Assessment of pine aboveground biomass within Northern Steppe of Ukraine using Sentinel-2 data

Viktoriia Lovynska1*, Yuriy Buchavyi2, Petro Lakyda3, Svitlana Sytnyk4, Yuriy Gritzan1, Roman Sendziuk4

1Department of Parks and Gardens, Faculty of Agronomy, Dnipro State Agrarian and Economic University, Dnipro, Ukraine
2Department of Ecology and Technologies of Environmental Protection, National Technical University “Dnipro Polytechnic”, Dnipro, Ukraine
3Education and Research Institute of Forestry and Landscape Park Management, National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine
4Department of Forest Measurement and Forest Management, National University of Life and Environmental Sciences of Ukraine, Kyiv, Ukraine

*Corresponding author: glub@ukr.net


Abstract: The present study offers the results of the spectral characteristics, calculated vegetative indices and biophysical parameters of pine stands of the Northern Steppe of Ukraine region obtained using Sentinel-2 data. For the development of regression models with the prediction of the biomass of pine forests using the obtained spectral characteristics, we used the results of the assessment of the aboveground biomass by the method of field surveys. The results revealed the highest correlation relations between the parameters of the general and trunk biomass with the normalised difference vegetation index (NDVI) and transformed vegetation index (TVI) vegetative indices and the fraction of absorbed photosynthetic active radiation (FARAP) and fraction of vegetation cover (FCOVER) biophysical parameters.

To generate the models of determining the forest aboveground biomass (AGB), we used both the single- and two-factor models, the most optimum of which were those containing the NDVI predictor separately and in combination with the FCOVER predictor. The predicted values of the total AGB for the mentioned models equalled 32.5 to 236.3 and 39.9 to 253.4 t·ha\(^{-1}\). We performed mapping of the AGB of pine stands of the Northern Steppe using multi-spectral Sentinel-2 images, particularly the spectral characteristics of their derivatives (vegetative indices, biophysical parameters). This study demonstrated promising results for conducting an AGB-mapping of pine woods in the studied region using free-access resources.

Keywords: Pinus sylvestris L.; spectral indices; remote sensing; allometric regression; Steppe zone

Forest plants are the main bodies in terrestrial ecosystems and play a significant role in the absorption of carbon (C) from the atmosphere through photosynthesis. Forests ecosystem store the C in the biomass with a higher concentration in the components of the aboveground biomass (foliage, branches, trunks) (Schimel et al. 1995; Mohd Zaki, Abd Latif 2016). Biomass estimation for forests has received much attention in recent years because the change in the regional biomass is associated with important components of climate change (Nemani et al. 2003; Madugundu et al. 2008). The gross primary production (GPP) of forest ecosystems constitutes the largest C flux and is the basis for the production of wood, and, therefore, has important implications for humans. The estimation of the GPP and forest carbon storage dynamics has an important role in forecasting the sustainable de-
Development and can be monitored by the periodic mapping of the forest biomass (Castillo et al. 2017; Chen et al. 2019).

Optical and microwave remote sensing have been used to monitor the forest biomass over the last few decades. The integration of remote sensing data from forest biomass assessments on a regional level to wide country areas has been an invaluable contribution for the estimations of global ecosystem process models (Ghosh, Behera 2018; Chrysafis et al. 2019). Remote sensing has been widely used to study the forest aboveground biomass (AGB) and terrestrial C stocks, in addition to it, remote sensing data allow one to examine the impacts of the global climate change on the C dynamics (Saatchi et al. 2007).

The AGB is generally estimated by empirically or functionally relating it to satellite-derived variables based on empirical models (regression or machine learning approaches using neural networks) and physically-based allometric models (Hazarika et al. 2005; Chrysafis et al. 2017).

Satellite-derived remote sensing vegetation indices (VIs) (e.g., normalised difference vegetation index (NDVI), enhanced vegetation index (EVI)) have been traditionally used for the evaluation of the gross primary production of the forest ecosystem. The obtained VIs have effectively been used in mathematical models for quantifying the forest aboveground biomass and C stocks (Potter et al. 1993; Zhou et al. 2001; Zhang, Kondragunta 2006). Simple regression models are established by associating a single VI or spectral reflectance to field biomass measurements (Roy, Ravan 1996).

Although a simple model is widely applied to estimate the AGB in a local region, the model’s accuracy and optimal independent variables vary with the spectral variables and local environment (Foody et al. 2003). Furthermore, simple regression models could also correlate the biomass to the shadow fraction from high resolution imagery and LAI (leaf area index) in logarithmic relationships (Madugundu et al. 2008).

A multiple regression model is able to improve the biomass estimates by combining the surface reflectance, VIs, and biophysical variables (Zhang, Kondragunta 2006).

Recently, there has been an increasing need to carry out the investigations of the functional status and productivity of forest stands in the steppe. The Scots pine is one of the main forest-forming species of the forests of the steppe zone of Ukraine. In this region, pine plantations cover an area of 21472.9 ha, have a natural (20%) and an artificial origin, and predominantly perform sanitation and recreational functions (Hulchak et al. 2011).

For the automated and rapid identification of the structural components of forest ecosystems, it is most appropriate to use multiband imagers with geo-information system tools like a special and zonal statistics analysis.

The purpose of the study is to develop mathematical models for the estimation of the aboveground biomass of the pine stands in the Northern Steppe of Ukraine on the basis of the data decoding multiband images of Sentinel-2.

**MATERIAL AND METHODS**

**Study area.** The Northern Steppe of Ukraine is located in a moderate climatic belt in the area of a mild moderately continental climate (Figure 1). The values of the annual total solar radiation in the Northern Steppe of Ukraine change from north to south from 4200 to 4400 mJ·(m²)⁻¹, the sum of the active temperatures higher than 10 °C change from north to south from 2700 to 3400. Winter isotherms change from north southwards from –6.2 °C to –4.0 °C, summer isotherms change from +20.5 °C to +22.0 °C. The average temperatures in January, the coldest month, and July, the warmest month, are –3.6 °C and +22.1 °C, respectively. The average amount of precipitation is equal to 400–500 mm. The coefficient of moisture is 0.625–0.688, indicating an insufficient amount of moisture in the territory. The main soil-forming geological rock in the majority of the territory of the surveyed forestry area is loess based on loose carbonate rocks. The most common genetic group of soils is the common black soil.

**Field measurements.** A field study was conducted from 2015 to 2019 to examine the relationship between the remotely measured reflectance and the aboveground biomass for the Scots pine forest plantations. The project was carried out in twenty-five temporary sample plot areas (TSPs) (50 × 50 m) which were used to build and validate the AGB estimation models. The TSPs are within the responsibility of the Dnipropetrovsk Administration of Forest and Hunting Management. The distribution of the sampling plots was randomly generated.

At each TSP, the GPS coordinates were recorded, the diameter of the trees at breast height (DBH)
and the tree heights were measured. The mean stand height and DBH for all the TSPs were 19.4 m and 21.8 cm, respectively.

The average diameter at breast height (D), average height (H) and relative density of the stand (RD) were measured for the forest AGB calculation of the main characteristics of the forest stands within a TSP. The AGB from D, H, RD were previously calculated based on the allometric function of a dependent field-based forest (Lovynska et al. 2019).

**Satellite data processing.** The present study used Sentinel-2 satellite imagery that has 13 spectral bands from visible to infrared and systematically acquires the optical imagery at a high spatial resolution from 10 m to 60 m. The Biophysical Processor computes the Level-2B Biophysical products from the Sentinel-2 reflectance. It derives a set of biophysical variables from the top-of-canopy normalised reflectance data, namely:

(i) leaf area index (LAI),
(ii) fraction of absorbed photosynthetic active radiation (FAPAR),
(iii) fraction of vegetation cover (FCOVER),
(iv) Chlorophyll content in the foliage (CAB).

The Sentinel-2 SNAP Toolbox biophysical variable retrieval algorithm is based on specific radioactive transfer models associated with strong assumptions, particularly regarding the canopy architecture (turbid medium model). All the variables derived from such algorithms should be seen as effective, i.e., the variables that would correspond to the measured satellite signal reflected by the canopy verifying all the assumptions made through the radioactive transfer models. Depending on the variable, this may lead to differences with the ground values that may be accessed from the field measurements.

Furthermore, the algorithm is “generic”, i.e., it should apply to any type of vegetation with reasonable performances. However, to better match the specificities of the given canopies, either a simple correction could be calibrated, or a more specific algorithm could be developed.

One strong assumption embedded in any single pixel retrieval algorithm as this one, is that the pixel targeted belongs to a landscape patch presenting enough homogeneity (on a pixel scale) preventing the unexpected loss or gain of radiation fluxes. For forests with large crowns, or any pixel showing strong heterogeneity, such as pixels at the intersection between two different vegetation patches, the results may be uncertain. This also applies to pixels where the neighbouring ones are very different.
Specific algorithms should be developed to detect such situations and possibly propose. The proposed algorithm is based on methods that have already been proven to be efficient (Figure 2).

They have been implemented in order to generate biophysical products from VEGETATION, MERIS, SPOT, and LANDSAT sensors. It mainly consists in generating a comprehensive database of vegetation characteristics and the associated Sentinel-2 top of canopy reflectance (TOC). Neural networks are then trained to estimate the canopy characteristics from the TOC reflectance along with a set of corresponding angles defining the observational configuration.

The actual algorithm running in SNAP runs the prediction step of the neural network, from a set of pre-processing coefficients computed during the training phase.

Cloud-free Sentinel-2A images covering the study area were obtained on December 25, 2015, and downloaded from the Copernicus Scientific Data Hub (https://scihub.copernicus.eu/) as a Level-1C product. The Level-1C product is composed of 100 × 100 km² tiles.

The statistical relationship between the investigated values was conducted using non-linear regression models. The best fit models were reported with the coefficients of determination (R²) and F-test coefficients.

RESULTS

The data acquired in the field surveys on the aboveground biomass of pine stands ranged from 6.22 (DBH 5 cm) to 472.31 (DBH 30 cm) t·ha⁻¹. The results of the LAI of the forests revealed that this parameter ranges from 1.1 to 3.36. The average values of the calculated parameters of the biomass and LAI equalled 211.5 ± 21.7 t·ha⁻¹ and 1.92 ± 0.12 m²·(m³)⁻¹, respectively.

The pairwise Pearson’s product-moment correlation analysis was operated to determine the relationship of the field-based AGB with the multisensor-derived indices and collinearity between each type of variables. Comparing the results of the field measurements with the satellite-derived parameters of the multi-spectral analysis for the studied TSP revealed correlations of different strength between the aboveground biomass of the pine stands and the NDVI and TVI with the biophysical parameters LAI, CAB, CW, FAPAR, FCOVER (Table 1).

We determined that all the variables obtained from the satellite images by Sentinel-2 had a positive correlation with the AGB of the pine forests. The obtained spectral characteristics correlated with stronger correlations with the AGB for the trunks and the total AGB compared with the AGB of the crown. According to the correlation analysis, the closest relationship (\( P < 0.05 \)) was present for
the AGB and the parameters NDVI, TVI, FAPAR, LAI and FCOVER. It was these particular predictors that were selected for the following surveys.

The next step was determining the mathematical relationship of the AGB to the selected predictors. As an analytical form of the obtained results, we used an equation of a non-linear (allometric) type, where the initial arguments were NDVI, LAI, FAPAR, TVI, FCOVER parameters. As a result, the obtained regression models contained the coefficients of equations given in Table 2.

The highest values of the coefficient of determination parameters were seen in case of the dependence of the AGB on the NDVI index separately, and also in combination with other predictors (NDVI + TVI; NDVI + FCOVER). In such models, the determination coefficient of the total AGB of the pine forests reached 0.707. The lowest accuracy of approximation was determined in the case of the dependence of the AGB on the LAI. The adequacy of the generated models was also determined while testing them for the Fisher (F) criterion. The obtained actual F-values exceeded the table values both in the case of applying the single-factor (F\text{table} – 9.33), and two-factor (F\text{table} – 6.93) models.

These assessment results demonstrate that empirical models could be potentially used to develop a map of spatial distribution of the AGB of the strata of the pine forests, based on satellite-derived variables, specifically NDVI.

The graphic interpretation of the obtained dependencies of the AGB on the NDVI is shown in Figure 3. The mean value of the NDVI for the sur-

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|}
\hline
\textbf{Values} & \textbf{NDVI} & \textbf{TVI} & \textbf{FAPAR} & \textbf{FCOVER} & \textbf{LAI} & \textbf{CAB} & \textbf{CW} \\
\hline
\textbf{AGB\text{total}} & 0.833 & 0.771 & 0.801 & 0.789 & 0.826 & 0.783 & 0.721 \\
\textbf{AGB\text{stem}} & 0.824 & 0.780 & 0.835 & 0.806 & 0.837 & 0.799 & 0.790 \\
\textbf{AGB\text{crown}} & 0.215* & 0.130* & 0.030* & 0.103* & 0.133* & 0.103* & 0.130* \\
\hline
\end{tabular}
\caption{Pearson’s correlation coefficients between the field-measurement and satellite-obtained data}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|c|c|}
\hline
\textbf{Model} & \textbf{a} & \textbf{b} & \textbf{c} & \textbf{R}^2 & \textbf{F-test} \\
\hline
\text{Allometric } y = ax^b, y = a \times x_1 \times x_2^c & & & & & \\
\text{– with NDVI} & 563.620 & 1.251 & – & 0.698 & 30.72 \\
\text{– with LAI} & 240.241 & 0.265 & – & 0.544 & 12.68 \\
\text{– with NDVI + LAI} & 495.317 & 1.098 & 0.098 & 0.636 & 7.13 \\
\text{– with NDVI + FAPAR} & 550.474 & 1.140 & 0.096 & 0.698 & 14.22 \\
\text{– with NDVI + TVI} & 716.261 & 1.608 & -1.644 & 0.706 & 15.19 \\
\text{– with FCOVER} & 486.879 & 0.554 & – & 0.641 & 23.67 \\
\text{– with NDVI + FCOVER} & 566.648 & 1.936 & -0.350 & 0.707 & 14.49 \\
\end{tabular}
\caption{Empirically fitting coefficients (a–c) used non-linear estimation of the pine forest aboveground biomass}
\end{table}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Allometric relationship between the forest aboveground biomass (AGB) and normalised difference vegetation index (NDVI)}
\end{figure}
veyed pine forests in the region was 0.42. It seems that the biomass of the pine stands in the amount of 150 to 250 t·ha$^{-1}$ accumulates with an NDVI value ranging from 0.4 to 0.5.

The determined difference in the values of the vegetative indices obtained using satellite imagery is due to the nature of the reflected radiation, and also the structural peculiarities of the forest group, since the spectral-zonal data mostly reflect the parameters formed by the phototrophic part of the plants (Voronin 2006).

Use of the determined regression dependence allowed us to extrapolate the results of the assessment of the aboveground biomass of the stands of the Scots pine in the territory of the studied region with the development of the corresponding maps of their distribution. Based on the processing of the multiband images from Sentinel-2 satellite, we performed a transformation of the obtained data on the abovementioned biomass of the pine forests area to the forest massif in the structure of Dnipropetrovsk State Forestry (on the example of the Obuchiv forestry enterprise).

To determine the boundaries of the distinguished territory with a depiction of the corresponding gradation of the distribution of the AGB of the Scots pine stands, we used the shape-files of the forestry enterprises of the Dnipropetrovsk region (http://www.lisproekt.gov.ua/).

The map of the spatial distribution of the total AGB included four main classes with the following gradation of the biomass: to 20 (t·ha$^{-1}$); from 20 to 50 (t·ha$^{-1}$), 50 to 100 (t·ha$^{-1}$), 100 to 200 (t·ha$^{-1}$), and over 200 (t·ha$^{-1}$) (Figure 4).

The value of the biomass in the study area within the Obuchiv forestry enterprise varies from 0.1 to 309 t·ha$^{-1}$. The average value of the aboveground biomass of the pine stands in the areas covered with forest vegetation within the study region is 140 t·ha$^{-1}$. The total value of the aboveground live biomass of the pine stands in the Obuchiv forest management sub-unit, calculated on the basis of the geo-information data processing, is 195782.40 t·ha$^{-1}$, while the obtained biomass index, determined by the land-based measurements, is 210861.95 t·ha$^{-1}$, which is 7.2 % higher than the estimate using the

Figure 4. Forest aboveground biomass (AGB) distribution map within the Obuchiv forestry enterprise (Dnipropetrovsk region) from the allometric model using the normalised difference vegetation index (NDVI) from Sentinel-2
remote sensing data. Thus, the obtained differences in the estimation of the pine live biomass by direct (terrestrial) and indirect (using GIS technologies) methods are insignificant.

**DISCUSSION**

In this study, we performed an assessment of the AGB of pine forests for the Dnipropetrovsk region, the Steppe part of Ukraine, by combining the data of field surveys with the spectral and biophysical values obtained using satellite images from Sentinel-2. In the subsequent studies on the assessment of the bio-productive processes in the pine forests of the abovementioned region in the quantitative equivalent, based on the results of field surveys, we developed empirical regression models with a subsequent comparison of the spectral values obtained from the Sentinel-2 imagery.

Currently, the formation of an information basis for the distribution of the forest biomass on a global scale is becoming increasingly important, and conducting such studies by remote sensing methods is one of the most promising on national and regional levels.

In this case, the assessment of the AGB by remote sensing methods is usually based on regression models obtained through field studies (Muukkonen 2007; Shvidenko et al. 2007).

The existing information databases, formed on the basis of geographic information system (GIS) data, characterise both the current state of the forests and the dynamic transformations of the terrestrial ecosystem biomass over a certain period of time (Hansen et al. 2013; Liu et al. 2015).

So far, it has not been possible to develop a universal method for estimating the quantitative characteristics of the vegetation that could be applied directly, regardless of the type of satellite or aircraft sensors. With some exceptions (Santoro et al. 2011), it is usually an integral part of any remote sensing technology, in estimating the vegetation indices and biophysical parameters that are indirectly related to the AGB through an empirical relationship, to verify ground-based observations to take the peculiarity of the input data for specific regional conditions into account (Houghton et al. 2007).

Our results show a high level of correlation between the obtained characteristics of the field measurements and satellite images. This is consistent with the data obtained by Gallaun et al. (2017), who conducted large-scale studies of broadleaves and conifers of forest-forming groups within Europe.

The vegetation indices determined from the multispectral images of the Sentinel-2 satellite are considered to be the most sensitive to the photosynthetic parts of the plants (le Maire et al. 2008; Swatantran et al. 2011). The bands B4 (red) and B8 (near infrared) of the MSI Sentinel-2 with improved spatial resolution (10 m) are the most effective channels for distinguishing the vegetation indices (in this work, primarily for NDVI, FAPAR, LAI). The importance of these spectral channels in determining the indicators that are directly related to the quantitative characteristics of the stands, in particular, the aboveground live biomass and growing stock, has been tested and reported in scientific literature (Sibanda et al. 2015; Navarro et al. 2019).

However, it should be noted that there is a weak correlation between the vegetation indices presented in this paper (Table 1) and the biomass of the assimilation apparatus \((ABG_{crop})\). In this case, the data obtained are consistent with the results of Mcmorrow et al. (2011), whose publication states that NDVI has a weak relationship with the biomass in the context of a forest structure assessment. Mganga and Lyaruu (2015), studying the biomass of Tanzanian forests, also determined a low correlation coefficient between the photosynthetic part of the AGB and the NDVI, which reached a maximum value of 0.23. The low level of correlation between the photosynthetic component of the biomass and the values of the vegetation indices is primarily explained by the geometry (compositional structure) of the forest canopy, whose spectral profile may vary depending on its orientation and compositional structure.

On the other hand, our research has established a high correlation between the studied spectral characteristics and the non-photosynthetic part (stem biomass and total biomass) of the pine stands. This may be due to the high sensitivity of the optical sensors to changes in the structural composition within the vertical profile of the stand’s canopy (Foody et al. 2001).

In this study, in order to determine the biomass of the pine forests, we tested one-factor and two-factor non-linear regression models by integrating the results of the assessment of the aboveground forest biomass and with use of data obtained from the images from Sentinel-2. Single-, and multi-factor regression models are broadly used for the
assessment of the forest biomass (Lambert et al. 2005; Muukkonen et al. 2007).

The results revealed that the coefficients of determination of the single-factor ($R^2 = 0.698$) and two-factor ($R^2 = 0.707$) regression algorithm models with involvement of the NDVI and the NDVI combined with the FCOVER are quite high, and the models themselves – are the most optimum for predicting the AGB of the studied region.

Comparison of the obtained and predicted AGB values of the pine forests in the case of using the mentioned models is presented in Figure 5. According to the models, the range of the predicted biomass of the pine forests is narrower than the one obtained during the field observations, and the maximum value of the biomass was lower in both presented models. The mentioned models can be used for the assessment of the AGB of the pine forests according to the data from Sentinel-2 for areas with similar conditions. Furthermore, compared with other resources of aerospace imagery with free access, the data from Sentinel-2 have significant spatial characteristics and resolution both on regional and national levels. Such broad accessibility is extremely important, especially for developing countries, where resources available for buying such software and satellite images are extremely limited. The obtained results demonstrate the perspectives of the suggested method for the assessment of the aboveground biomass of the pine forests on
a regional level, because this method is expedient for this region, allowing the determination of the methods of management of forest resources in the territory of the Steppe. Moreover, in a global sense, mapping the spatial distribution of the aboveground biomass of forests has recently been a cause of serious concern, especially in developing countries, where such studies are not conducted to the full extent.

**CONCLUSION**

Data obtained with Sentinel-2 are considered an effective system of satellite imagery for improving the use of remote sensing for forest vegetation and studying the prediction data on the biomass for exploiting forests. The correlation analysis of the vegetative indices with the data on the biomass obtained in the field surveys revealed the strongest correlation relationships taking place between the aboveground biomass of the Scots pine and the NDVI and TVI. For the obtained biophysical parameters, the highest coefficients of correlation were observed for the FAPAR and FCOVER with the total aboveground biomass and trunk biomass. Modelling the biomass of the pine forests using vegetation indices, particularly NDVI, demonstrated high values in the determination coefficient parameters. Extending the models with arguments of biophysical parameters from Sentinel-2 (FAPAR, FCOVER) contributed to the accuracy of the developed models, emphasising the significance of their implementation.

The obtained spectral characteristics became the basis for the general assessment of the biomass of the pine forests of the Northern Steppe in Ukraine, and combining them with a map – the basis for the regional assessment of the distribution of the biomass of the pine stands. This study indicates the usefulness of an analysis of Sentinel-2 images which can be used for the assessment of the biomass of coniferous forests and mapping the distribution of the pine stands within the Dnipropetrovsk region.

**REFERENCES**


Received: March 23, 2020
Accepted: July 30, 2020