Article

Comparison of Machine Learning Methods for Mapping the Stand Characteristics of Temperate Forests Using Multi-Spectral Sentinel-2 Data

Kourosh Ahmadi [1], Bahareh Kalantar [2],[*], Vahideh Saeidi [3], Elaheh K. G. Harandi [4], Saeid Janizadeh [5] and Naonori Ueda [2]

[1]   Department of Forestry, Faculty of Natural Resources and Marine Sciences, Tarbiat Modares University, Tehran 15119-43943, Iran; kourosh.ahmadi@modares.ac.ir

[2]   RIKEN Center for Advanced Intelligence Project, Goal-Oriented Technology Research Group, Disaster Resilience Science Team, Tokyo 103-0027, Japan; naonori.ueda@riken.jp

[3]   Department of Mapping and Surveying, Darya Tarsim Consulting Engineers Co. Ltd., Tehran 15119-43943, Iran; saeidi@daryatarsim.com

[4]   Business, Computer Science and Applied Technologies Division, De Anza College, Cupertino, CA 95014, USA; kalantarghorashiharandielaheh@student.deanza.edu

[5]   Department of Watershed Management Engineering, College of Natural Resources, Tarbiat Modares University, Tehran 14115-111, Iran; janizadehsaeid@modares.ac.ir

[*]   Correspondence: bahareh.kalantar@riken.jp; Tel.: +81-362-252-482

check for updates

Abstract: The estimation and mapping of forest stand characteristics are vital because this information is necessary for sustainable forest management. The present study considers the use of a Bayesian additive regression trees (BART) algorithm as a non-parametric classifier using Sentinel-2A data and topographic variables to estimate the forest stand characteristics, namely the basal area (m^2/ha), stem volume (m^3/ha), and stem density (number/ha). These results were compared with those of three other popular machine learning (ML) algorithms, such as generalised linear model (GLM), K-nearest neighbours (KNN), and support vector machine (SVM). A feature selection was done on 28 variables including the multi-spectral bands on Sentinel-2 satellite, related vegetation indices, and ancillary data (elevation, slope, and topographic solar-radiation index derived from digital elevation model (DEM)) and then the most insignificant variables were removed from the datasets by recursive feature elimination (RFE). The study area was a mountainous forest with high biodiversity and an elevation gradient from 26 to 1636 m. An inventory dataset of 1200 sample plots was provided for training and testing the algorithms, and the predictors were fed into the ML models to compute and predict the forest stand characteristics. The accuracies and certainties of the ML models were assessed by their root mean square error (RMSE), mean absolute error (MAE), and R-squared (R^2) values. The results demonstrated that BART generated the best basal area and stem volume predictions, followed by GLM, SVM, and KNN. The best RMSE values for both basal area (8.12 m^2/ha) and stem volume (29.28 m^3/ha) estimation were obtained by BART. Thus, the ability of the BART model for forestry application was established. On the other hand, KNN exhibited the highest RMSE values for all stand variable predictions, thereby exhibiting the least accuracy for this specific application. Moreover, the effectiveness of the narrow Sentinel-2 bands around the red edge and elevation was highlighted for predicting the forest stand characteristics. Therefore, we concluded that the combination of the Sentinel-2 products and topographic variables derived from the PALSAR data used in this study improved the estimation of the forest attributes in temperate forests.

Keywords: machine learning; remote sensing; forest stand characteristics; Bayesian additive regression tree

1. Introduction

Forests cover around 30% of the Earth's surface and are one of the main sources of human supplies and services [1]; sustainable development, climate change mitigation, and bio-diversity preservation can be achieved by forestry and forest management [2]. Continuous forest management is crucial, which requires precise knowledge of forest characteristics through detailed information extraction [3]. Spatiotemporal change detection, logging, and evaluating the forest management regime in the region can bring more transparency into the management and ecosystem services in the forest [4].

The extensive advancements in the remote-sensing (RS) technologies as well as the geographic information system (GIS), computer science, and algorithms, allow not only for a rapid and up-to-date data collection, but also an accurate broad Earth observation and reliable information extraction, specifically related to forest inventory and management [5–7]. The Sentinel satellites continuously map and monitor vast forest regions using high spatial, spectral and temporal resolution data, but at low costs [5,8]. The operation of the Sentinel-2 satellite provides multi-spectral data in 13 bands, with a spatial resolution of 10 to 60 m, 10-day revisiting period, and 290 km swath width.

Forest management explicitly necessitates the use of remotely sensed data owing to its cost effectiveness and regional availability. These data has been investigated in terms of data accuracy and reliability in many applications including those of vegetation, agriculture, and forestry [8–10]. Therefore, research focus is being increasingly centred on mapping the quantitative distribution of forest stand characteristics (e.g., diameter, height, location, basal area, potential volume of wood, and tree species) as a forest management strategy [1,6]. The use of Sentinel-2 as medium resolution and freely accessible data in forestry have been explored by [11–13]. In study done by [14,15], they emphasized on the successful application of Sentinel-2 data for forest classification and monitoring. In this context, Frampton et al. [16] applied simulated Sentinel-2 data to estimate the bio-physical variables of vegetation (e.g., canopy chlorophyll content index, leaf area index, and leaf chlorophyll concentration). Similarly, Majasalmi and Rautiainen [17] estimated the bio-physical properties of vegetation (canopy) using Sentinel-2 data in a boreal forest; they demonstrated the potential of using Sentinel-2 data with its four narrow bands around the red edge for detecting the bio-physical properties and vegetation applications of the canopy.

The ability of Sentinel-2 data in combination with other sensors (e.g., Landsat 8) for vegetation monitoring and estimating the normalised difference vegetation index (NDVI), C-Test1 indices, green chlorophyll index, and red-edge chlorophyll index were confirmed by Addabbo et al. [18] in urban areas. Grabska et al. [10] reviewed several studies focusing on the Sentinel-2A data and forest applications, and concluded that further investigations should be conducted using more time series images (e.g., Sentinel-2A data) in a dense forest to comprehensively evaluate the overall accuracies of the classification. Hence, the use of Sentinel-2 data still requires extreme consideration.

Additionally, quantitative monitoring of the forest stand parameters (e.g., volume, basal area, biomass, vegetation density, tree height, etc.) plays a significant role in sustainable forest management [6,19]. Recently, Astola et al. [20] used the Sentinel-2A data to estimate forest variables such as stem volume, stem diameter, tree height, basal area, and tree species, and their results were compared with those obtained using the Landsat 8 data; they reported that Sentinel-2 mostly outperformed the Landsat 8 predictions. Table 1 shows a summary of applications of RS data in forestry and vegetation.

The studies listed in Table 1 predominantly, focused on tree species classification, while very few reviewed the mapping of the forest stand characteristics. Hence, further research is required to determine the other quantitative factors (i.e., stem density, basal area, and stem volume) in forests. In addition, the development of intelligent techniques and algorithms for classification and information extraction facilitates the use of the big RS data in their full potential [28]. The RS technologies (especially, the Sentinel program) supply huge volumes of raw data (big earth data) with various spectral, spatial, coverage, and multi-scale characteristics; such volumes of data require various algorithms and appropriate models to precisely extract and classify the information [29].

Table 1. Brief review on details and applications of RS data in forestry and vegetation.

Application	Data	Models	Reference
Estimation of bio-physical variables of vegetation	Sentinel-2	Vegetation indices assessment	[16]
		Physically-based reflectance model (PARAS)	[17]
Classification of agricultural and tree species	Sentinel-2	Random Forest (RF)	[21]
Land use/cover and forest detection	Sentinel-2	Object-based image analysis (OBIA)	[15]
Tree cover mapping (forest/non forest and broadleaved/coniferous forest)	Sentinel-2	k-means	[12]
Forest type mapping	Sentinel-2	RF	[14]
			[10]
	Sentinel-2		[22]
Classification of forest tree species		RF	[7]
			[13]
	Sentinel-2 and DEM		[23]
Vegetation monitoring	Sentinel-1 and 2 and Landsat 8	Vegetation indices assessment	[18]
Estimation of forest stand parameters	Sentinel-2 and Landsat 8	Multi-layer perceptron neural network and regression tree	[20]
Mapping of forest attributes	Sentinel-2 data, PALSAR, airborne laser scanner, DEM	Multiple linear regression and RF	[11]
Estimating the forest stand volume and basal area	Pleiades data and climate data	RF	[6]
Forest parameters estimations (e.g., stand age, aboveground biomass, leaf area index, tree height, crown diameter)	Quickbird	Classification and regression tree (CART), SVM, ANN, and RF	[24]
Classification/change detection of tree species	Landsat TM time series	SVM	[25]
	Hyperspectral data from HySpex VNIR-1800 and SWIR-384		[1]
Tree species compositional changes	Landsat TM time series	K-means and iterative self-organizing data analysis technique (ISODATA), maximum likelihood, and SVM	[26]
Relationships between forest stand parameters and vegetation indices (e.g., volume, basal area, biomass, vegetation density, tree height)	Landsat TM	Vegetation indices assessment	[19]
Estimation of the forest structural attributes (e.g., stand volume, basal area, and tree stem density)	Landsat-5 TM, ASTER, and Quickbird	CART	[27]

For example, the optical Sentinel-2A and B satellites alone produce already ~3.4 TB of data on average per day according to the acquisition plan [30] and the combined Sentinel-1, Sentinel-2 and Sentinel-3 fleet produce an estimated data volume of ~20 TB per day [31]. Many studies have been conducted to develop and apply different algorithms to the RS data; such algorithms were grouped accordingly. The common classification algorithms include parametric and non-parametric [32].

Parametric methods are defined by strong assumptions regarding the probability distribution of the variables [33], while the non-parametric approaches have limited or no assumptions concerning the probability distributions of the data [33,34]. Some examples of popular parametric classifiers involving RS data classification are maximum likelihood and logistic regression [35]. Similarly, the common non-parametric approaches are K-nearest neighbour (KNN), random forest (RF), decision trees (DT), SVM, artificial neural network (ANN), and Bayesian additive regression trees (BART) [27,36,37]. The development of these artificial intelligence techniques enables the extensive use of multi-variables and datasets. In this context, the non-parametric ML algorithms have demonstrated robust intelligence and learning strategies to handle complex non-linear variables and received more attention in RS big data classification [38]. In contrast with the high-performing ANN and deep learning (DL) models,

the simplicity of the traditional ML methods was demonstrated to be reliable and straightforward [39] for the classification of vast areas such as forests.

In vegetation applications, most researchers have classified the Sentinel-2 data using non-parametric RF [7,10,13,14,21–23]. Fragou et al. [25] explored Landsat TM time series images (1993, 2001, and 2010) for classification and change detection within nine classes (e.g., tree species of Aleppo pines and Cephalonian fire, grassland, agriculture etc). The images were classified by machine learning (ML) algorithm (i.e., support vector machine (SVM)) and the obtained overall accuracies higher than 89.85% proved the robustness of the ML algorithm for vegetation classification and tree species in details and landscape changes. Tree species compositional changes were investigated in deciduous forests [26] during 30-year period by exploiting k-means and iterative self-organizing data analysis technique (ISODATA) clustering techniques, maximum likelihood, and SVM over multi-seasonal Landsat image-stacks. SVM obtained the best accuracy to compare with other algorithms even with small size of training datasets and again the capability of ML algorithm was confirmed.

Noorian et al. [27] used the classification and regression trees (CART) algorithm for Landsat-5 thematic mapper (TM), ASTER (advanced spaceborne thermal emission and reflection radiometer), and Quickbird satellite data to estimate and compare the forest structural attributes (e.g., stand volume, basal area, and tree stem density). They reported that among the three different satellites bands, Quickbird data exhibited the best performance with an RMSE of 2.44 m^2/ha, 50.98 m^3/ha, and 125 n/ha for basal area, stand volume, and stem density, respectively. Zhao et al. [24] applied four ML models (CART, SVM, RF, and ANN) to the Quickbird images of a plantation site to map its parameters such as diameter at breast height (DBH), stand density, tree height, and leaf area index; they reported that RF exhibited the highest accuracy.

On the other hand, the non-parametric regression trees in the BART model provide the advantages of tree-based ML approaches; besides, this model does not tent to overfit the dataset and is suitable for small-sized training data [37]. Among the most important tree-based methods, the interpretation of uncertainty by BART is different from that of the RF model. Thus, BART leads to the retrieval of missing data and better performance over the lost data; in addition, its RMSEs are smaller than those of RF [37,40]. The BART uses a set of low-performance regression trees to create a robust model for prediction and classification [40]. Therefore, the application of BART for effective mapping the forest parameter could be explored.

Therefore, in order to consistent and frequent forest monitoring and management, using RS data and automated data analysis techniques are required. Herein, we have applied the above-mentioned techniques in the northern forests of Iran (Hyrcanian temperate forests), which is one of the most important and valuable ecosystems. To the best of our knowledge, there are no studies have applied and evaluated the performances of the BART model to characterise temperate forests using Sentinel data. With the focus on developing more precise and robust framework, our methodology could decrease the fieldwork in temperate forest with inaccessible steep terrain and lower the cost and time of forestry (to measure and estimate forest stand characteristics) in vast region.

Comparing frequent and accurate results from such framework, the habitat changes and forest productivity are timely measured. The presence of deciduous trees in temperate forests also calls for continuous and valid monitoring due to the different leaf colors, photosynthesis, and conditions in four seasons. According to UNESCO World Heritage, hosting hundreds tree/animal species including endangered mammals (e.g., Persian Leopard and wild goat) in the Hyrcanian Forests World Heritage makes the area as a great concern to be preserved form human activities, deforestation, and logging (https://whc.unesco.org/en/list/1584/). Providing such continuous and cost-time effective forest stand estimations might enhance and assure the habitat conservation in such forest.

Therefore, the objectives of the present study are as follows: (i) to analyse the significance of the different variables using Sentinel-2 bands, indices, and topographic features derived from PALSAR data in forest characteristic mapping; (ii) to exploit four ML methods (namely BART, generalised linear model (GLM), SVM, and KNN) for mapping the characteristics of temperate forests; and (iii) to

evaluate the accuracy of the BART method for modelling the basal area, stem volume, and stem density of the forest trees using the root mean square error (RMSE), mean absolute error (MAE), and R-squared (R^2) values.

2. Study Area and Materials

2.1. Study Area

The chosen study area was a portion of the forests in northern Iran that is adjacent to the Caspian Sea—one of the national environmental treasures that needs to be preserved. It is very significant to regularly and remotely monitor this entire region for sustainable forest and natural resources management. The area is located at 36°28′08″–36°14′18″N and 52°07′39″–51°22′29″E and covers a part of the forests of Noor County, Mazandaran Province, with altitude variations from 26 to 1636 m and the study area is about 287 km² (Figure 1).

The average annual rainfall in the region is about 997 mm; besides, the region has a humid climate, and the average recorded temperature here is 16.4 °C. Geologically, it is mostly covered by conglomerate rock, sandstone and limestone, and calcareous marlstone. The study area includes 15 districts, including mixed and uneven-aged specimens, of Oriental beech (*Fagus orientalis*), common hornbeam (*Carpinus betulus*), Persian maple (*Acer velutinum*), Persian ironwood (*Parottia persica* (DC.) C. A. Mey.), checker tree (*Sorbus torminalis* L.), chestnut-leaved oak (*Quercus castaneifolia*), Caucasian alder (*Alnus subcordata* C.A.Mey.), Cappadocian maple (*Acer leatum* C.A.Mey.), wild cherry (*Prunus avium* L.), peach (*Prunus persica*), wych elm (*Ulmus glabra* L.), and English yew (*Taxus baccata* L). In more than 90% of study area, we have different tree layer, for example the first layer covered by dominant tree including chestnut-leaved oak (*Quercus castaneifolia*), Oriental beech (*Fagus orientalis*) in over story and in second layer we have the peach (*Prunus persica*) and common box (*Buxus hyrcana*) in the third layer.
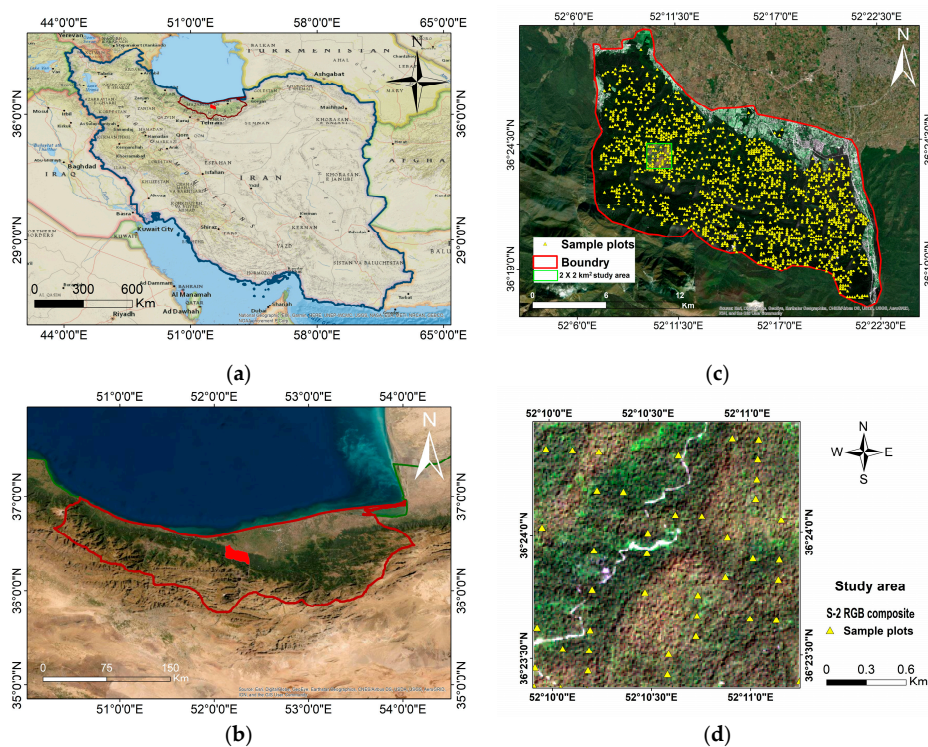

Figure 1. General location of the study area: (**a**) Iran; (**b**) location map; (**c**) The Hyrcanian forests including sample plots; (**d**) Sentinel-2 images (Red Green Blue (RGB) bands).

2.2. Ground Controls (GCs), Data, and Sample Plots

The forest inventory dataset provided by the Iranian Forests, Range and Watershed Management Organization for 287-km^2 area included 1200 sample plots, each with an area of 1000 m^2 was used. A sampling method, in northern Iran, with a 150 $*$ 200 m network has been conducted. In each plot, the diameter at breast of the trees above 7.5 cm and the type of species were measured and determined. The spatial error of each plot is about 3 m. Then, the stem density in each plot, which is expressed as the number of tress per unit area (ha), was estimated [41]. The basal area which defines the average area (m^2) per ha that was occupied by the tree stems at breast height was calculated from the DBH [27]. As the actual stem or the trunk of a tree is not an exact cylinder, we used a specific volume table to extract the stem volume or the volume of the wood/tree trunk according to the various tree species, locations, DBH, and height [42]. That table was also provided by the Iranian Forest, Range, and Watershed Management Organization. Finally, the sample plots were fed into the ML algorithms for training and testing [6]. Table 2 presents the minimum, maximum, and average measures of the sample plots for the forest stand variables.

Table 2. Mean, minimum, maximum, and standard deviation of the main forest variables for all plots in the study area.

Stand Variables	Descriptive Statistics			
	Minimum	Maximum	Mean	SD
Basal Area (m^2/ha)	63.35	125.29	95.42	11.29
Volume (m^3/ha)	138.4	371.10	256.07	41.67
Density (n/ha)	90	420.00	232.27	62.10

2.3. Remote Sensing Data

For this study, freely available Sentinel-2 images (Level-2A product, bottom-of-atmosphere reflectance, tile number T39SXA, dated: 21 August 2019) were downloaded from the Copernicus Open Access Hub (https://scihub.copernicus.eu/). As the study area is a part of the Hyrcanian mountain forests, most of the Sentinel-2 images include a high percentage of clouds. Therefore, it was very difficult to select an image with a low cloud percentage and for this reason, we only used the images accessed on the desired date. Initially, 11 spectral bands ranging from the visible (443 nm) to shortwave infrared (SWIR; 2190 nm) wavelengths were obtained from the Sentinel-2 data (Tables 3 and 4). The use of the elevation data along with the spectral bands proved to be effective and provided more accurate results in forest applications [6,23]. Therefore, the Advanced Land Observation Satellite (ALOS) Phased Array type L-band Synthetic Aperture Radar (PALSAR) DEM with a spatial resolution of 12.5 m was downloaded from the www.asf.alaska.edu website for the study area. ALOS is a Japanese satellite, Manufactured by NEC, Toshiba, and Mitsubishi Electric.

3. Methodology

3.1. Overview

The acquired image data was first pre-processed for the required image corrections. Then, the RS and ancillary datasets were used to compute the related features and indices using the workflow illustrated in Figure 2. Next, the training sample data was utilised by the ML algorithms (BART, GLM, SVM, and KNN) to model the most significant spectral bands and features using their feature selection procedures. Finally, the accuracy assessment was performed by 10-fold cross validation against the RMSE, MAE, and R^2 values.
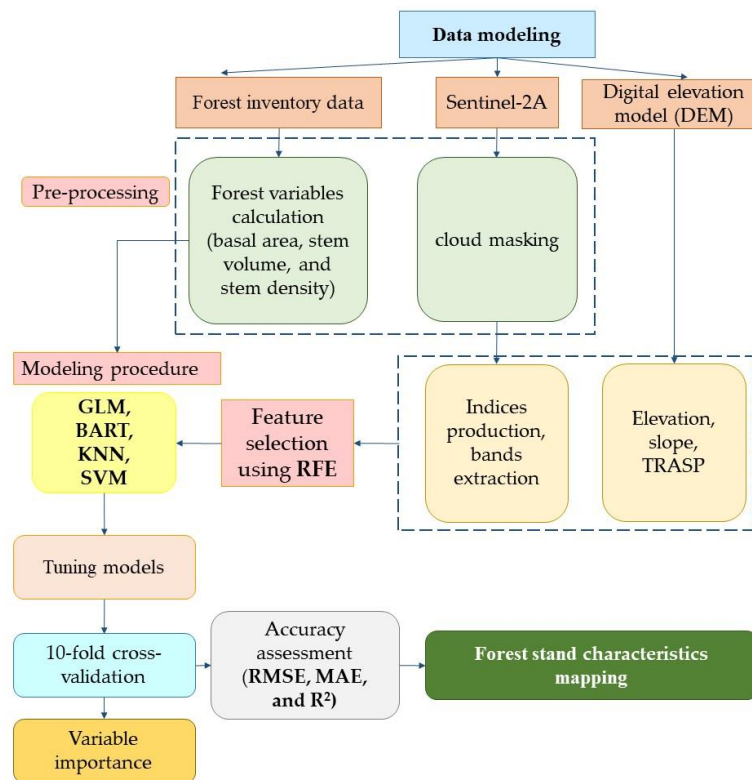
Figure 2. Framework of the proposed methodology.

3.2. Pre-Processing

Images with minimal cloud/haze from all bands (Sentinel-2 bands: 1, 2, 3, 4, 5, 6, 7, 8, 8a, 11, and 12) of 10, 20, and 60 m spatial resolutions were downloaded. The atmospheric correction of the Sentinel-2 data was implemented at Level-2A by the provider (ESA- Copernicus Scientific Data Hub). The Level-2A processing includes a scene classification and an atmospheric correction applied to Top-Of-Atmosphere (TOA) Level-1C orthoimage products. Level-2A main output is an orthoimage Bottom-Of-Atmosphere (BOA) corrected reflectance product (https://sentinel.esa.int/web/sentinel/ user-guides/sentinel-2-msi/processing-levels/level-2). Besides, the digital numbers of the image were converted to the TOA reflectance and the corrected reflectance product was obtained. Thereafter, the pixel values of all 11 bands were further processed and modelled using the R software (version 3.6.0). To resolve the problem of multi-scale datasets, all spectral bands were re-sampled to 10 m spatial resolution using nearest neighbours [43].

3.3. Feature Computation, Extraction, and Selection

For this research, some related attributes and features were extracted and calculated from spectral bands, indices, and variables (related formulas are listed in Table 3). Therefore, feature selection was utilised to select the most important contributors which could result in a more efficient classification and lower computation [44].

3.3.1. Topographic Feature Computation

The topographic factors such as elevation, slope, and topographic solar-radiation index (TRASP) were computed based on the 12.5 m DEM (Table 3). The TRASP was derived from the elevation based on the aspect map and the values of 0 and 1 were used to indicate the cool, north-facing slope and the hot and dry, south-facing slope, respectively [45,46]. All topographic variables resampled to 10m based on Sentinel-2A data.

3.3.2. Indices and Band Extraction

The vegetation indices are useful for comparison with individual bands in forestry and hence, the presence of multi-spectral sensors has revolutionised this concept [13]. Some vegetation indices (Table 3) such as NDVI and difference vegetation index (DVI) have proved to be effective for forest classification, mapping, and vegetation determination as a result of the reflectance enhancement in the near infrared (NIR) bands and chlorophyll absorption in the red band [12,13,16].

Considering the narrow bands around the red edge on the Sentinel data, other vegetation indices were applied by the researchers [16,17,47,48] including: transformed normalised difference vegetation index (TNDVI), weighted difference vegetation index (WDVI), normalised difference index 45 (NDI45), ratio vegetation index (RVI), infrared percentage vegetation index (IPVI), perpendicular vegetation index (PVI), inverted red-edge chlorophyll index (IRECI), and pigment specific simple ratio (PSSRA). The green-red vegetation index (GRVI) is another vegetation indicator that uses the green band instead of the NIR to balance the saturation problems of the NDVI in dense forest areas [13].

Overcoming the problem of soil reflectance in low or medium forest canopy regions is a challenge [36]. For the presence of bare soil without vegetation, the soil-adjusted vegetation indices (i.e., soil-adjusted vegetation index (SAVI), modified soil-adjusted vegetation index (MSAVI), and modified soil-adjusted vegetation index 2 (MSAVI2)) also provided promising results [13,22,47,49]. These indices utilise a soil-adjustment coefficient to compensate for the limitations of NDVI in various land covers. Therefore, this research also used the following indices for the initial step: DVI, NDVI, TNDVI, green normalized difference vegetation index (GNDVI), WDVI, NDI45, SAVI, MSAVI, MSAVI2, GRVI, RVI, IPVI, PVI, IRECI, and PSSRA. The calculations of all deployed indices are presented in Table 3 and were prepared using SNAP 5.0. Furthermore, the satellite images were imported as RasterStack layers and the data was further processed in the statistical environment R (version 3.6.0). Finally, the satellite images were clipped to the extent of the 2×2 km^2 study area for predicting the forest stand characteristics.

Table 3. The variables and predictors: Sentinel-2 bands, vegetation and soil indices, elevation derivatives, equations, and references.

Predictors	Description	Ref
B1	Coastal aerosol	-
B2	Blue	-
B3	Green	-
B4	Red	-
B5	Red-edge-1 (RE1)	-
B6	Red-edge-2 (RE2)	-
B7	Red-edge-3 (RE3)	-
B8	Near infrared (NIR)	-
B8A	NIR plateau (NIRp)	-
B11	Shortwave infrared (SWIR-1)	-
B12	SWIR-2	-
DVI	$NIR - Red$	[13]
NDVI	$\frac{NIR-Red}{NIR+Red}$	[13]
MSAVI	$\frac{NIR-Red}{NIR+Red+L} \times (1+L)$	[50]
MSAVI2	$\left(\left[2 \times NIR + 1 - sqrt\left((2 \times NIR + 1)^2 - 8 \times (NIR - Red)\right)\right]\right)/2$	[47]
GNDVI	$\frac{RE3-Green}{RE3+Green}$	[22]
IPVI	$\frac{NIR}{NIR+Red}$	[22]
IRECI	$\frac{RE3-Red}{RE1/RE2}$	[47]
NDI45	$\frac{RE1-Red}{RE1+Red}$	[47]
PSSRA	$RE3/Red$	[47]
PVI	$\sin(a) \times NIR - \cos(a) \times Red$	[47]
RVI	$\frac{NIR}{Red}$	[47]

Table 3. *Cont.*

Predictors	Description	Ref
SAVI	$\frac{(NIR-Red)\times 1.5}{NIR+Red+0.5}$	[22]
TNDVI	$\sqrt{\frac{NIR-Red}{NIR+Red}+0.5}$	[47]
WDVI	$NIR - Red \times 0.5$	[47]
Elevation	Digital elevation model	
Slope	$arc\tan\left(\frac{Vertical\ Height}{Horizontal\ Distance}\right)$	
TRASP	$1 - \frac{\cos\left[\left(\frac{\pi}{180}\right)(aspect-30)\right]}{2}$	[51]
Description:	$\alpha = 45°\ and\ L = 1 - 2.12 \times (NDVI \times WDVI)$	

3.3.3. Feature Selection

The uses of the various indices and factors increases the dimensionality within the dataset which necessitates proper feature selection to reduce the number of predictors and preserve only the relevant variables that ensure maximum accuracy [52,53]. Feature selection is a technique to remove noise and redundancy from the variables for timely, cost-efficient, and accurate performance, in addition to overcoming the overfitting problems [54]. In this context, the recursive feature elimination (RFE) or backward selection algorithm automatically filters the low-weight variables and removes them from the model by the repetitive modelling process [52,55]. This study uses the R package 'caret' [55] for feature selection and RFE, which is based on the Gini criterion with repeated 10-fold cross validation and calculation of the RF model to rank the variables in the order of their importance, i.e., from the most to least important predictors [52,53]. Then, the RFE results with the smallest error define the subset of the variables and the least important variables are iteratively removed. The quantitative error is measured against the percent increase in the mean square error and residual sum of squares (purities) in the nodes (trees) of the forest model [53,56]. Thus, only the optimal parameters remain.

3.4. Machine Learning Methods

Machine learning methods are regarded as popular and advanced models among the research communities. Their ability and flexibility of data modelling with a large set of variables together with their learning schemes and control over the non-linearity in the datasets have been tested and proved, especially in vegetation applications [38,57]. However, further investigations in terms of the various challenges and limitations of the ML algorithms is required. Herein, we have used four ML algorithms to detect the forest stand characteristics, which are described in the subsequent sections. In other words, each forest attributes were modelled using four different machine learning methods.

3.4.1. Generalised Linear Model (GLM)

The generalised linear model is a parametric statistical ML method which works with the common linear regression algorithm to handle linearity and simple relationship between the numeric datasets using assumptions based on normal and Gaussian distributions [58,59]. The model represents the continuous probability distribution for the random variables, and is still a widely used linear method with easy implementation; it often demonstrates better accuracy on relatively small-sized training datasets (observations) compared with the non-parametric algorithms [60]. Moreover, the GLM is sensitive to the existence of correlated variables, and the insignificant factors might affect its result, accuracy, and certainty [58]; GLM is explained by the following Equation (1):

$$f(\mathrm{y}) = C_0 + C_1 X_1 \ldots\ldots + C_n X_n \ (1) \tag{1}$$

where y is the estimation probability of the forest stand characteristics, C_i is the slope coefficient, X_i represents the predictors, and n is the number of total predictors used for the estimation [61,62].

3.4.2. K–Nearest Neighbour (KNN)

This is a popular and simple non-parametric ML method for classification and univariate/multivariate prediction that can be applied to a wide variety of non-linear variables [35,63]. To determine the k closest neighbours in a training dataset, the target class is assigned to the variable. Typically, the optimal k value for the datasets varies between 3 and 10 [64], and can be determined by cross-validation. During the modelling, the Euclidean distance and variable weighting are calculated according to the nearest target inversely to its distance [63]. At first, the Euclidean distance between the training data (O_i) and the feature to be predicted (P_i) are calculated using Equation (2) [64,65]:

$$d_j(P, O) = \sqrt{\sum_{i=1}^{d} (P_i - O_i)^2} \tag{2}$$

where d is the dimension of the feature space, and P_i and O_i are corresponding pixel or digital number of training samples and variable to be classified (predicted), respectively.

3.4.3. Support Vector Machine (SVM)

The SVM algorithm is a non-parametric classification and regression technique with a non-linear transformation that is based on the kernel function. It has been successfully used for forest species and vegetation classification by some researchers [35,38,48]. It uses a statistical learning mechanism to accurately handle the complexity and noise within the datasets [1,66]. Besides, SVM constructs a hyperplane and then, the optimal separating hyperplane of each class/observation is identified by the (small-sized) training dataset that mainly includes support vectors [1]. The hyperplane is defined as follows [66]:

$$\frac{y_i(w \times x_i + b)}{\geq 1} - \delta_i \tag{3}$$

where w denotes the coefficient vector defining the hyperplane orientation in the feature space, b defines the offset of the hyperplane from the origin and δ_i refers to the positive slack variables. Then, the optimisation is decided by the optimal hyperplane, as follows:

$$Minimise \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j y_i y_j (x_i x_j) \tag{4}$$

$$Subject \sum_{i=1}^{n} a_i y_j = 0, \ 0 \leq a_i \leq C \tag{5}$$

where a_i represents the Lagrange multipliers, y_j represents the predictions (stand variables: basal area, stem volume, and density), and C is the penalty parameter controlling the minimum error and maximum margin. Then, the kernel function is applied as follows:

$$K(x_i.y_i) = \left(-\gamma X_i - X_j \right).\gamma > 0 \tag{6}$$

where γ represents the gamma for radial basic function in the kernel, and X_i is an input vector of the predictors and variables (28 variables, e.g., B1, B2, NDVI).

3.4.4. Bayesian Additive Regression Trees (BART)

The BART is a non-parametric ML approach that offers flexibility in prediction and data modelling [40,67]. It has the ability to handle missing data and improve accuracy, either by modelling the missing data or adding a splitting criterion to deal with the missing values [37,67]. A prior distribution namely posterior distribution together with a likelihood function try to model the uncertainty and probability of the predictions [37,40]. The ensemble regression tree structure

in the Bayesian scheme combines the set of prior tree structures and leaf parameters defined by the hyperparameters to strengthen the model [37]. Equation (7) describes the BART algorithm [37]:

$$y \approx \sum_{i=1}^{m} T_i^M(x_i \ldots x_k) + \varepsilon, \ \varepsilon = N_n\left(0, \ \sigma^2 \ In\right) \tag{7}$$

The above equation shows that the sum of the trees model referring T^M as the regression tree structure; m defines the number of distinct trees formed by a set of x of k predictor variables; and M describes the set of leaf parameters at the terminal nodes of the quantity b_t such that the set of parameters are described as: $M_t = \left\{\mu_{t,1}, \mu_{t,2}, \ \ldots \ \mu_{t,b_t}\right\}$, and μ_{t,b_t} is assigned to x_i.

3.5. Model Evaluation

For accuracy assessment and evaluation of the models, three quantitative measures (RMSE, MAE, and R^2) with 10-fold cross validation were adopted. The k-fold cross-validation method randomly splits the inventory datasets into k number of equal folds or datasets [68]. According to the size of our dataset, the number of the folds was selected as 10: the 10-fold cross-validation method is commonly applicable to statistical learning algorithms with a computationally suitable learning fit [60]. Therefore, in this study, each fold included 120 samples (as the total number of sample plots is 1200): the first fold was considered as the validation data (for testing), and the remaining nine folds were used to train the models and the mean squared errors were calculated 10 times [60]. Then, the cross-validation process continued 10 times, such that every fold was utilised as the validation data and the remaining nine for the learning process. Eventually, the 10 mean squared errors of the 10 runs were averaged by each model for the forest characteristic estimations (stem volume, density, and basal area).

3.5.1. Root Mean Square Error (RMSE)

The RMSE measures the reliability of the estimations within the models and defines the error between the actual and predicted values [27]: it uses the predictions of the model and the observations from the inventory to compute the RMSE for an accurate assessment [69], using Equation (8):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{N}} \tag{8}$$

where O_i represents the observed values (samples), P_i represents the predicted values (i.e., basal area, stem volume, or stem density), and N is the total number of samples. A lower RMSE value signifies better performance of the model.

3.5.2. Mean Absolute Error (MAE)

The MAE was also used to evaluate the performance of the models and estimate the uncertainty of the prediction. It defines the difference between the prediction and observation (sample) as the mean [70]. It represents the average magnitude of the errors (mean absolute error) in a set of predictions and testing data. This is also less sensitive to outliers than RMSE; MAE is calculated using Equation (9) [71]:

$$MAE = \frac{\sum_{i=1}^{n}|O_i - P_i|}{N} \tag{9}$$

where O_i, P_i, and N denote the observations (actual data), predictions (output), and the total number of samples, respectively. A small difference between the prediction and observation in MAE certifies the certainty of the model.

3.5.3. R-Squared (R^2)

The Pearson coefficient squared (R^2) is another method that was used to evaluate the accuracy of the results [72]. The coefficient of determination is estimated using the following equation [20]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}\left(O_i - \overline{O}\right)^2} \tag{10}$$

where O_i, $\overline{O}$, and P_i represent the observed values from the inventory, the mean of the observed variables (mean of basal area, stem volume, and stem density), and predicted variables from the ML methods, respectively. The R^2 value varies from −1 to 1, indicating perfect negative correlation (uncorrelated) to perfect positive correlation, respectively, between the two variables (i.e., observation and prediction).

4. Results

We applied the four ML algorithms (BART, GLM, SVM, and KNN) to the most relevant datasets, and estimated three forest stand characteristics, in the study area, namely the basal area (m^2/ha), stem volume (m^3/ha), and stem density (number of trees per hectare). The RFE or backward selection algorithm used for variable selection. Table 4 presents the set (28 predictors) of spectral bands, ancillary variables, and selected parameters for each forest stand variable by the RFE method. For instance, to determine the basal area, B5, B6, B8A, B11, B12, IRECI, NDI45, PSSRA, elevation, slope, and TRASP were selected by RFE.

Table 4. Sentinel-2A and DEM predictors and variables used in forest stand characteristics modelling.

		Stand Characteristics		
Predictors	**Original Resolution (m)**	**Basal area**	**Volume**	**Density**
B1	60	-	-	-
B2	10	-	-	+
B3	10	-	+	+
B4	10	-	-	+
B5	20	+	+	-
B6	20	+	+	-
B7	20	-	+	+
B8	10	-	+	-
B8A	20	+	-	-
B11	20	+	+	+
B12	20	+	+	+
DVI	10	-	-	-
NDVI	10	-	+	-
MSAVI	10	-	-	+
MSAVI2	10	-	-	-
GNDVI	10	-	-	+
IPVI	10	-	-	-
IRECI	10	+	+	+
NDI45	10	+	-	+
PSSRA	10	+	+	+
PVI	10	-	-	-
RVI	10	-	+	+
SAVI	10	-	-	-
TNDVI	10	-	+	-
WDVI	10	-	-	+
Elevation	12.5	+	+	-
Slope	12.5	+	+	-
TRASP	12.5	+	+	-

The validations of the four ML algorithms are shown in Table 5. We observe that the BART algorithm presented a higher R^2 value (0.48) for basal area, followed by GLM (0.41), SVM (0.40), and KNN (0.36); however, all correlations exhibited a slightly low rate. For stem volume estimation, both GLM and BART achieved high values of 0.59 and 0.54 followed by SVM (0.44) and KNN (0.38), while for density prediction, GLM scored a slightly higher R^2 value (0.26), than BART ($R^2 = 0.22$), SVM ($R^2 = 0.19$), and KNN ($R^2 = 0.18$). Primarily, the minimum R^2 values for all stand characteristics were obtained by KNN, indicating lower positive correlation between the predicted and observed data. Unlike the R^2 value which defines a relative measure of fit, the RMSE provides an absolute measure of fit. Therefore, we calculated the RMSEs between the predictions and observations: the best RMSE was achieved by BART for basal area and stem volume estimation with the values 8.12 m²/ha and 29.28 m³/ha, respectively; meanwhile, GLM exhibited the best RMSE (53.12 n/ha) for stem density prediction. Once again, KNN generally exhibited the highest RMSEs for all stand variable predictions, indicating the worst performance. Finally, we also compared the differences between the predicted and observed values using the MAE: the lowest MAE was achieved by BART for the basal area (6.88 m²/ha) and stem volume (24.53 m³/ha) estimations, whereas the best MAE for stem density estimation was achieved by GLM (43.87 n/ha). Consistently, the KNN method again produced the highest MAE values for all predicted stand variables.

Table 5. Performance evaluation of the KNN, SVM, GLM, and BART.

Stand Variables		Models			
		KNN	SVM	GLM	BART
Basal Area (m²/ha)	R²	0.36	0.40	0.41	0.48
	RMSE	9.00	8.75	8.42	8.12
	MAE	7.55	7.31	7.18	6.88
	%RMSE	10.2	9.8	9.4	8.8
Stem Volume (m³/ha)	R²	0.38	0.44	0.59	0.54
	RMSE	31.74	31.43	29.32	29.28
	MAE	26.50	26.14	24.78	24.53
	%RMSE	12.1	12.01	11.9	11.9
Stem Density (n/ha)	R²	0.18	0.19	0.26	0.22
	RMSE	56.66	56.76	53.12	54.72
	MAE	46.88	45.91	43.87	45.08
	%RMSE	24.3	23.6	23.1	23.4

Figure 3 shows the scatterplot and relationship between the two sets of data (predicted versus observed); the scatterplot is the most commonly used method to graphically evaluate the model prediction [73]. We observe correlation between the predictions (i.e., basal area, stem volume, and stem density) on the Y-axis and the observations (sample/actual data) on the X-axis using BART, GLM, KNN, and SVM with 10-fold cross validation. Ideally, there should be no bias from the 1:1 regression line (diagonal), which indicates a perfect and accurate data modelling: the greater the deviation and scattered values from the regression line, the more the random errors induced into the predictions by the ML algorithms during the modelling. We observe that the patterns in the scatterplot mostly show slopes from the lower left to upper right, centralised to the 1:1 line, indicating a positive increasing correlation and linearity between the two sets of variables. A visual comparison of the scatterplots shows that the maximum correlation (best fit) is steadily exhibited by BART, followed by SVM and GLM in basal area detection. Meanwhile, the scatterplots of KNN suggested minimum correlation (weaker goodness-of-fit) especially, for the density prediction.
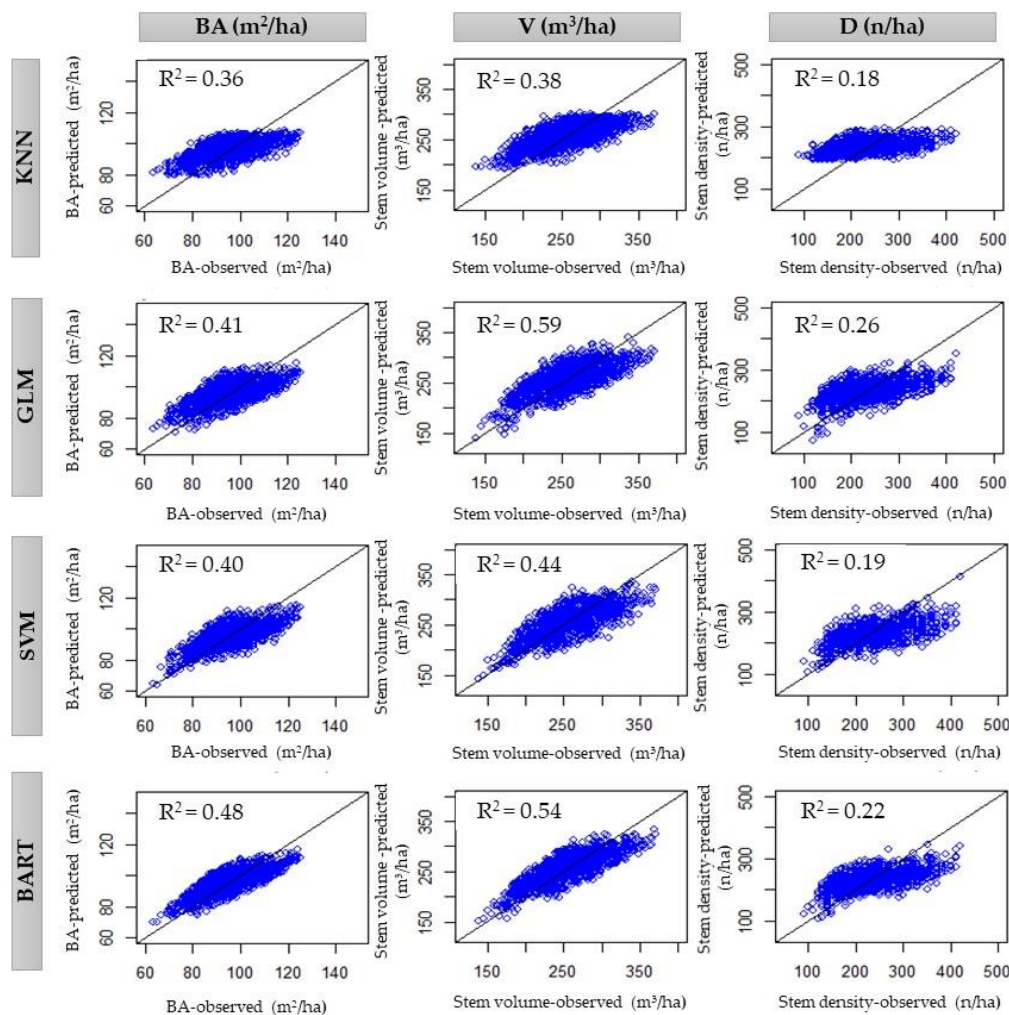
Figure 3. Scatterplots for the predicted vs observed values of basal area, volume, and density by 10-fold cross validation of BART, GLM, KNN, and SVM.

Figure 4 presents the significance of the variables for basal area, stem volume, and stem density estimations by each of the four ML models. Figure 4a highlights the significance of the elevation data and the insignificance of TRASP for estimating basal area by all models; we observe that while using GLM, the other factors exhibit a relatively lower influence on the basal area prediction, while the B6, B12, and B11 bands are the effective factors, after elevation for SVM and KNN. Similarly, Figure 4b indicates the importance of elevation for the stem volume estimation by the GLM, SVM, and KNN models, while B6 is the most important predictor for BART. Once again, TRASP is observed to be the least significant predictor for all models; besides, TNDVI and B7 also do not contribute much to the BART model. On the other hand, Figure 4c shows that the most important factor for stem density prediction by BART and GLM is the B12, followed by the B7 band. It is obvious that MSAVI, WDVI, and IRECI play significant roles while B2 and B4 are the least important predictors for stem density estimation by SVM and KNN modelling.
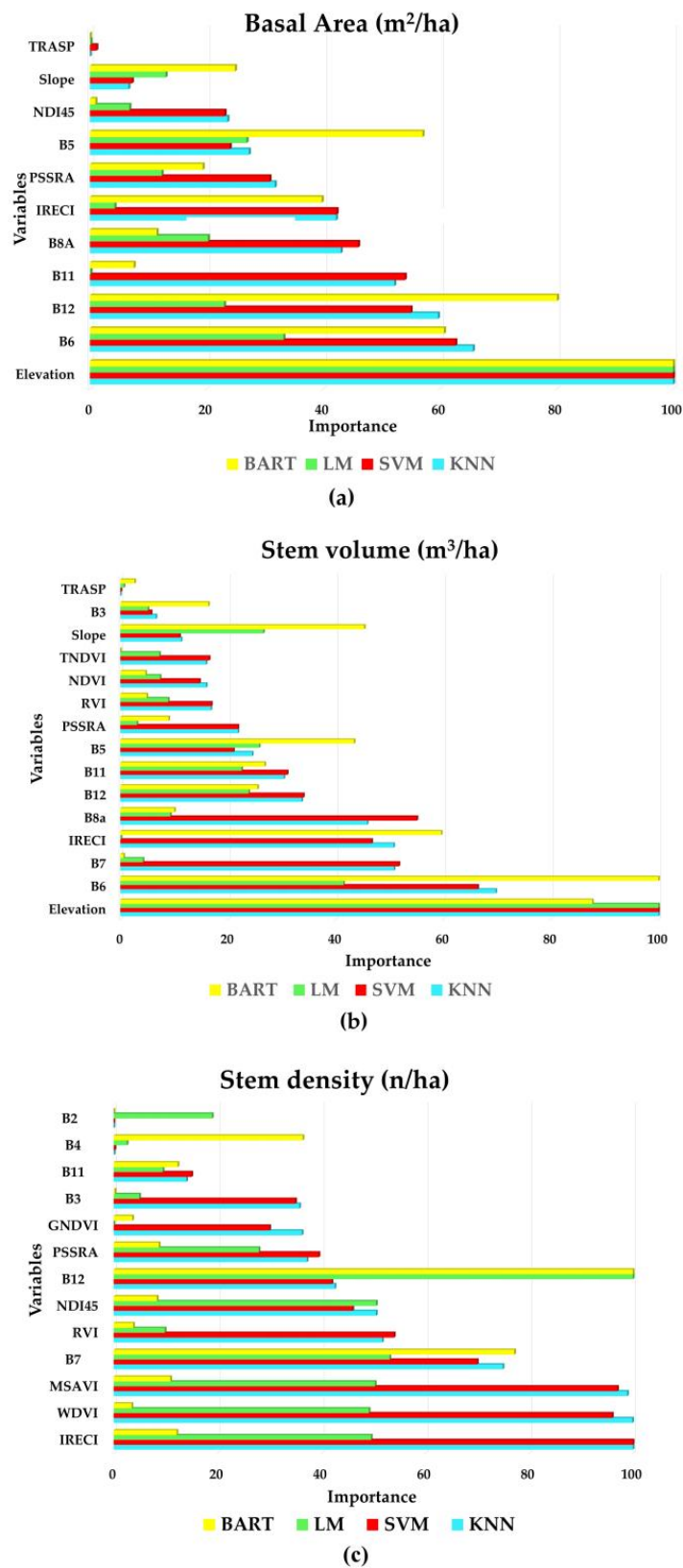
Figure 4. Significance of variables for each ML model for predicting: (**a**) basal area, (**b**) stem volume, and (**c**) stem density.

Figure 5 shows the mapping of the forest stand characteristics by BART, GLM, KNN, and SVM. We observe that the BART and SVM models exhibit a similar pattern in basal area detection, while KNN does not produce coverage values less than 60 m^2/ha. On the other hand, GLM achieves more maximum and minimum values of the basal area, stem volume, and stem density than the other three models. Overall, we observe that the stand characteristics detection and distribution by BART, GLM, and SVM are similar, while that of KNN has a very different appearance.
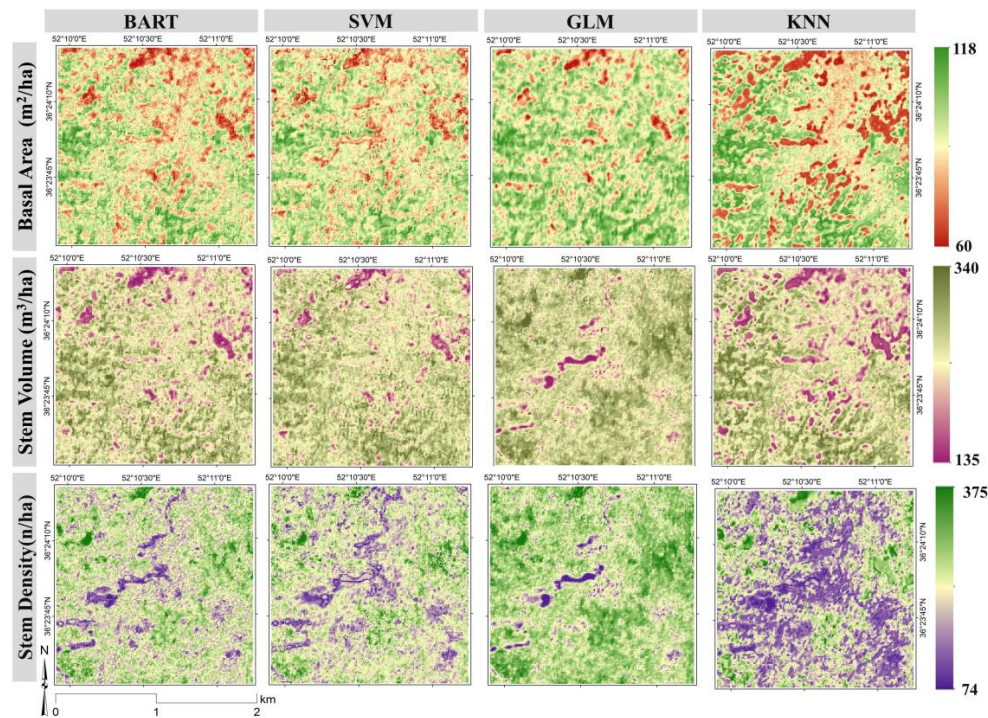
Figure 5. Mapping of the forest stand characteristics using BART, GLM, KNN and SVM in a 2 X 2 km^2 study area.

5. Discussion

The main objective of the present study was to determine the potential of the multi-spectral sensors in Sentinel-2 and the topographic variables derived from PALSAR in predicting the most common forest stand variables required for sustainable forest management. The RFE dimensionality reduction was used to quantitatively measure the variables. The Gini regression automatically calculated the weight of variables according to the relevance of spectral bands, their derivatives, and predictors based on the training dataset. Each forest stand attribute takes advantages of the most informative predictors in different way thus, the RFE weighting system recorded different results (predictors) for the feature selection. Four ML algorithms namely the BART, KNN, SVM, and GLM were used to model the forest characteristics. The results showed that the multi-spectral sensors of Sentinel-2 and topographic variables derived from PALSAR data are valuable sources for mapping the stand volume, basal area, and stem density in temperate forests. The results of the accuracy assessment using R^2, RMSE, and MAE for the four ML methods showed that BART was the most accurate and reliable algorithm to predict the basal area and stem volume. Besides, GLM performed slightly better than BART in stem density prediction considering the three aforementioned evaluation criteria. A comparison of these results with other studies would provide an understanding of the technical efficiency, although the wide range of variations in the forest landscapes, forest structures, and scale of studies must also be considered. We compared our results with those reported by Noorian et al. [27] which showed similar accuracy levels among the forest structural attributes; of these, the estimation of the basal area was the most reliable and certain (exhibiting the lowest RMSE) followed by stem volume and finally, tree density. Although

they used datasets from various satellites, i.e., from 0.68 (Quickbird panchromatic band) to 30 m (Landsat-5 TM and ASTER) resolution, their accuracies ranged in a pattern similar to ours. Moreover, the authors had used the summer data from three different satellites together with CART modelling. As mentioned earlier, this could also raise the argument over the time specification of the imagery for precise forest monitoring and attribute estimations. In terms of using combined satellite images for forestry applications, Mauya et al. [74] reported the high performance of Sentinel-1 (SAR), Sentinel-2, and their combinations to predict the growing stock volume in the small-scale forest plantations of Tanzania. Similarly, Vafaei et al. [75] provided reasonable results for the above-ground bio-mass estimation in the Hyrcanian forests of Iran, using different ML algorithms and a combination of the Sentinel-2A and the ALOS-2 PALSAR-2 data.

Regarding the accuracy assessment, our results indicated the different accuracies for the estimation of the forest stand attributes by the different ML methods. Based on the R^2 values, GLM and BART were more accurate only in stem volume estimation, while the other ML methods exhibited a lower goodness-of-fit. In addition, the lowest R^2 value was generated by the KNN method for all three stand characteristics; this confirms the existence of noise in the sample data, to which KNN is more sensitive [23]—a reason for the weakness of the KNN model could be the altitude variations in the study area (ranging from 26 to 1636 m). Essentially, for all ML models, the goodness-of-fit level decreased from stem volume to basal area and finally, to stem density, suggesting that data modelling is generally more fit for stem volume estimations than the other forest variables. The low R^2 value were observed by several studies and reported by [27]. Valbuena et al. [76] quoted from several researchers that lower R^2 value is not always an indicator for lower accuracy of predictions. Fatehi et al. [77] also experienced very low R^2 value especially for stem density estimation using digital terrain model of 1-m grid (airborne laser scanning) and multi-spectral image of 30-m resolution (imaging spectroscopy) to predict tree density and forest productivity in a heterogeneous Alpine landscape. The authors came to conclusion that low R^2 was due to the presence of small and diverse tree species and mentioned stem density was dependent on the mixture of different species, structures, and non-homogenous canopy. Using this kind of dataset is challenging to estimate stem density and it might be more applicable to obtain higher accuracy within homogenous forest area. The presence of various species [27] in our study area (around 80 woody species such as Fagus, Carpinus, Tilia, Parrotia persica etc.) could be another reason behind low R^2 value, and it suggest to examine it in other places.

Regarding the classifier performance in terms of RMSE, the BART algorithm exhibited the best performance for estimating basal area and stem volume, with values of 8.12 (8.8%) and 29.28 (11.9%), respectively, while GLM exhibited a higher RMSE (8.42 m^2/ha (9.4%) for basal area and 29.32 m^3/ha (11.9%) for stem volume) than BART, but lower than those of SVM and KNN. However, GLM generated the lowest RMSE values (53.12 n/ha (23.1%)) for stem density predictions, followed by BART (54.72 n/ha (23.4%), KNN (56.66 n/ha (24.3%)), and SVM (56.76 n/ha (23.6%)). The same pattern was also observed for the MAE evaluation method. Therefore, BART can be considered as a certain model and the best fit for forest stand characteristics, followed by GLM and SVM; consequently, KNN was the method with least accuracy and certainty. On the other hand, the BART method took greater advantage of some predictors (i.e., B5, B6, B12, IRECI, and slope, as shown in Figure 4a,b) than GLM, SVM, and KNN, and hence, achieved higher accuracy. Apart from the different modelling strategy of BART, the different weighting of the predictors and variables also contributed to its higher accuracy in basal area and stem volume prediction. Furthermore, the choice of parametric algorithms such as GLM demonstrated better performance and interpretation of this linear method against some of the non-parametric algorithms [60]. Hence, GLM (as a parametric algorithm) had outperformed the non-parametric algorithms such as SVM and KNN, but not BART. This could be because there was not enough sampling and inventory data to train the SVM and KNN models, as non-parametric methods inevitably require more sample data to obtain higher accuracy [60]. Thus, while Maponya et al. [8] and Vafaei et al. [75] claimed that the SVM was a great choice for vegetation classification, owing to its ability to handle high dimensional data with less training sample

plots, our research confirmed its weakness against the BART model. At the same time, it also proved the ability of non-parametric BART algorithm to handle and model the variables, even with small-sized training datasets, while SVM and KNN underestimated the predictions of the forest characteristics. Moreover, the BART algorithm is insensitive to multi-collinearity and can simultaneously model a large number of predictors. Thus, we demonstrated the advantages and capabilities of the BART model and compared it with the other ML methods to explore the strengths and limitations of both parametric and non-parametric approaches. The magnitude of the reliability among all ML models confirmed that the predictors and datasets were not sufficient for stem density prediction. This raises more arguments and concerns regarding the choice of datasets for stem density prediction and counting the number of trees per acre. Our findings suggest further exploration of the stem density prediction and its lower accuracy. During August, the forests in Iran have very low visible reflectance values because of the closed canopy and limited background and soil reflectance [36]. As the leaf density decreases in the fall season, the reflectance values would increase in November. However, the Sentinel data we used was dated 21/8/2019, when the canopy was full and dense; therefore, the number of trees per hectare could not be properly estimated there, and as a result, the stem density estimation was the least certain prediction by all ML algorithms. It recommends to acquire and stack all seasonal images for more complete information extraction as it was deployed by [26].

The variables were ranked according to their significance, and the results indicated that the role of elevation data was outstanding for both basal area and stem volume predictions. This is consistent with the research by Luther et al. [11] that emphasised on the importance of topographic data for better performance and accuracy of predictions. In addition to elevation data, the B12, B5, and B6 bands were the most influential predictors for the BART model. Similarly, the significance of the B6 and B12 bands was also highlighted by the SVM and KNN models. On the other hand, TRASP was the least significant attribute, and the prediction of the basal area and stem volume using this factor (the cool, north-facing slopes to hot, dry south-facing slopes) was not as accurate as the other factors.

Besides, the different weighting strategies of the ML algorithms could be responsible for the different rankings of the predictors, as suggested by Sothe et al. [13]: they used the RF algorithm to rank the significance of the variables (Sentinel-2 bands) and reported slight differences in the contribution of each predictor, which was different from our ML ranking system. The significance of the variables was similar in SVM and KNN, suggesting parallel modelling by both methods. Therefore, adopting a proper ranking system and modelling algorithm could directly affect the final estimations, and our findings emphasised on exploiting and comparing different algorithms for one particular application. We also observed conflicts in the significance of the variables for the stem density estimation by the four ML algorithms. This indicates that MSAVI, WDVI, and IRECI which were the most important predictors for SVM and KNN seemed to be insignificant and moderate factors for BART and GLM, respectively. Similarly, the B12 band had the most significant impact on the BART and GLM methods, while it was ranked as a moderate predictor by the SVM and KNN algorithms. For stem density, GLM moderately weighted almost all the predictors, which could be a reason for its being the most successful algorithm for this prediction. However, the differences between the variable ranking amongst the ML algorithms while predicting stem density suggested that the priority of the predictors could not be clearly determined.

Considering all vegetation and soil indices for this study, the efficiency and effectiveness of the predictors still requires further investigation, examination, and negotiation, as there was no considerable ranking of their significance (except IRECI) for the accuracy and performance of any algorithm. The lower contribution of the vegetation indices as data which is complementary to that from Sentinel-2 was in agreement with the study by Sothe et al. [13]. As mentioned earlier, IRECI was calculated from NIR, Red, Red-edge-1, and Red-edge-2, and our findings reflected the effectiveness of the chlorophyll index and consequently, the narrow bands of the Sentinel around the red edge for vegetation application; once again, this finding was consistent with the research by Sothe et al. [13]. Furthermore, the time of the imagery as well the season were observed to contribute to the insignificance

of the soil indices, because the area was mostly covered by the vegetation, while soil and bare land were not substantial. Among the Sentinel bands, we observed that B12 (SWTR-2) and B6 (Red-edge-2) for basal area, B6 for stem volume, and B12 and B7 (Red-edge-1) for stem density were the most significant bands, although they display a 20-metre spatial resolution, which is coarser than the other bands. In general, these are extremely weighted (by more than 60%) and significantly contribute to the modelling and predictions. In other words, the use of Sentinel-2 bands was successful for the forest stand predictions.

Thus, our findings suggested the use of BART rather than KNN for this specific application; this was in agreement with the study by Varvia et al. [57], which reported that the Bayesian method exhibits better RMSE value than KNN in the total basal area and volume estimation. Hence, the BART algorithm seems to be a perfect option in forestry, where heterogeneity and spatial autocorrelation structures exist and might involve inconsistencies in the linearity, homogeneity, and independency of the variables [41]. Varvia et al. [57] also mentioned the notable bias in Bayesian modelling in the stem density estimation; this could be another reason for the lower certainty in stem density prediction by BART. Further investigation is required to comprehensively link the lower stem density accuracies of all aforementioned ML algorithms. To put it briefly, the use of very high resolution RS data might provide more accurate estimation of forest characteristic [27] while open source, freely available, and medium resolution (e.g., Landsat 5 TM with 30 m-resolution) data proved to be popular, comparative and effective in estimating the forest attributes for sustainable management [27,48]. Almost all forests occupy the wide region and excessive urbanization and climate change necessitate continuous and regular observation of the forest characteristics; using very high spatial resolution, airborne and drone based hyper-spectral and Light Detection and Ranging (LiDAR) data are functional but not cost and time effective neither reachable by every sector, planner, and decision maker. With respect to time revisiting and the level of accessibility and reliability of Sentinel-2 data time series, this mission seems viable.

6. Conclusions

Out of the 28 initial predictors in this study, six variables were filtered and eliminated by RFE to avoid duplication, poor quality, and irrelevant contribution to the modelling and the specific forest application. According to RFE, the ML models exhibited different choice patterns for the different forest stand characteristic predictions. The use of computer science and Earth observation for forest applications, which involve vast areas and various datasets, mainly requires an accelerated process. Therefore, the application of RFE before data modelling is suggested not only to avoid heavy calculations involving the available big RS data but also to improve the performance of sensitive algorithms (e.g., GLM) that use the correlated datasets. A large number of features suggest that the usefulness of other robust feature selection techniques (e.g., neural networks based on removing input layers and keeping the most relevant features) to be compared with RFE to measure its efficiency during training for higher performance, and data and time reduction. An analysis of the factor significance for the ML models revealed that elevation data was the most important factor for both basal area and stem volume prediction by all ML models. We also observed that KNN and SVM exhibited an almost similar pattern in ranking the variables for stem density prediction. Furthermore, IRECI resulted in the enhancement of the vegetation reflectance and was confirmed to be effective in detecting both canopy bio-physical properties and stem volume using the BART algorithm. Our findings also emphasise the effectiveness of the narrow Sentinel bands around the red edge for the forest stand characteristics, especially considering the Sentinel data availability; besides, the precise choice of the imagery date plays a major role in improving the final accuracy of the variable estimation. In addition, we demonstrated that the use of higher resolution bands might not necessarily improve the estimation accuracy; instead, the use of the most informative bands related to the application as well as proper algorithm and variable modelling are outstanding and notable for better prediction. Owing to the promising accuracy of BART, the stem volume prediction exhibited a perfect positive correlation with the testing dataset, while the relatively lower R^2 values indicated that the sample

data (training and testing) was noisy and required more precise sampling and inventory for obtaining superior results, which is a suggestion for future works. It is expected to obtain higher R^2 in less heterogeneous forest where the more homogeneity is represented within the tree species. The result of our research with basic forest measurements could be beneficial at a broader scale for change detection after natural hazards (e.g., wildfire), ecosystems change, forest inventory, forest productivity monitoring, growth estimation, and wildlife habitat and nesting studies in vast forest especially mono-dominant forests to promote fast decision making, timely treatment, and proper mitigation. Such data and results have the potential application for forest categories and diversity studies at global scale. Categorizing the forest with large patches dominated by a single family tree species might raise concern regarding human activities and wildlife change in that region. Lower R^2 values form similar study might lead the future research on productive dispersal and spread of heavy seed by the animals within a forest in every corner and could be an indicator for forest species diversity, natural wildlife activities, and biodiversity restoration [78]. Besides, the Sentinel data was obtained in August, a time of the year when the tree coverage is densest; hence, the trees were not accurately separable to facilitate their proper counting. This might raise concerns regarding the data preparation prior to the application and indicate the importance of selecting the appropriate time of imagery for obtaining more accurate estimations of the forest stand characteristics. The lower visible reflectance values during August and the increase in the reflectance values during November (i.e., the fall season) when the leaf density decreases is a debatable concept. Therefore, further investigations are required to fully determine the significance of the vegetation and soil indices for forest variable predictions. Future works could investigate Sentinel time series during different seasons to compare the accuracies of the forest stand characteristics and identify the best season for obtaining more certain and accurate predictions. Temperate Forests are more often misty and cloudy especially in higher altitude where cooler temperature exists. The open Earth data observation provided by the European Space Agency (ESA) and Sentinel satellites is frequently delivered at satisfactory narrow spectral bands, spatial and temporal resolution particularly for forest management. Long term forest monitoring at low cost and the time of computation on medium resolution imagery would enable us to use more time series data to have cloud free and more informative images from the temperate regions. The robust algorithm such as BART which proved its ability to handle wide range of datasets and variable to estimate stand forest characteristics accurately seems a reliable and fast option. Yet, the BART algorithm should be comprehensively examined over other variables and data to understand its limitations and strengths. Therefore, our future investigations would focus on additional observed characteristics and spatial heterogeneity in forests by exploiting and developing the BART method.

Author Contributions: K.A. and S.J. acquired the data; K.A. and B.K. conceptualized and performed the analysis; B.K. and V.S. wrote the manuscript, discussion and analyzed the data; N.U. supervised; K.A., B.K. and E.K.G.H. provided technical sights, as well as edited, restructured, and professionally optimized the manuscript. All authors have read and agreed to the published version of the manuscript.

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

References

1. Wietecha, M.; Jełowicki, Ł.; Mitelsztedt, K.; Miścicki, S.; Stereńczak, K. The capability of species-related forest stand characteristics determination with the use of hyperspectral data. *Remote Sens. Environ.* **2019**, *231*, 111232. [CrossRef]
2. Nabuurs, G.J.; Masera, O.; Andrasko, K.; Benitez-Ponce, P.; Boer, R.; Dutschke, M.; Elsiddig, E.; Ford-Robertson, J.; Frumhoff, P.; Karjalainen, T.; et al. *Forestry. Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: New York, NY, USA, 2007.

3. Soares, P.; Tomé, M.; Skovsgaard, J.P.; Vanclay, J.K. Evaluating a growth model for forest management using continuous forest inventory data. *For. Ecol. Manag.* **1995**, *71*, 251–265. [CrossRef]

4. Kitayama, K.; Fujiki, S.; Aoyagi, R.; Imai, N.; Sugau, J.; Titin, J.; Nilus, R.; Lagan, P.; Sawada, Y.; Ong, R.; et al. Biodiversity observation for land and ecosystem health (BOLEH): A robust method to evaluate the management impacts on the bundle of carbon and biodiversity ecosystem services in tropical production forests. *Sustainability* **2018**, *10*, 4224. [CrossRef]

5. Roy, P.S.; Behera, M.D.; Srivastav, S.K. Satellite remote sensing: Sensors, applications and techniques. *Proc. Natl. Acad. Sci. India Sect. A Phys. Sci.* **2017**, *87*, 465–472. [CrossRef]

6. Zahriban Heasari, M.; Fallah, A.; Shataee, S.; Kalbi, S.; Persson, H. Estimating the forest stand volume and basal area using pleiades spectral and auxiliary data. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 1131–1136. [CrossRef]

7. Kumar, M.; Singh, M.P.; Singh, H.; Dhakate, P.M.; Ravindranath, N.H. Forest working plan for the sustainable management of forest and biodiversity in India. *J. Sustain. For.* **2020**, *39*, 1–22. [CrossRef]

8. Maponya, M.G.; Van Niekerk, A.; Mashimbye, Z.E. Pre-Harvest classification of crop types using a Sentinel-2 time-series and machine learning. *Comput. Electron. Agric.* **2020**, *169*, 105164. [CrossRef]

9. Ji, C.; Li, X.; Wei, H.; Li, S. Comparison of different multispectral sensors for photosynthetic and non-photosynthetic vegetation-fraction retrieval. *Remote Sens.* **2020**, *12*, 115. [CrossRef]

10. Grabska, E.; Hostert, P.; Pflugmacher, D.; Ostapowicz, K. Forest stand species mapping using the sentinel-2 time series. *Remote Sens.* **2019**, *11*, 1197. [CrossRef]

11. Luther, J.E.; Fournier, R.A.; Van Lier, O.R.; Bujold, M. Extending ALS-based mapping of forest attributes with medium resolution satellite and environmental data. *Remote Sens.* **2019**, *11*, 1092. [CrossRef]

12. Ottosen, T.-B.; Petch, G.; Hanson, M.; Skjøth, C.A. Tree cover mapping based on Sentinel-2 images demonstrate high thematic accuracy in Europe. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *84*, 101947.

13. Sothe, C.; De Almeida, C.M.; Liesenberg, V.; Schimalski, M.B. Evaluating Sentinel-2 and Landsat-8 data to map sucessional forest stages in a subtropical forest in Southern Brazil. *Remote Sens.* **2017**, *9*, 838.

14. Puletti, N.; Chianucci, F.; Castaldi, C. Use of Sentinel-2 for forest classification in Mediterranean environments. *Ann. Silvic. Res* **2018**, *42*, 32–38.

15. Szostak, M.; Hawryło, P.; Piela, D. Using of Sentinel-2 images for automation of the forest succession detection Using of Sentinel-2 images for automation of the forest succession detection. *Eur. J. Remote Sens.* **2018**, *51*, 142–149.

16. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogramm. Remote Sens.* **2013**, *82*, 83–92.

17. Majasalmi, T.; Rautiainen, M. The potential of Sentinel-2 data for estimating biophysical variables in a boreal forest: A simulation study. *Remote Sens. Lett.* **2016**, *7*, 427–436.

18. Addabbo, P.; Focareta, M.; Marcuccio, S.; Votto, C.; Ullo, S.L. Contribution of Sentinel-2 data for applications in vegetation monitoring. *Acta IMEKO* **2016**, *5*, 44–54.

19. Lu, D.; Mausel, P.; Brondízio, E.; Moran, E. Relationships between forest stand parameters and Landsat TM spectral responses in the Brazilian Amazon Basin. *For. Ecol. Manag.* **2004**, *198*, 149–167.

20. Astola, H.; Häme, T.; Sirro, L.; Molinier, M.; Kilpi, J. Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. *Remote Sens. Environ.* **2019**, *223*, 257–273.

21. Immitzer, M.; Vuolo, F.; Atzberger, C. First experience with Sentinel-2 data for crop and tree species classifications in central Europe. *Remote Sens.* **2016**, *8*, 166.

22. Bolyn, C.; Michez, A.; Gaucher, P.; Lejeune, P.; Bonnet, S. Forest mapping and species composition using supervised per pixel classification of sentinel-2 imagery. *Biotechnol. Agron. Soc. Environ.* **2018**, *22*, 172–187.

23. Hościło, A.; Lewandowska, A. Mapping forest type and tree species on a regional scale using multi-temporal Sentinel-2 data. *Remote Sens.* **2019**, *11*, 929.

24. Zhao, Q.; Yu, S.; Zhao, F.; Tian, L.; Zhao, Z. Comparison of machine learning algorithms for forest parameter estimations and application for forest quality assessments. *For. Ecol. Manag.* **2019**, *434*, 224–234. [CrossRef]

25. Fragou, S.; Kalogeropoulos, K.; Stathopoulos, N.; Louka, P.; Srivastava, P.; Karpouzas, S.; Kalivas, D.P.; Petropoulos, G.P. Quantifying land cover changes in a mediterranean environment using Landsat TM and Support Vector Machines. *Forests* **2020**, *11*, 750. [CrossRef]

26. Galgamuwa, G.A.P.; Wang, J.; Barden, C.J. Expansion of eastern redcedar (*Juniperus virginiana* L.) into the deciduous woodlands within the forest-prairie ecotone of Kansas. *Forests* **2020**, *11*, 154.

27. Noorian, N.; Shataee-Jouibary, S.; Mohammadi, J. Assessment of different remote sensing data for forest structural attributes estimation in the Hyrcanian forests. *For. Syst.* **2016**, *25*, 1–15. [CrossRef]

28. Huang, Y.; Chen, Z.X.; Yu, T.; Huang, X.; Gu, X. Agricultural remote sensing big data: Management and applications. *J. Integr. Agric.* **2018**, *17*, 1915–1931.

29. Sedona, R.; Cavallaro, G.; Jitsev, J.; Strube, A.; Riedel, M.; Benediktsson, J.A. Remote sensing big data classification with high performance distributed deep learning. *Remote Sens.* **2019**, *11*, 3056. [CrossRef]

30. Copernicus Space Component Mission Management Team. Sentinel High Level Operations Plan (HLOP): COPES1OP-EOPG-PL-15-0020. Available online: https://earth.esa.int/documents/247904/685154/Sentinel_High_Level_Operations_Plan (accessed on 17 August 2020).

31. Sudmanns, M.; Tiede, D.; Lang, S.; Bergstedt, H.; Trost, G.; Augustin, H.; Baraldi, A.; Blaschke, T. Big Earth data: Disruptive changes in Earth observation data management and analysis? *Int. J. Digit. Earth* **2020**, *13*, 832–850. [CrossRef]

32. Lu, D.; Weng, Q. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.* **2007**, *28*, 823–870. [CrossRef]

33. Andersen, R. Nonparametric methods for modeling nonlinearity in regression analysis. *Annu. Rev. Sociol.* **2009**, *35*, 67–85. [CrossRef]

34. Noi, P.T.; Kappas, M. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel-2 imagery. *Sensors* **2018**, *18*, 18.

35. Gajardo, J.; García, M.; Riaño, D. Applications of airborne laser scanning in forest fuel assessment and fire prevention. In *Forestry Applications of Airborne Laser Scanning Concepts and Case Studies*; Maltamo, M., Naesset, E., Vauhkonen, J., Eds.; Springer: Dodrecht, The Netherlands, 2014; ISBN 9789401786621.

36. Safari, A.; Sohrabi, H.; Powell, S.; Shataee, S. A comparative assessment of multi-temporal Landsat 8 and machine learning algorithms for estimating aboveground carbon stock in coppice oak forests. *Int. J. Remote Sens.* **2017**, *38*, 6407–6432. [CrossRef]

37. McCord, S.E.; Buenemann, M.; Karl, J.W.; Browning, D.M.; Hadley, B.C. Integrating remotely sensed imagery and existing multiscale field data to derive rangeland indicators: Application of Bayesian Additive Regression Trees. *Rangel. Ecol. Manag.* **2017**, *70*, 644–655. [CrossRef]

38. Ali, I.; Greifeneder, F.; Stamenkovic, J.; Neumann, M.; Notarnicola, C. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. *Remote Sens.* **2015**, *7*, 16398–16421. [CrossRef]

39. Bayat, M.; Ghorbanpour, M.; Zare, R.; Jaafari, A.; Thai Pham, B. Application of artificial neural networks for predicting tree survival and mortality in the Hyrcanian forest of Iran. *Comput. Electron. Agric.* **2019**, *164*, 104929. [CrossRef]

40. Chipman, H.A.; George, E.I.; McCulloch, R.E. BART: Bayesian additive regression trees. *Ann. Appl. Stat.* **2010**, *4*, 266–298. [CrossRef]

41. Izadi, S.; Sohrabi, H.; Khaledi, M.J. Estimation of coppice forest characteristics using spatial and non-spatial models and Landsat data. *J. Spat. Sci.* **2020**, 1–14. [CrossRef]

42. Clark, A.I.; Souter, R.A. *Stem Cubic-Foot Volume Tables for Tree Species in the South*; US Department of Agriculture, Forest Service, Southeastern Forest Experiment Station: Asheville, NC, USA, 1994; p. 252.

43. Al-Najjar, H.A.H.; Kalantar, B.; Pradhan, B.; Saeidi, V.; Halin, A.A.; Ueda, N.; Mansor, S. Land cover classification from fused DSM and UAV images using convolutional neural networks. *Remote Sens.* **2019**, *11*, 1461. [CrossRef]

44. Shanableh, A.; Al-Ruzouq, R.; Gibril, M.B.A.; Flesia, C.; Al-Mansoori, S. Spatiotemporal Mapping and Monitoring of Whiting in the Semi-Enclosed Gulf Using Moderate Resolution Imaging Spectroradiometer (MODIS) Time Series Images and a Generic Ensemble Tree-Based Model. *Remote Sens.* **2019**, *11*, 1193. [CrossRef]

45. Ball, L.; Tzanopoulos, J. Interplay between topography, fog and vegetation in the central South Arabian mountains revealed using a novel Landsat fog detection technique. *Remote Sens. Ecol. Conserv.* **2020**, 1–16. [CrossRef]

46. Ahmadi, K.; Jalil Alavi, S.; Zahedi Amiri, G.; Mohsen Hosseini, S.; Serra-Diaz, J.M.; Svenning, J.C. Patterns of density and structure of natural populations of Taxus baccata in the Hyrcanian forests of Iran. *Nord. J. Bot.* **2020**, *38*, 1–10. [CrossRef]

47. Rozenstein, O.; Haymann, N.; Kaplan, G.; Tanny, J. Validation of the cotton crop coefficient estimation model based on Sentinel-2 imagery and eddy covariance measurements. *Agric. Water Manag.* **2019**, *223*, 105715. [CrossRef]

48. Shataee, S.; Kalbi, S.; Fallah, A.; Pelz, D. Forest attribute imputation using machine-learning methods and ASTER data: Comparison of k-NN, SVR and random forest regression algorithms. *Int. J. Remote Sens.* **2012**, *33*, 6254–6280. [CrossRef]

49. Adeyeri, O.E.; Akinsanola, A.A.; Ishola, K.A. Investigating surface urban heat island characteristics over Abuja, Nigeria: Relationship between land surface temperature and multiple vegetation indices. *Remote Sens. Appl. Soc. Environ.* **2017**, *7*, 57–68. [CrossRef]

50. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [CrossRef]

51. Hammer, E.S.; Walsh, S.J. Canopy Structure in the Krummholz and Patch Forest Zones. In *The Changing Alpine Treeline*; Butler, D.R., Malanson, G.P., Wals, S.J., Fagre, D.B., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Volume 12, ISBN 9780444533647.

52. Kuhn, M.; Johnson, K. *Applied Predictive Modeling*; Springer: New York, NY, USA, 2013; ISBN 9781461468493.

53. Lebedev, A.V.; Westman, E.; Van Westen, G.J.P.; Kramberger, M.G.; Lundervold, A.; Aarsland, D.; Soininen, H.; Kłoszewska, I.; Mecocci, P.; Tsolaki, M.; et al. Random forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *NeuroImage Clin.* **2014**, *6*, 115–125. [CrossRef]

54. Yan, K.; Zhang, D. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. *Sens. Actuators B Chem.* **2015**, *212*, 353–363. [CrossRef]

55. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. [CrossRef]

56. Li, W.; Niu, Z.; Chen, H.; Li, D.; Wu, M.; Zhao, W. Remote estimation of canopy height and aboveground biomass of maize using high-resolution stereo images from a low-cost unmanned aerial vehicle system. *Ecol. Indic.* **2016**, *67*, 637–648. [CrossRef]

57. Varvia, P.; Lähivaara, T.; Maltamo, M.; Packalen, P.; Seppänen, A. Gaussian process regression for forest attribute estimation from airborne laser scanning data. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 3361–3369.

58. Li, X.; Wang, Y. Applying various algorithms for species distribution modelling. *Integr. Zool.* **2013**, *8*, 124–135. [PubMed]

59. Youssef, A.M.; Pradhan, B.; Pourghasemi, H.R.; Abdullahi, S. Landslide susceptibility assessment at Wadi Jawrah Basin, Jizan region, Saudi Arabia using two bivariate models in GIS. *Geosci. J.* **2015**, *19*, 449–469.

60. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning—With Applications in R*; Springer: Berlin/Heidelberg, Germany, 2013; ISBN 9781461471370.

61. Myers, R.H.; Montgomery, D.C.; Vining, G.G.; Robinson, T.J. *Generalized Linear Models with Applications in Engineering and the Sciences*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2002; Volume 95, ISBN 9780470454633.

62. Naghibi, S.A.; Pourghasemi, H.R. A Comparative Assessment Between Three Machine Learning Models and Their Performance Comparison by Bivariate and Multivariate Statistical Methods in Groundwater Potential Mapping Learning Models and Their Performance Comparison by Bivariate and Multivaria. *Water Resour. Manag.* **2015**, *29*, 5217–5236.

63. McRoberts, R.E. Forest ecology and management estimating forest attribute parameters for small areas using nearest neighbors techniques. *For. Ecol. Manag.* **2012**, *272*, 3–12.

64. Khan, M.; Ding, Q.; Perrizo, W. K-Nearest neighbor classification on spatial data streams using P-trees. In *Lecture Notes in Computer Science, Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining, Taipei, Taiwan, 6-8 May 2002*; Springer: Berlin/Heidelberg, Germany, 2002; Volume 2336, pp. 517–528.

65. Lang, M.; Arumäe, T.; Lükk, T.; Sims, A. Estimation of standing wood volume and species composition in managed nemoral multi-layer mixed forests by using nearest neighbour classifier, multispectral satellite images and airborne lidar data. *For. Stud.* **2014**, *61*, 47–68.

66. Kalantar, B.; Pradhan, B.; Naghibi, S.A.; Motevalli, A.; Mansor, S. Assessment of the effects of training data selection on the landslide susceptibility mapping: A comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN). *Geomat. Nat. Hazards Risk* **2018**, *9*, 49–69.

67. Tan, Y.V.; Roy, J. Bayesian additive regression trees and the general BART model. *Stat. Med.* **2019**, *38*, 5048–5069.

68. Ghorbanzadeh, O.; Rostamzadeh, H.; Blaschke, T.; Gholaminia, K.; Aryal, J. A new GIS-based data mining technique using an adaptive neuro-fuzzy inference system (ANFIS) and k-fold cross-validation approach for land subsidence susceptibility mapping. *Nat. Hazards* **2018**, *94*, 497–517.

69. Lindberg, E.; Hollaus, M. Comparison of methods for estimation of stem volume, stem number and basal area from airborne laser scanning data in a hemi-boreal forest. *Remote Sens.* **2012**, *4*, 1004–1023.

70. Khosravi, K.; Sartaj, M.; Tsai, F.T.C.; Singh, V.P.; Kazakis, N.; Melesse, A.M.; Prakash, I.; Tien Bui, D.; Pham, B.T. A comparison study of DRASTIC methods with various objective methods for groundwater vulnerability assessment. *Sci. Total Environ.* **2018**, *642*, 1032–1049. [CrossRef] [PubMed]

71. Opitz, D.; Blundell, S. An AFE approach for combining LIDAR and color imagery background: Feature analyst and LIDAR analyst. In Proceedings of the ASPRS 2008 Annual Conference, Portland, OR, USA, 28 April–2 May 2008.

72. Zhao, P.; Gao, L.; Gao, T. Extracting forest parameters based on stand automatic segmentation algorithm. *Sci. Rep.* **2020**, *10*, 1–13. [CrossRef] [PubMed]

73. Piñeiro, G.; Perelman, S.; Guerschman, J.P.; Paruelo, J.M. How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecol. Modell.* **2008**, *216*, 316–322. [CrossRef]

74. Mauya, E.W.; Koskinen, J.; Tegel, K.; Hämäläinen, J.; Kauranne, T.; Käyhkö, N. Modelling and predicting the growing stock volume in small-scale plantation forests of tanzania using multi-sensor image synergy. *Forests* **2019**, *10*, 279. [CrossRef]

75. Vafaei, S.; Soosani, J.; Adeli, K.; Fadaei, H.; Naghavi, H.; Pham, T.D.; Bui, D.T. Improving accuracy estimation of forest aboveground biomass based on incorporation of ALOS-2 PALSAR-2 and Sentinel-2A imagery and machine learning: A case study of the Hyrcanian forest area (Iran). *Remote Sens.* **2018**, *10*, 172. [CrossRef]

76. Valbuena, R.; Hernando, A.; Manzanera, J.A.; Görgens, E.B.; Almeida, D.R.A.; Silva, C.A.; García-Abril, A. Evaluating observed versus predicted forest biomass: R-squared, index of agreement or maximal information coefficient? *Eur. J. Remote Sens.* **2019**, *52*, 1–14. [CrossRef]

77. Fatehi, P.; Damm, A.; Leiterer, R.; Bavaghar, M.P.; Schaepman, M.E.; Kneubühler, M. Tree density and forest productivity in a heterogeneous alpine environment: Insights from airborne laser scanning and imaging spectroscopy. *Forests* **2017**, *8*, 212. [CrossRef]

78. Arévalo-Sandi, A.; Bobrowiec, P.E.D.; Chuma, V.J.U.R.; Norris, D. Diversity of terrestrial mammal seed dispersers along a lowland Amazon forest regrowth gradient. *PLoS ONE* **2018**, *13*, e0193752. [CrossRef]