNESCO Biosphere Reserve in June 2010. The KNA focuses on research, collection, classification, and conservation of forest species, as well as forest environment education. It is also committed to the restoration and conservation of rare and native plants. Analysis of meteorological data from Dongducheon, which is geographically close to Gwangneung Forest, shows that the average annual rainfall, temperature, and wind speed between 1981 and 2010 were 1503 mm, 11.2 °C, and 1.6 m/s, respectively [23]. According to data collected between 2011 and 2015, the average annual rainfall was 1378 mm, the average temperature was 15.7 °C, and the average wind speed was 1.6 m/s. Gwangneung Forest is mainly distributed on weathered and metamorphic rocks, including gneiss and schist [23].

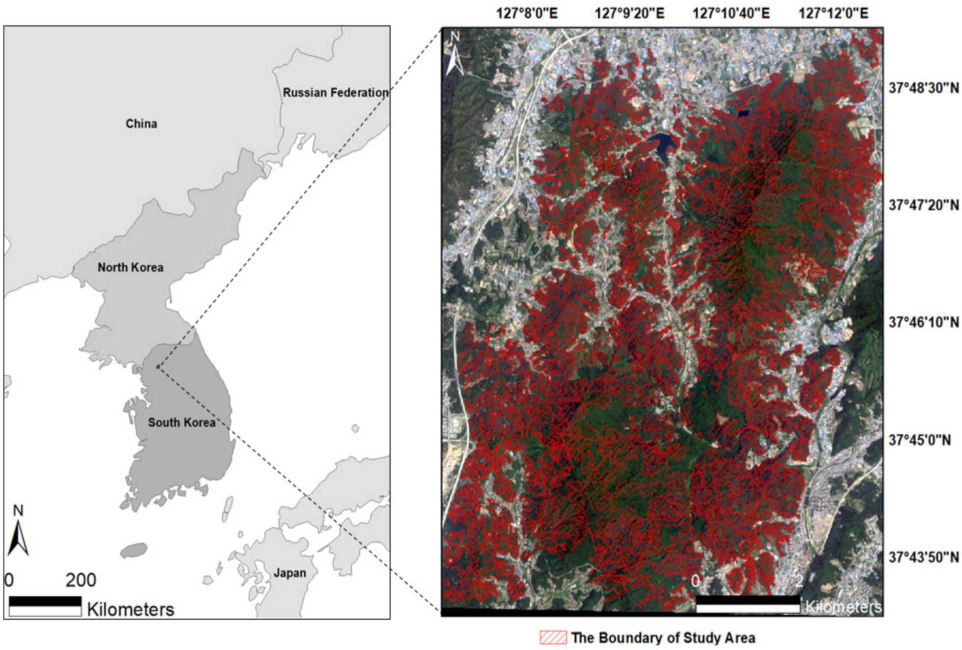

Figure 2. Boundary of study area.

2.2. Pre-Processing of Satellite Imageries

Tree species classification was performed using satellite imageries from RapidEye and Sentinel-2. Sentinel images can be downloaded for free through the Copernicus Open Access Hub of the European Space Agency (ESA) "https://scihub.copernicus.eu/ (accessed on 3 April 2023)". However, RapidEye images must be purchased through contracted vendors as they are commercial and not freely available. The usability of CAS500-4, which has similar specifications to RapidEye, was confirmed through its use in the study, with a total of five spectral bands and a spatial resolution of 5 m. Integrated bands were also created using Sentinel-2 images with WVP and SWIR bands, which were not available in CAS500-4, but were expected to have a significant impact on the accuracy of tree species classification. Given that broad-leaved forests often exhibit large differences in the presence or absence of leaves and shape depending on the season, multi-temporal images were selected based on the growing season for each species in the broad-leaved forests (Table 1). For Sentinel-2 and RapidEye images, geometric and orthometric corrections were performed based on a refined forest-type map and digital elevation map. In addition, in order to equalize the spatial resolution, the resampling and snap functions were used to adjust the spatial resolution of Sentinel-2 to RapidEye.

Table 1. Comparison and information of satellites' specification.

Satellite	Spatial Resolution (m)	Time Resolution (Days)	Swath Width (km)	Spectral Bands	Center Wavelength (nm)	Band Width (nm)
CAS500-4	5	1~3	125	Blue	490	65
				Green	560	35
				Red	665	30
				Red Edge	705	15
				NIR	842	115
RapidEye	5	5.5	77	Blue	475	70
				Green	555	70
				Red	657.5	55
				Red Edge	710	40
				NIR	805	90
	* Acquisition date of RapidEye imageries: 8 March 2019, 26 May 2019, 4 July 2019, 24 September 2019, 5 December 2019.					
Sentinel-2	60	5	290	Coastal aerosol	443	20
	10			Blue	490	65
				Green	560	35
				Red	665	30
	20			Red Edge	705	15
				Red Edge	740	15
				Red Edge	783	20
	10			NIR	842	115
	20			Red Edge	865	20
	60			Water vapor	945	20
				SWIR-Cirrus	1375	30
	20			SWIR	1610	90
				SWIR	2190	180
	* Acquisition date of Sentinel-2 imageries: 27 March 2019, 23 May 2019, 30 September 2019.					

To ensure the accuracy of the imageries, orthometric corrections were performed for both Sentinel-2 and RapidEye images based on a refined forest-type map and digital elevation map. Moreover, to equalize the spatial resolution of the images, the resampling and snap functions were employed to adjust the spatial resolution of Sentinel-2 to that of RapidEye.

The RapidEye images used in this study were obtained at Level 1B and underwent pre-processing stage consisting of radiation and geometric corrections. In order to perform atmospheric correction, the quick atmospheric correction (QUAC) model was utilized, which is a method for atmospheric correction of hyper-spectral and multi-spectral images from visible light to near-infrared and short-wave infrared regions [24].

QUAC is a simple and effective method that uses the spectral reflectance and characteristics of the image itself, without requiring any physical characteristics of the sensor or prior information at the time of shooting. The method involves dividing the image into several end members, each representing an independent field of view. Equation (1) shows the mathematical formula used for this process:

$$p = \frac{(p_1 + p_2 + \cdots + p_n)}{n} \tag{1}$$

where n indicates the number of end members.

For atmospheric correction of RapidEye, the wavelength values of the already atmospherically corrected Sentinel-2 Level-2A images were used as a comparison target, and the corrected RapidEye images were compared to these values. After atmospheric correction, the difference in wavelength range of the simultaneous RapidEye image was corrected to less than 5%. Figure 3a,b show the wavelength range before and after atmospheric correction of RapidEye, while Figure 3c shows the wavelength range of the already atmospherically corrected Sentinel-2 images.

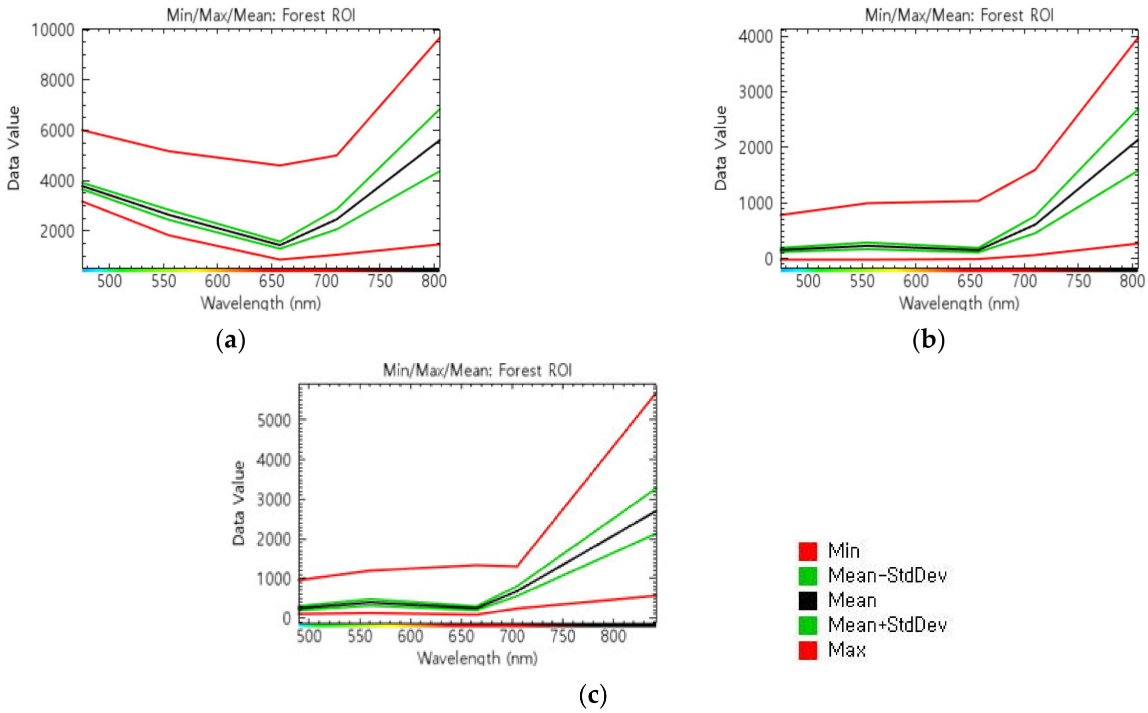

Figure 3. Comparison of the Statistics of Wavelengths for Atmospheric Correction. (**a**) RapidEye: before atmospheric correction; (**b**) RapidEye: after atmospheric correction; (**c**) Sentinel-2: level 2a image.

2.3. Texture Analysis

The gray-level co-occurrence matrix (GLCM) is a remote sensing technique that quantitatively reflects the texture characteristics of tree species [25,26]. GLCM, developed by [27], is a well-known technique in remote sensing for producing texture information by considering the relationship between pixels. GLCM has been widely used in tree species and land cover classification using texture [7,18,28,29].

GLCM represents the distance and angle relationship of a specific area of an image of a specified size. Texture is quantitatively expressed by measuring the simultaneous occurrence frequency of grayscale pixel brightness value pairs in a user-defined moving window. In this study, seven types of GLCM were extracted according to statistics: mean, variance, homogeneity, contrast, dissimilarity, entropy, and angular second moment (Equations (2)–(8)). GLCM is a matrix indicating the frequency represented by the corresponding pixel pair with gray-level (i, j) for the distance and direction in the image. Contrast, dissimilarity, and homogeneity are related to the contrast and homogeneity of brightness, and entropy and angular second moment are related to regularity. The window size used in the generation of texture information inevitably affects the accuracy of tree species classification [30–32].

$$\text{MEAN} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} i \times h_c(i,\ j) \tag{2}$$

$$\text{VARIANCE} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} (i-\mu)^2 \times h_c(i,\ j) \tag{3}$$

$$\text{HOM} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} \frac{1}{1+(i-j)^2} \times h_c(i,\ j) \tag{4}$$

$$\text{CON} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} (i-j)^2 \times h_c(i,\ j)^2 \tag{5}$$

$$\text{DIS} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i,\ j)^2 \times |i-j| \tag{6}$$

$$\text{ENT} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i,\ j) \times \log[h_c(i,\ j)] \tag{7}$$

$$\text{ASM} = \sum_{i=0}^{quant_k} \sum_{j=0}^{quant_k} h_c(i,\ j)^2 \tag{8}$$

where $quant_k$ is the quantization level of band k (e.g., 28 = 0 to 255) and $h_c(i,\ j)$ is the $(i,\ j)$th entry in one of the angular brightness value spatial-dependency matrices.

2.4. Vegetation Indices

The vegetation indices were utilized as inputs for deep learning to investigate their influence on tree species classification. Ten vegetation indices were computed using multispectral data acquired by RapidEye and Sentinel-2 sensors. The RapidEye-derived vegetation indices were difference vegetation index (DVI), green normalized difference vegetation index (GNDVI), infrared percentage vegetation index (IPVI), normalized difference index (NDI34), normalized difference vegetation index (NDVI), ratio vegetation index (RVI), and transformed normalized difference vegetation index (TNDVI). The Sentinel-2-derived vegetation indices include global vegetation moisture index (GVMI), normalized burn ratio (NBR), and simple MIR/NIR ratio drought index (RDI). Table 2 provides details of the vegetation indices used in this study.

Table 2. Vegetation Indices Derived from RapidEye and Sentinel-2.

Satellite	Wavelength	Indices	Formula	Reference
RapidEye	Visible and Near Infrared (VNIR)	Difference Vegetation Index (DVI)	$B_{NIR} - B_{Red}$	Pettorelli et al., 2005 [33]
		Green Normalized Difference Vegetation Index (GNDVI)	$\frac{B_{NIR}-B_{Green}}{B_{NIR}+B_{Green}}$	Buschmann and Nagel, 1993 [34]
		Infrared Percentage Vegetation Index (IPVI)	$\frac{B_{NIR}}{B_{NIR}+B_{Red}}$	Crippen, 1990 [35]

Table 2. *Cont.*

Satellite	Wavelength	Indices	Formula	Reference
		Normalized Difference Index (NDI34)	$\frac{B_{Red\ edge}-B_{Red}}{B_{Red\ edge}+B_{Red}}$	Delegido et al., 2011 [36]
		Normalized Difference Vegetation Index (NDVI)	$\frac{B_{NIR}-B_{Red}}{B_{NIR}+B_{Red}}$	Rouse et al., 1974 [37]
		Ratio Vegetation Index (RVI)	$\frac{B_{NIR}}{B_{Red}}$	Major et al., 1990 [38]
		Transformed Normalized Difference Vegetation Index (TNDVI)	$\sqrt{\frac{B_{NIR}-B_{Red}}{(B_{NIR}+B_{Red})+0.5}}$	Senseman et al., 1996 [39]
Sentinel-2	Visible and Near Infrared (VNIR), Short-wave Infrared (SWIR)	Global Vegetation Moisture Index (GVMI)	$\frac{(B_{WVP}+0.1)-(B_{SWIR}+0.02)}{(B_{WVP}+0.1)+(B_{SWIR}+0.02)}$	Ceccato et al., 2002 [40]
		Normalized Burn Ratio (NBR)	$\frac{B_{WVP}-B_{SWIR}}{B_{WVP}+B_{SWIR}}$	Key et al., 2002 [41]
		Simple Ratio MIR/NIR Ratio Drought Index (RDI)	$\frac{B_{SWIR}}{B_{WVP}}$	Ill and McLeod, 1992 [42]

2.5. Spectral Separability and Similarity Analysis

The Jeffries–Matusita Distance (JMD) was utilized to assess the spectral separability between different tree species in the input data. JMD is a distance metric that ranges from 0 (identical distributions) to 1.414 (complete dissimilarity) and is frequently used to quantify the degree of separation [43–45]. Equation (9) was employed to compute the JMD between the probability distributions of the classification categories.

$$J_{i,\,j} = \int_x \sqrt{p(x|w_i)} - \sqrt{p\,(x|w_j)}^2 dx \tag{9}$$

where i and j represent two classification categories, and $p(x\,|\,w_i)$ represents the probability density function of variable x in category i. $p(x\,|\,w_j)$ represents the probability density function of variable x in category j.

The JMD is a measure of the average distance between the density functions of the two categories. If the classification category follows a normal distribution, the expression can be expressed in terms of the Bhattacharyya distance (Equations (10) and (11)).

$$J_{i,\,j} = 2\left(1 - e^{-B}\right) \tag{10}$$

$$\mathrm{B} = \frac{1}{8}\left(m_i - m_j\right)^t \frac{\sum_i + \sum_j \left(m_i - m_j\right)}{2^{-1}} \frac{1}{2} \ln \frac{\left|\left(\sum_i + \sum_j /2\right|\right.}{\left|\sum_i\right|^{1/2}\left|\sum_j\right|^{1/2}} \tag{11}$$

where m_i and m_j represent the means of variable x in categories i and j, respectively, and $\sum_{\mathrm{I}}$, $\sum_j$ represent the covariance matrices of variable x in categories i and j, respectively. ln represents the natural logarithm.

JMD is a method for measuring the difference between two probability distributions, which can be used to understand the spectral characteristics of each tree species and select variables that can be used as independent variables in a random forest model. To comprehend the spectral characteristics of each tree species, the wavelength reflectance data was extracted from all points within the study area. The JMD values were then computed by taking the average of the spectral values for each species. In order to utilize an independent variable in the random forest models, there must be a discernible difference between the factors. If the machine learning model is trained with the same or similar input dataset, the reliability of the input may decrease, resulting in a reduction in classification accuracy. Thus, separate variables derived from vegetation indices should be utilized as independent variables for training the model only when the degree of separation between variables varies significantly.

2.6. Refined Forest Type Map

The forest type map was utilized as a primary data source for tree species classification. The forest type map is a commonly used forest map in the Republic of Korea, which provides various attributes such as tree type, tree species, diameter, age, and crown density, and is one of the major theme maps produced by national institutions on a national scale. The forest type map, with a scale of 1:5000, was field surveyed from 2006 to 2019, and was used to set the true value of the tree species in the random forest models. The forest type map was investigated in units of stands and was corrected by considering the field information. Additional investigations were conducted by [46] to identify changes in misclassified species, which were then corrected with data from field surveys using electronic equipment, such as global positioning system (GPS) and Vertex. The field surveys were conducted on trees with a diameter at breast height (DBH) of 6 cm or higher for the major tree species and species with a high occupancy rate. The area of one standard plot was 0.04 ha (20 m × 20 m), and a total of 513 sample plots were selected based on the status of the tree species from the precise field survey.

The study area, Gwangneung region, has a forest cover mainly consisting of nine tree species, namely *P. densiflora*, *P. koraiensis*, *L. kaempferi*, *Pinus rigida*, *A. holophylla*, *Quercus acutissima*, *Quercus* (*aliena*, *dentata*, *serrata*), *C. crenata*, and *R. pseudoacacia* (Figure 4). These nine tree species occupy more than 55% of the total forest area distributed throughout the Republic of Korea [23,47]. The total distribution area of tree species in the study site was 28,200 ha, with *P. densiflora* covering 2702 ha (9.58%), *P. koraiensis* covering 6682 ha (23.70%), *L. kaempferi* covering 4063 ha (14.41%), *P. rigida* covering 4342 ha (15.40%), *A. holophylla* covering 786 ha (2.79%), *Q. acutissima* covering 296 ha (1.05%), *Quercus* (*aliena*, *dentata*, *serrata*) covering 8309 ha (29.46%), *C. crenata* covering 840 ha (2.98%), and *R. pseudoacacia* covering 180 ha (0.64%).

Pinus densiflora

Pinus koraiensis

Larix kaempferi (Lamb.) Carrière

Figure 4. *Cont.*

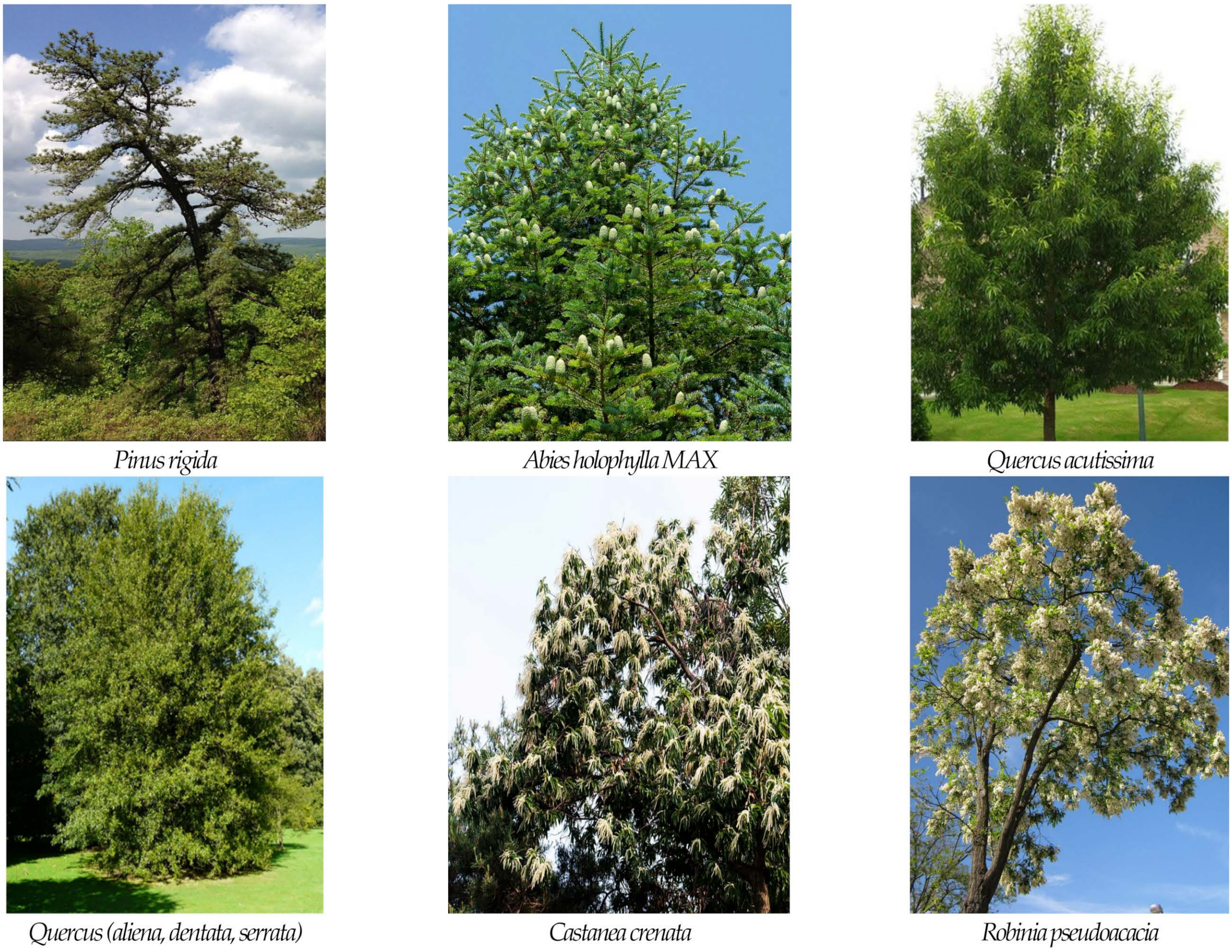

Figure 4. Photographs of Nine Tree Species Used in the Study (Wikipedia, 2023).

2.7. Random Forest Model for Tree Species Classification

According to [48], the Random Forest (RF) model is a machine learning technique designed to explain the spatial relationships between variables by generating multiple decision trees. RF creates an ensemble of decision trees, forming a forest of multiple decision trees, and performs predictions using the majority rule or average values. Unlike linear or logistic regression models, RF does not provide information on the statistical significance of individual variables. Instead, RF determines variable importance indirectly in three steps:

I. Calculating the out-of-bag (OOB) error from the raw data set.
II. Calculating the OOB error for the dataset in which the values of specific variables are randomly mixed.
III. Determining individual variable importance by considering the mean and variance of the difference between OOB errors in steps 1 and 2.

The use of RF in tree species classification models offers several advantages [48]. Firstly, RF exhibits high accuracy even in the presence of multicollinearity between variables, and there is no need to remove unnecessary variables as it considers the interaction between variables in decision-making. Secondly, RF uses ensemble learning to avoid overfitting by combining multiple decision trees, leading to higher prediction accuracy compared to models that use a single decision tree. Thirdly, RF can evaluate variable importance, making it useful for feature selection and extraction. Fourthly, RF can evaluate model performance through OOB error, which is more effective than performing cross-validation as it uses data that was not used in model training. Finally, many studies have shown that

RF outperforms other classification techniques, making it an effective tool for achieving high classification accuracy in tree species classification models.

Variable importance can be measured using the mean decrease Gini (MDG), which measures the extent to which a variable reduces the Gini impurity metric in a particular class [48,49]. RF also provides an OOB estimate of the error rate, which can be used to select the best model [16,50,51]. Previous studies have shown that RF outperforms other classification techniques [52–55]. To create the RF model in this study, bootstrap sampling with a sample size of n was performed on the training set, and an optimal classifier was selected at the end of this process. A final ensemble model of decision trees was created, and the calculation of the class of each decision tree is called classification, and the resulting average prediction value is called regression [56]. For nine species based on the refined tree-type map, 5000 points per species were extracted, and a total of 45,000 sample points were randomly generated. Of all sampled points, 70% were used as the training data set, and the remaining 30% were set as OOB, which was used to evaluate the model without using it for training.

3. Results

3.1. Gray-Level Co-Occurrence Matrix (GLCM)

In the GLCM analysis, the variability of GLCM statistics can be significantly influenced by the size of the analysis target area and the window size, which should be considered based on the spatial resolution and texture of the input data. To select an appropriate window size for the study area, the number of windows was increased by 4 pixels to account for the influence of each tree species, and the number was increased 13 times from 3×3 pixels to 51×51 pixels, based on the window size.

Seven types of GLCM statistics were calculated using the NIR band of multi-temporal RapidEye imageries. The x-axis of Figure 5 represents the increasing pixels by 4, while the y-axis indicates the variability of each statistic by normalization to the range of 0–1. The lower the variability value in the graph, the higher the similarity of the pixel value for each species. Overall, the graph shows that when the window size is mostly below 11×11 pixels, the variability value per pixel is largely large for all species. However, increasing the window size indiscriminately leads to the loss of raw data due to the padding that fills the space with 0, and the excessive blurring of image texture, making it difficult to differentiate between tree species. In this study, a window size of 31×31 pixels was chosen as the appropriate size, as it minimizes variability by tree species and reduces data loss to a minimum.

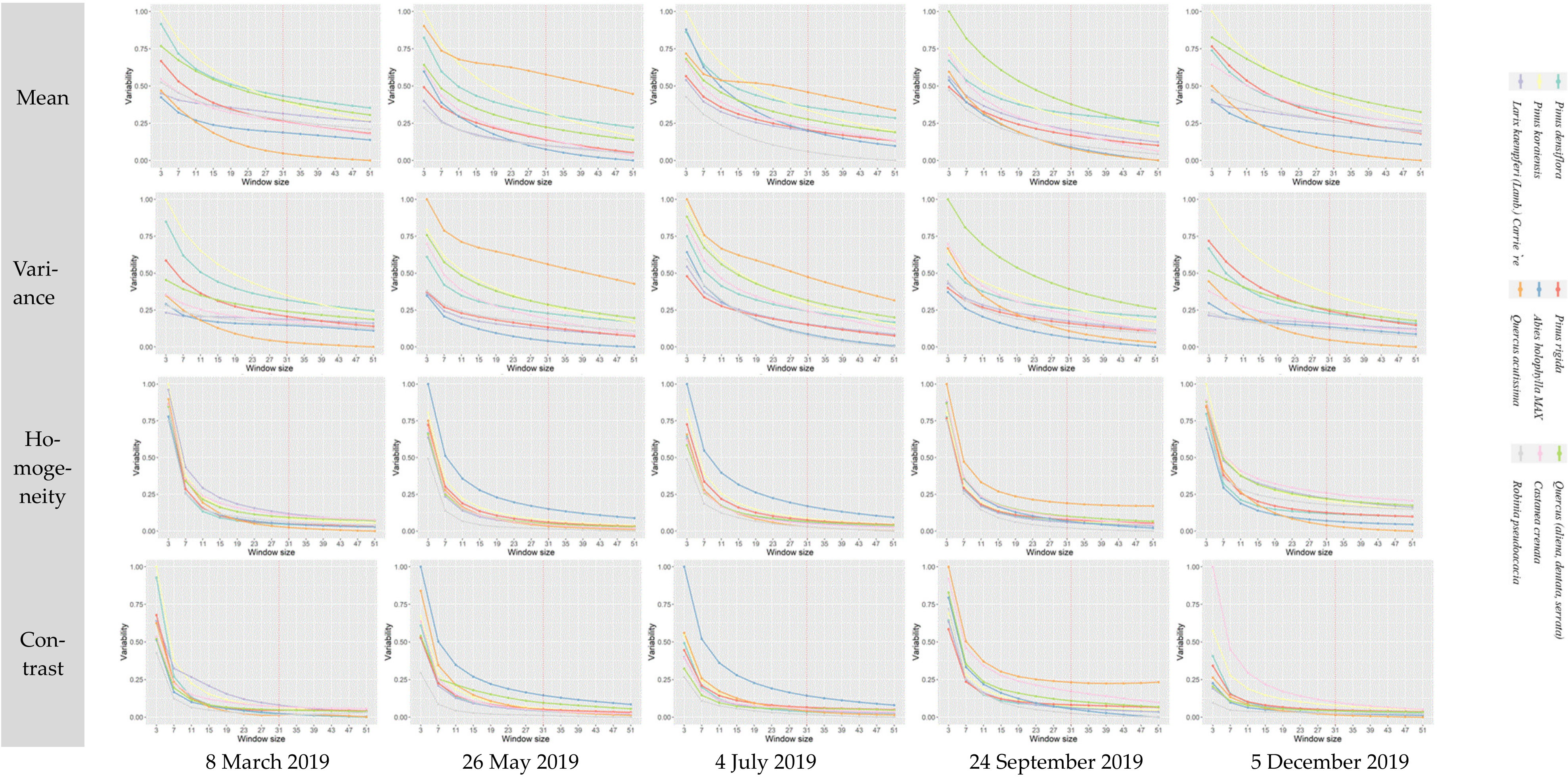

Figure 5. *Cont*.

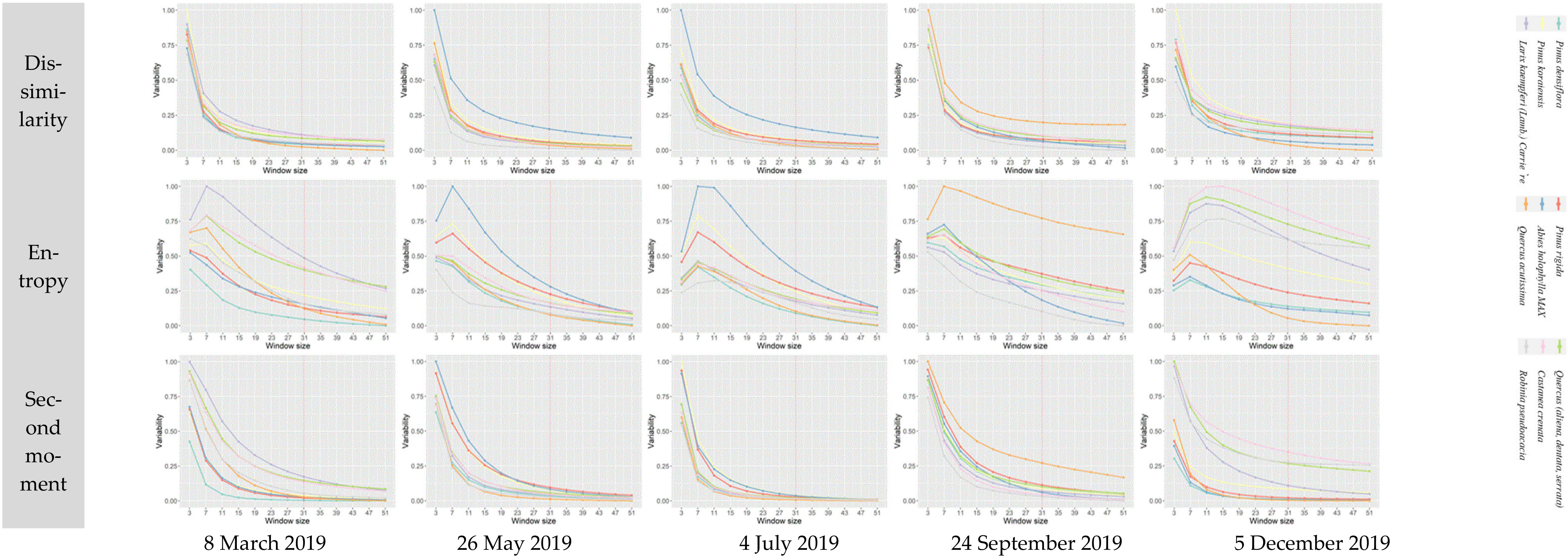

Figure 5. Variability of GLCM Statistics by Tree Species.

3.2. Spectral Separability and Similarity

The JMD (Jeffries–Matusita Distance) value, which represents the similarity between the spectral values of different tree species, was calculated by taking the average of the spectral values for each species. The separation value between *P. koraiensis* and *L. kaempferi* was found to be 0.009, which was very small, indicating a high degree of similarity between these two species. On the other hand, *A. holophylla* and *Q. acutissima* exhibited the most distinct spectral patterns, with a separation value of 0.287, which was in line with the characteristic differences between needle-leaved and broad-leaved trees. The average degrees of separation between broad-leaved and needle-leaved trees were found to be 0.060 and 0.040, respectively. This suggests that there is a spectral difference between these two types of trees. While the overall JMD value for all tree species appeared to be low at 0.110 on average, the difference by wavelength band was notable, particularly in the NIR and SWIR bands (Table 3; Figure 6).

Table 3. Jeffries–Matusita Distance (JMD) Values of Wavelengths in Tree Species.

	S1	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**
S2	0.016							
S3	0.023	0.009						
S4	0.039	0.024	0.023					
S5	0.134	0.119	0.115	0.096				
S6	0.158	0.173	0.176	0.196	0.287			
S7	0.116	0.130	0.133	0.154	0.245	0.046		
S8	0.107	0.121	0.125	0.145	0.237	0.053	0.010	
S9	0.084	0.099	0.102	0.122	0.215	0.076	0.032	0.023

S1: *Pinus densiflora*, S2: *Pinus koraiensis*, S3: *Larixkaempferi* (*Lamb.*) *Carrie 're*, S4: *Pinus rigida*, S5: *Abies holophylla MAX*, S6: *Qeurcus acutissima*, S7: *Quercus* (*aliena*, *dentata*, *serrala*), S8: *Castanea crenata*, S9: *Robinia pseudoacacia*.

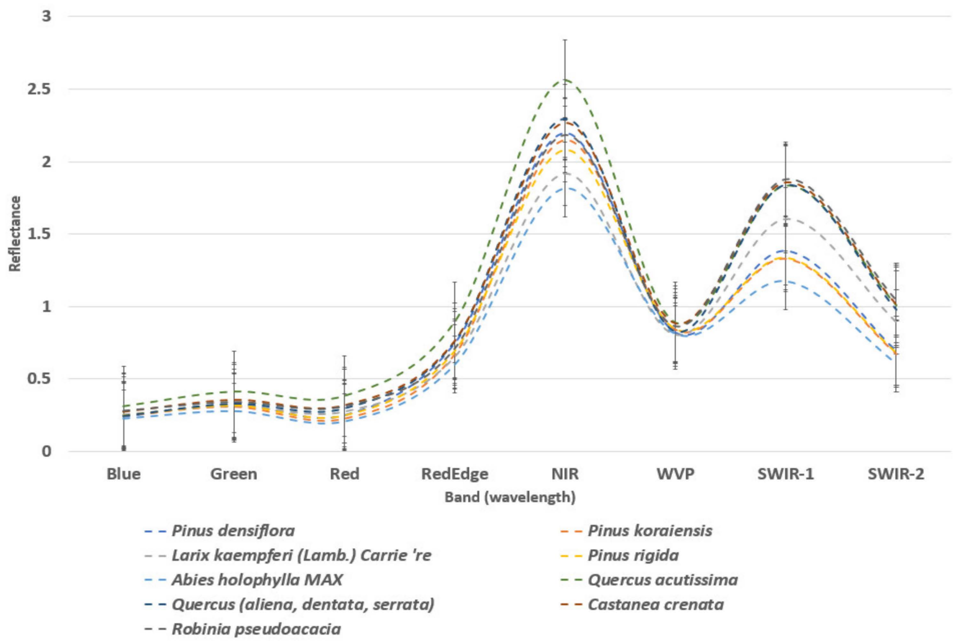

Figure 6. Spectral Distribution according to the Tree Species.

The average JMD values of RapidEye and Sentinel-2 reflectance, vegetation indices, and GLCM statistics were 1.233, 1.329, and 1.125, respectively. The smallest JMD value was 0.668, which indicated the separation between SWIR-1 and SWIR-2 in the wavelength reflectance. The separation between SWIR-1 and red-edge was also relatively small, with a difference of 0.699. However, these values still showed sufficient separation among the statistics. The vegetation indices exhibited high JMD values, indicating sufficient separation between the indices. Although the dissimilarity and angular second moment of GLCM statistics had the lowest JMD value of 0.175, their separation with other GLCM statistics was still sufficiently high, and therefore, all GLCM statistics were included as input data. Overall, the independent variables demonstrated a sufficiently high degree of separation for all indices, and therefore, they were deemed suitable for use in model evaluation.

3.3. Random Forest based Tree Classification using Multi-Temporally Integrated Satellite Imageries

The Random Forest (RF) algorithm was employed to classify tree species in different scenarios using various combinations of remote sensing data. In Scenario I, using only the spectral wavelengths of RapidEye, the tree species classification accuracy was 64.0% (kappa value: 0.595). In Scenario II, incorporating both RapidEye's spectral wavelengths and GLCM statistics, the classification accuracy improved significantly to 83.2% (kappa value: 0.812). The inclusion of GLCM information alone increased the accuracy by 19.2%, highlighting the importance of texture information in tree species classification. Despite the difficulty in visually interpreting the morphological characteristics of tree crowns in satellite imageries, numerous studies have demonstrated the statistical differences between tree species using GLCM [7,14–18].

In Scenario III, where vegetation indices produced by RapidEye were added to Scenario II, the accuracy decreased by 3.6% to 79.6% (kappa value: 0.796), suggesting that adding various vegetation indices with similar patterns may lower the accuracy of tree species classification, although each vegetation index may have an impact on the model [21,22]. In Scenario IV, where the spectral wavelengths of Sentinel-2 were added to Scenario II, the model achieved the highest accuracy of 84.5% (kappa value: 0.825). Compared to Scenario II, the accuracy increased by 1.3%, confirming the positive effects of the SWIR and WVP bands on the classification [12,13].

Adding vegetation indices produced by RapidEye in Scenario V, where Scenario IV was used as the base, resulted in a slightly decreased accuracy of 81.6% (kappa value: 0.793), similar to the result of Scenario III. In Scenario VI, where all the factors were used, the inclusion of Sentinel-2's vegetation indices in Scenario V increased the accuracy to 83.3% (kappa value: 0.801). The tree classification accuracy of RF ranged from 64.0% (kappa value: 0.595) to 84.5% (kappa value: 0.825), demonstrating that the classification accuracy varied significantly with the combination of input data. The difference between the maximum and minimum accuracies for each scenario was 20.5%, indicating a significant improvement in tree classification. Among the scenarios tested, Scenario IV was identified as the best one (Table 4; Figure 7).

Table 4. Overall Results of Random Forest Model by Scenarios.

Scenario	User Accuracy	95% CI	*p*-Value	Kappa Value
I	0.640	(0.631, 0.649)	$<2.2 \times e^{-16}$	0.595
II	0.832	(0.825, 0.839)	$<2.2 \times e^{-16}$	0.812
III	0.796	(0.788, 0.803)	$<2.2 \times e^{-16}$	0.770
IV	0.845	(0.838, 0.851)	$<2.2 \times e^{-16}$	0.825
V	0.816	(0.809, 0.823)	$<2.2 \times e^{-16}$	0.793
VI	0.833	(0.826, 0.840)	$<2.2 \times e^{-16}$	0.801

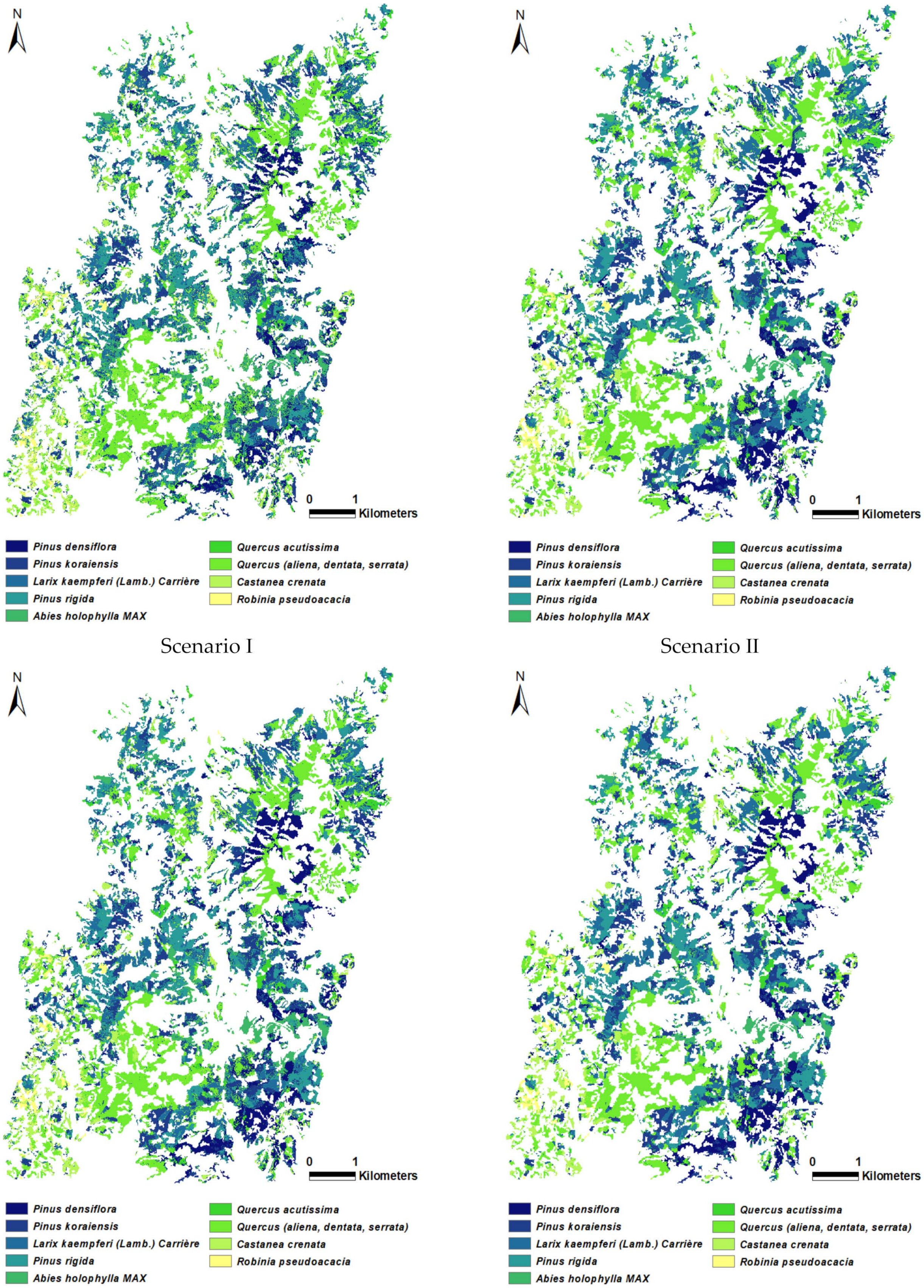

Figure 7. Cont.

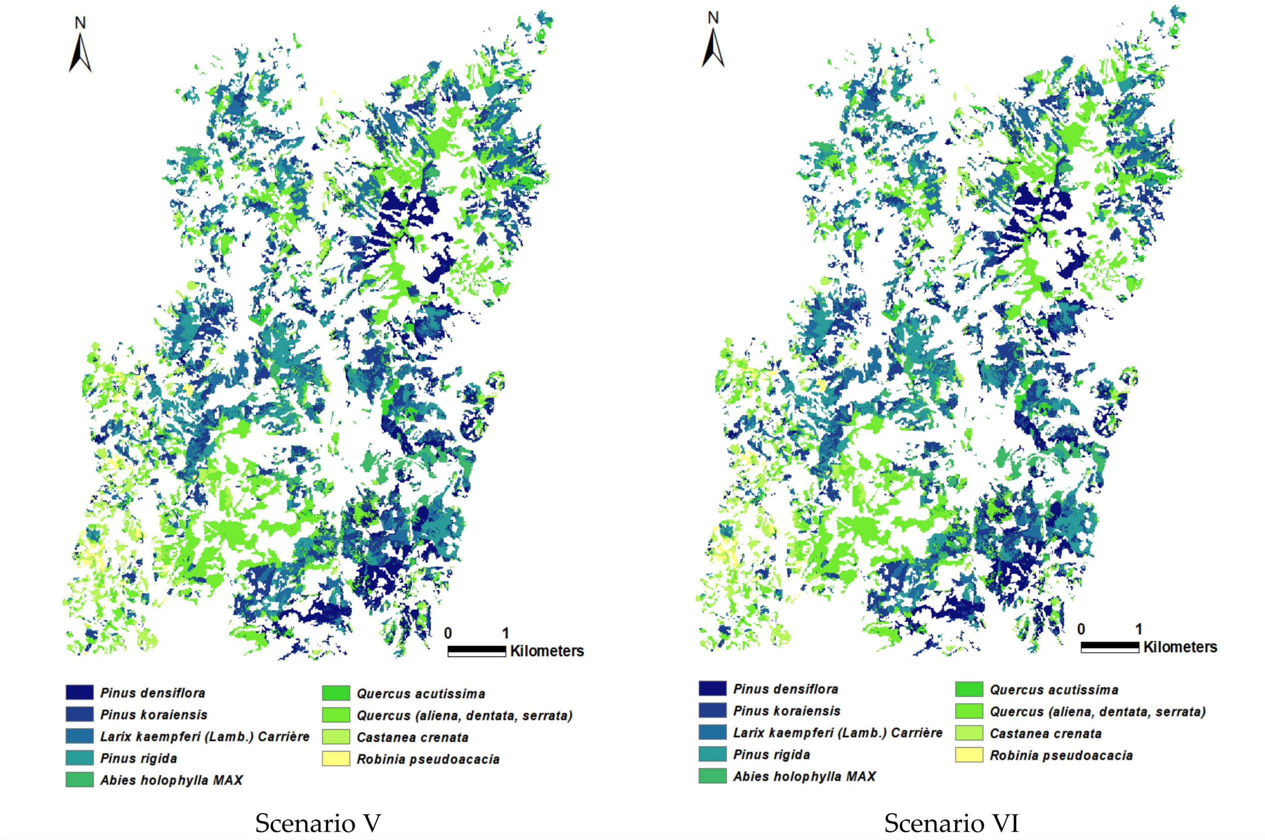

Figure 7. Map of tree species by scenarios using Random Forest.

In the Random Forest (RF) results, the variable importance was expressed by the mean decrease in the Gini value. The integration of Sentinel-2's SWIR band was found to have the greatest influence on tree species classification, confirming the importance of image integration. Additionally, the variable importance was relatively high in the WVP band and all GLCM statistics (Figure 8). Although the influence of blue, green, and red was relatively low, it is unlikely that the low variable importance significantly reduced the classification accuracy. Further analysis, such as the effect of each variable on classification accuracy using linear or multiple regression methods, is required to prove this. This study used a combination of input datasets that have been shown to sufficiently affect tree species classification in previous studies. However, the quantitative impact range of the input dataset requires additional verification from other models.

The classification accuracy for each species ranged from 79.8% for *P. koraiensis* to 96.9% for *R. pseudoacacia*. The average classification accuracies for needle-leaved and broad-leaved trees were 86.2% and 91.1%, respectively (Table 5). The results showed that the classification accuracy was higher for broad-leaved trees, in which the time-series effect was prominent due to the use of multi-temporal imageries.

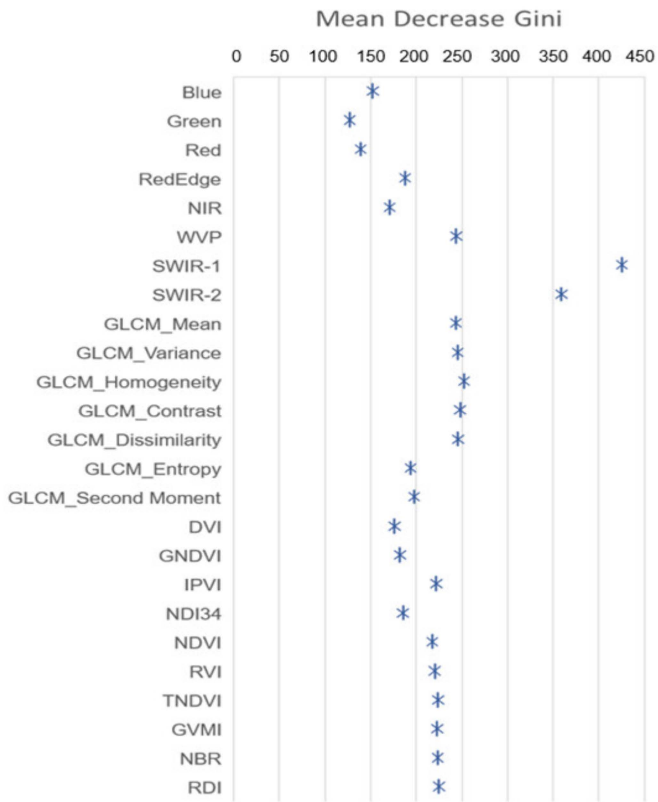

Figure 8. Variable Importance Plot from Random Forest.

Table 5. Comparison of Classification Accuracy by Tree Species using Random Forest.

Scenario	S1	S2	S3	S4	S5	S6	S7	S8	S9
I	0.745	0.741	0.830	0.776	0.846	0.852	0.748	0.737	0.901
II	0.874	0.811	0.903	0.862	0.957	0.983	0.842	0.943	0.985
III	0.849	0.798	0.873	0.851	0.936	0.969	0.823	0.898	0.975
IV	0.881	0.808	0.915	0.877	0.960	0.984	0.858	0.949	0.989
V	0.857	0.802	0.896	0.861	0.944	0.975	0.837	0.924	0.981
VI	0.863	0.807	0.899	0.866	0.953	0.977	0.835	0.928	0.985
Average	0.845	0.795	0.886	0.849	0.933	0.956	0.824	0.896	0.969
Rank	7	9	5	6	3	2	8	4	1

S1: *Pinus densiflora*, S2: *Pinus koraiensis*, S3: *Larixkaempferi* (*Lamb.*) *Carrie 're*, S4: *Pinus rigida*, S5: *Abies holophylla MAX*, S6: *Qeurcus acutissima*, S7: *Quercus* (*aliena*, *dentata*, *serrala*), S8: *Castanea crenata*, S9: *Robinia pseudoacacia*.

4. Discussion

Tree species classification using satellite imagery is a challenging task due to the complex nature of the forest ecosystem and the inherent variability in satellite imagery data. However, recent advancements in satellite technology, including high-resolution spatial and temporal data, multispectral data, and machine learning algorithms, have made accurate tree species classification feasible. In this paper, we compared the accuracy of tree species classification across different scenarios using multi-temporal integrated satellite imagery and achieved a maximum accuracy of 84.5% (kappa value: 0.825). The current state-of-the-art methods for tree species classification using satellite imagery primarily rely on machine learning algorithms, including deep learning, random forests, and support vector machines. These algorithms use spectral, textural, and contextual features extracted from satellite imagery to classify trees into different species. Despite the advancements in tree species classification using satellite imagery, there are still several challenges and limitations that need to be addressed. One of the key challenges is the lack of ground-truth data for training and validating machine learning algorithms. In this regard, future research should focus on collecting high-quality ground-truth data and developing methods for incorporating this data into machine learning algorithms. Another challenge is the need to account for the variability in satellite imagery data due to atmospheric conditions, sensor noise, and image acquisition parameters. In this regard, future research should focus on developing methods for pre-processing satellite imagery data to account for these sources of variability. Moreover, future research should focus on the integration of multi-temporal imageries to capture the temporal changes in the forest ecosystem, which can help in identifying and mapping different forest types, tracking individual tree growth and health, and capturing seasonal changes in the ecosystem. This information can also be used to classify trees based on their phenology and to monitor forest disturbances such as forest fires and insect outbreaks. Future research should also focus on the development of novel machine learning algorithms that can effectively integrate spectral, textural, and contextual features from satellite imagery data for accurate and comprehensive tree species classification. Tree species classification using satellite imagery has the potential to support forest inventory, management, and conservation.

Tree classification using satellite imagery is a significant area of research in remote sensing, forestry, and ecology. Accurate tree classification is essential for forest inventory, monitoring, and management, particularly under climate change. Multi-temporal imagery is a crucial factor in achieving accurate tree classification. It provides information about the temporal changes in the forest ecosystem and can help identify and map different forest types, track individual tree growth and health, and capture seasonal changes. This information can also be used to classify trees based on their phenology and monitor forest disturbances such as fires and insect outbreaks. Furthermore, the integration of multi-temporal satellite imagery provides several advantages for achieving accurate tree classification. Firstly, it provides high spatial and temporal resolution data that can capture the variability in the forest ecosystem. Secondly, multispectral data provided by satellite imagery can capture the spectral response of vegetation in different wavelengths, providing information on health, growth, and species, thereby improving the accuracy of tree classification. Thirdly, satellite imagery provides a large coverage area, making it possible to identify and map different forest types. Fourthly, satellite imagery provides an objective way of collecting data, reducing errors and biases associated with traditional ground-based methods. Finally, satellite imagery is cost-effective and allows for repeatable measurements over time, which can provide insights into the long-term changes in the forest ecosystem. Overall, the integration of satellite imagery is critical for accurate tree classification and offers several advantages, including high spatial and temporal resolution, multispectral data, large coverage area, objectivity, and cost-effectiveness, which are essential for forest inventory, monitoring, and management.

The utilization of texture information such as Gray-Level Co-occurrence Matrix (GLCM) is crucial for performing tree classification using satellite imagery. In this study, the optimal window size of 31×31 was extracted. Window size plays an important role in calculating the Gray-Level Co-occurrence Matrix (GLCM). It determines the size of the window used to calculate GLCM. When the window size is small, GLCM reflects local image features and is sensitive to noise. On the other hand, when the window size is large, GLCM reflects overall image features and may miss detailed information. Moreover, even with the same window size, the GLCM values may vary depending on the location in the image. This is because the pixel values within the window vary depending on the size and location of the object in the image. To reduce this error range, multiple window sizes can be used to calculate GLCM values and then averaged. Therefore, when extracting GLCM, it is important to consider the error range caused by the window size and select an appropriate window size. Using multiple window sizes and averaging them can also improve the accuracy of the extraction results.

Various factors can affect the accuracy of tree species classification. These factors include spectral variability due to differences in canopy structure, pigment content, and phenological cycles; environmental factors such as soil type, topography, and climate; and confusion with other species that may have similar spectral characteristics. For instance, coniferous trees such as pine and spruce may be challenging to differentiate based on their spectral signatures. To improve the accuracy of tree species classification, several approaches can be considered. One possible direction is to incorporate other sources of information such as LiDAR data or hyperspectral imagery, which can provide additional information on the three-dimensional structure and spectral characteristics of the forest. Additionally, machine learning algorithms such as Neural Networks can be used to classify trees, considering multiple variables and interactions between them. Moreover, further research can be conducted to investigate the potential of integrating data from different satellite sensors, such as combining optical and radar data, to improve the accuracy of tree species classification. Lastly, it is important to continue exploring the potential of advanced image processing techniques such as object-based image analysis and deep learning, which may provide more robust and accurate results for tree species classification in the future.

5. Conclusions

This study aimed to evaluate the feasibility of using satellite imagery for tree species classification, which is essential information for the forest sector under climate change. While the method has limitations such as lower spatial resolution and shadow effects, it provides a quick and efficient way to analyze tree species information nationwide. In this study, various multi-temporal imageries, GLCM, and vegetation indices were used to improve accuracy. The study resulted in the establishment of an optimal range of window size calculation methodology and the proposal of an optimal input dataset for tree species classification. The accuracy was improved by approximately 20.5% using multi-temporally integrated satellite imageries in the RF model. However, there are limitations to this study, such as the misclassification that may occur with additional tree species and complex heterogeneous forests. Subsequent studies can investigate the possibilities of utilizing more advanced remote sensing technologies and machine learning algorithms to improve the precision and versatility of tree species classification and forest monitoring across diverse ecosystems and regions. Moreover, further research is necessary to refine the approach and develop a more precise machine learning model, incorporating supplementary input variables such as topography, climate factors, and LiDAR data for various ecosystems and forest conditions. Additionally, more experiments and tests are necessary to broaden the study area and verify the feasibility of this method. Overall, this study contributes to improving the performance of the random forest model to a level that can be used in the field of forestry by continuously accumulating a large amount of training data for each tree species and applying it to the classification model.

Author Contributions: S.C. and J.L. developed overall concept and methodology. S.C. conducted analysis, J.L. and J.Y. processed data. K.K. and W.-K.L. supervised overall research. W.-K.L. supported fund for the research. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA). Grant funded by the Ministry of Land, Infrastructure, and Transport (23UMRG-B158194-04).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Acknowledgments: The authors would like to thank the anonymous reviewers and editorial staff for their comments, which greatly improved this manuscript. This study was developed in National Institute of Forest Science (Project No. 'FM0103-2021-01-2023' and 'FM0103-2021-04-2023').

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

References

1. Crisco, W.A. *Interpretation of Aerial Photographs*; US Department of the Interior, Bureau of Land Management: Anchorage, AK, USA, 1983; Volume 287, p. 38.
2. van Leeuwen, W.J.D.; Huete, A.R.; Laing, T.W. MODIS Vegetation Index Compositing Approach: A Prototype with AVHRR Data. *Remote Sens. Environ.* **1999**, *69*, 264–280. [CrossRef]
3. Guo, Q.; Zhang, J.; Guo, S.; Ye, Z.; Deng, H.; Hou, X.; Zhang, H. Urban Tree Classification Based on Object-Oriented Approach and Random Forest Algorithm Using Unmanned Aerial Vehicle (UAV) Multispectral Imagery. *Remote Sens.* **2022**, *14*, 3885. [CrossRef]
4. Ali, J.; Khan, R.; Ahmad, N.; Maqsood, I. Random forests and decision trees. *Int. J. Comput. Sci. Appl.* **2012**, *9*, 272.
5. Wang, L.; Sousa, W.P.; Gong, P.; Biging, G.S. Comparison of IKONOS and QuickBird images for mapping mangrove species on the Caribbean coast of Panama. *Remote Sens. Environ.* **2004**, *91*, 432–440. [CrossRef]
6. Puliti, S.; Talbot, B.; Astrup, R. Tree-Stump Detection, Segmentation, Classification, and Measurement Using Unmanned Aerial Vehicle (UAV) Imagery. *Forests* **2018**, *9*, 102. [CrossRef]
7. Immitzer, M.; Atzberger, C.; Koukal, T. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sens.* **2012**, *4*, 2661–2693. [CrossRef]
8. Queiroz, G.L.; McDermid, G.J.; Castilla, G.; Linke, J.; Rahman, M.M. Mapping Coarse Woody Debris with Random Forest Classification of Centimetric Aerial Imagery. *Forests* **2019**, *10*, 471. [CrossRef]
9. Lim, J.; Kim, K.-M.; Jin, R. Tree Species Classification Using Hyperion and Sentinel-2 Data with Machine Learning in South Korea and China. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 150. [CrossRef]
10. Pu, R.; Landry, S. Mapping urban tree species by integrating multi-seasonal high resolution pléiades satellite imagery with airborne LiDAR data. *Urban For. Urban Green.* **2020**, *53*, 126675. [CrossRef]
11. Masoumi, F.; Eslamkish, T.; Abkar, A.A.; Honarmand, M.; Harris, J.R. Integration of spectral, thermal, and textural features of ASTER data using Random Forests classification for lithological mapping. *J. Afr. Earth Sci.* **2017**, *129*, 445–457. [CrossRef]
12. Zhou, J.; Qiu, Y.; Chen, J.; Chen, X. A geometric misregistration resistant data fusion approach for adding red-edge (RE) and short-wave infrared (SWIR) bands to high spatial resolution imagery. *Sci. Remote Sens.* **2021**, *4*, 100033. [CrossRef]
13. Maillard, P. Comparing Texture Analysis Methods through Classification. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 357–367. [CrossRef]
14. 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]
15. 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]
16. 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.* **2014**, *43*, 101–107. [CrossRef]
17. Kukunda, C.B.; Duque-Lazo, J.; González-Ferreiro, E.; Thaden, H.; Kleinn, C. Ensemble classification of individual Pinus crowns from multispectral satellite imagery and airborne LiDAR. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *65*, 12–23. [CrossRef]
18. Kwon, S.K.; Kim, K.M.; Lim, J. A Study on Pre-evaluation of Tree Species Classification Possibility of CAS500-4 Using RapidEye Satellite Imageries. *Korean J. Remote Sens.* **2021**, *37*, 291–304.
19. Lloyd, D. A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *Int. J. Remote Sens.* **1990**, *11*, 2269–2279. [CrossRef]
20. Cha, S.-Y.; Pi, U.-H.; Yi, J.-H.; Park, C.-H. Identification of two common types of forest cover, Pinus densiflora(Pd) and Querqus mongolica(Qm), using the 1st harmonics of a Discrete Fourier Transform. *Korean J. Remote Sens.* **2011**, *27*, 329–338. [CrossRef]
21. Tucker, C.J.; Vanpraet, C.L.; Sharman, M.J.; Van Ittersum, G. Satellite remote sensing of total herbaceous biomass production in the senegalese sahel: 1980–1984. *Remote Sens. Environ.* **1985**, *17*, 233–249. [CrossRef]
22. Loveland, T.; Merchant, J.; Brown, J.; Ohlen, D. Development of a land-cover characteristics database for the conterminous U.S. *Photogramm. Eng. Remote Sens.* **1991**, *57*, 1453–1463.
23. Korea Forest Service (KFS). *The Statistical Yearbook of Forestry*; KFS: Daejeon, Republic of Korea, 2020; ISBN 11-1400000-000001-10.
24. Bernstein, L.S.; Jin, X.; Gregor, B.; Adler-Golden, S.M. Quick atmospheric correction code: Algorithm description and recent upgrades. *Opt. Eng.* **2012**, *51*, 111719-1–111719-11. [CrossRef]
25. Shojanoori, R.; Shafri, H.Z.M.; Mansor, S.; Ismail, M.H. Generic rule-sets for automated detection of urban tree species from very high-resolution satellite data. *Geocarto Int.* **2017**, *33*, 357–374. [CrossRef]
26. Kavzoglu, T.; Tonbul, H.; Colkesen, I. Evaluation of atmospheric correction methods for Sentinel-2 imagery in the spectral identification of poplar (Populus Deltoides Bartr.) species. In Proceedings of the International Symposium on Applied Geoinformatics (ISAG-2021), Riga, Latvia, 2–3 December 2021.
27. Hall-Beyer, M. GLCM texture: A tutorial. National Council on Geographic Information and Analysis. *Remote Sens. Core Curric.* **2000**, *3*, 75.
28. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *6*, 610–621. [CrossRef]
29. Wallner, A.; Elatawneh, A.; Schneider, T.; Knoke, T. Estimation of forest structural information using RapidEye satellite data. *For. Int. J. For. Res.* **2014**, *88*, 96–107. [CrossRef]

30. Mohammadpour, P.; Viegas, D.X.; Viegas, C. Vegetation Mapping with Random Forest Using Sentinel 2 and GLCM Texture Feature—A Case Study for Lousã Region, Portugal. *Remote Sens.* **2022**, *14*, 4585. [CrossRef]
31. Ferro, C.J.S. *Scale and Texture in Digital Image Classification*; West Virginia University ProQuest Dissertations Publishing: Morgantown, WV, USA, 1998; p. 1394100.
32. Lan, Z.; Liu, Y. Study on Multi-Scale Window Determination for GLCM Texture Description in High-Resolution Remote Sensing Image Geo-Analysis Supported by GIS and Domain Knowledge. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 175. [CrossRef]
33. Pettorelli, N.; Vik, J.O.; Mysterud, A.; Gaillard, J.M.; Tucker, C.J.; Stenseth, N.C. Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends Ecol. Evol.* **2005**, *20*, 503–510. [CrossRef]
34. Buschmann, C.; Nagel, E. In vivo spectroscopy and internal optics of leaves as basis for remote sensing of vegetation. *Int. J. Remote Sens.* **1993**, *14*, 711–722. [CrossRef]
35. Crippen, R. Calculating the vegetation index faster. *Remote Sens. Environ.* **1990**, *34*, 71–73. [CrossRef]
36. Delegido, J.; Verrelst, J.; Alonso, L.; Moreno, J. Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content. *Sensors* **2011**, *11*, 7063–7081. [CrossRef]
37. Rouse, J.W., Jr.; Haas, R.H.; Deering, D.W.; Schell, J.A.; Harlan, J.C. *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural vegetation*; ScienceOpen, Inc.: Burlington, MA, USA, 1974; p. E75-10354.
38. Major, D.J.; Baret, F.; Guyot, G. A ratio vegetation index adjusted for soil brightness. *Int. J. Remote Sens.* **1990**, *11*, 727–740. [CrossRef]
39. Senseman, G.M.; Bagley, C.F.; Tweddale, S.A. Correlation of rangeland cover measures to satellite-imagery-derived vegetation indices. *Geocarto Int.* **1996**, *11*, 29–38. [CrossRef]
40. Ceccato, P.; Gobron, N.; Flasse, S.; Pinty, B.; Tarantola, S. Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1: Theoretical approach. *Remote Sens. Environ.* **2002**, *82*, 188–197. [CrossRef]
41. Key, C.H.; Benson, N.; Ohlen, D.; Howard, S.; McKinley, R.; Zhu, Z. The normalized burn ratio and relationships to burn severity: Ecology, remote sensing and implementation. In Proceedings of the Ninth Forest Service Remote Sensing Applications Conference, San Diego, CA, USA, 8–12 April 2002.
42. Ill, J.E.P.; McLeod, K.W. Indications of relative drought stress in longleaf pine from Thematic Mapper data. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 495–501.
43. Davis, S.M.; Swain, P.H. *Remote Sensing: The Quantitative Approach (No. BOOK)*; McGraw-Hill International Book Company: New York, NY, USA, 1978.
44. Jensen, J.R.; Lulla, K. *Introductory Digital Image Processing: A Remote Sensing Perspective*, 4th ed.; Pearson: Glenview, IL, USA, 2015.
45. Deur, M.; Gašparović, M.; Balenović, I. Tree Species Classification in Mixed Deciduous Forests Using Very High Spatial Resolution Satellite Imagery and Machine Learning Methods. *Remote Sens.* **2020**, *12*, 3926. [CrossRef]
46. National Institute of Forest Science (NIFoS). *Forest Resource Survey of Gwangneung Test Area*; NIFoS: Deajeon, Republic of Korea, 2016; ISBN 11-1400377-000940-01.
47. George, R.; Padalia, H.; Kushwaha, S. Forest tree species discrimination in western Himalaya using EO-1 Hyperion. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *28*, 140–149. [CrossRef]
48. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
49. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
50. 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]
51. 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]
52. Pal, M. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.* **2005**, *26*, 217–222. [CrossRef]
53. 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]
54. 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]
55. 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]
56. Cutler, A.; Cutler, D.R.; Stevens, J.R. *Random Forests*; Ensemble Machine Learning Springer: Boston, MA, USA, 2012; pp. 157–175.