Ann. For. Res. 65(1): 165-182, 2022
https://doi.org/10.15287/afr.2022.2390

ANNALS OF FOREST RESEARCH
www.afrjournal.org

Above-ground biomass estimation in a Mediterranean sparse coppice oak forest using Sentinel-2 data

Fardin Moradi[1], Seyed Mohammad Moein Sadeghi[2], Hadi Beygi Heidarlou[3], Azade Deljouei[2], Erfan Boshkar[4], Stelian Alexandru Borz[2]✉

Moradi F., Sadeghi S. M. M., Heidarlou H. B., Deljouei A., Boshkar E., Borz S. A., 2022. Above-ground biomass estimation in a Mediterranean sparse coppice oak forest using Sentinel-2 data. Ann. For. Res. 65(1): 165-182.

Abstract Implementing a scheduled and reliable estimation of forest characteristics is important for the sustainable management of forests. This study aimed at evaluating the capability of Sentinel-2 satellite data to estimate above-ground biomass (AGB) in coppice forests of Persian oak (Quercus brantii var. persica) located in Western Iran. To estimate the AGB, field data collection was implemented in 80 square plots (40×40 m, area of 1600 m^2). Two diameters of the crown were measured and used to calculate the AGB of each tree based on allometric equations. Then, the performance of satellite data in estimating the AGB was evaluated for the area of study using the field-based AGB (dependent variable) as well as the spectral band values, spectrally-derived vegetation indices (independent variables) and four machine learning (ML) algorithms: Multi-Layer Perceptron Artificial Neural Network (MLPNN), k-Nearest Neighbor (kNN), Random Forest (RF), and Support Vector Regression (SVR). A five-fold cross-validation was used to verify the effectiveness of models. Examination of the Pearson's correlation coefficient between AGB and the extracted values showed that IPVI and NDVI vegetation indices had the highest correlation with AGB (r = 0.897). The results indicated that the MLPNN algorithm was the best ML option (RMSE = 1.71 t ha^{-1}; MAE = 1.37 t ha^{-1}; relative RMSE = 24.75%; R^2 = 0.87) in estimating the AGB, providing new insights on the capability of remotely sensed-based AGB modeling of sparse Mediterranean forest ecosystems in an area with limited number of field sample plots.

Keywords: Sentinel-2, Above-ground biomass, Optical satellite data, Machine learning, Modeling, Estimation, Accuracy, Sparse-Coppice.

Addresses: [1]Department of Forestry and Forest Economics, Faculty of Natural Resources, University of Tehran, Karaj, Iran| [2]Department of Forest Engineering, Forest Management Planning and Terrestrial Measurements, Faculty of Silviculture and Forest Engineering, Transilvania University of Brasov, Romania| [3]Department of Forestry, Faculty of Natural Resources, Urmia University, Iran| [4]Natural Resources Department, Faculty of Agriculture, Razi University, Kermanshah, Iran.

✉ **Corresponding Author:** Stelian Alexandru Borz (stelian.borz@unitbv.ro).

Manuscript: received March 03, 2021; revised June 29, 2022; accepted June 29, 2022.

Introduction

Forests are among the most important terrestrial resources that can be used to manage global warming because they absorb large amounts of carbon dioxide through photosynthesis (Favero et al. 2020). As an essential source of information for better understanding and estimating the global carbon cycle, forest biomass is a critical component in forest management and national development planning (Khan et al. 2021, Santoro et al. 2021). Knowledge on the amount of forest biomass is essential for carbon trading and sustainable forest management (Knoke et al. 2021), assessment of habitat status, and its production capacity (Kovac et al. 2020). Awareness on the dynamics in biomass quantity and the background processes triggering its change is also required to shape national policies for forest management and carbon sequestration (Yao et al. 2018, Lamont et al. 2020).

Estimation of forest above ground biomass (AGB) can be done by destructive and non-destructive methods. Using destructive methods such as field-based sampling is the most accurate approach to estimating forest AGB (Vashum & Jayakumar 2012). However, this usually takes time and it is costly, while sometimes it may not be practical for large-scale study areas and poorly accessible forests. These limitations have led scholars to test and use non-destructive methods such as remote sensing (e.g., Pandit et al. 2018, Pham et al. 2018, 2020).

Capabilities such as quick processing and analysis, compatibility with Geographical Information Systems (GIS) and continuity, made satellite data an excellent provider of forest information (Pande et al. 2021). However, for accuracy reasons, past studies have shown that satellite data should not be used as a single source, but complemented with non-spectral data such as that coming from field-data sampling (Holmgren et al. 2000). Although radar, LIDAR, and high-resolution satellite imagery data can estimate biomass at increased accuracy, its wide use is still prevented by high operating costs, lack of international protocols, and low terrestrial coverage (Kelsey & Neff 2014). At least for these reasons, researchers in the field of forest sciences increasingly consider the use of freely available data at medium resolution.

Sentinel-2 data are being widely used in precision agriculture (Sagarra et al. 2020), land-use management (Pazur et al. 2021), and other fields due to high accuracy, multiple imaging bands, fast updating speed and free availability (Forkuor et al. 2020).

Sentinel-2 satellites are equipped with a Multi-Spectral Instrument (MSI), which provides a spatial resolution of 10 to 60 meters, a short revisit time (five days with two satellites under cloud-free conditions), and 13 spectral bands that range from the visible range to the shortwave, making it suitable for monitoring the vegetation conditions (Korhonen et al. 2017).

In recent years, many researchers studied the capabilities of optical satellite data to estimate forest AGB (Torabzadeh et al. 2019, Lovynska et al. 2020, Safari & Sohrabi 2020, Li et al. 2021). Due to the complex relations between remotely sensed variables and AGB (Zhao et al. 2016, Zhu et al. 2017), machine learning algorithms, such as the artificial neural networks (ANNs), k-nearest neighbor (kNN), random forest (RF), and support vector regression (SVR) were found to provide promising results, mainly due to the fact that they can handle nonlinear relations. This has led to an increasing popularity of these algorithms in the past decade as well as to a widespread application to AGB estimation in temperate (Moradi et al. 2022), Mediterranean (Puletti et al. 2020), tropical (Ghosh & Behera 2018), subtropical (Castillo et al. 2017), and boreal (Neumann et al. 2012) forest ecosystems.

ANN algorithms, for instance, have the advantage of distributed parallel processing and nonlinear and adaptive learning in AGB estimation (Foody et al. 2001). The kNN method calculates the spectral distance between plot locations and the estimated pixels and then computes a weighted average of their forest AGB values (Zhang et al. 2019). The algorithm gives more weight to the ground truth values which are similar to those of the pixel estimations and kNN can be used to predict continuous as well as discrete forest variables without assuming a normal distribution or linear relations between variables (Gao et al. 2018). Classical bagging algorithms such as RF have been considered to be among the best options of machine learning (Li et al. 2021). SVR algorithms, which can deal with both linear and nonlinear high dimensionality problems, are widely used in spectral analysis (Clevers et al. 2007, Zhang et al. 2020).

Zagros forests of Iran provide many economic and social benefits (Beygi Heidarlou et al. 2019). With an area of around 5 million ha (Roozitalab et al. 2018), these forests are spread over 11 provinces and play a critical role in improving the climate condition, water supply, economic and social balance of the region. Zagros oak forests are regarded as typical semi-arid Mediterranean forests (Fathizadeh et al. 2021), being open mixed stands dominated by Persian oak (*Quercus brantii* var. *persica*) (Pourhashemi & Sadeghi 2020).

Among all the studies on AGB estimation in oak forests, few focused on the *Q. brantii* dominated forests, particularly on those managed by coppicing (Safari & Sohrabi 2020). For example, Torabzadeh et al. (2019) estimated the AGB of coppice oak forests using optical data of the Sentinel-2 satellite. They used the RF algorithm for modeling by keeping the AGB as a dependent variable and the spectral band values and spectral-derived vegetation indices as independent variables, with a spatial resolution of 10 and 20 m. The model using all spectral bands and the derived vegetation indices provided the best estimates of AGB (R2 = 0.87 and RMSE = 10.75 t ha^{-1}). Their results proved the capability of these data in estimating the biomass at a lower cost compared to the data provided by active and hyperspectral sensors. Safari & Sohrabi (2020) used Sentinel-1 data to estimate the AGB of two sites with different levels of human intervention in the Zagros coppice oak forests. Their results showed that the genetic algorithm and removal-based variable screening techniques performed better; also, RF and multiple linear regression methods produced the best results.

In a similar study from northern Ukraine, Lovynska et al. (2020) examined the capability of the Sentinel-2 satellite data to estimate the amount of AGB in pine forests. Their study used the Normalized Difference Vegetation Index (NDVI), Transformed Vegetation Index (TVI), Fraction of Absorbed Photosynthetic Active Radiation (FARAP), and Fraction of Vegetation Cover (FCOVER) biophysical parameters as independent variables and the amount of biomass obtained from field surveys as a dependent variable. Their results showed a high correlation between the remotely-sensed independent variables and the AGB field estimates. Nathammachot et al. (2018) used Sentinel-2 spectral data to estimate the AGB in Indonesia. They used seven vegetation indices obtained from the satellite data and biomass data derived from 45 plots for stepwise regression modeling. Results showed that the NDI45 index (Normalized Difference Index using bands No. 4 and 5 of Sentinel-2) had the highest correlation with AGB (R2 = 0.79). The regression model was found to be a good option for characterizing the dependence between field-estimated AGB and the vegetation indices (R^2=0.81). Past studies, on the other hand, emphasized that there is no best single machine learning method for estimating AGB in different forest ecosystems (Ali et al. 2015).

Considering the importance of oak forests in Iran and other Mediterranean regions, and the role of these forests in the general environment, in preventing soil erosion (Attarod et al. 2017, Fathizadeh et al. 2021), providing wildlife habitats, providing various kinds of products and living conditions for forest dwellers, information on their condition becomes extremely important. On the other hand, these forests are typically non-commercial and the high cost of receiving non-free satellite data will make it difficult to manage them. Furthermore, it is unclear which machine-learning algorithm could produce the most accurate results for a particular study area and set of remotely sensed data. Hence, a comparative analysis needs to be carried out to check the performance of different modeling algorithms to estimate the AGB (Gao et al. 2018). Moreover, there are rare reports on comparative analysis of modeling algorithms under the conditions of coppice oak forests.

The use of freely available Sentinel-2 data has provided good opportunities to monitor forests and to carry on inventories for a sustainable management of these Mediterranean forests.

In this paper, we evaluate the capability of Sentinel 2 freely available data to estimate the AGB in coppice oak forests located in western Iran. Specifically, we attempt to answer two research questions: i) is the optical data provided by the Sentinel-2 satellite useful in estimating the AGB of coppice oak forests? and ii) which machine learning technique (ANN, kNN, RF, and SVR) is able to provide a suitable model for estimating AGB? Through this study, we can better understand the AGB modeling mechanisms by finding the suitable remotely sensed variables and modeling algorithms under the conditions of Mediterranean oak coppice forests.

Materials and Methods

Study area

The Zagros oak forests of western Iran are divided into several classes based on climate, tree and plant species: northern, central, and southern Zagros forests. The Gahvareh region, which is our study area, is located in Dalahu forests, Kermanshah Province and central Zagros, at an altitude ranging from 1850 to 2100 m, between 34˚31´55´´ to 34˚33´01´´ N and 46˚10´52´´ to 46˚12´19´´ E (Figure 1). The region's average annual precipitation is 498 mm, with winter and summer being the seasons of the most and least annual rainfall, respectively. The average annual temperature of the region is 15 °C. In most of the Zagros forests, the climate is Mediterranean (Attarod et al. 2017). The main tree species in Gahvareh is *Q. brantii*, which grows in coppiced form. Other species include gall oak (*Q. infectoria*), hawthorns (*Crataegus aronia*), wild pear (*Pyrus glabra*), and wild pistachio (*Pistacia atlantica*). All forest stands are characterized by low densities and heterogeneous canopies. Typically, forest stands are composed of isolated trees.

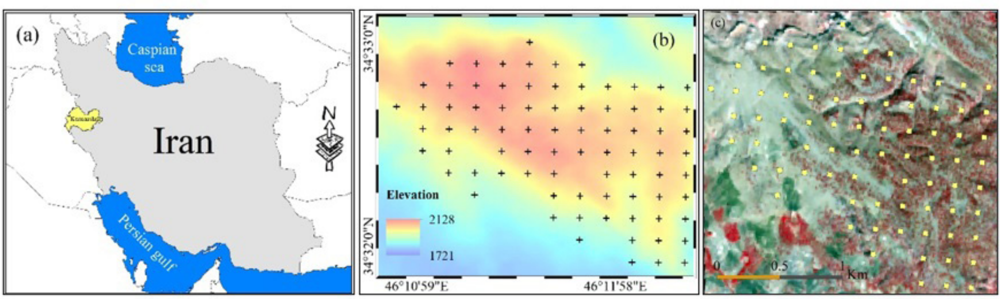

Figure 1 The geographic location of study area. (a) Kermanshah province of western Iran, (b) DEM map, and (c) sampling plots.

Field data collection

In order to obtain field data, 80 georeferenced square plots with an area of 1600 m^2 (40×40 m) each were established in the late summer of 2020. The plots were chosen based on the systematic method; in each plot, diameters at the breast height (DBH) of all trees thicker than 5 cm were measured (Karlson et al. 2015). The crown diameter of each tree was calculated as the average of two values measured along with two perpendicular directions from the location of each tree. Additionally, the species name of each tree was recorded in all 80 plots. Tree-level AGB was estimated using allometric equations (Eq. 1 and 2) developed for Zagros oak trees forests from Iran (Iranmanesh 2013). These were developed for calculating the AGB (t/ha) of single trees (Eq. 1) and sprout-clumps (Eq. 2) based on crown diameters (X, in m), respectively, being widely used in the Zagros Forest of Iran.

$$AGB=0.425\times X^{3.230} \quad (1)$$

$$AGB=1.275\times X^{2.362} \quad (2)$$

Satellite data

A single tile (cloud-free) standard Sentinel-2 product acquired on 21 December 2020 was downloaded from the Copernicus Data and Information Access Service (DIAS) (www.scihub.copernicus.eu). This data was the closest to the date of the field survey with per-pixel radiometric measurements and 13 spectral bands ranging from the visible to the shortwave infrared (SWIR). In this study, we excluded bands no. 1, 9, and 10 to extract spectral values corresponding to ground sample plots and statistical analysis. This study used cloud-free images produced through the Sen2Cor atmospheric correction processor on the Level-2A product (Castillo et al. 2017). Following the atmospheric correction, 9 bands were composited and clipped to the study area (Table 1).

Before analysis, we converted images into surface reflectance (SR) values using the FLAASH tool of the ENVI software (Ver. 5.3) (FLAASH 2009). In addition, 2D available 1:50,000 scale topographic maps and vector layers of roads and rivers from the Gahvareh region were used to correct geometric errors (Beygi Heidarlou et al. 2019). Table 2 lists the vegetation indices used in this study. We used the IDRISI Taiga software for all processing performed on the image's bands and for the calculation of considered indices. In order to extract values corresponding to each sample plot, the coordinates of 4 points around the plots were entered in ArcGIS (Ver. 10.8) to create a polygon for each plot. Then, the created shape files were exported to IDRISI Terrset software (Ver. 19.0.6), and after digitizing the plots, digital values of all bands were extracted and transferred to Statistica software (Ver. 12.0) for modeling.

The workflow implemented to model the AGB using Sentinel-2 data is shown in Figure 2, and it included four major steps: (1) preparation of data from different sources; (2) selection of the variables from Sentinel-2 data; (3) development of AGB estimation models using different algorithms; and (4) comparison and evaluation of the AGB modeling results.

Table 1 Spectral bands of the Sentinel-2 satellite imagery.

Spectral band	Center wavelength (nm)	Band width (nm)	Spatial resolution (m)
B2	490	65	10
B3	560	35	10
B4	665	30	10
B5	705	15	20
B6	740	15	20
B7	783	20	20
B8	842	115	10
B8a	865	20	20
B11	1610	90	20
B12	2190	180	20

Data analysis

ANN, kNN, RF, and SVR algorithms were used for modeling. All models were optimized by a five-fold cross-validation (Pham et al. 2020) to avoid overfitting (Pandit et al. 2018).

Table 2 Vegetation indices used in this study.

Vegetation Index	Equation	Reference
Inverted Red-Edge Chlorophyll Index (IRECI)	(NIR–RED)/(RED/RED)	(Frampton et al. 2013)
MERIS Terrestrial Chlorophyll Index (MTCI)	(NIR–RE)/(RE-RED)	(Dash & Curran 2007)
Modified Chlorophyll Absorption in Reflectance Index (MCARI)	[(RE–RED) – ((0.2)×(RE–GREEN))] × (RE × RED-1)	(Daughty et al. 2000)
Second Modified Soil-Adjusted Vegetation Index (MSAVI2)	(1 + 0.5) × (NIR − RED)/(NIR + RED + L); L= 0.5	(Huete 1988)
Soil-Adjusted Vegetation Index (SAVI)	[(NIR – RED)/(NIR + RED) + 0.5] × 0.5	(Clevers et al. 2002)
Normalized Difference Vegetation Index with Band 4 & 5 (NDI45)	NIR/(NIR + RED)	(Crippen 1990)
Sentinel-2 Red-Edge Position (S2REP)	705 + 35 × (RED + RE3)/(2−RE1)/(RE2 − RE1)	(Guyot & Baret 1988)
Green Normalized Difference Vegetation Index (GNDVI)	(NIR − GREEN)/(NIR + GREEN)	(Gitelson et al. 1996)
Normalized Difference Vegetation Index (NDVI)	(NIR − RED)/(NIR + RED)	(Rouse 1973)
Infrared Percentage Vegetation Index (IPVI)	0.5× (NDVI + 1)	(Crippen 1990)
Difference Vegetation Index (DVI)	NIR − RED	(Tucker 1979)
Pigment Specific Simple Ratio (PSSRA)	NIR/RED	(Roujean & Breon 1995)
Red-Edge Inflection Point (REIP)	705 + 35 × (RED – RE3)/(2 – RE1)/RE2-RE1)	(Hermann et al. 2011)
Ratio Vegetation Index (RVI)	NIR/RED	(Jordan 1969)

Note: NIR: Near Infrared and RE: Red-edge bands;
The value of n is equal to: n = (2 × (NIR2 - RED2) + 1.5× NIR + 0.5 × RED)/(NIR + RED + 0.5)

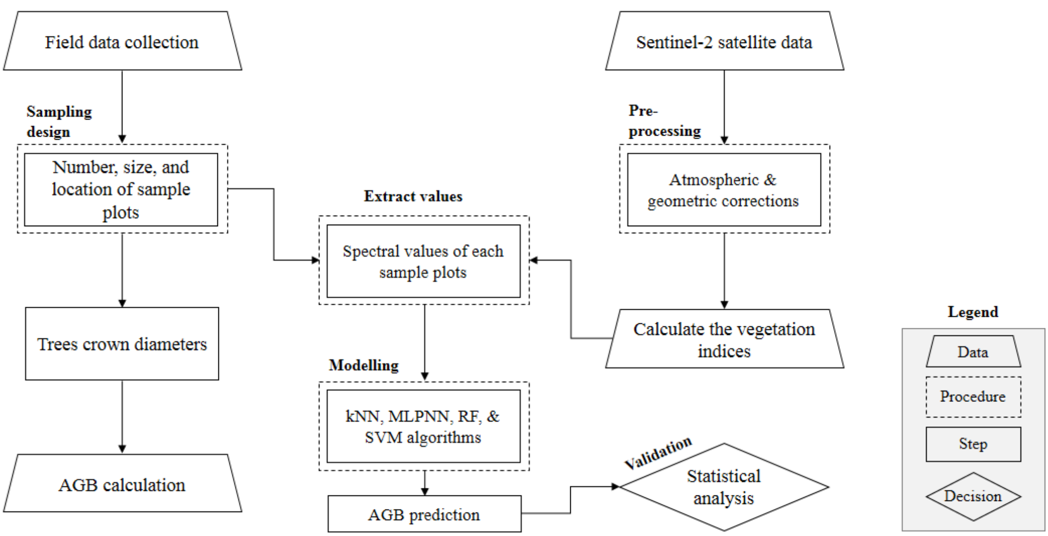

Figure 2 Research flowchart.

Multi-Layer Perceptron ANN algorithm

For non-linear, complex relations such as those specific to forest AGB modeling, an ANN consists of several nodes that are interconnected using mathematical algorithms (Li et al. 2020). Multi-Layer Perceptron ANNs (hereafter MLPNNs) are the most commonly used neural network algorithms in environmental modeling (Hu & Weng 2009, Hirigoyen et. al. 2021), and in forest AGB estimation (Mas 2004, Pham et al. 2017, 2018, Long et al. 2021). Each layer within a MLPNN is made up of several nodes or neurons.

There are typically three layers in such a model: the input, hidden, and output layers. In the input layer, the neurons receive the values of explanatory variables, whereas the number of neurons in the hidden layers should be tuned based on the characteristics of the data. For regression purposes, the output layer contains one neuron that holds the value of the predicted AGB.

MLPNN's efficacy is significantly affected by the connection weights between the input and hidden layers, as well as between the hidden and output layers (Moradi et al. 2022). A back-propagation algorithm (Pham et al. 2017) is used to adjust the weights in the training phase so as to lower the difference between the AGB values predicted by the MLPNN model and those from field inventory; the algorithm it is repeated until a predefined accuracy level or a specified number of iterations is reached.

To construct the MLPNNs model, the number of hidden neurons that significantly impact AGB estimation (Haykin 1998, Mas 2004) was selected based on the findings of previous studies (Bui et al. 2016). We then developed the best MLPNNs models by varying the number of neurons against the root mean square error (RMSE). The initial weights were assigned randomly, and when developing the network, the interconnection weights were adjusted to minimize the prediction error.

kNN algorithm

In forestry studies and particularly for those aiming at forest stand structural parameter estimation, the kNN algorithm is widely used (Gao et al. 2018, Zhang et al. 2019, Moradi et al. 2022). In the kNN, the similarity between observed and predicted values is measured using one- or a multiple-variable approach. The selection of k value and distance metric are critical factors affecting the estimation accuracy of the kNN algorithm. The optimal k value of the kNN algorithm was considered to stay between one and 50, according to the results of similar studies such as those of McRoberts (2008) and Shataee et al. (2012). Optimal k value was selected for the instance in which the prediction accuracy reached the maximum. There are four types of distance metrics commonly used in the kNN algorithm: the Euclidean, squared Euclidean, Manhattan, and Chebychev (Yazdani et al. 2020, Moradi et al. 2022) (Equations 3-6) distances.

$$\text{Euclidean} = \sqrt{(x-p)^2} \quad (3)$$

$$\text{Squared Euclidean} = (x-p)^2 \quad (4)$$

$$\text{Manhattan} = Abs(x-p) \quad (5)$$

$$\text{Chebychev} = \max(|x-p|) \quad (6)$$

where x and p are the target and reference units, respectively.

RF algorithm

RF is a robust machine learning algorithm known for its high accuracy; it is based on bagging and random feature selection (Breiman 2001, Belgiu & Dragut 2016). RF can capture complex non-linear relations and it can deal with multicollinearity problems (Anderson et al. 2018, Dalla Corte et al. 2020, Torre-Tojal et al. 2022). RF algorithm is fast and easy to use, and it can make accurate predictions even when highly correlated variables are present, such as when predicting biomass data (Torre-Tojal et al. 2022). In the past decade, this algorithm has been widely used in creating models that relate forest characteristics to variables derived from multi-source data (Tian et al. 2014, Anderson et al. 2018, Dalla Corte et al. 2020, Yadav et al. 2021, Torre-Tojal et al. 2022, Xi et al. 2022).

The RF picks random subsets of explanatory variables and builds the tree up to a certain point. The Classification and Regression Trees (CART) algorithm is used to first construct multiple decision trees, and then to combine the predictions from each tree to produce an ensemble response (Gisalson et al. 2006). RF can be used for classification problems (the output of the RF represents the class selected by the majority of trees) or regression problems (the individual trees return the mean or average prediction). In regression problems, RF is used to vote responses to be combined (averaged) to estimate the dependent variable by an arbitrary number (ensemble) of simple trees (a subset of independent variables) (Shataee et al. 2012). To generate a forest of regression trees, randomly sampled data and variables can be bagged and bootstrapped iteratively (Stevens et al. 2015). Implementing RF requires regularizing a decision tree and setting stopping parameters. Finally, the RF model is constructed by grouping base-decision trees into a forest (Pham et al. 2019).

Seventy percent of the total plots from the training dataset (i.e., in bag data), should be contained in these bootstrap datasets, and the rest of the datasets (i.e., 30%; out of bag data) are used to evaluate the performance of the RF model (Rodriguez-Galiano et al. 2012). Past research has found that the number of base-decision trees should be carefully selected since the RF model's performance is dependent on this parameter (Stevens et al. 2015, Pham et al. 2019). Three parameters are used in the architecture of the RF algorithm: the number of trees (Ntree), the minimum number of observations per leaf (mtry), and the number of repetitions in the calculation of importance (nperm). A total of 500 trees (i.e., Ntree) were used for this study to ensure the stability of RF model results (Stevens et al. 2015, Moradi et al. 2022). The number of trees that produced a stable error was kept as the optimal number of trees in the RF model. We modified both mtry and Ntree by rerunning the settings, and then compared the average squared error between test samples (Gao et al. 2018).

SVR algorithm

The SVR has been recognized as one of the most efficient machine learning algorithms (Smola & Scholkopf 2004), including in the estimation of forest AGB (Wu et al. 2016, 2018). Literature reviews showed that the performance of the SVR model is significantly influenced by the selection of the kernel type (Vafaei et al. 2018, Wu et al. 2018). In this study, four kernels were considered for the SVR algorithm, including Linear, Polynomial, Radial Basis Function (RBF), and Sigmoid (Ronoud et al. 2021). To estimate the AGB, an SVR function is built according to Eq. 7.

$$AGB = \sum_{i=1}^{n} a_i k(x_i; x) + b \qquad (7)$$

where $k(x_i; x)$ is the kernel function; x_i is the training vector; α represents the LaGrange multiplier; and b denotes the bias term in the regression.

Statistical analysis and model performance assessment

A sensitivity analysis was performed to determine the most effective model parameters (Baloly et al. 2018). The Kolmogorov-Smirnov test was used to check if samples came from a population with a normal distribution. We used the Pearson correlation coefficient (*r*) to check the association between the AGB of each sample plot (dependent variable) and the corresponding spectral values (independent variables). Statistical analysis was implemented in the Statistica software (Ver. 12.0).

Cross-validation was used to evaluate the model's generalization capability as a strategy to deal with a small number of ground-truth sample points and to avoid overfitting of the model (Pandit et al. 2018). We used four performance metrics (Equations 8-11), namely the root mean square error (*RMSE*), the mean absolute error (*MAE*), relative

RMSE (*RMSE*%), and the coefficient of determination (R^2) (Nazari, et al. 2020, Nazari, et al. 2020), to compare the performance of the selected machine learning techniques in AGB estimation, where low *RMSE* and *MAE* values, on the one hand, and high R^2 values, on the other hand, characterize better predictive capabilities of the models.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{\left(AGB_p - AGB_i\right)^2}{n}} \tag{8}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|AGB_p - AGB_i\right| \tag{9}$$

$$RMSE(\%) = \frac{AGB_o}{RMSE} \times 100$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(AGB_i - AGB_p\right)^2}{\sum_{i=1}^{n}\left(AGB_i - AGB_o\right)^2} \tag{10}$$

where AGB_p and AGB_i are the predicted and observed AGB per i^{th} plot, respectively, n is the total number of testing sample plots, and AGB_o is the average of the testing sample plots.

Results

Field Survey

Table 3 shows the descriptive statistics of the stand structural parameters and AGB. Field measurements indicated that the minimum, maximum, and mean values of the AGB were of 0.03, 19.76, and 6.91 t/ha, respectively, with a standard error of 0.60 t/ha (Table 3).

Correlation between AGB and spectral bands

Based on Pearson's correlation coefficient, a significant negative association was found between the spectral band information (i.e., B2, B3, B4, B5, B6, B7, B8, B8a, B11, and B12) and in situ AGB (Table 4); there was a significant positive relation between AGB and vegetation indices ($p < 0.01$; Table 4). IPVI and NDVI vegetation indices of the Sentinel-2 data outputted the highest correlation with in situ AGB ($r = 0.897$; Table 4). The sensitivity

Table 3 Characteristics of the sparse coppice oak forest from the study site.

Statistical Parameter	Mean			
	Tree Height (m)	DBH (cm)	Crown Diameter (m)	AGB (t/ha)
Maximum	8.17	70.06	6.55	19.76
Minimum	1.45	7.14	1.30	0.04
Average	4.07	23.97	3.86	6.91
Standard Deviation (±)	1.54	11.15	1.15	5.36
Standard Error (±)	0.17	1.25	0.13	0.60
Coefficient of Variation (%)	37.85	46.54	29.89	77.63

Table 4 Pearson correlation coefficient between spectral variables and aboveground biomass (AGB).

Input Variable	Correlation Coefficient (*r*)	Input Variable	Correlation Coefficient (*r*)
B2	-0.712**	IPVI	0.897**
B3	-0.729**	IRECI	0.843**
B4	-0.758**	MCARI	0.618**
B5	-0.715**	MSAVI2	0.890**
B6	-0.508**	MTCI	0.896**
B7	-0.412**	NDI45	0.726**
B8	-0.386**	NDVI	0.897**
B8a	-0.361**	PSSRA	0.886**
B11	-0.674**	REIP	0.854**
B12	-0.741**	RVI	0.890**
DVI	0.776**	S2REP	0.705**
GNDVI	0.885**	SAVI	0.890**

Notes: ** Significance level: 0.01.

analysis results showed that the variables B8a, RVI, and SAVI had the greatest impact on modeling, respectively.

Models for AGB estimation

The results of the four machine learning models are summarized in Table 5. In general, the MLPNN algorithm produced the best results for AGB estimation in our study area, (the lowest RMSE, Relative RMSE, MAE and the highest R^2), while the worst results were those outputted by the kNN algorithm. The best result was obtained when the MLPNN machine learning model had three hidden layers. Among the distance metrics used for the kNN model, the squared Euclidean distance produced the most accurate results (i.e., RMSE = 2.32 t ha^{-1}, MAE = 1.75 t ha^{-1}, relative RMSE = 33.58%, and R^2 = 0.79). The best kernel type for the SVR algorithm was the linear one, which returned the lowest RMSE and MAE and the highest R^2.

Figure 3 shows the average squared error rates plotted against the number of trees used for estimating AGB by the RF algorithm during training and testing. The optimal number of trees was determined as the point where increasing the number of trees does not cause changes in the error rate (Figure 3). Improvement in error rates from 160 trees onward was low, so this number was used as the optimum number of trees (Figure 3).

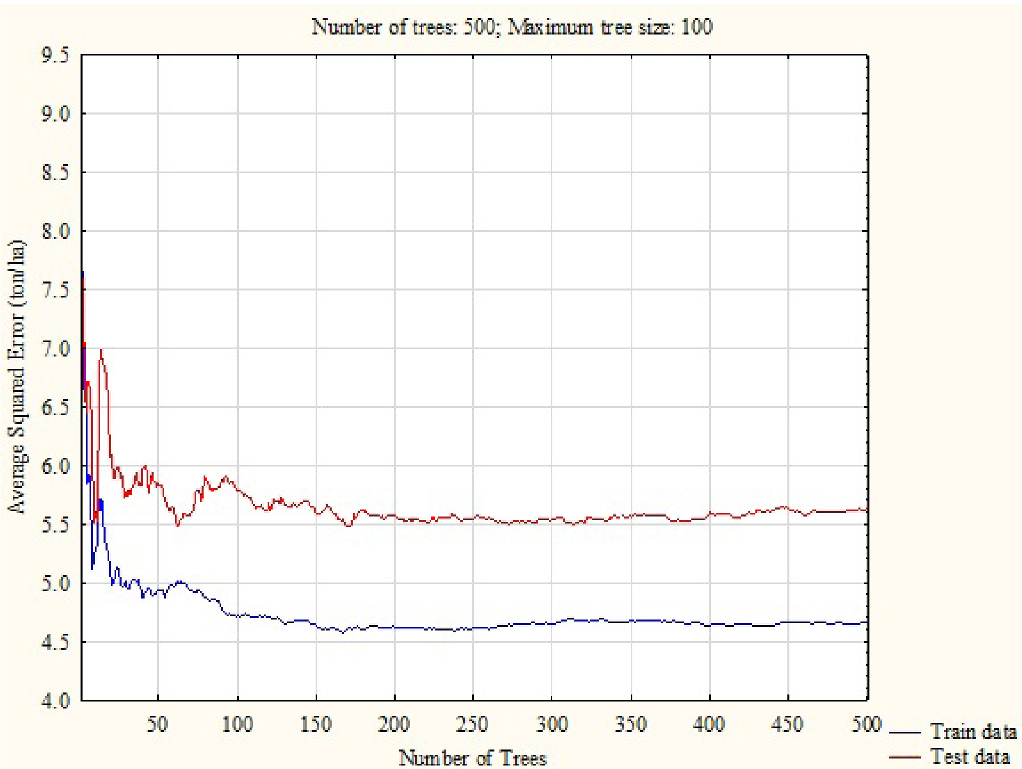

Figure 3 Random Forest (RF) error testing graph - the average squared error of above ground biomass (AGB) outputted by the RF algorithm, plotted against the number of trees using the training and testing datasets.

Table 5 Performance of the tested machine learning algorithms.

Algorithm	Parameters		RMSE ($t\ ha^{-1}$)	MAE ($t\ ha^{-1}$)	Relative RMSE (%)	R^2
MLPNN	3 hidden layers (first layer = 24 neurons; second layer = 18 neurons; third layer = 1 neuron)		1.71	1.37	24.75	0.86
Algorithm	**Distance Variable**	**Optimal k**	**RMSE ($t\ ha^{-1}$)**	**MAE ($t\ ha^{-1}$)**	**Relative RMSE (%)**	**R^2**
kNN (*k* range = 1-50)	Euclidean	34	2.46	1.88	35.60	0.78
	Squared Euclidean	38	2.32	1.75	33.58	0.79
	Manhattan	32	2.44	1.79	35.31	0.77
	Chebyshev	34	2.68	2.08	38.78	0.75
Algorithm	**Optimal k**		**RMSE ($t\ ha^{-1}$)**	**MAE ($t\ ha^{-1}$)**	**Relative RMSE (%)**	**R^2**
RF	7		2.08	1.64	30.10	0.82
Algorithm	**Kernel Type**		**RMSE ($t\ ha^{-1}$)**	**MAE ($t\ ha^{-1}$)**	**Relative RMSE (%)**	**R^2**
SVR	Linear		2.03	1.49	29.38	0.86
	Polynomial		2.41	1.78	34.88	0.80
	Radial Basis Function (RBF)		2.13	1.53	30.83	0.84
	Sigmoid		2.14	1.55	30.97	0.83

Discussion

Monitoring changes in forested area as well as its biomass are important activities for the management and sustainable development of forest ecosystems (Cakir et al. 2008, Li et al. 2021). Therefore, monitoring forests and obtaining up-to-date and accurate information will play a significant role in improving the management of forest resources. Capabilities such as the ability to quickly process large data sets, acceptable spatial and temporal coverage, free satellite image data, and the accuracy of the obtained information stand as good sources of information for a sustainable resource management. On the other hand, uncertainty of predictions based on remotely sensed data was always discussed by experts. Such uncertainties may come from an incorrect ground survey, improper size and an insufficient number of sample plots, heterogeneous conditions of the study areas, data processing tools and techniques, statistical methods, and a considerable time interval between the acquisition of remotely sensed data and field data collection. This study aimed at investigating the capabilities of the optical Sentinel-2 satellite data in estimating the AGB of Mediterranean coppice oak forests located in Iran based on Sentinel-2 bands with a spatial resolution of 10 and 20 m, 15 vegetation indices and four common non-parametric machine learning algorithms (MLPNN, kNN, SVR and RF).

All of the vegetation indices derived from the Sentinel-2 datasets showed a significant correlation with forest AGB in these Mediterranean forests. Thus, by increasing forest AGB value, the reflectance will also increase, which is in line with the results of other studies (Ronoud et al. 2021). The Pearson correlation coefficient between input variables and AGB showed that IPVI and NDVI indices had the highest correlation with the field-sampled AGB. In previous studies, NDVI and the infrared band showed a high correlation with forest characteristics, such as forested area, density, and AGB (Luo et al. 2013, Karlson et al. 2015; Mohamed et al. 2016, Safari et al. 2017, Barakat et al. 2018, Eskandari et al. 2020, Pirotti et al. 2020). This is due to the high sensitivity of the infrared spectrum to the forest structure. The amount of reflection is also reduced by reducing the biomass, especially in the near-infrared spectrum. Decreased reflection prevents the occurrence of the saturation phenomenon (Gasparri et al. 2010), which leads to better performance of the near-infrared band and of the NDVI index, which is highly effective by the near-infrared band. The amount of forest biomass is directly related to forest management. The lower average value of forest biomass in this study as compared to previous studies conducted in the Zagros forests (Safari et al. 2017, Torabzadeh et al. 2019) can be due to cutting tree branches and removing of leaves by local people to ensure their fuel needs. The area was also heavily affected by the *Loranthus europaeus* (a hemiparasitic plant), so the Department of Natural Resources and the villagers cut down the infected branches to combat oak decline (Pourhashemi & Sadeghi 2020), which reduced the canopy density and eventually its AGB. Reducing the density of tree canopies increases the interference brought by soil reflection which combines with that of tree cover at the pixel level. Mixed pixels were found to significantly affect the accuracy of estimates such as carbon sequestration in low canopy forests (Calvao & Palmeirim 2004, Eisfelder et al. 2012).

The machine learning algorithm selection is of first importance as it may significantly affect the accuracy of estimations (Vafaei et al. 2017, Ronoud et al. 2021, Moradi et al. 2022). It is unclear which algorithm will produce the most accurate results for a given study area and set of remotely sensed data (Safari et al. 2017, Vafaei et al. 2017). Hence, a comparative study of different AGB estimation algorithms is required. Most machine learning algorithms do not make any assumptions about the model or the distribution of input data, which makes it possible to explicitly describe the nonlinear relation between forest AGB and remotely sensed data (Liu et al. 2017, Pandit et al. 2018). Among the four machine learning techniques used in this study, the MLPNN model provided the highest estimation accuracy (highest R^2, and the lowest RMSE, Relative RNSE, and MAE performance metrics). The Sentinel-2-derived AGB model using the MLPNN algorithm yielded a RMSE% of 24.7%. Considering the heterogeneity of the study area and the small number of available field measurements, the obtained results can be considered satisfactory. The MLPNN algorithm learns quickly, generalizes well, and has high self-learning ability levels (Camargo et al. 2019). MLPNN consists of an input layer, which receives the signal, an output layer, which makes a decision or prediction about the input, and an arbitrary number of hidden layers (or none at all) in between (Mas 2004, Pham et al. 2017, Camargo et al. 2019). The MLPNN model has been the most commonly-applied ANN model for reliable forest ABG prediction (Foody et al. 2001, Englhart et al. 2011, Ozcelik et al. 2017, Pham et al. 2017). For example, Pham et al. (2017) concluded that the MLPNN algorithm achieved the best performance ($R^2 = 0.78$) for estimating the AGB of mangrove apple (*Sonneratia caseolaris*) in Vietnam as compared to SVR, RF, radial basis function neural networks (RBFNN), and Gaussian process (GP). In another study, Masjedi et al. (2018) pointed out that the MLPNN models have a more precise result than RF to estimate AGB of *Sorghum bicolor* in Iran.

The kNN presented the highest accuracy and lower errors when using the squared Euclidean distance metric (Table 5). These results confirmed the outcomes of previous research, which reported that kNN implementations with the squared Euclidean metric had a consistently smaller error and larger accuracy than those with other distance metrics (Shataee et al. 2012). Literature research showed that the squared Euclidean metric lowers the bias of the estimates (e.g., Altman 1992, Tuominen et al. 2010). The type of the kernel function had a significant effect on the performance of SVR models, and we found that the best one was the linear kernel (Table 5). It requires the fewest parameters and it is less susceptible to overfitting compared to other kernel types (Axelsson et al. 2013, Marabel & Alvarez-Taboada 2013).

There are some uncertainties in our AGB estimation procedure. First, we did not have access to the species-related allometric equations for our study area, therefore we used equations (i.e., Equations 2 and 3) which are general for all the Zagros forests. In addition, GPS positional errors considerably affect the results obtained from remote sensing studies (Ronoud et al. 2021) while all of the machine learning algorithms are site-specific in the sense that their results might fluctuate by the study area. As such, generalization of our results should be made with caution and based on site-specificity validation.

Conclusion

This study investigated the capability of Sentinel-2 spectral data to estimate the AGB of oak coppice stands in a part of Kermanshah province (located in a Mediterranean Zagros region, Iran). These sparse-covered forests are being currently managed mainly for their environmental value. Among the four-tested ML algorithms, our findings indicate a better performance of the MLPNN in estimating the AGB from sparse coppice oak forests. The MLPNN algorithm used in our research can be replicated over other coppice oak forests sharing similar characteristics, stand structures and biophysical parameters, standing for an effective way of estimating the AGB in low-access forests or in forests lacking the resources for a detailed ground-based sampling of AGB.

Acknowledgements

Seyed Mohammad Moein Sadeghi and Azade Deljouei's research at the Transilvania University of Brasov, Romania, has been supported by the program "Transilvania Fellowship for Postdoctoral Research/Young Researchers." The authors acknowledge the support of the Department of Forest Engineering, Forest Management Planning and Terrestrial Measurements, Faculty of Silviculture and Forest Engineering, Transilvania University of Brasov.

Ethical statement

All ethical practices have been followed in relation to the development, writing, and publication of the paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Ali I., Greifeneder F., Stamenkovic J., Neumann M., Notarnicola C., 2015. Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data. Remote Sensing 7(12): 16398-16421. https://doi.org/10.3390/rs71215841

Altman N.S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician 46: 175–185. https://doi.org/10.2307/2685209

Anderson K.E., Glenn N.F., Spaete L.P., Shinneman D.J., Pilliod D.S., Arkle R.S., 2018. Estimating vegetation biomass and cover across large plots in shrub and grass dominated drylands using terrestrial lidar and machine learning. Ecological Indicators 84: 793-802. https://doi.org/10.1016/j.ecolind.2017.09.034

Attarod P., Sadeghi S., Pypker T., Bayramzadeh V., 2017. Oak trees decline; a sign of climate variability impacts in the west of Iran. Caspian Journal of Environmental Sciences 15(4): 373-384.

Axelsson C., Skidmore A.K., Schlerf M., Fauzi A., Verhoef W., 2013. Hyperspectral analysis of mangrove foliar chemistry using PLSR and support vector regression. International Journal of Remote Sensing, 34: 1724-1743. https://doi.org/10.1080/01431161.2012.725958

Baloloy A., Blanco A., Candido C., Argamosa R., Dumalag J., Dimapilis L., 2018. Estimation of mangrove forest aboveground biomass using multispectral bands, vegetation indices and biophysical variables derived from optical satellite imageries: Rapideye, Planetscope and Sentinel-2. ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences 4(3): 29-36. https://doi.org/10.5194/isprs-annals-IV-3-29-2018

Barakat A., Khellouk R., El Jazouli A., Touhami F., Nadem S., 2018. Monitoring of forest cover dynamics in eastern area of Béni-Mellal Province using ASTER and Sentinel-2A multispectral data. Geology, Ecology, and Landscapes 2(3): 203-215. https://doi.org/10.1080/24749508.2018.1452478

Belgiu M., Drăguţ L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing 114: 24-31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Beygi Heidarlou H., Shafiei A.B., Erfanian M., Tayyebi A., Alijanpour A., 2019. Effects of preservation policy on land use changes in Iranian Northern Zagros forests. Land Use Policy 81: 76-90. https://doi.org/10.1016/j.landusepol.2018.10.036

Breiman L., 2001. Random forests. Machine Learning 45(1): 5-32. https://doi.org/10.1023/A:1010933404324

Bui D.T., Tuan T.A., Klempe H., Pradhan B., Revhaug I., 2016. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides 13(2): 361-378. https://doi.org/10.1007/s10346-015-0557-6

Çakir G., Sivrikaya F., Keleş S., 2008. Forest cover change and fragmentation using Landsat data in Maçka State Forest Enterprise in Turkey. Environmental Monitoring and Assessment 137(1): 51-66. https://doi.org/10.1007/s10661-007-9728-9

Calvao T., Palmeirim J., 2004. Mapping Mediterranean scrub with satellite imagery: biomass estimation and spectral behaviour. International Journal of Remote Sensing 25(16): 3113-3126. https://doi.org/10.1080/01431160310001654978

Camargo F.F., Sano E.E., Almeida C.M., Mura J.C., Almeida, T., 2019. A comparative assessment of machine-learning techniques for land use and land cover classification of the Brazilian tropical savanna using ALOS-2/PALSAR-2 polarimetric images. Remote Sensing 11(13): 1600. https://doi.org/10.3390/rs11131600

Castillo J.A.A., Apan A.A., Maraseni T.N., Salmo III S.G., 2017. Estimation and mapping of above-ground biomass of mangrove forests and their replacement land uses in the Philippines using Sentinel imagery. ISPRS Journal of Photogrammetry and Remote Sensing 134: 70-85. https://doi.org/10.1016/j.isprsjprs.2017.10.016

Clevers J., De Jong S., Epema G., Van Der Meer F., Bakker W., Skidmore A., 2002. Derivation of the red edge index using the MERIS standard band setting. International Journal of Remote Sensing 23(16): 3169-3184. https://doi.org/10.1080/01431160110104647

Clevers J., Van der Heijden G., Verzakov S., Schaepman M.E., 2007. Estimating grassland biomass using SVM band shaving of hyperspectral data. Photogrammetric Engineering & Remote Sensing 73(10): 1141-1148. https://doi.org/10.14358/PERS.73.10.1141

Crippen R.E., 1990. Calculating the vegetation index faster. Remote Sensing of Environments 34(1): 71-73. https://doi.org/10.1016/0034-4257(90)90085-Z

Dalla Corte A.P., Souza D.V., Rex F.E., Sanquetta C.R., Mohan M., Silva C.A., 2020. Forest inventory with high-density UAV-Lidar: Machine learning approaches for predicting individual tree attributes. Computers and Electronics in Agriculture 179: 105815. https://doi.org/10.1016/j.compag.2020.105815

Dash J., Curran, P., 2007. Evaluation of the MERIS terrestrial chlorophyll index (MTCI). Advances in Space Research 39(1): 100-104. https://doi.org/10.1016/j.asr.2006.02.034

Daughtry C.S., Walthall C., Kim M., De Colstoun E.B., McMurtrey Iii J., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. Remote Sensing of Environment 74(2): 229-239. https://doi.org/10.1016/S0034-4257(00)00113-9

Eisfelder C., Kuenzer C., Dech S., 2012. Derivation of biomass information for semi-arid areas using remote-sensing data. International Journal of Remote Sensing 33(9): 2937-2984. https://doi.org/10.1080/01431161.2011.620034

Englhart S., Keuck V., Siegert F., 2011. Modeling aboveground biomass in tropical forests using multi-frequency SAR data—A comparison of methods. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5(1): 298-306. https://doi.org/10.1109/JSTARS.2011.2176720

Eskandari S., Jaafari M., Oliva P., Ghorbanzadeh O., Blaschke T., 2020. Mapping land cover and tree canopy cover in Zagros forests of Iran: Application of Sentinel-2, Google Earth, and field data. Remote Sensing 12(12): 1912. https://doi.org/10.3390/rs12121912

Fathizadeh O., Sadeghi S.M.M., Pazhouhan I., Ghanbari S., Attarod P., Su L., 2021. Spatial variability and optimal number of rain gauges for sampling throughfall under single oak trees during the leafless period. Forests 12(5): 585. https://doi.org/10.3390/f12050585

Favero A., Daigneault A., Sohngen B., 2020. Forests: Carbon sequestration, biomass energy, or both?. Science Advances 6(13): 6792. https://doi.org/10.1126/sciadv.aay6792

FLAASH., 2009. Atmospheric correction module: QUAC and FLAASH user's guide. Version, 4, 44.

Foody G.M., Cutler M.E., McMorrow J., Pelz D., Tangki H., Boyd D.S., 2001. Mapping the biomass of Bornean tropical rain forest from remotely sensed data. Global Ecology and Biogeography 10(4): 379-387. https://doi.or/10.1046/j.1466-822X.2001.00248.x

Forkuor G., Zoungrana J.B.B., Dimobe K., Ouattara B., Vadrevu K.P., Tondoh J.E., 2020. Above-ground biomass mapping in West African dryland forest using Sentinel-1 and 2 datasets-A case study. Remote Sensing of Environment 236: 111496. https://doi.org/10.1016/j.rse.2019.111496

Frampton W.J., Dash J., Watmough G., Milton E.J., 2013. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. ISPRS Journal of Photogrammetry and Remote Sensing 82: 83-92. https://doi.org/10.1016/j.isprsjprs.2013.04.007

Gao Y., Lu D., Li G., Wan G., Chen Q., Liu L., 2018. Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. Remote Sensing 10(4): 627. https://doi.org/10.3390/rs10040627

Gasparri N.I., Parmuchi M.G., Bono J., Karszenbaum H., Montenegro C.L., 2010. Assessing multi-temporal Landsat 7 ETM+ images for estimating above-ground biomass in subtropical dry forests of Argentina. Journal of Arid Environments 74(10): 1262-1270. https://doi.org/10.1016/j.jaridenv.2010.04.007

Ghosh S.M., Behera M.D., 2018. Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest. Applied Geography 96: 29-40. https://doi.org/10.1016/j.apgeog.2018.05.011

Gislason P.O., Benediktsson J.A., Sveinsson J.R., 2006. Random forests for land cover classification. Pattern Recognition Letters 27(4): 294-300. https://doi.org/10.1016/j.patrec.2005.08.011

Gitelson A.A., Kaufman Y.J., Merzlyak M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environments 58(3): 289-298. https://doi.org/10.1016/S0034-4257(96)00072-7

Guyot G., Baret F., 1988. Utilisation de la haute resolution spectrale pour suivre l'etat des couverts vegetaux. In Spectral Signatures of Objects in Remote Sensing (Vol. 287, p. 279).

Haykin S., 1998. A comprehensive foundation. Neural Networks 2(2004): 41. https://doi.org/10.1142/S0129065794000372

Herrmann I., Pimstein A., Karnieli A., Cohen Y., Alchanatis V., Bonfil D., 2011. LAI assessment of wheat and potato crops by VENμS and Sentinel-2 bands. Remote Sensing of Environment 115(8): 2141-2151. https://doi.org/10.1016/j.rse.2011.04.018

Hirigoyen A., Acosta-Muñoz C., Ariza Salamanca A.J., Varo-Martinez M.Á., Rachid-Casnati C., Franco J., Navarro-Cerrillo R., 2021. A machine learning approach to model leaf area index in Eucalyptus plantations using high-resolution satellite imagery and airborne laser scanner data. Ann. For. Res. 64(2): 165-183. https://doi.org/10.15287/afr.2021.2073

Holmgren J., Joyce S., Nilsson M., Olsson H., 2000. Estimating stem volume and basal area in forest compartments by combining satellite image data with field data. Scandinavian Journal of Forest Research 15(1): 103-111. https://doi.org/10.1080/02827580050160538

Hu X., Weng Q., 2009. Estimating impervious surfaces from medium spatial resolution imagery using the self-organizing map and multi-layer perceptron neural networks. Remote Sensing of Environment 113(10): 2089-2102. https://doi.org/10.1016/j.rse.2009.05.014

Huete A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment 25(3): 295-309. https://doi.org/10.1016/0034-4257(88)90106-X

Iranmanesh Y., 2013. Assessment on biomass estimation methods and carbon sequestration of *Quercus brantii* Lindl. Chaharmahal & Bakhtiari forests. PhD Dissertation. Tehran: Tarbiat Modares University, 107.

Jordan C.F., 1969. Derivation of leaf-area index from quality of light on the forest floor. Ecology 50(4): 663-666. https://doi.org/10.2307/1936256

Karlson M., Ostwald M., Reese H., Sanou J., Tankoano B., Mattsson E., 2015. Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using Landsat 8 and random forest. Remote Sensing 7(8): 10017-10041. https://doi.org/10.3390/rs70810017

Kelsey K.C., Neff J.C., 2014. Estimates of aboveground biomass from texture analysis of Landsat imagery. Remote Sensing 6(7): 6407-6422. https://doi.org/10.3390/rs6076407

Khan I.A., Khan W.R., Ali A., Nazre M., 2021. Assessment of above-ground biomass in Pakistan forest ecosystem's carbon pool: A review. Forests 12(5): 586. https://doi.org/10.3390/f12050586

Knoke T., Kindu M., Schneider T., Gobakken T., 2021. Inventory of forest attributes to support the integration of non-provisioning ecosystem services and biodiversity into forest planning—from collecting data to providing information. Current Forestry Reports 7(1): 38-58. https://doi.org/10.1007/s40725-021-00138-7

Korhonen L., Packalen P., Rautiainen M., 2017. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. Remote Sensing of Environment 195: 259-274. https://doi.org/10.1016/j.rse.2017.03.021

Kovac M., Gasparini P., Notarangelo M., Rizzo M., Cañellas I., Fernández-de-Uña L., Alberdi I., 2020. Towards a set of national forest inventory indicators to be used for assessing the conservation status of the habitats directive forest habitat types. Journal for Nature Conservation 53: 125747. https://doi.org/10.1016/j.jnc.2019.125747

Lamont K., Saintilan N., Kelleway J.J., Mazumder D., Zawadzki A., 2020. Thirty-year repeat measures of mangrove above-and below-ground biomass reveals unexpectedly high carbon sequestration. Ecosystems 23(2): 370-382. https://doi.org/10.1007/s10021-019-00408-3

Li C., Zhou L., Xu W., 2021. Estimating aboveground biomass using Sentinel-2 MSI data and ensemble algorithms for grassland in the Shengjin Lake Wetland, China. Remote Sensing 13(8): 1595. https://doi.org/10.3390/rs13081595

Li Y., Li M., Li C., Liu Z., 2020. Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. Scientific Reports 10(1): 1-12. https://doi.org/10.1038/s41598-020-67024-3

Liu K., Wang J., Zeng W., Song J., 2017. Comparison and evaluation of three methods for estimating forest above ground biomass using TM and GLAS data. Remote Sensing 9(4): 341. https://doi.org/10.3390/rs9040341

Long S., Zeng S., Wang G., 2021. Developing a new model for predicting the diameterdistribution of oak forests using an artificialneural network. Ann. For. Res. 64(2): 3-20. https://doi.org/10.15287/afr.2021.2060

Lovynska V., Buchavyi Y., Lakyda P., Sytnyk S., Gritzan Y., Sendziuk R., 2020. Assessment of pine aboveground biomass within Northern Steppe of Ukraine using Sentinel-2 data. Journal of Forest Science 66(8): 339-348. https://doi.org/10.17221/28/2020-JFS

Luo X., Chen X., Xu L., Myneni R., Zhu Z., 2013. Assessing performance of NDVI and NDVI3g in monitoring leaf unfolding dates of the deciduous broadleaf forest in Northern China. Remote Sensing 5(2): 845-861. https://doi.org/10.3390/rs5020845

Marabel M., Alvarez-Taboada F., 2013. Spectroscopic determination of aboveground biomass in grasslands using spectral transformations, support vector machine and partial least squares regression. Sensors 13: 10027-10051. https://doi.org/10.3390/s130810027

Mas J., 2004. Mapping land use/cover in a tropical coastal area using satellite sensor data, GIS and artificial neural networks. Estuarine, Coastal and Shelf Science 59(2): 219-230. https://doi.org/10.1016/j.ecss.2003.08.011

Masjedi A., Zhao J., Thompson A.M., Yang K.W., Flatt J.E., Crawford M.M., 2018. Sorghum biomass prediction using UAV-based remote sensing data and crop model simulation. IEEE pp. 7719-7722. https://doi.org/10.1109/IGARSS.2018.8519034

McRoberts R.E., 2008. Using satellite imagery and the k-nearest neighbors technique as a bridge between strategic and management forest inventories. Remote Sensing of Environments 112(5): 2212-2221. https://doi.org/10.1016/j.rse.2007.07.025

Mohamed M., Yusup S., Bustam M.A., 2016. Synthesis of CaO-based sorbent from biomass for CO2 capture in series of calcination-carbonation cycle. Procedia Engineering 148: 78-85. https://doi.org/10.1016/j.proeng.2016.06.438

Moradi F., Darvishsefat A.A., Pourrahmati M.R., Deljouei A., Borz S.A., 2022. Estimating aboveground biomass in dense Hyrcanian forests by the use of Sentinel-2 data. Forests 13(1): 104. https://doi.org/10.3390/f13010104

Neumann M., Saatchi S.S., Ulander L.M., Fransson J.E., 2012. Assessing performance of L-and P-band polarimetric interferometric SAR data in estimating boreal forest above-ground biomass. IEEE Transactions on Geoscience and Remote Sensing 50(3): 714-726. https://doi.org/10.1109/TGRS.2011.2176133

Nuthammachot N., Phairuang W., Wicaksono P., Sayektiningsih T., 2018. Estimating aboveground biomass on private forest using Sentinel-2 imagery. Journal of Sensors https://doi.org/10.1155/2018/6745629

Özçelık R., Diamantopoulou M.J., Eker M., Gürlevık N., 2017. Artificial neural network models: an alternative approach for reliable aboveground pine tree biomass prediction. Forest Sciences 63(3): 291-302. https://doi.org/10.5849/FS-16-006

Pande C.B., Moharir K.N., Singh S.K., Varade A.M., Elbeltagi A., Khadri S.F.R., Choudhari P., 2021. Estimation of crop and forest biomass resources in a semi-arid region using satellite data and GIS. Journal of the Saudi Society of Agricultural Sciences 20(5): 302-311. https://doi.org/10.1016/j.jssas.2021.03.002

Pandit S., Tsuyuki S., Dube T., 2018. Estimating above-ground biomass in sub-tropical buffer zone community forests, Nepal, using Sentinel 2 data. Remote Sensing 10(4): 601. https://doi.org/10.3390/rs10040601

Pazúr R., Huber N., Weber D., Ginzler C., Price B., 2021. A national extent map of cropland and grassland for Switzerland based on Sentinel-2 data. Earth System Science Data Discussions 14: 295-305. https://doi.org/10.5194/essd-14-295-2022

Pham T.D., Le N.N., Ha N.T., Nguyen L.V., Xia J., Yokoya N., 2020 Estimating mangrove above-ground biomass using extreme gradient boosting decision trees algorithm with fused sentinel-2 and ALOS-2 PALSAR-2 data in can Gio biosphere reserve, Vietnam. Remote Sensing 12(5): 777. https://doi.org/10.3390/rs12050777

Pham T.D., Xia, J., Ha N.T., Bui D.T., Le N.N., Tekeuchi W., 2019. A review of remote sensing approaches for monitoring blue carbon ecosystems: Mangroves, seagrassesand salt marshes during 2010–2018. Sensors 19(8): 1933. https://doi.org/10.3390/s19081933

Pham T.D., Yoshino K., Le N.N., Bui D.T., 2018. Estimating aboveground biomass of a mangrove plantation on the Northern coast of Vietnam using machine learning techniques with an integration of ALOS-2 PALSAR-2 and Sentinel-2A data. International Journal of Remote Sensing 39(22): 7761-7788. https://doi.org/10.1080/01431161.2018.1471544

Pham T.D., Yoshino, K., Bui, D.T., 2017. Biomass estimation of *Sonneratia caseolaris* (l.) Engler at a coastal area of Hai Phong city (Vietnam) using ALOS-2 PALSAR imagery and GIS-based multi-layer perceptron neural networks. GIScience & Remote Sensing 54(3): 329-353. https://doi.org/10.1080/15481603.2016.1269869

Pirotti F., Kutchartt E., Csaplovics E., 2020. Assessment of volume and above-ground biomass in Araucaria forest through satellite images, comparing different methods in the south of Chile. IEEE pp. 579-584. https://10.1109/LAGIRS48042.2020.9165668

Pourhashemi M., Sadeghi S.M.M., 2020. A review on ecological causes of oak decline phenomenon in forests of Iran. Ecology of Iranian Forest 8(16): 148-164.

Puletti N., Grotti M., Ferrara C., Chianucci F., 2020. Lidar-based estimates of aboveground biomass through ground, aerial, and satellite observation: a case study in a Mediterranean forest. Journal of Applied Remote Sensing 14(4): 044501. https://doi.org/10.1117/1.JRS.14.044501

Rodriguez-Galiano V.F., Ghimire B., Rogan J., Chica-Olmo M., Rigol-Sanchez J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing 67: 93-104. https://doi.org/10.1016/j.isprsjprs.2011.11.002

Ronoud G., Fatehi P., Darvishsefat A.A., Tomppo E., Praks J., Schaepman M.E., 2021. Multi-sensor aboveground biomass estimation in the broadleaved Hyrcanian forest of Iran. Canadian Journal of Remote Sensing 47(6): 818-834. https://doi.org/10.1080/07038992.2021.1968811

Roozitalab M.H., Siadat H., Farshad A., 2018. The soils of Iran. Springer. https://doi.org/10.1007/978-3-319-69048-3

Roujean J.L., Breon F.M., 1995. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. Remote Sensing of Environments 51(3): 375-384. https://doi.org/10.1016/0034-4257(94)00114-3

Rouse R.A., 1973. Conformational analysis of saturated heterocycles substituted final ozonides. International Journal of Quantum Chemistry 7(7): 289-294. https://doi.org/10.1021/ja00792a003

Safari A., Sohrabi H., 2020. Integration of synthetic aperture radar and multispectral data for aboveground biomass retrieval in Zagros oak forests, Iran: an attempt on Sentinel imagery. International Journal of Remote Sensing 41(20): 8069-8095. https://doi.org/10.1080/01431161.2020.1771789

Safari A., Sohrabi H., Powell S., Shataee S., 2017. A comparative assessment of multi-temporal Landsat 8 and machine learning algorithms for estimating aboveground carbon stock in coppice oak forests. International Journal of Remote Sensing 38(22): 6407-6432. https://doi.org/10.1080/01431161.2017.1356488

Santoro M., Cartus O., Carvalhais N., Rozendaal D., Avitabile V., Araza A., Willcock S., 2021. The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations. Earth System Science Data 13(8): 3927-3950. https://doi.org/10.5194/essd-13-3927-2021

Segarra J., Buchaillot M.L., Araus J.L., Kefauver S.C., 2020. Remote sensing for precision agriculture: Sentinel-2 improved features and applications. Agronomy 10(5): 641. https://doi.org/10.3390/agronomy10050641

Shataee S., Kalbi S., Fallah A., Pelz D., 2012. Forest attribute imputation using machine-learning methods and ASTER data: comparison of k-NN, SVR and random forest regression algorithms. International Journal of Remote Sensing 33(19): 6254-6280. https://doi.org/10.1080/01431161.2012.682661

Smola A.J., Schölkopf B., 2004. A tutorial on support vector regression. Statistics and Computing 14(3): 199-222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Stevens F.R., Gaughan A.E., Linard C., Tatem A.J., 2015. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. PloS one 10(2): e0107042. https://doi.org/10.1371/journal.pone.0107042

Tian X., Li Z., Su Z., Chen E., van der Tol C., Li X., 2014. Estimating montane forest above-ground biomass in the upper reaches of the Heihe River Basin using Landsat-TM data. International Journal of Remote Sensing 35(21): 7339-7362. https://doi.org/10.1080/01431161.2014.967888

Torabzadeh H., Moradi M., Fatehi P., 2019. Estimating aboveground biomass in Zagors forest, Iran, using sentinel-2 data. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 42: 1059-1063. https://doi.org/10.5194/isprs-archives-XLII-4-W18-1059-2019

Torre-Tojal L., Bastarrika A., Boyano A., Lopez-Guede J.M., Graña M., 2022. Above-ground biomass estimation from LiDAR data using random forest algorithms. Journal of Computational Science 58: 101517. https://doi.org/10.1016/j.jocs.2021.101517

Tucker C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment 8(2): 127-150. https://doi.org/10.1016/0034-4257(79)90013-0

Tuominen S., Eerikäinen K., Schibalski A., Haakana M., Lehtonen A., 2010. Mapping biomass variables with a multi-source forest inventory technique. Silva Fennica 44(1): 109-119. https://doi.org/10.14214/sf.458

Vafaei S., Soosani J., Adeli K., Fadaei H., Naghavi H., Pham T.D., 2018 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 Sensing 10(2): 172. https://doi.org/10.3390/rs10020172

Vashum K.T., Jayakumar S., 2012. Methods to estimate above-ground biomass and carbon stock in natural forests-a review. Journal of Ecosystem & Ecography 2(4): 1-7. http://doi.org/10.4172/2157-7625.1000116

Wu C., Shen H., Shen A., Deng J., Gan M., Zhu J., 2016. Comparison of machine-learning methods for above-ground biomass estimation based on Landsat imagery. Journal of Applied Remote Sensing 10(3): 035010. https://doi.org/10.1117/1.JRS.10.035010

Wu C., Tao H., Zhai M., Lin Y., Wang K., 2018. Using nonparametric modeling approaches and remote sensing imagery to estimate ecological welfare forest biomass. Journal of Forestry Research 29(1): 151-161. https://doi.org/10.1007/s11676-017-0404-9

Xi Z., Xu H., Xing Y., Gong W., Chen G., Yang S., 2022. Forest canopy height mapping by synergizing ICESat-2, Sentinel-1, Sentinel-2 and topographic information based on machine learning methods. Remote Sensing 14(2): 364. https://doi.org/10.3390/rs14020364

Yadav S., Padalia H., Sinha S.K., Srinet R., Chauhan P., 2021. Above-ground biomass estimation of Indian tropical forests using X band Pol-InSAR and Random Forest. Remote Sensing Applications: Society and Environment 21: 100462. https://doi.org/10.1016/j.rsase.2020.100462

Yao Y., Piao S., Wang T., 2018. Future biomass carbon sequestration capacity of Chinese forests. Science Bulletin 63(17): 1108-1117. https://doi.org/10.1016/j.scib.2018.07.015

Yazdani M., Jouibary S.S., Mohammadi J., Maghsoudi Y., 2020. Comparison of different machine learning and regression methods for estimation and mapping of forest stand attributes using ALOS/PALSAR data in complex Hyrcanian forests. Journal of Applied Remote Sensing 14(2): 024509. https://doi.org/10.1117/1.JRS.14.024509

Zhang J., Tian H., Wang D., Li H., Mouazen A.M., 2020. A novel approach for estimation of above-ground biomass of sugar beet based on wavelength selection and optimized support vector machine. Remote Sensing 12(4): 620. https://doi.org/10.3390/rs12040620

Zhang L., Shao Z., Liu J., Cheng Q., 2019. Deep learning based retrieval of forest aboveground biomass from combined LiDAR and Landsat 8 data. Remote Sensing 11(12): 1459. https://doi.org/10.3390/rs11121459

Zhao P., Lu D., Wang G., Wu C., Huang Y., Yu S., 2016. Examining spectral reflectance saturation in Landsat imagery and corresponding solutions to improve forest aboveground biomass estimation. Remote Sensing 8(6): 469. https://doi.org/10.3390/rs8060469

Zhu J., Huang Z., Sun H., Wang G., 2017. Mapping forest ecosystem biomass density for Xiangjiang River Basin by combining plot and remote sensing data and comparing spatial extrapolation methods. Remote Sensing 9(3): 241. https://doi.org/10.3390/rs9030241