Technical Support to Forest Policy Development and National Forest Inventory Implementation

Concept Study: Remote Sensing Based National Forest Inventory Support

unique land use

Technical Support to Forest Policy Development and National Forest Inventory Implementation

Concept Study: Remote Sensing Based National Forest Inventory Support

Client
Ukrainian – German project on Sustainable Forestry Implementation (SFI)
Project-No.: W-UKR 21-01

Authors
Dr. Axel Weinreich
Maximilian Sperlich
Victor Myroniuk
Yuriy Farion

Date
March 31, 2023

Table of contents

List of tables

List of figures

List of abbreviations

BOA	Bottom of Atmosphere
CROBAS	Tree growth and CROwn BASe from carbon balance
DCP	Data Collection Plots
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
EO	Earth Observation
ESA	European Space Agency
FMP	Forest Management Planning
F-TEP	Forestry Thematic Exploitation Platform
GIS	Geographic Information System
HR	High Resolution
NFI	National Forest Inventory
PREBAS	PRELES+CROBAS
PRELES	PREdict Light-use efficiency, Evapotranspiration and Soil water
RS	Remote Sensing
RMSE	Root-Mean-Square Error
SFE	State Forest Enterprise
SFI	Sustainable Forestry Implementation
SRTM	Shuttle Radar Topography Mission
SWIR	Short Wave Infrared
TAO	Top Of Atmosphere
VHR	Very High Resolution

Summary

The national forest inventory (NFI) of the Ukraine started in 2019 and is planned as a strictly field based inventory. The use of remote sensing imagery was only meant for pre-clarification purposes, to determine, if the planned inventory plots are located within a forest or not.

However, as a result of the war in Ukraine, the ongoing national forest inventory cannot be conducted as planned, however. Aside from other impeding factors, such as the reduced staff or the unstable power grid, vast parts of the forest are inaccessible due to life threatening risks for the field teams caused by occupation, mines, or ongoing armed conflicts. Thus, several planned inventory plots cannot be visited by field teams. It was therefore decided that remote sensing methods should be applied as an auxiliary method to the terrestrial campaign, to estimate forest structural variables, specifically for the inaccessible forest areas, but also to obtain a more detailed estimations of the spatial distribution of the variables for the entire country of Ukraine.

The aim of this study is (i) to provide a decision basis on which remote sensing analysis method to choose, (ii) suggest necessary steps for reference data acquisition and preparation, and (iii) propose an implementation plan of the steps needed to estimate the forest structural variables using remote sensing data. Following these main ideas, the document is structured into three parts, subsequently depending on the outcome of the previous one. The outcome of each part will be produced in close correspondence with and in dependence on the agreement of the Ukrainian project partners:

Part I: General concept outline

- Overview of methods and data requirements for applying remote sensing data in a forest inventory context
- Recommendation of appropriate method and remote sensing data
- Definition of reference data requirements

Part II: Acquisition, description, and analysis of available reference and ancillary data

- Description of available reference and ancillary data
- Concept for reference data preparation

Part III: Implementation plan

Part I: General concept outline

Using remote sensing data analysis for forest inventory applications can benefit the process in multiple ways. Here, the focus is on predicting forest structure variables for inaccessible forest areas. The original plans for the field based NFI will not be substantially altered. The results from the remote sensing analysis aims to create additional valuable information for the NFI, such as wall-to-wall mapping of the forest structure variables, with an increased spatial distribution, for the entire country. This is especially relevant for areas inaccessible to the inventory teams.

2.1. Introduction

Remote sensing can provide several benefits for forest inventory applications where forests can be large and difficult to access. One of the key benefits is that it allows for the rapid and cost-effective acquisition of data over large areas. Remote sensing can also provide high-resolution data that can be used to accurately map and classify different types of vegetation. This can be helpful for understanding the structure and composition of a forest, for monitoring changes like forest cover, vitality, and health over time. Additionally, remote sensing can be useful for detecting and mitigating the impacts of natural hazards and other threats.

There are several benefits to using remote sensing to support field surveys in forest inventories. One of the key benefits is that remote sensing can provide a broad overview of the area being studied, which can be useful for identifying areas that may be worth further investigating in the field. If a regular grid of sampling plots is used, remote sensing can help determine which of the plots are located within the forest and can help assess the plot's accessibility. Remote sensing data can also provide detailed information on the structure and composition of the forest, which can be used to guide field surveys and to ensure that data is collected in a consistent and systematic manner. This can help to improve the accuracy and precision of the inventory.

Remote sensing data from satellites can be used to distinguish different types of vegetation within a forest based on their unique spectral signatures, which are light wavelengths that are reflected or emitted by different types of vegetation. Different types of vegetation have different spectral characteristics due to differences in their chemical composition, structure, and photosynthetic pigments which arise due to their varying foliage, health, vitality etc. In addition to these spectral characteristics, satellite data can also be used to identify other characteristics of vegetation, such as canopy structure, canopy cover, tree height, tree density, and species composition. The specific variables that can be estimated using remote sensing will depend on the characteristics of the forest being studied and the type & resolution of the satellite data available.

The accuracy of forest structure data estimated using remote sensing can vary depending on several factors, including the type of remote sensing technology used, the resolution of the data, and the type of forest being studied. In general, remote sensing can provide highly accurate data on forest structure when Very High Resolution (VHR) & High Resolution (HR) imagery is used and when the forest being studied has relatively homogenous characteristics. However, the accuracy of remote sensing data can decrease when dealing with forests that have complex structures or when the resolution of the data is not sufficient. Additionally, the accuracy of remote sensing data can be affected by factors such as area coverage, cloud cover and rapid changes in the forest over time.

2.2. Data requirements and methodological options for RS based NFI support

2.2.1. Remote sensing platforms and sensors

For the remote sensing surveys of forests, three main platforms are distinguished: Uncrewed aerial vehicles (UAV), aircrafts and satellites (compare Figure 1). These platforms can be equipped with various sensors which acquire data with different spectral and temporal characteristics.

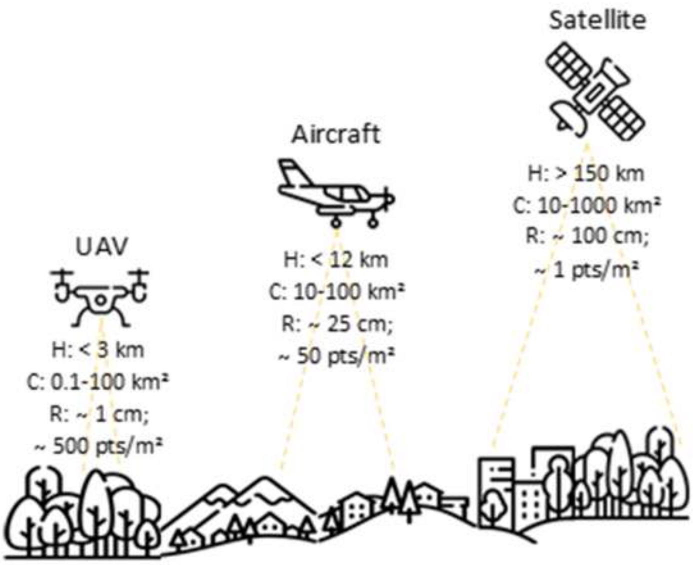

Figure 1 Comparison of flight height (H), image area coverage (footprint) (C) and spatial resolution (R; both ground sampling distance (GSD) for multispectral images and point density for LiDAR point clouds) of remote sensing platforms UAV, Aircraft and Satellite

Sensors that record and measure electromagnetic energy fall within two groups: passive sensors and active sensors. Passive sensors such as cameras, multispectral or thermal scanners, mainly rely on the sun as an external source of energy and are hence, dependent on the sun's illumination for data acquisition. Active sensors, such as light detection and ranging (LiDAR) or radio detection and ranging (RADAR), provide their own energy and are hence, not dependent on the sun's illumination for data acquisition.

Deployed passive sensors in the context of forest inventories are usually multispectral sensors, recording multiple spectral bands (wavelengths of electromagnetic energy) in the range of visible light and infrared. Multispectral sensors can measure up to 15 spectral bands (e.g. compare Worldsview-2 in Figure 2) and are distinguished from hyperspectral sensors, which can provide more than 100 bands (e.g. compare EO-1 Hyperion in Figure 2).

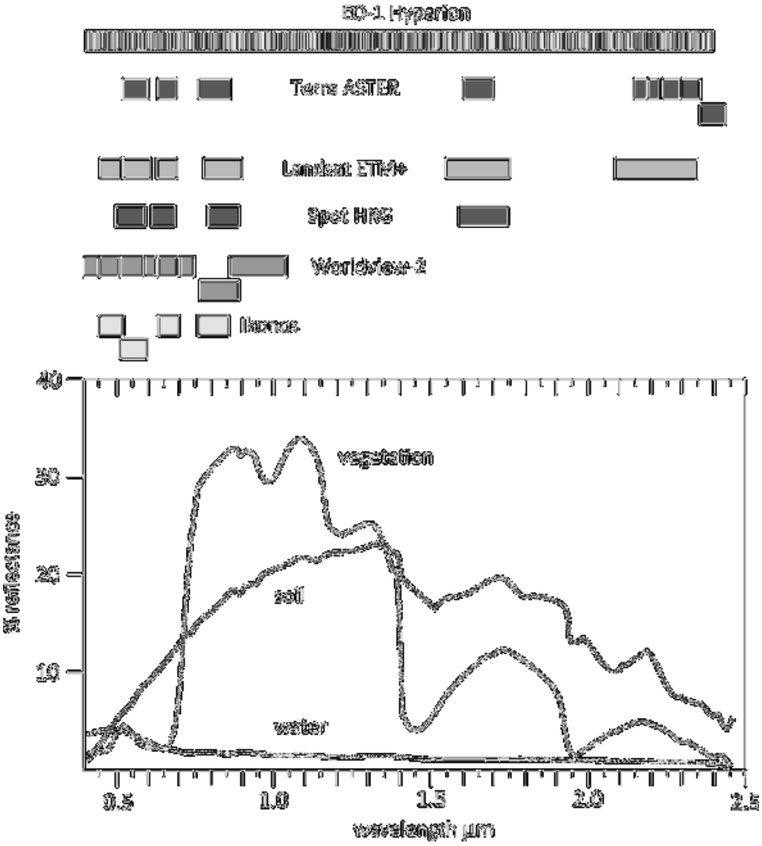

Figure 2 Spectral reflectance characteristics (below) in the visible (0.38 – 0.75 µm), near- (0.75 - 1.4 µm), and short wave infrared (1.4 - 3 µm) range for three common land cover types shown together with locations of spectral bands of common sensors (above), some of which have additional bands extending further into the infrared range not depicted here (Richards, 2013, S. 11)

Of the active sensors, LiDAR which stands for Light Detection and Ranging, is mainly used to derive forest metrics. As an active sensor system, a LiDAR system generates its own energy (light) or laser, which is emitted in pulses (burst of light energy) and records the returned signals. The time it takes the signal to return to the sensor is used to calculate the distance travelled. The LiDAR system uses the speed of light to calculate the distance between the top of the object and the platform used. Based on the orientation of the platform and the distance the light travels, a 3D point cloud is generated, with each point representing a strength of the returned signal. A distinguishing factor between LiDAR systems is the way the return signal is recorded (compare Figure 3). Discrete Return LiDAR systems identify peaks in the returned signal and record these points as returns. The number of returns recorded per pulse usually lies between 1 and 4, mainly recording the top of the canopy and the ground. Full Waveform LiDAR systems capture the distribution of light energy returned to the sensor and display them as waveforms (figure 3). While the latter requires complex processing, the information gained in comparison to discrete return systems increases the possibility of tree and shrubs detection in lower forest layers and improves the possibility of tree species classification.

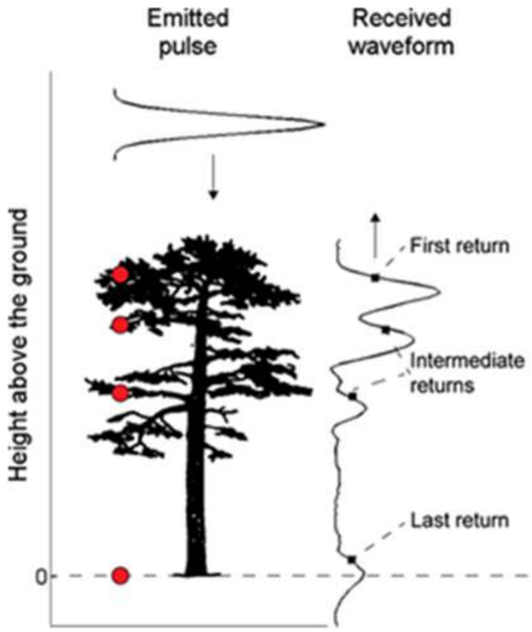

Figure 3 Returned discrete signal (red dots) and waveform of an emitted pulse from a LiDAR system (Lindberg & Holmgren, 2017)

The decision on the platform used and sensor combination depends on the required spatial and temporal resolution and the cost of data acquisition. While UAV or aircrafts equipped with hyperspectral or LiDAR sensors facilitate the analysis of the data on a single tree level, using these in a NFI context for a wall-to-wall mapping of the country will lead to relatively high costs and acquisition time compared to satellite imagery. Moreover, considering the ongoing war in Ukraine, collecting data over the conflict areas using UAV or aircrafts is not recommended due to the risk of the platforms being attacked leading to a loss of equipment or even causing harm to the pilots. Instead of covering the entire country, UAV- or aircraft-based data could also be collected using a sampling-based approach, like a field campaign. As this study's main purpose is to offer suggestions on how to derive estimates for the now inaccessible forest areas using remote sensing data, and without substantially altering the original NFI concept, the following chapters will thus focus on the use of satellite imagery.

Other active sensors that are often used for remote sensing are Synthetic Aperture Radar (SAR) sensors that use microwaves to create 2D images of the Earth. SAR sensors on board satellites provide high resolution, daylight and weather independent data that can be used for a wide range of applications. The weather and daylight independence makes SAR sensors the most suited for monitoring, hazard mapping and other applications that require images independent of cloud cover. Until now, built on the principal of a RADAR (Range Detection and Ranging), it involves sending a signal of high power to a target and measuring the strength (amplitude) and time (phase) of the returned echoes from the

Earth’s surface. The transmitted signal interacts with the Earth surface and only a portion of it is backscattered back to the sensor. SAR sensors use the forward motion of their platforms (aircrafts or satellites) to synthesize larger apertures thereby being able to provide higher resolution images. Unlike optical images, visualizing raw SAR data does not give any useful information about the imaged scene and signal processing steps are required to produce a 2D grayscale image (Moreira, 2013). While SAR amplitude can provide information about the composition and characteristics of the objects in the area imaged, SAR phase can provide information about ground and terrain deformation. SAR sensors operate in the microwave range of the electromagnetic spectrum of light between 0.3 Ghz and 40 Ghz and are hence, not hindered by clouds and other atmospheric phenomena. Commonly used SAR wavelengths are X band (3.1 cm), C band (5.6 cm), L band (23 cm) and S band (8-15 cm).

SAR data are usually amplitude data and phase data. SAR amplitude data consist of SAR backscattering information displayed in 2D images where the SAR backscatter from objects on the Earth’s surface depend on the different geometric properties and chemical properties of those objects. In the context of forest applications, SAR backscatter from vegetation will depend on vegetation content, soil moisture, foliage, and dielectric properties of the targets. The amount the transmitted signals will penetrate forest canopies will depend on the wavelength used. For forest inventory applications, SAR amplitude has proven to be beneficial as it can provide information about Above Ground Biomass and for monitoring vegetation. SAR sensors are also very sensitive to water and hence can be effective for soil moisture analysis. SAR polarimetry is also commonly used for land cover application which requires an extensive understanding of the SAR signal and how it interacts with the objects in the image scene and requires complicated processing steps. The use of SAR for forest and biomass applications has been reported extensively in the NASA SAR handbook for forest monitoring and Biomass (SARHB_FullRes.pdf (servirglobal.net))

2.2.2. Satellite data requirements

Remote sensing using multispectral & SAR satellites provide a cost and time effective alternative to LiDAR data and field surveys especially bearing in mind the current instability in Ukraine. Satellite data can be acquired by passive and active sensors i.e. sensors that measure the amount of sun’s radiation reflected back to them, or those that generate their own power and measure the strength of the return signals. Since the launch of ERS-1 in 1972 which marked the beginning of the Earth Observation era, many satellites, both active and passive have been launched for a variety of applications including forest and land cover monitoring. While passive sensors on board satellites are dependent on daylight conditions and can be hindered by the presence of clouds, haze and other weather phenomena, active sensors such as RADAR, LiDAR and SAR work independent of weather and daylight conditions.

Multispectral Satellites

For multispectral satellite data to be useful for forest inventory applications, it should have high spectral, temporal, and spatial resolutions and must not be hindered by weather conditions and clouds.

- Spectral resolution refers to the number of different wavelengths of light that a satellite sensor can detect and measure. For forest monitoring, it is important to be able to measure light reflected in a wide range of wavelengths (spectral bands) so that a wide range of features within the forest, such as different types of vegetation, bare soil, and water bodies can be detected.
- Temporal resolution refers to the how often the satellite acquires data. For forest inventory, to estimate forest structural data, it is important to have a high temporal resolution so that imagery can be acquired for a point in time as close as possible to the date of the reference data. Additionally, through the analysis of time series (many images acquired over a large time period), tree species (compositions) in forests are more easily distinguished.
- Spatial resolution refers to the size of the smallest feature/object on the ground (pixel) that a satellite can detect. For forest monitoring, it is important to have a high spatial resolution so that small features within the forest, such as individual trees, can be accurately detected and monitored.
- In addition to these technical requirements, satellite data for forest monitoring should also be collected under consistent and stable atmospheric conditions to ensure the accuracy and reliability of the data. Data collected with large amount of cloud cover or other atmospheric phenomena are not suitable for such a study.

As shown by Astola, Häme, Sirro, Molinier, & Kilpi (2019) in the comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction, freely available imagery can already meet the above-mentioned requirements. Sentinel-2 and Landsat 8 (or the recently launched Landsat 9) are two satellites with multispectral sensors on board and hence, the next section focuses only on multispectral satellite imagery.

SAR satellites

There are currently many SAR satellites that provide open access to their data. For forest inventory applications, two main groups of characteristics must be considered: sensor and target characteristics. Sensor characteristics include SAR wavelength, SAR polarization. Incidence angle of the SAR signal and the orbit direction of the satellite. The target characteristics are forest biomass and structural complexity.

- SAR Wavelength: The wavelength of the sensor determines the penetration depth of the transmitted signal into the vegetation or ground layer of the Earth. The longer the wavelength, the deeper the penetration can be. In forest applications,

- X-band (3.1 cm) SAR sensors will mainly measure the backscatter returned from the top layer of the canopy,
- C band (5.6 cm) SAR sensors will measure the backscatter from the top and other branches and stems of the trees
- L-Band (23 cm) SAR sensors will be able to measure the backscatter from the upper vegetation through to the ground surface

Figure 4 Pentration of SAR signals according to the SAR wavelength used (modified from NASA SAR handbook)

- SAR polarization: SAR data can collect signals in different polarizations by collecting and controlling the polarization during both transmission and reception. Polarization refers to the orientation of the plane of the transmitted SAR signal, Horizontal (H) and vertical (V) polarizations being the most used. The orientation of the transmitted SAR signal may get altered due to interactions with different elements on the ground such as vegetation, water, soil etc.. Hence, the ability of the SAR sensor to measure this change in polarization can be useful when trying to understand tree species, crop type, vegetation vitality etc. Sensors like Sentinel-1 are only capable of sending signals in one polarization, either H or V and are able to receive signals in both directions thereby providing dual polarized images in VV (Vertical sent, Vertical received) and VH (Vertical sent, Horizontal received) or HH and HV. Other SAR satellites like TerraSAR-X or ALOS-PALSAR can send in both polarizations and receiving both thereby providing quad polarized data in HH, VV, HV and VH. By a number of complex decomposition methods, information about tree, crop species may be extracted using a technique called SAR polarimetry.
- SAR incidence angle: The incidence angle is the angle between the SAR sensor and the going and the surface normal of the illuminated Earth's surface. SAR backscatter is strongly influenced by this angle and this angle determines scattering in the crown layers, trunks and interactions with the ground. Objects tilted towards the sensor tend to return a higher backscatter than those tilted away from the sensor.
- SAR orbit: SAR sensors are side looking and not Nadir pointing (like optical sensors). The look direction of the SAR refers to the direction the SAR antenna points to when emitting and receiving. Most commonly, SAR sensors tend to be right looking sensors which means that the part of the Earth being imaged by them will be to their right. When SAR satellites are in Ascending orbits, they are flying from the south pole to the north pole

and imaging areas to their right, whereas the same area may be imaged when the satellite is in its descending orbit (north pole to the south. The objects tilted to the sensor during ascending acquisitions may return a stronger backscatter than during descending acquisitions.

2.2.3. Feature requirements

Most methods for estimating forest structure variables from remote sensing imagery rely on machine learning techniques. In machine learning, a property of the object, which is being analyzed, in our case the remote sensing data, is called a feature. There are several features that can be extracted from satellite imagery for the classification of forest types and forest structural variables. These features can be used to represent the spectral, spatial, and temporal characteristics of the trees, and can be used to train machine learning algorithms to distinguish between the spectral and spatial characteristics. The features relevant for forest related remote sensing data analysis can be grouped into three categories:

- Spectral features: These are features that represent the spectral characteristics of the trees, such as the reflectance of light at different wavelengths, the normalized difference vegetation index (NDVI), and the greenness index.
- Spatial features: These are features that represent the spatial characteristics of the trees, such as the size, shape, and texture of the trees, as well as the arrangement of the trees within the forest.
- Temporal features: These are features that represent the temporal characteristics of the trees, such as the phenological stages of the trees (e.g., leaf emergence, leaf senescence) and the seasonal changes in the trees.

In addition to these features derived from the main satellite imagery, other features, such as topographic features (e.g., elevation, slope), meteorological features (e.g., temperature, humidity), and land use features (e.g., soil type, land cover) can also be extracted from additional satellite data and used to improve the accuracy of the analysis.

2.2.4. Methods for remote sensing estimation of forest structure variables

Apply published methods

The estimation of forest structure variables from remote sensing data is an actively researched field with new or updated methods being published regularly. Especially advances in the field of machine learning, specifically artificial neural networks, and new sources of satellite imagery are the main driving forces behind the innovative developments in this field.

Many of the studies yield very promising results. However, the studies are often confined to specific study areas, and therefore specific forest types, and are not tested for more general applications. Any method adopted from publications needs to be checked for its applicability and modified to fit the requirements specific to the forest types in Ukraine. Ideally, methods should be selected, which were developed specifically for Ukraine forests, such as published by Myroniuk, Bell, Gregory, Vasylyshyn, & Bilous (2022).

While describing the developed methods, many published papers do not supply access to the code used. To apply the described methods, one would need to reverse engineer them. This can be achieved either by writing the described code in a local development environment, usually based on R or Python, involves a lot of time and effort and the success of the reverse engineering depends on how well the methodology was described in the publication. Processing the data for an entire country comes at a high local processing and storage cost, which might limit this option, making a cloud-based solution, such as using the Google Earth Engine, preferable.

To reduce the amount of time and cost spent on writing code when reproducing the methods described in the publications, they can also be implemented using no-code services for remote sensing data analysis, such as Earth Blox. However, these services are limited in the available function blocks and thus might not bring the necessary functionality to reproduce complex workflows, such as deep neural networks with very specific hidden layers.

Use existing services

As often is the case in research, the main focus is on developing and testing novel methods rather than producing industry ready tools, which is why only few ready-made tools for the estimation of forest structural variables from remote sensing data are available. One of which, proposed as the most adequate approach and described in more detail below, is *Forest Flux* (https://cordis.europa.eu/project/id/821860). The service resulted from an EU funded project, to which unique contributed and was able to gather experience in multiple pilot study areas. It is also under further development as part of the Forest Carbon Monitoring project.

Forest Flux provides cloud-based services for the prediction of forest structural variables and carbon assimilation via freely available high-resolution (HR) Earth Observation (EO) data (Häme, et al., 2021). Outputs are generated in the form of digital maps and statistical information. Copernicus satellite images provided with no cost by the European Space Agency (ESA) are the main source for EO data. *Forest Flux* follows a holistic approach implemented in a single processing chain. The *Forest Flux* services are implemented on the Forestry Thematic Exploitation Platform (F-TEP) and is fully functional since 2022 (Häme, et al., 2022). F-TEP offers a range of services to support commercial, governmental and research institutions in the forestry sector. Via F-TEP, versatile processing approaches are accessible online to generate forest information products for forest management and monitoring.

Chapter 2.3 describes the processing steps for *Forest Flux*. The steps are general enough to be followed if the platform is decided against.

2.3. Recommended processes

In the following sections the recommended processing chain for estimating forest structural variables via RS is described and the relevant subsystems to be modified are outlined (based on the recommendation to utilize and adapt the *Forest Flux* services available on the F-TEP platform). Hereafter, we focus only on using multispectral data from Sentinel-2 and VHR data from other missions such as PlanetScope. For a more detailed description of the approaches, underlying algorithms, and their application, please refer to the final report of *Forest Flux* (Häme, et al., 2022).

2.3.1. Overall processing chain

A schematic overview of the entire processing chain to be implemented is shown in Figure 5. The main processing chain constitutes of three subsystems:

1. EO data pre-processing (HR and very high resolution (VHR) RS-images),
2. Extraction/computation of forest structural variables, and
3. Compilation of the expected end products.

To derive the desired map products (output), target area specific input and reference data are required.

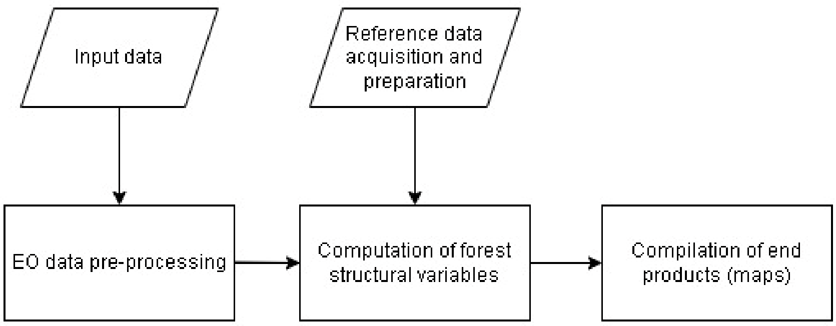

Figure 5 Schematic overview of the processing chain of the applied Forest Flux services (Häme, et al., 2022)

In the **"EO data pre-processing"** subsystem EO data sets (Sentinel-2) are processed to the level required by the next processing subsystem. The default pixel size is 10 m, which is the highest spatial resolution of Sentinel-2 data. Ancillary datasets will also be used here. Target specific input data are required to be input into this subsystem. The subsystem "**Computation of forest structural variables**" produces the forest inventory product layers. After this subsystem has completed processing, the product layers are ready for final product compilation.

This subsystem also includes **"reference data acquisition and preparation"** from VHR satellite images, aerial images or open VVHR-image data like Google Maps, Bing Maps etc..

The third subsystem "**Compilation of end products**" compiles the product layers in the final formats for delivery. It includes styling, formatting or color coding based on the defined look up tables and format conversions according to the national requirements.

The main expected end products (output maps) for the RS based Forest Inventory for the Ukraine have been derived from the list of results defined for the terrestrial NFI. The expected products are listed in Table 1 along with possible sources of reference data according to level of availability. Reference data can be obtained on three levels:

- Level 1: Estimates of attributes based on visual interpretation of very high resolution (VHR) imagery (Aerial images, Google Earth or Bing Maps images). The Level 1 will always be necessary to improve the quality of results or to verify and delineate reference data from Level 2 and Level 3 sources.
- Level 2: Reference polygons containing ground and stand based forest structural attributes from recent (< 5 years) Forest Management Plan (FMP) data.
- Level 3: List of ground attributes sampled in the field from NFI plots (DCPs) Administrative data (roads, rivers)

Table 1 Overview of the expected end products and sources of reference data to train the RS analysis model

Expected end products - resulting map layers	Targeted use of reference data		
	Level 1: Attributes sampled by visual interpretation of VHR data	**Level 2: Attributes sampled from actual (< 5 years old) FMP data + VHR images**	**Level 3: List of ground attributes necessary for RS mapping (ground truthing)**
Forest mask (forest cover map at 10 m spatial resolution)	Forest / non-forest	Forest /non-forest	Forest /non-forest
Forest Types	Forest Types	Forest Type & Tree species composition from stand level improved by analyzing VHR images	Forest Type & Forest species composition on plot level
Tree species distributions (raster maps of tree species presence and distribution at 10 m spatial resolution)	Partly aplicable – dominant species	Tree species % from stand level data improved by analyzing VHR images	Forest tree species % on plot level
Mean tree height [m]	NA	Mean tree height from stand level data improved by analyzing VHR images	Mean tree height on plot level
Mean DBH [cm]	NA	See above	See above
Density – number of trees [N/ha]	NA	See above	See above
Basal area [m²/ha]	NA	See above	See above
Mean age [years]	NA	See above	See above
Mean growing stock [m³/ha]	NA	Growing stock /ha from stand level improved by analyzing VHR images	Growing stock /ha on plot level
Increment / ha	NA	Increment /ha from stand level data from stand level improved by analyzing VHR images	Increment /ha on plot level
Above and below ground biomass / ha	NA	Above and below ground biomass /ha from stand level data from stand level improved by analyzing VHR images	Above and below ground biomass/ha on plot level
Carbon stock [t/ha]	NA	See above	See above

2.3.2. Input data

2.3.2.1. High resolution satellite imagery – Sentinel-2

Input data consist of HR EO data and ancillary data. The main recommended data source of EO are Copernicus Sentinel-2 multispectral images which cover 13 spectral bands. Four of these bands have a 10 m spatial resolution (blue, green, red and near infrared), six bands have a 20 m resolution, and three bands have a 60 m spatial resolution. The 10 m bands serve as the primary data source, but other bands can be applied as needed. In addition to the 10 m bands, the main 20 m bands to be used are the short-wave infrared (SWIR) bands and the red edge band 5 (Table 2). Sentinel-2 data are delivered as 100 km by 100 km tiles and are available at two processing levels: Level-1C orthoimages and Level-2A atmospherically corrected orthoimages. This means that Level-1C images contain top of atmosphere (TOA) reflectance values and Level-2A data bottom of atmosphere (BOA) reflectance values. Sentinel-2 Level-2A data will constitute the main data source. From the scene classification maps, cloud and cloud shadow mask are compiled which will be based on data with a high level of processing. Table 2 shows the applicable Sentinel-2A spectral bands with their respective spatial resolution.

Table 2 Applicable Sentinel-2A bands.

Band		Central wavelength (µm)	Spatial resolution (m)
B2	Blue	0.492	10
B3	Green	0.560	10
B4	Red	0.665	10
B5	Vegetation Red Edge	0.704	20
B6	Vegetation Red Edge	0.741	20
B7	Vegetation Red Edge	0.783	20
B8	NIR	0.833	10
B8A	Vegetation Red Edge	0.865	20
B11	SWIR	1.614	20
B12	SWIR	2.202	20

2.3.2.2. Ancillary data

Ancillary data includes all other RS-data sets needed for production of the output maps. The main ancillary data needed for EO data pre-processing are: a) digital elevation models (DEM) and b) topographic maps. It is not expected that any orthorectification is required for the HR satellite images. For processing of VHR images orthorectification with DEM is however often needed. So, the optimal available DEM resources area asked now to be provided for the RS-NFI from Ukrainian partners. In case that no better suitable national DEM is available, the global DEM derived by the Shuttle Radar Topography Mission (SRTM) (https://www2.jpl.nasa.gov/srtm/) provides a good global alternative.

Further, the global 3D elevation model from the TanDEM-X mission (https://geoservice.dlr.de/web/dataguide/tdm90/) is freely available for scientific use. The role of topographic maps is mainly quality assurance of data sets, particularly when VHR RS-images are used. In addition to map data, e.g. Google Maps, Google Earth, Bing Maps and Open Street Map data can be used if official national data are not available.

2.3.3. Reference data

The main data source for the subsystem 2) computation of forest structural variables (cf. Figure 5) are the of pre-processed HR satellite images as outputs of sub-system 1). In addition, field reference data of the variables to be estimated is required. As the temporal, geometric and thematic accuracy of the field reference data affects the accuracy of the respective product, reference data should be highly accurate and available for each target variable. Consequently, the availability of reference data may limit the selection of output products. The reference data can be geocoded sample plot data, stand vector data derived from FMP or other maps. Where required, computation and format conversions will be applied.

The variables necessary to train a remote sensing data classifier for forest structural variables include information both on the site and the forest structure. It is important, that the reference data is provided in a standardized matter, including a short description on the collection process and context (e.g., FMP, NFI plots). The following Table 3 lists the variables together with the expected format/unit and a short description with the most important ones highlighted in bold.

Table 3 Forest structure variable reference data

Variable name	Format/ Unit	Description
Year of data	Integer	Year of data acquisition
Data type	Categorical variable	Type of reference data, integer as category: 1. circular plot 2. stand 3. relascope plot 4. other
Plot center east	m	X-coordinate in UTM coordinate reference system The EPSG code of the UTM coordinate reference system should be included in the description.
Plot center north	m	Y-coordinate in UTM coordinate reference system The EPSG code of the UTM coordinate reference system should be included in the description.
Area of plot or stand	m^2	If multiple-radius plot, write area corresponding to the largest radius. Value -1 for a relascope plot.

Variable name	Format/ Unit	Description
Soil type	Categorical variable	Specific soil type definition 1. mineral 2. organic Legend for categories should be given. If no detailed information is available (e.g., from soil maps), at least, it should be distinguished between mineral and organic soil (peat).
Site index	m	h100: Height (hypothetical) of forest in meters at the age of 100 years with ideal tree species If h100 is not available, similar site index classes can be used.
Layers	Categorical variable	Number of vertical layers in the forest, integer as category: 1. even aged 2. two layers 3. three layers 4. uneven aged If defined units cannot be reported, similar classes can be used.
Forest type	Categorical variable	Specific forest type definition, integer as category: If no detailed or up to date information is available in form of maps, at least, the following ecoregion categories (Dinerstein, et al., 2017) could be applied: • Central European mixed forests • Crimean Submediterranean forest complex • East European forest steppe • Pannonian mixed forests • Carpathian montane conifer forests • Pontic steppe Legend for categories should be given. A detailed map of the forest types for the entire Ukraine may also be generated using RS-based methods.
Tree species distribution:		
Species 1	%	Proportion of tree species 1 by basal area including all layers Define tree species and coding with the delivery of reference data. Species proportions should sum up to 100 %.
Species 2	%	Proportion of tree species 2 by basal area including all layers. Define tree species and coding with the delivery of reference data. Species proportions should sum up to 100 %.
Species n	%	Proportion of tree species n by basal area including all layers Define tree species and coding with the delivery of reference data. Species proportions should sum up to 100 %.
Mean height	m	Basal area weighted mean height including all species and layers
Mean diameter	cm	Basal area weighted mean diameter at breast height including all species and layers
Density	n/ha	Number of stems per hectare including all tree species and layers
Stem basal area	m^2/ha	Basal area of stems per hectare for the whole plot including all species and layers
Mean age	years	Stand or plot level mean for main layer or dominant species
Growing stock volume	m^3/ha	Growing stock volume per hectare including all species and layers

Variable name	Format/ Unit	Description
Increment	m³/ha/a	Increment of stock volume per hectare and year including all species and layers
Biomass	t dry/ha	AGB and BGB per hectare including all species and layers
Carbon stock	t/ha	C-content of biomass or CO_2 equivalent per hectare
Crown closure	%	Projected area of tree crowns as the proportion of the plot or stand area

2.3.4. Pre-processing of satellite images

The objective of the **pre-processing subsystem** is to produce Sentinel-2 satellite images that are ready for processing and analysis. The applied methods of the recommended *Forest Flux* service for the processing chain produces image data in a format that is directly compatible with the processing chain, along with masks that identify pixels that are not useful for further processing (i.e. cloud and cloud shadows mask). As the processing chain of the *Forest Flux* service use ER Mapper image as internal format, primarily ER Mapper images will be produced as output. In case the cloud and cloud shadow mask needs to be revised manually, an additional GeoTiff file is produced which can be processed in any local GIS software.

The HR Sentinel-2 images are selected interactively using the F-TEP search tool. The selected image is then taken as input by the EnvimonS2 tool (Figure 6), implemented in F-TEP. The EnvimonS2 service unpacks the Sentinel-2 image from zip archive to ER Mapper format (.ers). Available pixel size options are 10 m, 20 m and 60 m. For Level 2A images, the cloud mask is extracted from the 20 m resolution scene classification image produced by the Level 2A processing. The output is a GeoTiff file with pixel values of 255 for the classes saturated/defective, cloud shadows, clouds (low/medium/high probability) and cirrus and value 0 for other classes. The output of the pre-processing step is a zip file that contains the Sentinel-2 image bands in the selected pixel size and one cloud mask. A flow chart of the HR data pre-processing is provided in Figure 6.

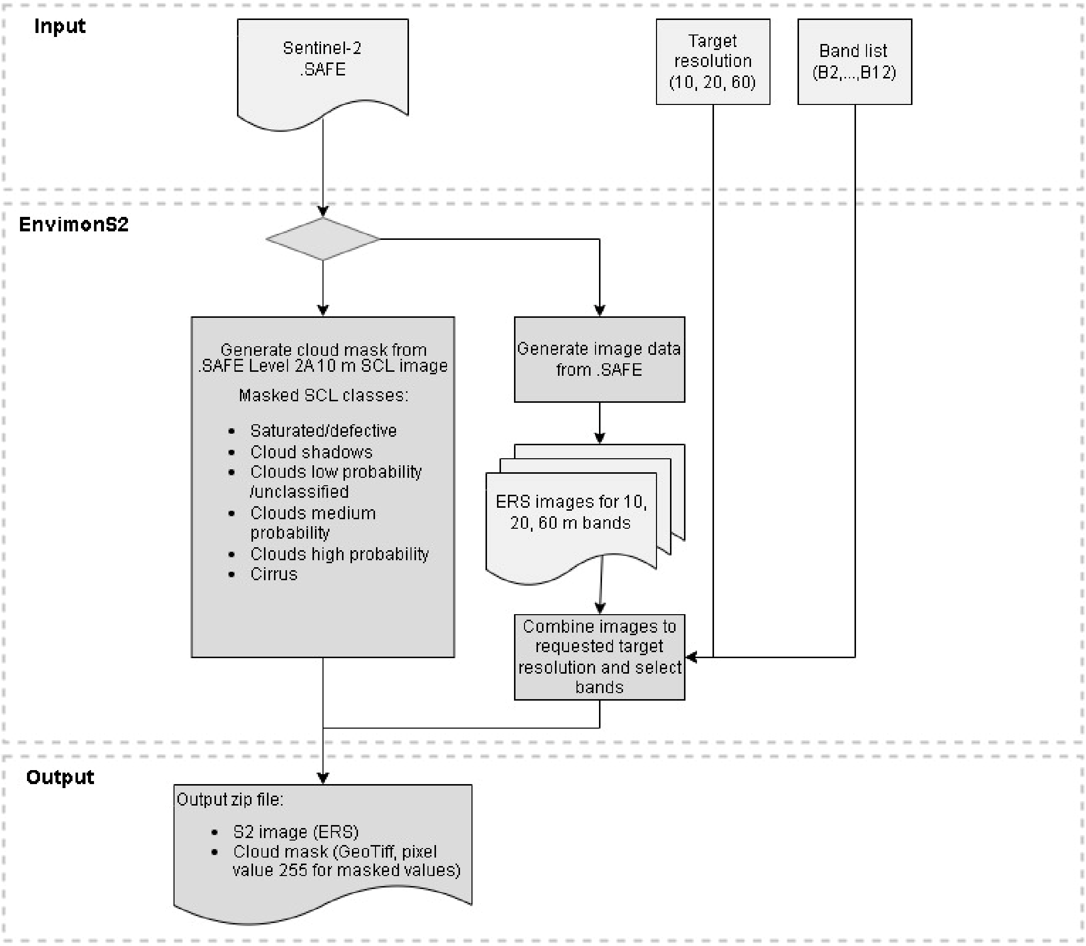

Figure 6 Flow chart for the pre-processing of HR satellite images (Häme, et al., 2022)

As alternative we can implement the pre-processing steps using the Sen2Cor1 processing tool provided by the ESA. Sen2Cor also produces indicators for cloud and snow probabilities and a cloud and cloud shadow mask. The mask and the cloud probabilities are combined into an initial cloud mask candidate, which is then interactively evaluated in a GIS program and manually fine-tuned. For visual validation, false and true color images are compiled from every Sentinel-2 image.

2.3.4.1. Pre-processing of VHR images

The level 1 source of reference data, which is used as additional source of information in combination with level 2 and level 3 data is especially relevant also for all inaccessible areas with insufficient level 2 and level 3 reference data. In this preprocessing step the use of these VHR remote sensing images like aerial images, sat-images, open data images from Google or Bing Maps are described. The VHR images are used by visual interpretation of targeted attributes like forest / non forest or forest type, age classes etc. and allow to select and delineate homogenous reference polygons from level 2 and level 3 sources.

[1] https://step.esa.int/main/third-party-plugins-2/sen2cor/

The aim of pre-processing is to create a dataset from which reference data can be extracted for both land cover classification and change monitoring. This requires that the images are correctly orthorectified and that they are suitable for visual interpretation.

Figure 7 provides an overview of the related processing chain from VHR images to the respective reference data.

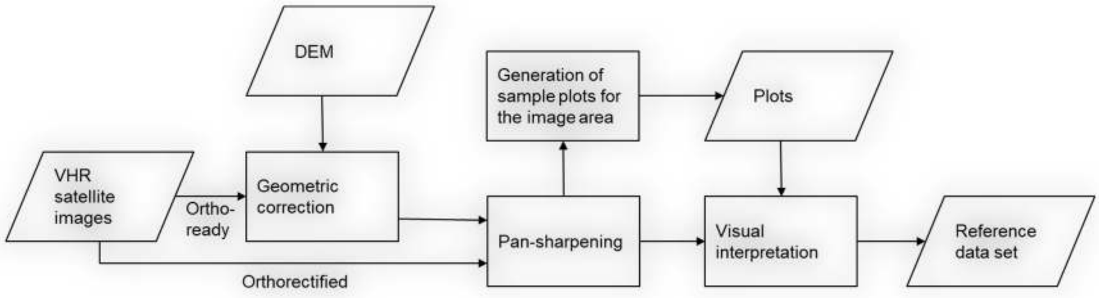

Figure 7 Pre-processing chain for VHR satellite images (Häme, et al., 2022)

For images that are received as orthorectified, true and false color images is computed using pan-sharpening. For images that are received as ortho-ready, ortho-rectification using available DEM and nearest neighbor resampling is applied before pan-sharpening. After pre-processing, reference data are collected via visual interpretation. Attributes are recorded remotely for each plot according to an expert assessment using VHR images and GIS. This process is facilitated by means of a support tool which allows rapid navigation between and evaluation of samples and easy recording of the respective attribute values.

2.3.4.2. Output of EO data pre-processing

The outputs from the subsystem EO data pre-processing consist of:

- Pre-processed wall-to-wall images
- True color mosaic from wall-to-wall data
- False color mosaic from wall-to-wall data
- True color mosaic from VHR images
- False color mosaic from VHR images
- Reference data sets collected from the VHR images

The pixel size of the wall-to-wall images is in the case of Sentinel-2 is 10 meters and for VHR satellite images the pixel size of the respective panchromatic band of the image.

2.3.5. Computation of forest structural variables

The main components of the subsystem 2) Computation of forest structural variables including land cover classification, are algorithms of the *Probability chain* software (*Proba_cluster*, *Proba_model, Proba_estimates*) that define a non-linear mapping from a multi-dimensional input space (reflectance values) to a multi-dimensional output variable space classification, using EO and reference data. The *Probability chain* software has been provided based on a development of VTT Technical Research Centre of Finland and is still maintained by VTT. The computation chain produces probability estimates for all forest variables for which reference information is available. In addition to the *Probability chain* software tools, the chain includes a post processing step to apply masks to the estimate images and in case of category classification, compile the final classification from the continuous estimate layers. Figure 8 provides a schematic overview of the respective subsystem.

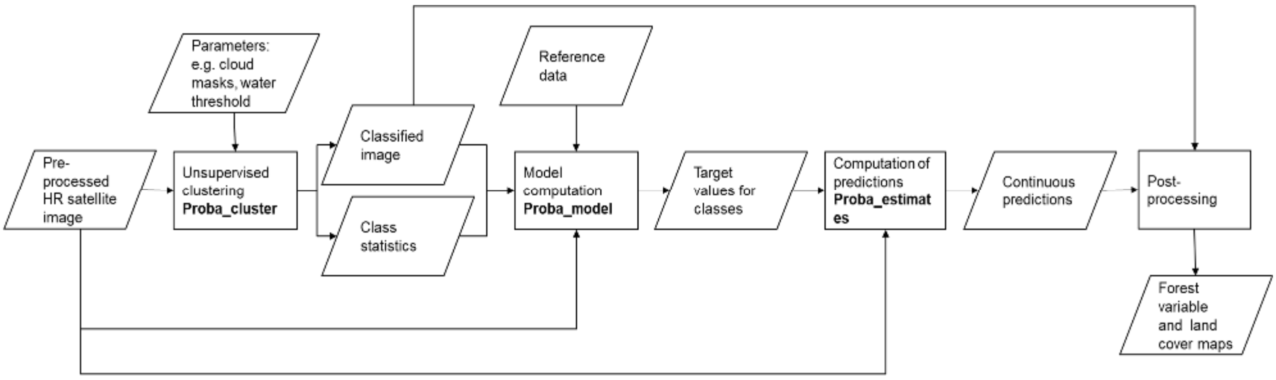

Figure 8 Computation chain of forest variable prediction and land cover classification (Häme, et al., 2022)

The probability estimation is based on the algorithms described in Häme et al., 2001. First, an unsupervised clustering is performed in the image feature space using a k-means algorithm via the *Proba_cluster* tool. Consequently, feature classes are produced (typically between 30 – 50), along with the respective statistics and normal distribution parameters for each class. For each feature class, the values of the target variables are assigned using the reference data and the *Proba_model* tool. The *Proba_model* can use reference data in stand, plot, or image format. In the final step of the probability estimation, the variable predictions are computed as weighted sum of target values of all feature classes with the *Proba_estimates* tool. The assigned weights correspond to the probabilities of a pixel to belong to each feature class. The basic principle of the probability estimation is that all variables are treated as continuous. However, the method can be used also for land cover classification tasks. When the probability estimation method is applied for classification, the reference data category classes are first converted to continuous variables, resulting in e.g., pixel values of a forest and non-forest reference map of 0 or 100%. At the final stage, the category classification is produced from the continuous predictions using a rule-based classification. It is also possible to directly label the feature classes produced by *Proba_cluster* to category classes using reference data or interactively using visual interpretation of the source data or VHR satellite images. The tool *proba_estimates* computes a prediction for each target variable for every pixel in the image. E.g., for forest cover maps prediction probability values ranging from 0 to 100% are assigned to every pixel in the respective image. As a result, each pixel is assigned to the class, whose proportion is the highest. If several classes are used in category classification, a hierarchical approach is applied. In the case of the growing stock volume every pixel is assigned the predicted stem volume value as cubic meters per hectare. Also, pixels covered by clouds, water and agricultural land are given a forest variable prediction, if they have not been masked out and received a NULL value in the pre-processing. For the continuous forest variable products, the post processing module masks all classes that should not be considered for the forest variable prediction, *i.e.,* non-forest pixels.

2.3.6. Compilation of end products (maps)

The estimates of the expected forest structural variables are produced by the "forest inventory tool". The forest inventory tool produces one layer for each defined forest variable or category for which reference data is available. By default, the pixel size of the produced maps is 10 m for Sentinel-2 data. The value of each pixel corresponds directly to the estimated value of the variable or the coded category class.

As set of products the following targeted continuous forest variable spatial layers (maps) can be processed for the whole area of the Ukraine – each with a resolution of 10 m:

- Forest mask or forest cover map
- Tree species distribution maps
- Forest Type map
- Mean tree height [m]
- Tree mean diameter [cm]
- Stem number/density [n/ha]
- Basal Area [m2/ha]
- Average Age [yrs]
- Mean growing stock [m^3 / ha]
- Increment [m^3/ ha]
- Above and below ground biomass [t dry/ ha]
- Carbon stock [t / ha]
- Crown closure / Total relative stocking [%]

Examples of final products and levels of potential accuracy

In the following we provide some samples for the final products and describe which levels of accuracy could be achieved based on experiences from our Forest Flux project, but also from previous research studies mainly from Scandinavia. According to the experiences with the *Forest Flux* processing service, overall accuracies of around 70-95% have been observed for discrete variables, depending on the number of classes latest in the frame of the pilot studies during the Forest Flux project but also in research projects in Scandinavia before. Figure 9 Examples land cover and tree species classification in Finland (left) and species-wise growing stock volume estimation (right, Forest management units outlines in red) (Häme, et al., 2022) provides examples of discrete land cover and tree species classifications based on Sentinel-2 data from Finland. For continuous variables, Root Mean Square Error (RMSE) levels of around 30-60% of the mean have been reached on plot level, depending on the variable in question (Astola et al. 2019). In practice, the plot level errors tend to balance each other out, resulting in increasingly accurate estimates as the area of interest gets larger. The bias is typically very low, indicating that on average the estimates are very close to reality. This forms a firm basis for reliable estimates of forest structural variables for the area of interest.

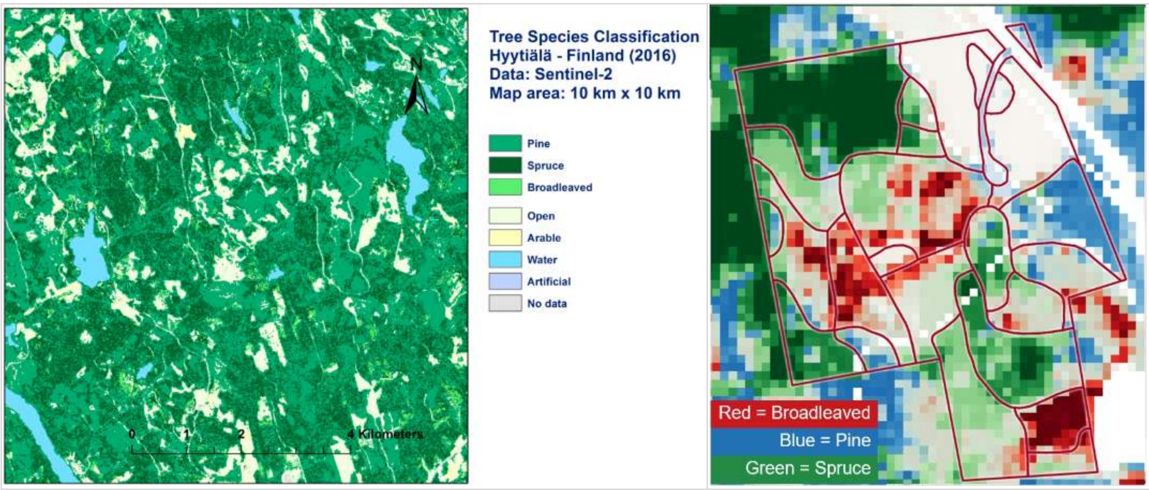

Figure 9 Examples land cover and tree species classification in Finland (left) and species-wise growing stock volume estimation (right, Forest management units outlines in red) (Häme, et al., 2022)

The right image in Figure 9 Examples land cover and tree species classification in Finland (left) and species-wise growing stock volume estimation (right, Forest management units outlines in red) (Häme, et al., 2022)

shows growing stock volume estimates per tree species for pine (*Pinus sylvestris*), Spruce (*Picea abies*) and broadleaved trees, respectively. For reference, forest management units are outlined in red, showing a good match with the detected variability of growing stock volume per tree species.

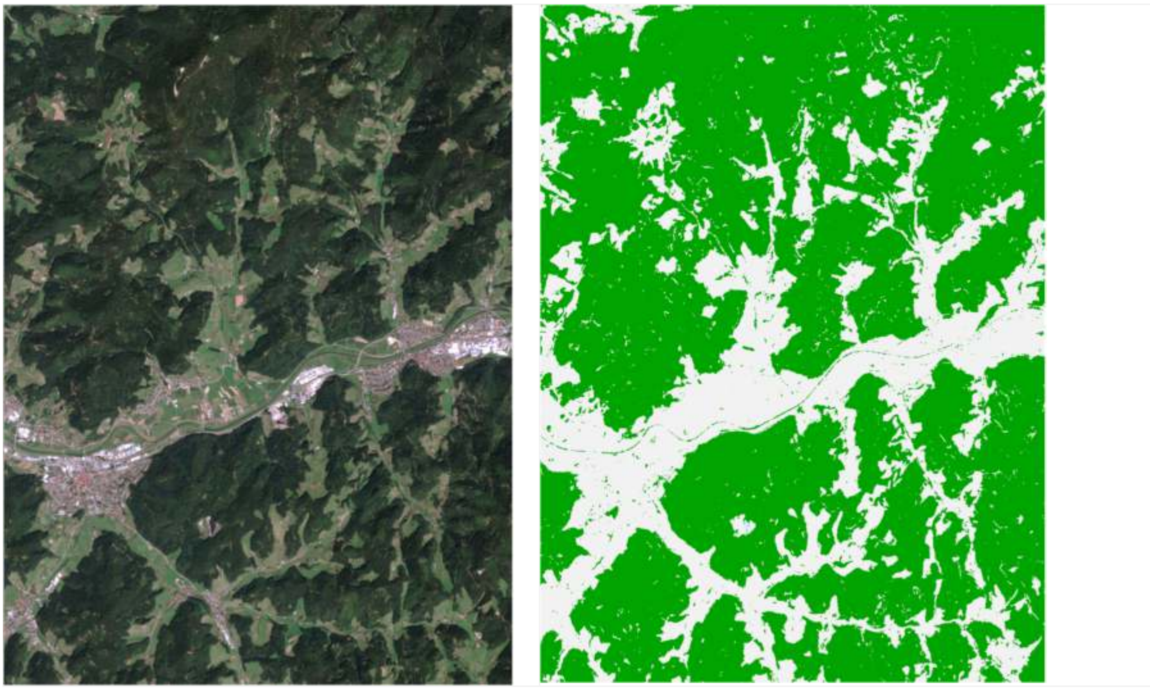

Figure 10 Example of forest cover (green) and non-forest (white) map in Germany (Häme, et al., 2022)

Figure 10 shows an exemplary extract of a forest cover map for an area of approximately 13 km x 9 km in Germany. Here, the accuracy computed using visually interpreted sample plots was 95%. However, accuracies obtained for continuous variable estimates at plot level range between RMSEs of 30 to 60%, which is in line with accuracies obtained in other projects (e.g. Astola et al. 2019). However, accuracies vary between structural variables, being highest for diameter, height, and stem basal area (**Fehler! Verweisquelle konnte nicht gefunden werden.**). Since bias is typically rather low, it is assumed that on average the estimates were very close to reality. Results also vary between sites.

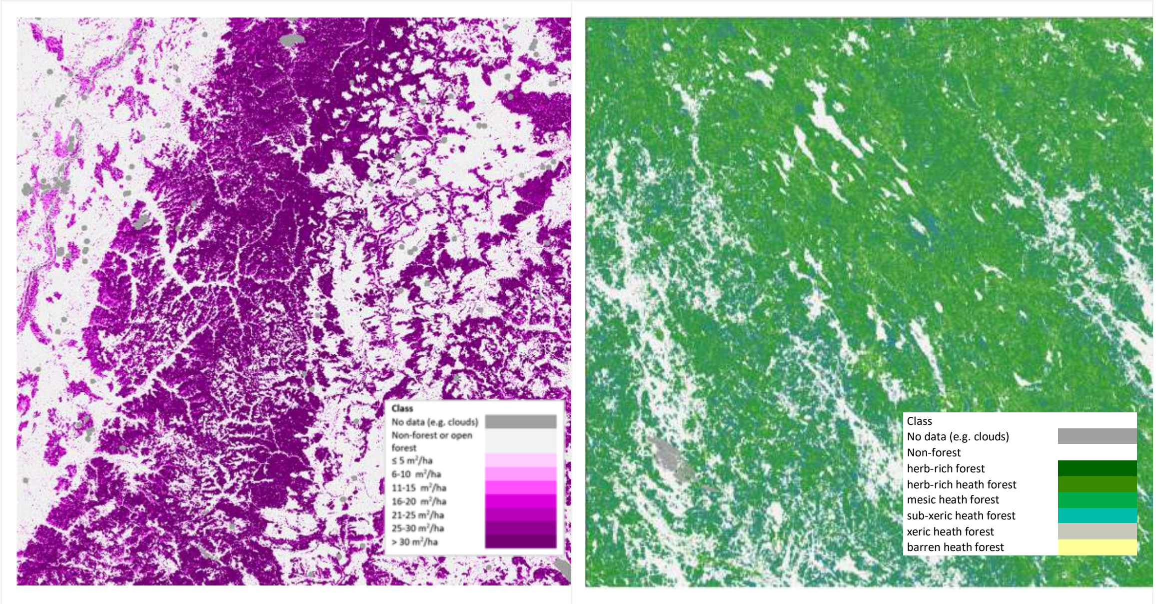

Figure 11 Mean estimated basal area/ha (Germany, left) and Forest type classification (Finland, right) (Häme, et al., 2022)

For stem volume (Figure 12 Tree height (left) and stem volume products (right) in Romania (Häme, et al., 2022)

), the RMSEs are typically above 50%.

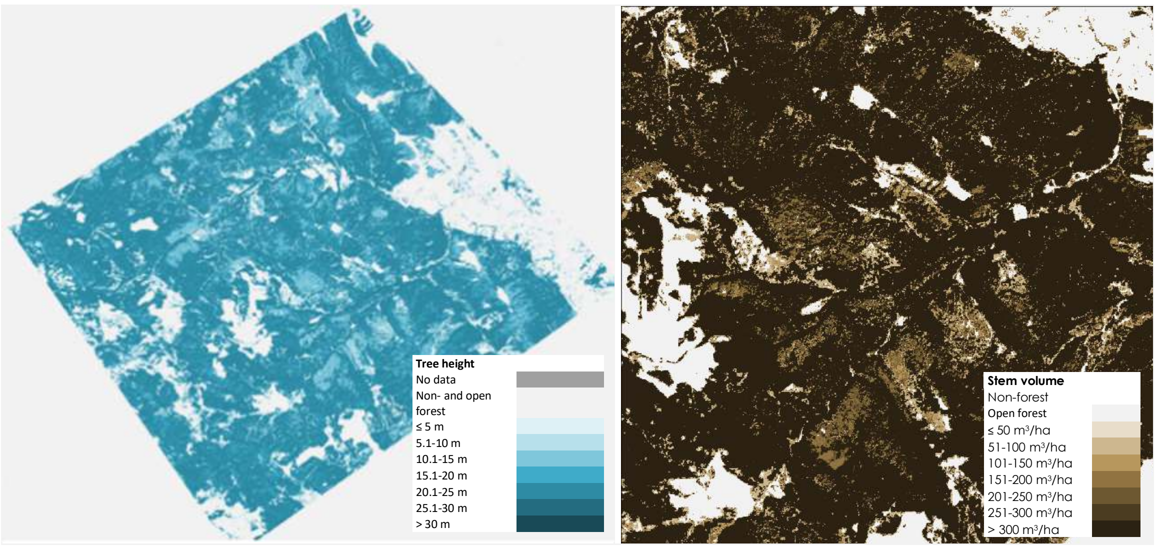

Figure 12 Tree height (left) and stem volume products (right) in Romania (Häme, et al., 2022)

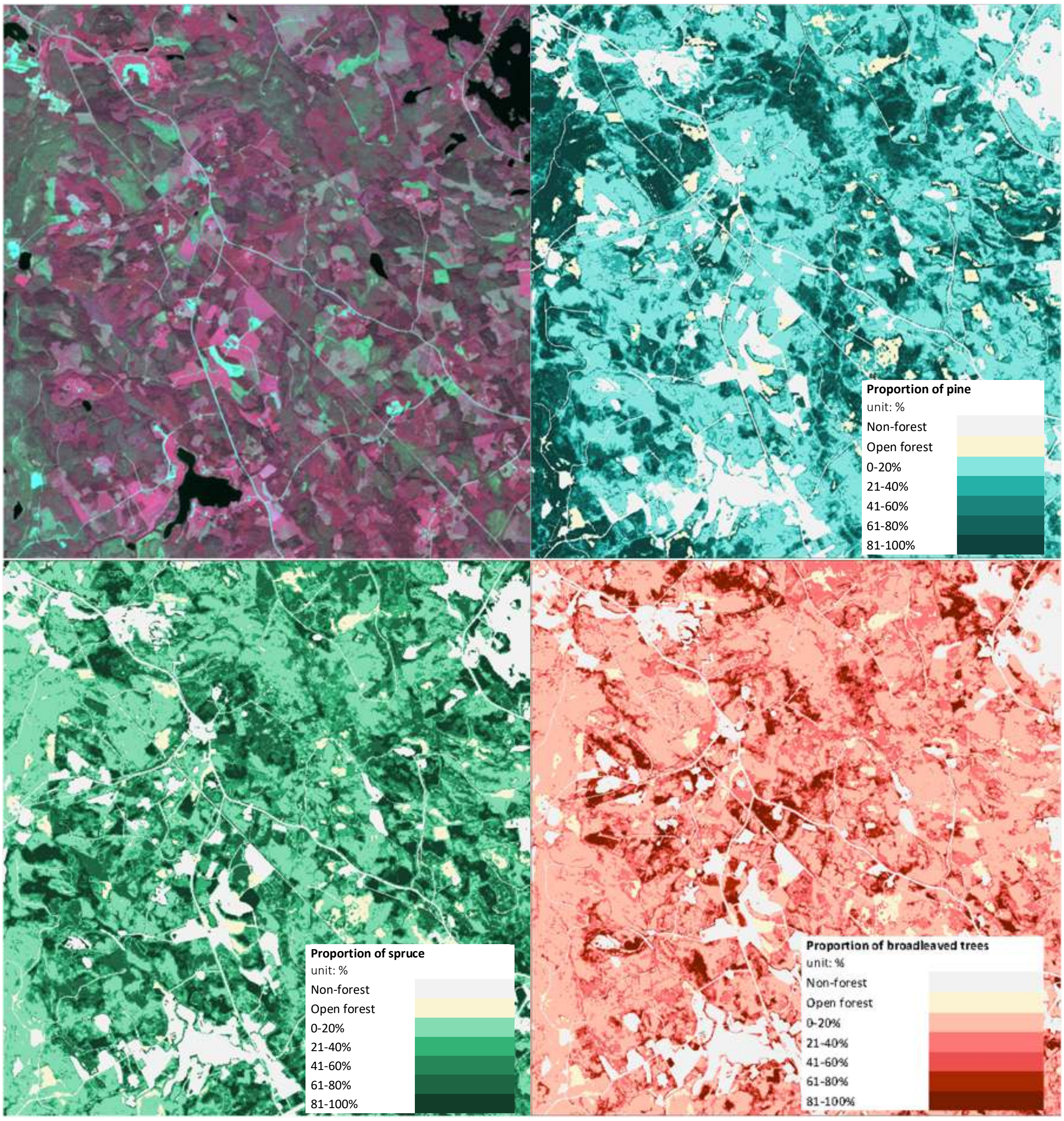

Figure 13 False color composite and tree species proportions in Finland (Häme, et al., 2022)

For tree species, results vary remarkably depending on the species and area of interest. E.g. in Finland (Figure 13 False color composite and tree species proportions in Finland (Häme, et al., 2022)

), map accuracies were estimated as relative RMSE of 54% for pine (bias -0.3%, top right), relative RMSE of 63% for spruce (bias -0.6%, bottom left) and relative RMSE of 88% (bias 1.1%, bottom right) for broadleaves, respectively.

Part II: Acquisition, description, and analysis of reference data

The following sections describe the necessary input data and method to determine the required amount and relevant reference data from forest management planning (FMP) and national forest inventory (NFI).

The estimation of forest structure variables, using remote sensing methods, relies on sufficient and qualitative reference data. The acquisition of such data is often limited by available resources. Hence, our goal is to make use of as much reference data as possible. While ensuring the quality of the data used, we also apply methods to determine the minimum required data to achieve the pursued accuracy.

To achieve the predefined accuracy for the classification of remote sensing images, we prepare reference data by first selecting relevant data, estimating additional field work that may be required and preprocessing steps required. To determine if data is relevant, we consider the eligibility of available datasets and estimate the additional field work that will be required. We then evaluate the added value of including more data and the resources required to achieve the agreed accuracy goal. The reference data selection and acquisition can therefore be split into three stages:

1. Eligibility assessment:
    - analyze and determine if any readily available data could be used as reference data in addition to the NFI field data collection (e.g., previously collected NFI data or FMP data)
    - analyze any available data for its recency, accuracy and if the required attributes listed under 2.2.3 can be derived from the data (no processing required)
2. Relevance assessment:
    - determine the amount of refence data necessary to achieve a given accuracy
    - define spatial grouping variables (e.g., forest types) and delineate their spatial boundaries to preselect relevant data sets and field plots
3. Preparation:
    - select the reference data sets (based on 1. and 2.) from the preselected available data sets
    - add as many field survey plots as possible to fill any gaps in data availability

3.1. Required data

3.1.1. Reference data

FMP data

- **For eligibility assessment:** attributes are collected
- **For relevance assessment:** Polygons or points indicating positions for where FMP data is available including information on the method and date of data acquisition and possibly information on the accuracy of the estimated stand level data if it can be included without any processing of data (any other stand level information, such as species or age composition, would also be beneficial, if it does not require any processing)
- **For reference data preparation:** Polygons or points with all stand level data necessary to prepare the data as required under 2.3.3

NFI data

- **For eligibility assessment:** attributes are collected with coordinates
- **For relevance assessment:** Coordinates of all NFI data collection plots (DCP), including the already measured plots and the planned plots for 2023 and later
- **For reference data preparation:** Coordinates with plot level data collected in the field

3.1.2. Ancillary data

- **Inaccessible areas:** Polygons of areas not accessible due to the ongoing war
- **Forest types:** Polygons for the entire forest area of all of Ukraine delineating coherent forest types

3.2. Eligibility assessment

For reference data to be considered eligible it needs to meet the following criteria:

- it shall enable the preparation of all required attributes listed under 2.2.3
- it shall be collected within the last 5 years
- if it consists of sample based estimates, it shall optimally have a relative standard error below 10 % at a 95 % confidence level
- it shall describe plots or stands homogeneous in species and age composition

3.3. Relevance assessment

Data grouping

To reduce the number of required samples, the population (entire forest area), which is being sampled, is partitioned into subpopulations (forest types). By grouping together coherent forest areas composed of similar forest types, the variance of the corresponding forest structure variables is reduced, and thus less samples are needed to make significant estimates for the subpopulations. The following considerations need to be taken into account for the final definition of applied forest types:

- If there is no forest type data available for the entire area of the country and no literature/publication is available, an alternative could be using a combination of ecoregions (Central European mixed forests, Crimean Submediterranean forest complex, East European forest steppe, Pannonian mixed forests, Temperate coniferous forests, Carpathian montane conifer forests, Temperate grasslands, savannas and shrublands, Pontic steppe) or climate zones and administrative districts.
- If country wide forest type data is available, and the forest type data is very nuanced, it should be grouped by similarity, to produce up to approx. 5 forest types per ecoregion / climate zone. The similarity is determined based on growth characteristics.
- It further needs to be determined, if there are any forest types, which are completely inaccessible (e.g. Crimean Submediterranean forest complex ecoregion). In such cases, representative forest types need to be determined.

Required sample size per group

The required sample sizes are then calculated for each forest type separately, using the averages and standard deviations of one of the main attributes (e.g., growing stock) available from existing data (e.g., forest management planning data).

The required sample size is determined using the formula for continuous data developed by Cochran (1977):

$$n = \left(\frac{Z\sigma}{e}\right)^2$$

With:

- n as the required sample size
- Z as the critical value for the desired confidence level (two-tailed)
- σ as the estimated standard deviation (e.g., determined through preceding study)
- e as the accepted margin of error (in the unit used for the standard deviation)

FMP data coverage

Once the required sample sizes for each forest type are determined, it is checked, how many FMP data sets are available per forest type. As not all FMP data sets may be applicable as training data, the data needs to meet the following prerequisites:

- Recent: collected within the last 3 years (2020-2022), maximum 5 years
- Consistent: the described stand should be relatively homogeneous in species and age composition
- Accurate: if the stands are measured using a sample-based approach, the relative standard error of the estimates of the forest structure variables should be lower than 20 % at a 95 % confidence level

NFI plot recommendation

NFI plots should be selected for these forest types, which are not represented sufficiently by the FMP data, with the aim of filling any gaps in the reference data. If all forest types are already sufficiently covered through FMP data, the NFI plots should be distributed between the forest types in a way to get the distribution of reference data between the forest types closer to the distribution of the calculated required sample sizes.

3.4. Preparation of reference data

3.4.1. FMP data

3.4.1.1. Description of FMP data as reference data source

FMP data in Ukraine represent detailed stand attributes that are updated usually in 10-year intervals through field surveys. The surveys are conducted within all forests managed by branches of the State Forest Enterprise (SFE) "Forests of Ukraine" or other users, e.g., municipal forest enterprises, protected areas, etc.. Thus, the FMP data over Ukrainian forests were collected in different periods.

Figure 14 Example for FMP polygons, color-coded by dominant tree species

During FMP-surveys, forest stands are visited by field crews to evaluate their characteristics. Boundaries of forest polygons, i.e., forest stands, temporally unforested areas (harvested areas, burnt areas, dead forests, unstocked forests, etc.), and non-forested areas (meadows, bogs, unproductive areas, waters, etc.), are updated in advance using the most recent high-resolution remote sensing imagery. These boundaries can be corrected during field visits. Site conditions are evaluated for all forested and temporally unforested areas. There are two indices used to characterize site quality: forest site conditions and forest type. The forest site condition describes soil fertility (four levels, A - the poorest fertility, D - the most fertile soils) and soil moisture (6 levels, 0 - very dry to 5 -very wet). The forest type is a more complex characteristic that apart from soil fertility and soil moisture provides also a list of species representing a native forest stand in this area. The forest type classification in Ukraine is rather complex. It includes more than 100 forest types that are specific for the Carpathians Mountains, a non-montane territory with some modifications for the Right-Bank (Dnipro river) and Left-Bank Ukraine.

Trained field crews use mostly ocular methods to estimate forest stand attributes which can be combined with elementary tree measurements (measuring heights and diameters of average trees) or sampling. However, the use of point-sampling is obliged within premature, mature, and overmature stands to obtain precise estimations of basal area. For younger stands, only relative stocking is visually evaluated as a fraction of the stand occupancy in comparison with the normal stand within a given site. This is somewhat subjective and largely depends on a surveyor's experience. The stand characteristics include a list of forest attributes that describe each tree species within stand’s layer: origin, age, mean diameter, mean height, relative stocking or basal area, and percentage of merchantable trees. Site index is provided only for dominant species. Growing stock volume is estimated using yield tables for a given age, height, and relative stocking.

For the full list of stand level attributes, please see the FMP attribute structure in the annex A 2.

3.4.1.2. Preparation of FMP data as reference data source

Spatial and attribute FMP data are stored in two separate data sets that can be joined using a unique stand ID (A 2). It can be a combination of the following fields of the FMP database:

- KALG - forest management unit (branch of the SFE "Forests of Ukraine")
- KAIG - forest management sub-unit
- KAWN - forest block
- KAVN - forest polygon (stand)

The FMP attribute information is structured in a relational database that links tables to describe different forest polygon characteristics (Figure 15).

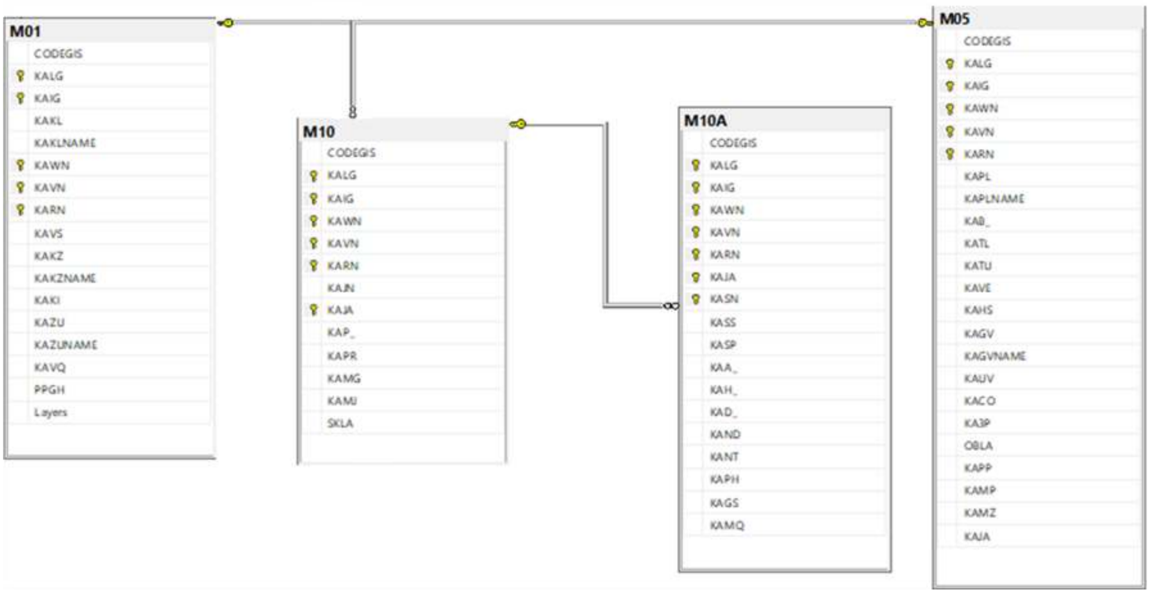

Figure 15 Relationships of tables in the FMP database

The following information can be derived from the database and used as the reference data source in the study:

- Table M01 – general information on forest polygon: stand area (KAVS), land use (KAKZ)
- Table M05 – general information on forest stand: main species (KAPL), site index (KAB_), forest type (KATL), forest site conditions (KATU)
- Table M10 – characteristics of stand layers: layer number (KAJA), relative stocking (KAP_), growing stock volume per ha (KAMG)

- Table M10A – tree species characteristics: tree species (KASP), fraction of species abundance (KASS), age (KAA_), mean diameter (KAD_), mean height (KAH_), basal area (KAGS).

Thus, forest polygons can be joined with associated forest attributes. In the study, the stand-level information should be used in a two-way table in which rows represent unique forest stands and columns represent forest attributes.

The following table presents the list of targeted forest structural variables on the one side and the related FMP stand level attributes, which can directly be used or based on them the target variable can be derived for the RS based NFI:

Table 4 Forest structure variable reference data

Target variable	Format/Unit	FMP attribute(s) – Abbreviation	FMP Attribute and explanation of its processing
Year of data	integer	KAVQ	Year of forest management planning inventory survey
Area of plot or stand	m²	KAVS	Forest sub-polygon area
Soil type	Categorical variable	KATU	Forest site conditions
Site index	m	KAB_	Site Index
Layers	Categorical variable	KAJA (calculated)	Calculated from variable "Layer"
Forest type	Categorical variable	KATL	Forest Type
Tree species distribution:			
Species 1	%	KASP + KASS (calculated)	Tree species combined with Species abundance (in % of BA)
Species 2	%	KASP + KASS (calculated)	Tree species combined with Species abundance (in % of BA)
Species n	%	KASP + KASS (calculated)	Tree species combined with Species abundance (in % of BA)
Mean height	m	KAH_ (calculated)	Mean height of tree species in layer 1
Mean diameter	cm	KAD + KASP + KANT (calculated)	Mean DBH of tree species in layer 1 weighted by N of trees per ha
Density	n/ha	KANT + KASP (calculated)	N of trees for all tree species in all layers per ha
Stem basal area	m²/ha	KAGS + KASP	BA of all tree species in all layers per ha
Mean age	years	KAA + KAGS (calculated)	Mean age of all tree species in layer 1 weighted by N of trees per ha
Growing stock volume	m³/ha	KAMQ + KAGS (calculated)	Volume of all tree species in all layers per ha
Increment	m³/ha/a	KAMZ + KAGS (calculated)	Increment of all tree species in all layers per ha
Biomass	t dry/ha	KAMQ + KAGS (calculated)	Above and below-ground biomass derived from volume and respective biomass conversion factors

Target variable	Format/Unit	FMP attribute(s) – Abbreviation	FMP Attribute and explanation of its processing
Carbon stock	t/ha	KAMQ + KAGS (calculated)	Derived from biomass and respective conversion factors
Crown closure	%	KAPP	Total relative stocking (1+2+3 layers)

3.4.2. NFI data

3.4.2.1. Description of NFI data as reference data source

NFI data is collected within forest inventory plots of the same configuration and spatial arrangement across all of Ukraine. The sampling frame of the NFI was designed using a 5 x 5 km regular reference grid. Within this grid, clusters of four sample plots are located randomly excluding a buffer zone of 250 m along the perimeter of the reference grid cells (Figure 16).

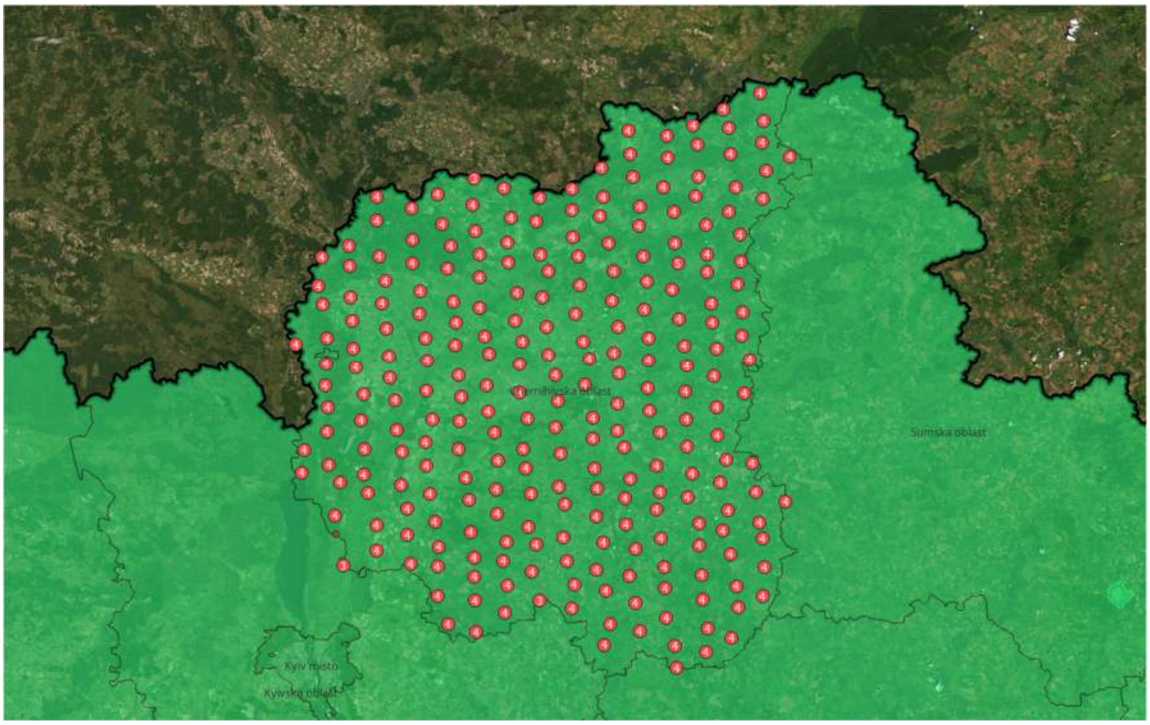

Figure 16 Clusters of four sample plots per grid cell for one oblast

Within this grid, clusters of four sample plots were randomly located excluding a buffer zone of 250 m along the perimeter of the reference grid cells (Figure 14). Sample plots are distributed with 420 m spacing in corners of the square cluster (**Fehler! Verweisquelle konnte nicht gefunden werden.**). The clusters are measured on five-annual rotation scheme.

Figure 17 One Cluster including four sample plots with a survey radius of 12.62m, each

The NFI sample plot consists of four nested circular plots of fixed area (Figure 17). The main plot with a radius of 12.62 m (500 m^2) is used to measure trees with a diameter at breast height greater than 26 cm. Subplots of 8.92 m (250 m^2), 3.98 m (50 m^2), and 0.56 m (1 m^2) are designed to measure smaller trees using the following diameter thresholds: 14-26 cm, 6-14 cm, and 2-6 cm, respectively (Figure 18). Nested sample plots that straddle stands boundary are divided into segments.

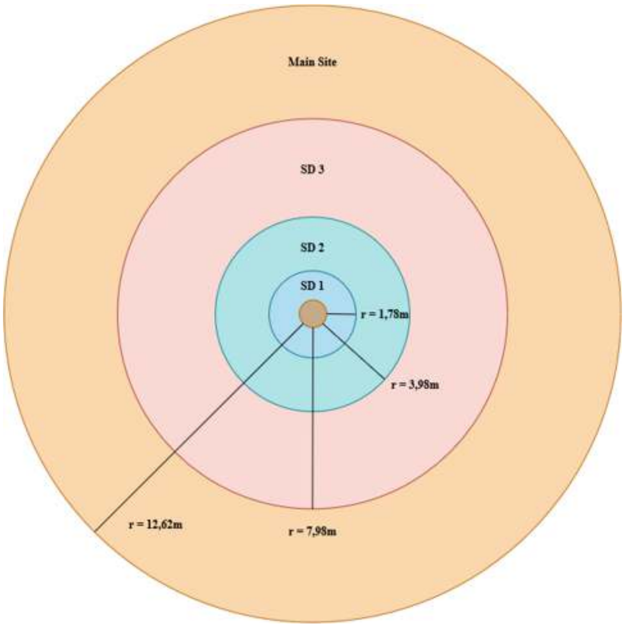

Figure 18 Plot design - concentric circle structure

Trees are mapped within the relevant plot radius. Diameters are measured for all tally trees, while heights are measured only for sample trees. The sample trees are selected for each species using their diameter distributions. More than four heights are measured for dominant species selecting trees within sample plots from different diameter classes. For the rest species, at least one (<20% in species composition) or two (>20% in species composition) average sample trees are selected. Additionally, two average tariff trees sampled outside sample plots are used to characterize stand age and increment using bores.

3.4.2.2. Field measurements are automatically stored in a database which controls the quality of the collected data.Preparation of NFI data as reference data source

The NFI data collected on nested sample plots should be processed to obtain per-hectare information on key forest attributes (see Table 3 Forest structure variable reference dataTable 3). There are following issues need to be considered in the calculations:

- Information is assigned to individual segments of a sample plot, thus this needs to be considered in data processing collected on the corresponding nested plot.
- Plot-level variables (e.g., land use, forest type conditions, site index, etc.) independently characterize each plot's segment.
- Per-hectare estimates for tree-level variables are obtained using varying tree factors that depend on the plot radius and area of the corresponding segment.

Processed NFI information can be joined with earth observations (Sentinel 2 spectral profiles) and ancillary data using plot-level estimates. Respectively, plots that straddle different land use categories (i.e., forested and unforested) need to be excluded from further RS-based analysis. Potentially, it can cause mixed pixels problems, especially if the coordinates of plot centers are not precisely calculated on the field (marginal errors can make up to 10 m). It is also necessary to exclude from forest attribute predictions all completely forested plots that represent forest stands of different ages. **Fehler! Verweisquelle konnte nicht gefunden werden.** represent a set-up for visual interpretation of VHR imagery to determine if level 2 and 3 data is located within a homogenous area.

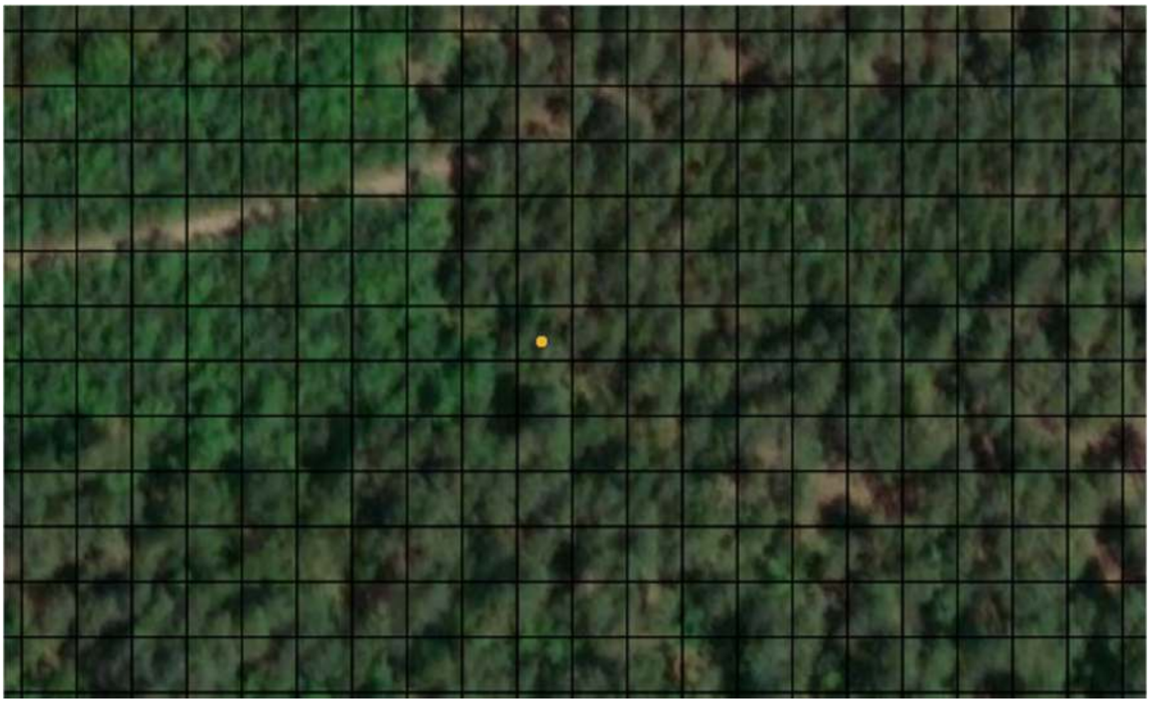

Figure 19 Example of a level 1 data (VHR base-map) overlayed with a 10 x 10 m grid representing the pixel size of HR Sentinel-2 imagery to determine if a sample plot (center-point) is located in a suitable homogenous area for reference polygon delineation

However, the final decision on plot inclusion in the analysis can be made directly in the modeling using leave-one-out predicted values of the corresponding attribute.

With the help of the following set of spatial data layers the reference polygons can be evaluated and delineated (Table 5).

Table 5: Spatial Data Layers used inside a GIS for evaluation and delineation of NFI plots

Spatial Data	Usage
VHF satellite images, preferably true and false color composites	Main source for forest cover information within a NFI plot
NFI plot locations	Location of NFI survey plot, Used to create buffer (see below)
Buffer of 12,62m around NFI plot locations	Information on the survey plots area
Grid with 10m x 10m cells in line with the pixels of Sentinel 2 imagery	Cells that intersect with the buffer of the survey plots are used to store information for a 10x10m polygon that can be used as reference data
Climatic zones polygons	Additional Information on forest tree species composition in a given area

Part III: Implementation Plan

In this part III the implementation of the above described concept for a RS based NFI is described in the form of a rough work- and time plan. It also contains the share of available resources in form of expert workdays. The budget in time and financial resources was fixed and concluded earlier in autumn 2022 and just gave the frame to consider the share for the different tasks and processing steps described in the following table.

Table 6: Proposed work- and time plan for the implementation of the RS based NFI

Tasks	ISTE subtasks	Proposed n of work days	NSTE subtasks	Proposed n of work days	Due date
Method description	Description of methods applied throughout the analysis	10		0	31.12. 2023
	Building capacities with local experts enabling future autonomous analysis	10	Training of national forest inventory experts	15	
Data acquisition	Consulting the local experts in the requirements for satellite data, reference data and any ancillary data necessary to conduct the analysis	15	Acquire raw data relevant for reference data preparation (all three levels: FMP & NFI data + aerial imagery)	5	31.03. 2023
Data preparation	Consulting the local experts in the process of preparing the satellite data, reference data (including QA/QC) and any necessary ancillary data	15		0	15.10. 2023
		0	Determine minimum sample size per group (forest type)	3	01.04. 2023
		0	Determine eligible FMP stand data	3	
	Support reference data validation work process	5	Set up reference data validation work process	5	
		0	Prepare FMP data for the use as RS reference data	36	

Tasks	ISTE subtasks	Proposed n of work days	NSTE subtasks	Proposed n of work days	Due date
		0	Prepare previously collected & 2023 NFI data as RS reference data	15	
	Selection and preprocessing source data (Sentinel 2 images)	10		0	
Model training and derivation of forest structural variables	Supporting the production of RS based estimation of forest structure variables using ForestFlux on the F-TEP platform, including the calculation of performance metrics	13	Production of RS based estimation of forest structure variables using ForestFlux on the F-TEP platform, including the calculation of performance metrics	15	15.11. 2023
Processing of RS and field inventory output into maps and tables structured for NFI reporting	Supporting the aggregation of the RS analysis output into reporting tables	10	Aggregation of the RS analysis output into reporting tables	25	15.12. 2023
		88		122	

A 1. Use of remote sensing in forest inventory and mapping in Ukraine

A 1.1. Recent advances improving the RS-based forest monitoring capacity in Ukraine

Mapping forest cover using dense satellite time series

The forests in Ukraine are heavily disturbed due to intensive human activities (logging, infrastructure development) and natural factors (wildfires, insects). For example, recent beetle-caused pine forest dieback has dramatically impacted forest landscapes and resulted in more intensive salvage logging in the most forested northern regions. The forest decline and accumulation of dead fuels in line with the climate change have sufficiently increased the flammability of the landscapes which led to catastrophic wildfires. Still, the illegal amber mining remains a very specific type of disturbance in northern Ukraine where about 3,000 ha of forested lands have been degraded (Myroniuk et al., 2020). Thus, national forest policy requires continuously updated information to maintain biodiversity, economic value of forest resources, and support international reporting on forests at state level.

Russian invasion of Ukraine has limited the ability of the National Forest Inventory because large areas of Ukraine are not available for field data collection. From this perspective, remote sensing technologies can provide significant support to the traditional methods of sample-based forest inventory. Apart from many developed countries that utilize remote sensing technologies at operational levels, Ukraine has not enough experience. However, recent studies including those conducted in Ukraine indicate the great potential of the RS-NFI.

The use of remote sensing technologies to monitor forest ecosystem dynamics over a range of spatial and temporal scales is well established. Retrospective analyses of forest dynamics has become a prominent component of modern forest monitoring as the availability of time series of satellite observations has been increased (Banskota et al., 2014; Gómez et al., 2016). Advances in dense time-series processing and implementation of the relevant algorithms on the Google Earth Engine (GEE) cloud-based platform (Gorelick et al., 2017) has improved the capacity to monitor forests over large spatial scales. Spectral trajectories of satellite time series have proven useful for detecting long-lasting and non-stand replacing disturbance (Bullock et al., 2020; Chen, 2021; Coops et al., 2020). Various temporal segmentation algorithms have been proposed to describe these trajectories (Hermosilla et al., 2015a; Huang et al., 2002; Kennedy et al., 2010; Zhu & Woodcock, 2014). Moreover, studies have shown a potential to detect casual agents of forest disturbance using spectral change metrics (e.g., duration, magnitude) extracted from the temporal trajectories and use them in more accurate forest mapping (Nguyen et al., 2018a; Schroeder et al., 2017). Free access to satellite data has promoted efforts to developing methods of near-real time monitoring land cover (Brown et al., 2022; Francini et al., 2020).

The historically deep and data-rich Landsat archive has been recognized as an important resource for large-area forest monitoring. Benefiting from a free access policy, wide spatial coverage, moderate resolution of imagery, and over four decades of data acquisition, Landsat imagery has found extensive application in various monitoring programs (Wulder et al., 2012, 2019). Over the last decade, Landsat time series have been used as the basis for forest disturbance detection (Cohen et al., 2017; Nguyen et al., 2018a; Schroeder et al., 2017), and as sources providing more stable spectral trajectories for forest classification (Hermosilla et al., 2018; Matsala, Bilous, Myroniuk, Holiaka, et al., 2021), including nearest neighbor imputation mapping (Bell et al., 2015, 2021; Kennedy et al., 2018). The LandTrendr (Landsat-based detection of Trends in Disturbance and Recovery) algorithm (Kennedy et al., 2010) has been used worldwide for annual change detection (Kennedy et al., 2012; Myroniuk et al., 2020; Nguyen et al., 2018a; Rathnayake et al., 2020) and improved multi-year forest mapping (Matsala, Bilous, Myroniuk, Holiaka, et al., 2021). Similarly, the Vegetation Change Tracker (VCT) (Huang et al., 2010) and the Composite-to-Change (C2C) approach (Hermosilla et al., 2015a) are mostly focused on detecting annual changes and producing corresponding maps of forest cover. Given the limitation for characterizing spectral variation caused by vegetation phenology using annual composites, the Continuous Change Detection and Classification (CCDC) algorithm has become extremely useful to derive intra-annual vegetation phenology and gradual inter-annual vegetation growth or abrupt change (Zhu & Woodcock, 2014). In addition to their ability to capture vegetation phenology, harmonic regression coefficients extracted from time series of spectral data have proven to be better predictors for key forest attributes than composite image variables. Particularly, Wilson et al (2018) achieved about two- to threefold increase in the coefficient of determination for models of continuous variables (i.e., number of trees, basal area, biomass) using harmonic models terms versus those using composite imagery. Further evaluation of seasonal composite and harmonic regression approaches identified better discrimination between forest types utilizing spectral trends and seasonality information quantified with harmonic inputs (Adams et al., 2020). Moreover, Derwin et al. (2020) showed that adding harmonic terms to spectral data and other explanatory variables (e.g., terrain variables) improves the performance of predictive models. The CCDC algorithm is also recommended as an effective means to deal with missed observations that appear due to the presence of snow or cloudiness (Awty-Carroll et al., 2019; Wilson et al., 2018). More recent publications have demonstrated the potential for improvement of the CCDC algorithm to detect even subtle changes in canopy (Ye, Rogan, Zhu, Hawbaker, et al., 2021) and near-real-time forest monitoring (Ye, Rogan, Zhu, & Eastman, 2021; Zhu et al., 2020).

Use of RS-derived covariates for mapping forest inventory attributes

The increased demand on spatially explicit information along with technological changes of data collection has led to wide application of LTS in support of forest inventories (Lister et al., 2020). However, national forest inventory programs were not initially designed to support mapped output, which is a key information need for most monitoring programs. Based on the seminal work of Tomppo (1990), many of these programs today use nearest neighbor modeling techniques to relate ground-based observations from a sparse network of sample plots to spectral properties from remotely sensed observations in order to produce wall-to-wall maps of forest characteristics (Chirici et al., 2016; Eskelson et al., 2009; McRoberts et al., 2010).

Within the forest inventory community, nearest neighbor imputation refers to a class of methods that estimate characteristics of target pixel based on k reference observations (i.e., field plots) most similar in multivariate space of ancillary remote sensed, topographic, climatic, and other environmental variables (Eskelson et al., 2009; McRoberts, 2012; McRoberts et al., 2010). Nearest neighbor imputation has emerged as effective technique for spatial prediction of forest attributes based on similarities between pixels associated with field plots and those without such association. This technique is recognized to be useful for spatial prediction because it is nonparametric and multivariate, and thus can be applied for simultaneously mapping multiple forest attributes (Henderson et al., 2014). The similarity between target pixel and reference observations can be determined using common distance metrics (e.g., Euclidean), or quantified using a multivariate ordination technique such as canonical correlation analysis (used in most similar neighbor (Moeur & Stage, 1995)) or canonical correspondence analysis (used in gradient nearest neighbor (GNN) (Ohmann & Gregory, 2002)). An optimal choice of the number of neighbors (k) is always about the trade-off between accuracy and preservation of covariance among response variables. In particular, maps of forest characteristics produced with k = 1 nearest neighbor imputation model maintain existing combinations of attributes that can also be observed in the field (Ohmann et al., 2014). McRoberts (2009) demonstrated that covariance between forest attributes is preserved when only one nearest neighbor is used in imputation models. Conversely, nearest neighbor imputation with k > 1 usually minimizes the root mean square error of prediction (Hudak et al., 2008; Wilson et al., 2012) and higher values of k can be applied if there are large number of reference observations in the database (Eskelson et al., 2009).

Forest mapping using nearest neighbor imputation methods is commonly used in national- and regional-scale forest monitoring projects (Beaudoin et al., 2018; Chirici et al., 2020; Wilson et al., 2012). More recent applications of these techniques are associated with advances in image processing and availability of free satellite imagery that has created opportunities for integration of temporal segmentation methods with nearest neighbor imputation techniques. Such approaches provide an important insight into long-term monitoring of forest attribute dynamics. Specifically, Ohmann et al. (2012) first used yearly covariates derived from the LandTrendr temporal segmentation algorithm (Kennedy et al., 2010) as input to the GNN model to map yearly forest vegetation attributes in the Pacific Northwest USA. The LandTrendr temporal segmentation approach was also used to create temporally smoothed synthetic imagery, which was utilized to map forest biomass dynamics in Australia using nearest neighbor imputations (Nguyen et al., 2018b, 2020). Bell et al. (2021) incorporated an ensemble LandTrendr disturbance mapping algorithm (Cohen et al., 2018) to produce temporally smoothed LTS and assess trends in large live trees and snags using GNN imputation mapping. The C2C approach (Hermosilla et al., 2015b) for image compositing was incorporated to predict forest structural attributes and aboveground biomass using nearest neighbor imputation in Canada (Matasci et al., 2018; Zald et al., 2016).

Implications for forest monitoring in Ukraine

Traditionally, studies devoted to forest cover mapping in Ukraine were mostly focused at the Carpathians Mountains (Kuemmerle et al., 2009, 2011) or larger European territories including Ukraine (Griffiths et al., 2014; Potapov et al., 2015; Senf & Seidl, 2021). The authors reported on forest changes, but not on forest characteristics such as species composition and growing stock volume. Conversely, the few existing application of the nearest neighbor techniques to map forest characteristics within local test sites in Ukraine (Bilous et al., 2017; Matsala, Bilous, Myroniuk, Diachuk, et al., 2021) have not relied on data collected via a national forest inventory sampling design.

Apart from many countries, the wide application of information extracted from remote sensing data in forest inventory is under development in Ukraine. Preliminary results of regional-scale forest inventories that were conducted in Sumy and Ivano-Frankivsk oblasts (regions) during 2008¬–2015 and 2009–2015, respectively, demonstrated great potential for supporting decision making, and updating the national forest statistics (Storozhuk & Polley, 2017). However, a detailed analysis of the collected data in combination with satellite time series has been provided in limited studies (Myroniuk et al., 2022). To the best of our knowledge, there were no more applications in Ukraine to monitor changes of forest characteristics in a spatially explicit way using data obtained through the national forest inventory program (i.e., wall-to-wall mapping). Thus, the current international experience can contribute to advance methodology of the RS-NFI in Ukraine.

A 1.2. Implementation of Sentinel 2 imagery to map forest cover in Ukraine

Existing sources on the spatial distribution of forest cover in Ukraine and the need to create and maintain an up-to-date forest mask for the entire territory of Ukraine

The need for up-to-date, content-uniform, cartographically presented information on the distribution of forests throughout the territory of Ukraine is obvious in many areas of scientific and practical activity. Despite the long experience of forest research and the significant amount of information obtained, this issue has not yet been fully resolved in Ukraine. The last State Forest Assessment in Ukraine was conducted back in 2010. It was mainly statistical in nature and was not supported by cartographic materials.

The experience of forest management works in Ukraine has allowed to collect a huge amount of information on forest cover, including its spatial distribution. However, the disadvantage of this information is its low “temporal resolution” and its temporal inconsistency (basic forest inventory is conducted once every ten years at different times for different forest users), as well as the fact that it is conducted at the expense of forest users and therefore does not cover forests that are not in official forest use. An important source of information for studying the spatial distribution of Ukraine's forests can be the materials of the national forest inventory, but in Ukraine these works have so far been implemented only on the example of certain regions (Myroniuk et al., 2022).

Another source of information on the spatial distribution of Ukraine's forests can be various global forest products (in particular, Global Forest Change (GFC), Landsat Tree Cover Continuous Fields (LTCCF), etc.). The advantage of such products is that they cover the entire territory of Ukraine, but they are not always accurate enough and are not always kept up to date. In addition, the issues of the spatial distribution of forest cover in Ukraine are considered in a number of scientific papers that contain both the justification of the methodology of such studies and the results of its practical application (e.g., Lesiv et al, 2019; Lyalko et al, 2019; Myroniuk, Kutia et al., 2020). However, for objective reasons, such studies cover only certain time periods and/or certain parts of the territory of Ukraine.

Thus, despite the certain triviality of this task, there is still a need in Ukraine to create an up-to-date forest mask that would cover the entire territory of the country, be created according to a single methodology, be constantly and promptly updated, integrate other sources of information on the spatial distribution of forest cover, and be available for wide use for practical and scientific purposes (including through integration into the National Geospatial Data Infrastructure). Obviously, the basis for creating such a forest mask should be remote sensing data, and the main its advantages are high enough speed and relative cheapness of creation and updating (but at the same time, relatively lower accuracy compared to forest management materials). Such a forest mask could also be of some use for the national forest inventory, starting from the planning stage of inventory works and ending with spatial representation and analysis of inventory data.

The work on the creation of such a forest mask was carried out and continues to be carried out at the State Enterprise "Forestry Innovation Research Center" with the support of State Forest Resources Agency of Ukraine (Public report, 2022).

Methodology for creating a forest mask of Ukraine based on Sentinel-2 data

First of all, when creating a forest mask, there is a need for a clear definition of "forest" concept, as there are its different interpretations for different purposes (Chazdon et al., 2016, etc.). In our work, we used the definition of a forest as a land plot of at least 0.1 hectare covered with woody vegetation, with crown coverage of at least 30% of the plot area. The "forest" category included all areas within Ukraine that met this definition, including those that, according to Ukrainian legislation, do not belong to the forest fund (for example, green spaces within settlements).

As the basic data for creating the forest mask there were used Sentinel-2 Level 2A imagery. Traditionally, in such works Landsat data have been widely used and continue to be successfully used. At the same time, Sentinel-2 data have some advantages over Landsat data (which does not diminish the importance of the latter):

- better spatial resolution of Sentinel-2 data (10 m and 20 m) compared to Landsat 8 data (30 m); this allows for more accurate identification of land cover types and their boundaries, including linear tree stands;
- better temporal resolution of Sentinel-2 data – 5 days compared to 16 days in Landsat 8;
- a larger number of spectral channels in Sentinel-2 imagery (in particular, the presence of Band 5, Band 6, Band 7, Band 8A channels, which are absent in Landsat 8 data).

One of the main advantages of Landsat imagery compared to Sentinel-2 imagery is a much longer time series of data, which reaches 50 years. At the same time, since the start of the Sentinel-2 mission in 2015, a certain data series (over 7 years) has also already been formed, which allows using various methods of spatio-temporal analysis for Sentinel-2 data.

At the next stage of the work, cloudless mosaics were created for the warm season (April – October) for each of the regions of Ukraine. To create seasonal cloud-free mosaics, we selected images in the appropriate time ranges with no more than 30% cloud cover. A cloud mask based on the Sentinel-2: Cloud Probability product was applied to all the selected images (the cloud probability value above which the data was excluded from further analysis was 10%). The seasonal mosaics were formed by determining the median value for each channel among all cloud-free values of each pixel for the corresponding time period.

As the spatial basis for the training sample creation there was used the previously developed sampling network for the national forest inventory (Storozhuk, 2019). This will allow for a better combination of the national forest inventory data with remote sensing data in the future. The training plots were interpreted mainly with the help of open high-resolution imagery available in Google Earth Pro, using the OpenForis Collect Earth software (Bey et al., 2016). An additional source for interpreting the training plots was the high-resolution SuperView imagery available for certain regions of Ukraine.

The training plots were classified into two classes – "forest" and "non-forest". Training plots for which it was impossible to unambiguously determine which land cover class prevails within their boundaries were not included in the training sample. For each of Ukraine's regions, from 1123 to 5389 training plots were interpreted (depending on the area of a region). In total, over 77 thousand training plots were interpreted for the entire territory of Ukraine. For some oblasts in the southern part of Ukraine, where too few forest training plots were interpreted, additional forest plots, that did not belong to the sampling network of the national forest inventory, were added to the training sample.

The cloudless mosaics were classified in the Google Earth Engine environment. For all cloudless mosaics, the values of Band 2, Band 3, Band 4, Band 5, Band 6, Band 7, Band 8, Band 8A, Band 11, Band 12 channels, NDVI index, and tasseled cap transformation channels (brightness, greenness, wetness) were used as variables. The classification was carried out using Random Forest classifier. The forest masks were generated separately for each region of Ukraine for 2020. For each of the obtained regional forest masks, the values of training overall accuracy and validation overall accuracy were at least 0.99. At the final stage, objects with an area of less than 0.1 hectares were removed from the resulting tree cover masks.

Results of creating a forest mask of Ukraine based on Sentinel-2 data

Currently, a forest mask of Ukraine for 2020 has been created (Fig. 1). The resolution of this forest mask is 10 m and it contains allotments covered with forest vegetation with an area of at least 0.1 hectares.

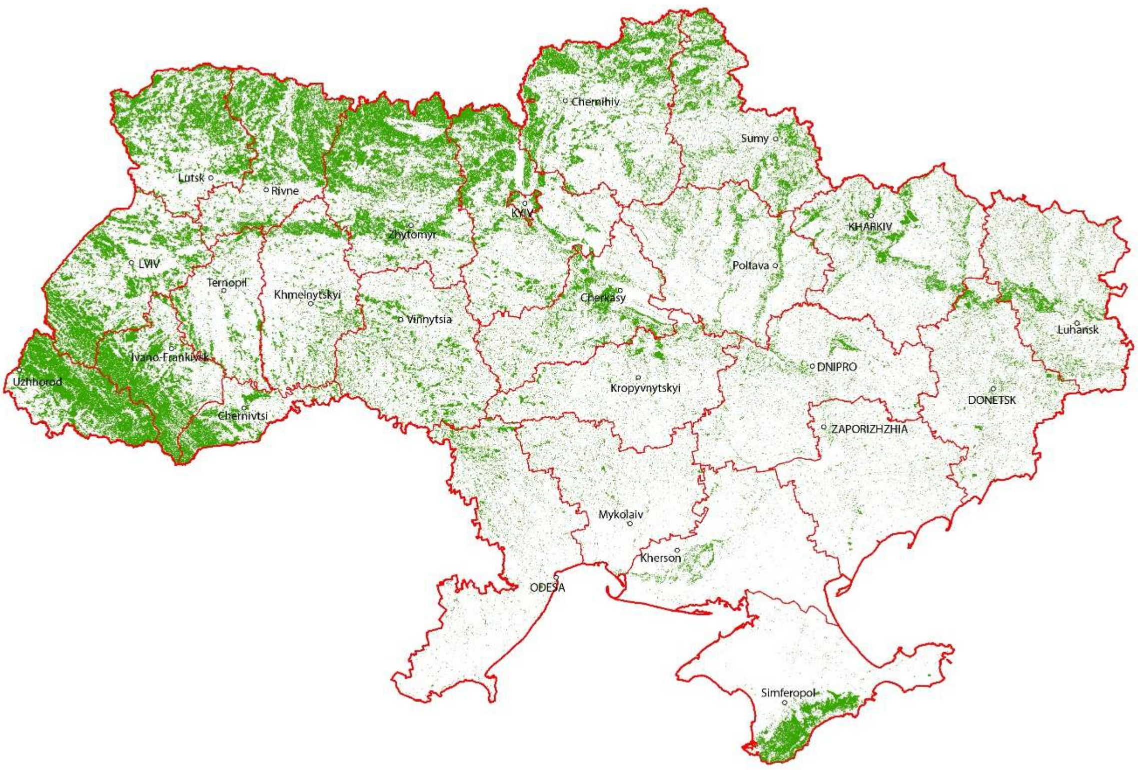

Figure 20 The forest mask of Ukraine for 2020 created based on Sentinel-2 data

Based on this forest mask, the forest area was calculated for each oblast and for the entire territory of Ukraine. The resulting forest area in 2020 for Ukraine as a whole and for most of its regions is higher than according to State Forest Assessment data as of January 1, 2011. In particular, the forest area in Ukraine as a whole according to 2020 data exceeds State Forest Assessment data as of January 1, 2011 by about 1.5 million hectares. One of the reasons for such differences may be that State Forest Assessment did not take into account data on some forest users, as well as data on forest areas that did not have official forest users or were not considered as forests (self-sown forests, green spaces within settlements, etc.). Thus, if all tree stands with an area of 0.1 hectares or more are classified as forests, Ukraine's forest area (in % of land area) would be about 2.5% higher than according to the official data.

The data obtained refute to some extent the popular perception in Ukraine of a significant reduction in forest area over the past decades. The forest mask data show that in the most forested regions of Ukraine (Polissya and the Carpathians), there has been no significant decrease in forest area. In this case, it would be more accurate to speak not so much of a quantitative reduction of forest area in Ukraine as of a qualitative deterioration in the forest's condition (both from forestry and biodiversity conservation perspective), due to the creation of monocultures that are less resistant to negative external influences. At the same time, in the eastern and southern regions of Ukraine, where the conditions for forest growth and regeneration are generally much less favorable, there is a noticeable decrease in forest area. This is especially true for coniferous stands on sandy terraces of rivers.

Among the main problems that arise when creating a forest mask based on remote sensing data, we can first mention the presence of certain classification errors – when non-forest areas are mistakenly classified as forest, or vice versa. The total percentage of such areas is generally quite small, but their presence negatively affects the accuracy of the data obtained. The main reason for such errors is the presence of land cover types with spectral characteristics similar to those of forests (in particular, shrubs, meadows, wetlands), which makes it difficult to distinguish between them. Another reason for classification errors is the complexity of vegetation cover, which does not always "fit" into the rigid and simplified framework of forest definitions. In some cases (e.g., in areas with sparse tree cover or in areas where natural reforestation is occurring), it can be difficult to clearly define which category a particular area belongs to – "forest" or "non-forest" – and to draw the line between "forest" and "non-forest" – even during field research.

Another problem is the limitations of the spatial resolution of the remote sensing data used, which are especially critical when identifying linear tree stands that are widespread in Ukraine. The capabilities of Sentinel-2 data in this regard are better than those of Landsat, which allows for fairly accurate identification of linear objects with a width of more than 20 meters. However, the task of identifying narrow linear tree stands cannot be fully solved even with Sentinel-2 data.

The main directions of further work on creating the forest mask of Ukraine:

- creation of forest masks for other years, spatio-temporal analysis and creation of a spatio-temporal forest mask for the entire period for which Sentinel-2 data are available, based on modern methods;
- use of new high spatial resolution data to form a more accurate and detailed training sample to be used for the classification;
- improving the classification methodology to more accurately identify "forest" and "non-forest" in problem areas and, accordingly, reduce the number of classification errors;
- integration of the obtained data with the national forest inventory; taking into account the data of the national forest inventory to improve the forest mask.

A 1.3. Role of satellite time series and nearest neighbor imputation technique in RS-based forest inventory of Ukraine

Through the Fulbright Program, data exchange between the Center of national forest inventory of Ukraine, and cooperation between the USDA Forest Service (USFS) and Ukrainian scientists, the workflow to map forest cover, tree species, and growing stock volume (1990–2020) was developed and tested within two regions: Sumy and Ivano-Frankivsk. This approach was based on historical forest inventory data collected during regional inventories between 2008 and 2015, Landsat time series, and GEE-based cloud computing algorithms (Myroniuk et al., 2022). The CCDC segmentation algorithm was used to produce cloud-free and temporally smoothed spectral data which were coupled with reference observations obtained for each sample plot locations.

The developed mapping workflow consisted of two consecutive phases: (i) forest/non-forest mapping, and (ii) predicting basal areas to map species and growing stock volume. The forest maps were developed using Random Forest classifier (Breiman, 2001), while basal areas were predicted barely within produced forest masks using the Gradient Nearest Neighbor (GNN) imputation technique (Ohmann & Gregory, 2002). Both models were independently trained within the regions. Species distributions were mapped using a threshold of 1.0 m2 ha-1 and the GNN model with k = 1 nearest neighbor. The approach achieved better accuracy of discrete presence-absence maps of species or species groups that belong to common environmental niches. The GNN model with k = 3 was used to get a prediction of growing stock volumes (basal areas) exclusively for species groups since the accuracy of individual species was low. In total, 3202 and 9860 observations were used to map forest/non-forest within Ivano-Frankivsk and Sumy regions respectively, while only data from 838 (Ivano-Frankivsk) and 1196 (Sumy) forested sample plots were utilized in nearest neighbor imputation.

The study showed positive dynamics of forest cover change between 1990 and 2020 for both regions (Figure 21). The accuracies of derived maps were high and stable over the study periods (OA > 0.96±0.02; UA > 0.94±0.03; PA > 0.93±0.04). Additionally, time series identified several hot spots of forest cover loss due to more intensive logging observed between 2000–2010 and 2010–2020 in the Carpathians Mountains. Analysis also documented two waves of forest regrowth on abandoned farmlands over the northern Ukraine occurred after collapse of the Soviet Union in 1991 and land reform in Ukraine (2000-2010).

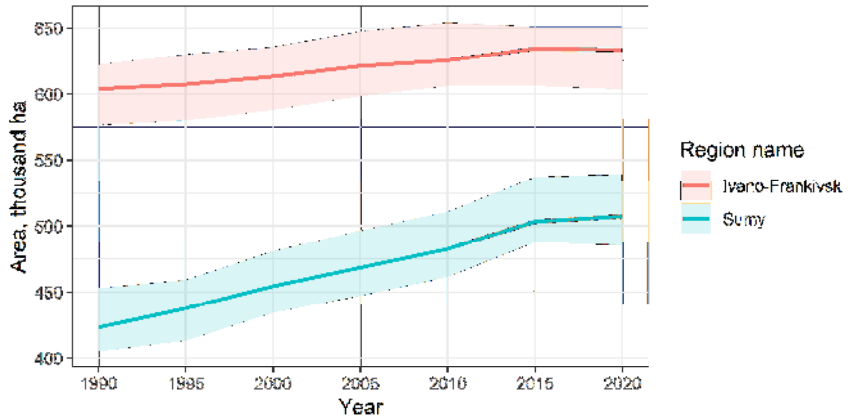

Figure 21 Dynamics of mapped forested area (solid lines) based on Landsat time series classification with 95% confidence intervals (ribbons) for two regions in Ukraine (Myroniuk et al., 2022).

The accuracy of GNN-derived presence-absence maps tended to decrease with species prevalence across the study regions. The model exhibited greater skill for Picea abies in the Ivano-Frankivsk (Cohen's kappa = 0.500) and Pinus sylvestris in the Sumy region (Cohen's kappa = 0.800) that had relatively higher mean values of basal areas and often grew in pure stands. In contrast, GNN predictions for some rare species was not much better than random prediction (e.g., Acer pseudoplatanus in Ivano-Frankivsk region, Cohen's kappa = 0.081). Deciduous species, which normally grow in mixed stands (e.g., Fraxinus excelsior, Acer platanoides, and Tilia cordata in the Sumy region), tended to exhibit similar map skill. The mapped distributions of the most common species (prevalence ≥ 20%) for both study regions were similar to the observed spatial patterns obtained from sample plots data at landscape-scales, represented in Figure 22 by 20-km hexagons.

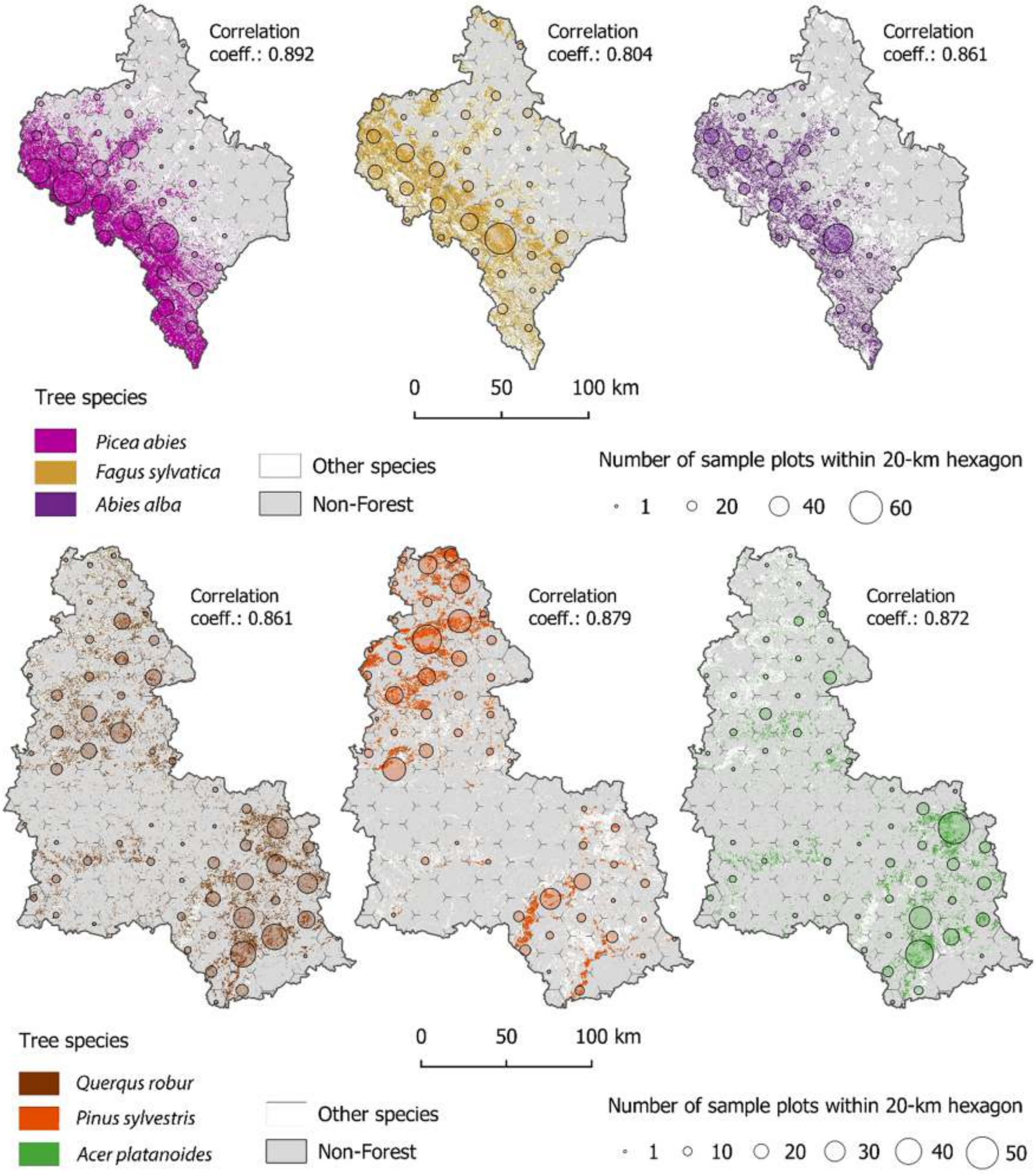

Figure 22 Distribution of individual species within the Ivano-Frankivsk (top panel) and Sumy (bottom panel) region based on GNN imputation (k = 1): correlation coefficients represent the rela-tionship between mapped area of species within 20-km hexagon and number

The GNN model with k = 3 nearest neighbors demonstrated ability to predict growing stock volume. R-squared values at Landsat 30-m spatial resolution (plot locations) varied for species groups between 0.33-0.53 (Ivano-Frankivsk) and 0.12-0.71 (Sumy), however, the performance of the developed imputation models increased at landscape level. For example, R-squared values increased to 0.50-0.74 (Ivano-Frankivsk) and 0.21-0.83 (Sumy) for aggregated at 5-km hexagon level. Similar tendency was observed for total basal area (growing stock volume) (Figure 23).

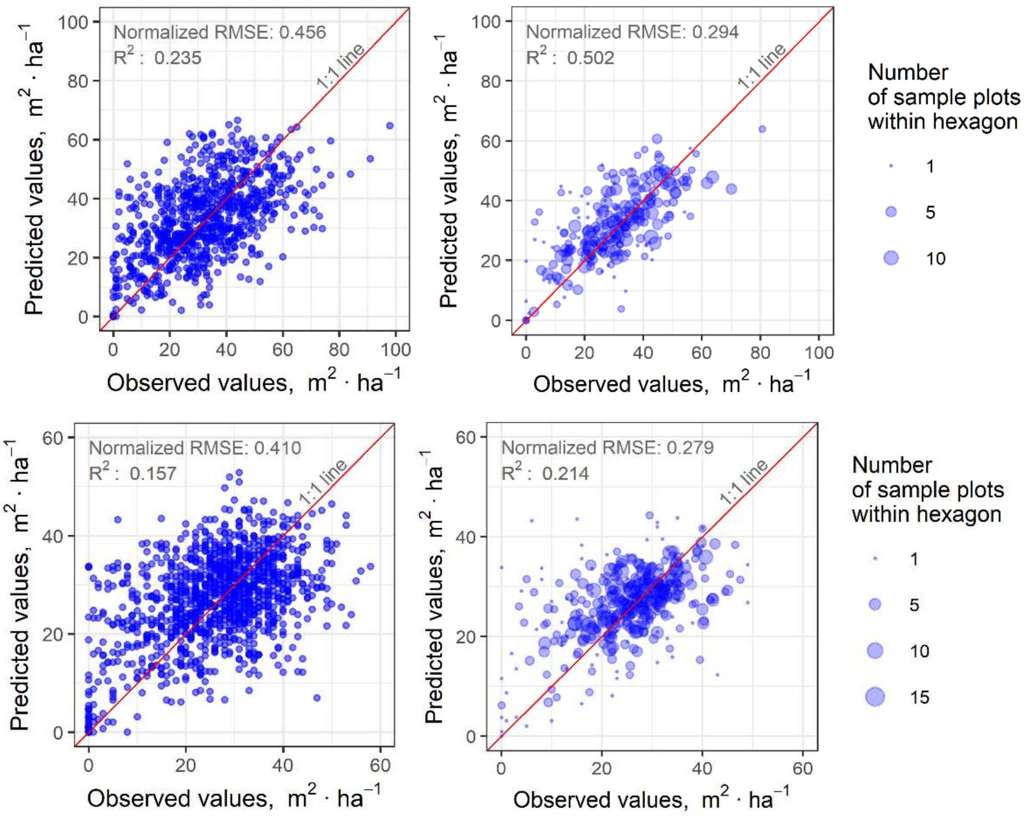

Figure 23 Predicted versus observed values of BA in Ivano-Frankivsk (top panel) and Sumy (bottom panel) region based on the GNN imputation model (k = 3) (Myroniuk et al., 2022).

A 1.4. Conclusions

From a methodological perspective, introduced by Myroniuk et al. (2022) a forest mapping and nearest neighbor imputations workflow in the GEE cloud-based platform enable users to follow this methodology over other regions of Ukraine with newly collected forest inventory observations and other types of satellite data (e.g., Sentinel 2). Fitted inter- and intra-annual trends in Landsat time series extracted using the CCDC segmentation algorithm strengthened multidecadal assessment of forest characteristics. This contributes to more reliable monitoring of forest cover and consistent estimation of forest area throughout time. Predictive performance of GNN imputation technique that utilized all available historical forest inventory data linked with the temporally normalized Landsat time series can be significantly improved when data are aggregated at higher than plot level domain. In addition to quality of RS data, spatial accuracy of field-sampled data is among key requirements to obtain accurate maps and estimates of forest inventory attributes.

A 2. FMP stand level attributes

The attributes assessed and aggregated on a forest stand level can be used to define reference polygons as described in chapter 3.4.1 above. As shown in Table 4 it is possible to derive and calculate all target forest structural variables based on the below listed set of stand variables for the purpose of the RS based NFI. FMP data in Ukraine represent detailed stand attributes that are updated usually in 10-year intervals through field surveys. The surveys are conducted within all forests managed by branches of the State Forest Enterprise (SFE) "Forests of Ukraine" or other users, e.g., municipal forest enterprises, protected areas, etc.. Thus, the FMP data over Ukrainian forests are collected in different periods.

CODE	NAME UKR	Name ENG	Definition/description of attribute
M00 - загальна інформація про квартал – General information on forest block			
KALG	*Лісогосподарське підприємство*	*State forest enterprise*	*Compartment adress*
KAIG	*Лісництво*	*Forest district*	
KAWN	*Квартал*	*Forest block (Compartment)*	
KAKL	Категорія захисності	Conservation status	
KARA	Адміністративний район	Administrative district of Ukraine	
KAGE	Тип рельєфу місцевості	Relief (mountain, flat land)	
MAPL	Площа кварталу	Forest block area (Compartment area)	Area in ha
MAIS	Джерело пожежної небезпеки	Source of fire danger	
MARI	Віддаль до джерела пожежної небезпеки	Distance to the source of fire danger	
MAIR	Радіаційне забруднення кварталу	Radioactive contamination level	
MATN	Табельний номер таксатора	Forest surveyor	
MAEK	Лісовпорядна експедиція	Forest survey	

CODE	NAME UKR	Name ENG	Definition/description of attribute
PPOV	Місцевий орган влади	Local authority body	
PPPZ	Природна зона	Climatic zone	
PPGL	Група лісів	Forests Groups (commercial, …)	
OBLA	Адміністративна область	Administrative oblast of Ukraine	
M01 - загальна інформація по виділу – General information on forest polygon			
KALG	*Лісогосподарське підприємство*	*State forest enterprise*	*Stand address*
KAIG	*Лісництво*	*Forest district*	
KAKL	*Категорія захисності*	*Conservation status*	
KAWN	*Квартал*	*Forest block (compartment)*	
KAVN	Виділ	Forest polygon (Forest Stand)	
KARN	Підвиділ	Forest sub-polygon[2] (Sub-stand)	
KAVS	Площа виділу (підвиділу)	Forest sub-polygon area	All attribute information is provided for sub-polygon (sub-stand), e.g., 1,2, … . However, more often polygons are not divided, so that there will be only one record
KAKZ	Категорія земель	Land use category	
KAKI	Ознака земель переданих в тимчасове користування	Sign of land transferred for temporary use	Information on who manages forests - permanent (state enterprise) or temporal (game management, recreation) users.
KAZU	Ознака особливо захисних ділянок лісів	A sign of particularly protective areas of forests	This refers to presence of some high conservation values (rare species)
KAVQ	Рік таксації	Year of forest management planning inventory survey	
PPGH	Господарська частина	Economic part	Forests are divided into four main categories (protective, historical, recreational, commercial). Within these

[2] If forest operations (logging) partially took place within a forest polygon (stand) it can be subdivided into several homogeneous sub-units

CODE	NAME UKR	Name ENG	Definition/description of attribute
			categories more detailed classification can be applied (i.g., forest-park zone of cities green belts)

M05 - лісогосподарські характеристики – General information on forest sub-stand by species

CODE	NAME UKR	Name ENG	Definition/description of attribute
KALG	*Лісогосподарське підприємство*	*State forest enterprise*	*Sub-stand address*
KAIG	*Лісництво*	*Forest district*	
KAWN	*Квартал*	*Forest block*	
KAVN	*Виділ*	*Forest polygon (stand)*	
KARN	*Підвиділ*	*Forest sub-polygon (sub-stand)*	
KAPL	Головна порода	Main species	
KAB_	Клас бонітету	Site index	Defined by height in a certain age
KATL	Тип лісу	Forest type	Example B2-OakPine
KATU	Тип лісорослинних умов	Forest site conditions	Example B2
KAVE	Ознака можливих для експлуатації лісів	Sign of possible forest exploitation	Explanation of possible forest exploitation: Related to PPGH. Indicates whether logging is allowed
KAHS	Господарська секція	Economic section	Synthetic category that combines stands that have similar composition, origin, productivity, harvesting age etc. Can be disconnected spatially.
KAGV	Група віку	Age group (young, middle-aged ..)	
KAUV	Код віку рубки	Harvesting age (code)	Age for a final harvest: pine - >80 in commercial forests ... = rotation period
KACO	Селекційна категорія	Breeding category	Reasonable or not for collecting seeds
KA3P	Цільова порода	Target breed	It related with KAHS. Forest crops can have different species composition from optimal within KAHS. So, it must be regu-lated during thinning.

CODE	NAME UKR	Name ENG	Definition/description of attribute
			For example, birch can dominated over pine at age of 10-15 years, however crops were planted to grow pine forest by age of harvesting. Pine is target species
OBLA	Адміністративна область	Administrative oblast of Ukraine	
KAPP	Сумарна повнота 1+2+3 ярусів	Total relative stocking (1+2+3 layers)	Relative stocking as a measure of stand volume in comparison to a yield table (fully stocked stands)
KAMP	Сумарний запас 1+2+3 ярусів	Total volume (1+2+3 layers)	
KAMZ	Середня зміна запасу	Annual increment	

M10 - розподіл за ярусами (підріст, підлісок) – Forest stand information by layers

KALG	*Лісогосподарське підприємство*	*State forest enterprise*	
KAIG	*Лісництво*	*Forest district*	
KAWN	*Квартал*	*Forest block (compartment)*	
KAVN	*Виділ*	*Forest polygon (stand)*	
KARN	*Підвиділ*	*Forest sub-polygon (sub-stand)*	
KAJN	Порядковий номер ярусу	Layer number #	
KAJA	Ярус	Layer name	
KAP_	Повнота ярусу	Relative stocking	Relative stocking using BA in comparison to a fully stocked stand defined in a yield table
KAPR	Приживлюваність незімкнутих лісових культур	Forest crops survival	Percentage of seedlings that survive during one vegetation cycle after crop planting
KAMG	Запас на 1 га	Volume in m^3/ha	
KAMJ	Запас ярусу на виділі	Volume in m^3	

M10A - таксаційна характеристика по породно – Forest stand information by species

CODE	NAME UKR	Name ENG	Definition/description of attribute
KALG	*Лісогосподарське підприємство*	*State forest enterprise*	
KAIG	*Лісництво*	*Forest district*	
KAWN	*Квартал*	*Forest block (compartment)*	
KAVN	*Виділ*	*Forest polygon (stand)*	
KARN	*Підвиділ*	*Forest sub-polygon (sub-stand)*	
KAJA	*Ярус*	*Layer*	
KASN	Порядковий номер деревної породи	Tree species number	
KASS	Коефіцієнт складу	Species abundance (in % of BA)	
KASP	Деревна порода	Tree species	
KAA_	Вік	Age (mean age)	
KAH_	Висота	Height (mean height (m))	
KAD_	Діаметр	Diameter (mean DBH (cm))	
KAND	Відсоток ділових стовбурів	Percentage of commercial trees	Commercial trees: BA per commercial species is measured using relascope in pre-mature, mature and overmature stands. For young and middle-aged forest, relative stocking can be estimated visually.
KANT	Кількість підросту	Number of seedlings/trees	N of trees per ha
KAPH	Походження	Origin	
KAGS	Сума площі перерізів	Basal Area (BA)	in m2/ha . *The field KAGS in most cases is empty. More reliable data on abundance would be KAMG - volume for stand layer. Then, species volume can be estimated using KAMG and KASS. Generally, there is no BA but relative stocking instead.*
KAMQ	Запас деревної породи на виділі	Volume within forest polygon	Volume of a tree species in a forest stand or sub-stand

References

Astola, H., Häme, T., Sirro, L., Molinier, M., & Kilpi, J. (2019). Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. *Remote Sensing of Environment, 223*, 257-273. doi:https://doi.org/10.1016/j.rse.2019.01.019

Dinerstein, E., Olson, D., Joshi, A., Vynne, C., Burgess, N., Wikramanayake, E., . . . Hansen, M. (2017). An ecoregion-based approach to protecting half the terrestrial realm. *BioScience, 67*(6), 534-545.

Häme, T., Kilpi, J., Ahola, H., Rauste, Y., Antropov, O., Rautiainen, M., . . . Bounpone, S. (2013). Improved Mapping of Tropical Forests With Optical and SAR Imagery, Part I: Forest Cover and Accuracy Assessment Using Multi-Resolution Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 6(1), 74-91. doi:http://dx.doi.org/10.1109/JSTARS.2013.2241019

Häme, T., Sirro, L., Dees, M., Mäkelä, A., Penttilä, J., Marin, G., & Tomé, M. (2021). Helping forest owners to manage forest carbon – the Forest Flux project. *GI_Forum,* 9(1), 137-142. doi:https://doi.org/10.1553/giscience2021_01_s137

Häme, T., Sirro, L., Seitsonen, L., Rauste, Y., Mõttus, M., Miettinen, J., . . . Barreiro, S. (2022). Forest Flux - Final Report. *VTT Technology, 403*, 48. doi:https://doi.org/10.32040/2242-122X.2022.T403

Häme, T., Stenberg, P., Andersson, K., Rauste, Y., Kennedy, P., Folving, S., & Sarkeala, J. (2001). AVHRR-based forest proportion map of the Pan-European area. *Remote Sensing of Environment, 77*(1), 76-91. doi:https://doi.org/10.1016/S0034-4257(01)00195-X

Lindberg, E., & Holmgren, J. (2017). Individual Tree Crown Methods for 3D Data from Remote Sensing. *Current Forestry Reports, 3*(1), pp. 19-31.

Myroniuk, V., Bell, D. M., Gregory, M. J., Vasylyshyn, R., & Bilous, A. (2022). Uncovering forest dynamics using historical forest inventory data and Landsat time series. *Forest Ecology and Management, 513*(120184). doi:https://doi.org/10.1016/j.foreco.2022.120184

Richards, J. A. (2013). *Remote Sensing Digital Image Analysis* (5 Ausg.). Heidelberg, Germany: Springer.

Sirro, L., Häme, T., Rauste, Y., Kilpi, J., Hämäläinen, J., Gunia, K., . . . Paz Pellat, F. (2018). Potential of Different Optical and SAR Data in Forest and Land Cover Classification to Support REDD+ MRV. *Remote Sensing, 10*(6), 26. doi:https://doi.org/10.3390/rs10060942

References Annex 1.

Adams, B., Iverson, L., Matthews, S., Peters, M., Prasad, A., & Hix, D. M. (2020). Mapping Forest Composition with Landsat Time Series: An Evaluation of Seasonal Composites and Harmonic Regression. *Remote Sensing*, *12*(4), 610. https://doi.org/10.3390/rs12040610

Awty-Carroll, K., Bunting, P., Hardy, A., & Bell, G. (2019). An Evaluation and Comparison of Four Dense Time Series Change Detection Methods Using Simulated Data. *Remote Sensing*, *11*(23), 2779. https://doi.org/10.3390/rs11232779

Banskota, A., Kayastha, N., Falkowski, M. J., Wulder, M. A., Froese, R. E., & White, J. C. (2014). Forest Monitoring Using Landsat Time Series Data: A Review. *Canadian Journal of Remote Sensing*, *40*(5), 362–384. https://doi.org/10.1080/07038992.2014.987376

Beaudoin, A., Bernier, P. Y., Villemaire, P., Guindon, L., & Guo, X. J. (2018). Tracking forest attributes across Canada between 2001 and 2011 using a *k* nearest neighbors mapping approach applied to MODIS imagery. *Canadian Journal of Forest Research*, *48*(1), 85–93. https://doi.org/10.1139/cjfr-2017-0184

Bell, D. M., Acker, S. A., Gregory, M. J., Davis, R. J., & Garcia, B. A. (2021). Quantifying regional trends in large live tree and snag availability in support of forest management. *Forest Ecology and Management*, *479*, 118554. https://doi.org/10.1016/j.foreco.2020.118554

Bell, D. M., Gregory, M. J., & Ohmann, J. L. (2015). Imputed forest structure uncertainty varies across elevational and longitudinal gradients in the western Cascade Mountains, Oregon, USA. *Forest Ecology and Management*, *358*, 154–164. https://doi.org/10.1016/j.foreco.2015.09.007

Bilous, A., Myroniuk, V., Holiaka, D., Bilous, S., See, L., & Schepaschenko, D. (2017). Mapping growing stock volume and forest live biomass: A case study of the Polissya region of Ukraine. *Environmental Research Letters*, *12*(10), 13. https://doi.org/10.1088/1748-9326/aa8352

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

Brown, C. F., Brumby, S. P., Guzder-Williams, B., Birch, T., Hyde, S. B., Mazzariello, J., Czerwinski, W., Pasquarella, V. J., Haertel, R., Ilyushchenko, S., Schwehr, K., Weisse, M., Stolle, F., Hanson, C., Guinan, O., Moore, R., & Tait, A. M. (2022). Dynamic World, Near real-time global 10 m land use land cover mapping. *Scientific Data*, 9(1), 251. https://doi.org/10.1038/s41597-022-01307-4

Bullock, E. L., Woodcock, C. E., & Olofsson, P. (2020). Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis. *Remote Sensing of Environment*, *238*, 110968. https://doi.org/10.1016/j.rse.2018.11.011

Chen, S. (2021). Monitoring temperate forest degradation on Google Earth Engine using Landsat time series analysis. *Remote Sensing of Environment*, 22.

Chirici, G., Giannetti, F., McRoberts, R. E., Travaglini, D., Pecchi, M., Maselli, F., Chiesi, M., & Corona, P. (2020). Wall-to-wall spatial prediction of growing stock volume based on Italian National Forest Inventory plots and remotely sensed data. *International Journal of Applied Earth Observation and Geoinformation*, *84*, 101959. https://doi.org/10.1016/j.jag.2019.101959

Chirici, G., Mura, M., McInerney, D., Py, N., Tomppo, E. O., Waser, L. T., Travaglini, D., & McRoberts, R. E. (2016). A meta-analysis and review of the literature on the k-Nearest Neighbors technique for forestry applications that use remotely sensed data. *Remote Sensing of Environment*, *176*, 282–294. https://doi.org/10.1016/j.rse.2016.02.001

Cohen, W. B., Healey, S., Yang, Z., Stehman, S., Brewer, C., Brooks, E., Gorelick, N., Huang, C., Hughes, M., Kennedy, R., Loveland, T., Moisen, G., Schroeder, T., Vogelmann, J., Woodcock, C., Yang, L., & Zhu, Z. (2017). How Similar Are Forest Disturbance Maps Derived from Different Landsat Time Series Algorithms? *Forests*, *8*(4), 98. https://doi.org/10.3390/f8040098

Cohen, W. B., Yang, Z., Healey, S. P., Kennedy, R. E., & Gorelick, N. (2018). A LandTrendr multispectral ensemble for forest disturbance detection. *Remote Sensing of Environment*, *205*, 131–140. https://doi.org/10.1016/j.rse.2017.11.015

Coops, N. C., Shang, C., Wulder, M. A., White, J. C., & Hermosilla, T. (2020). Change in forest condition: Characterizing non-stand replacing disturbances using time series satellite imagery. *Forest Ecology and Management*, *474*, 118370. https://doi.org/10.1016/j.foreco.2020.118370

Derwin, J. M., Thomas, V. A., Wynne, R. H., Coulston, J. W., Liknes, G. C., Bender, S., Blinn, C. E., Brooks, E. B., Ruefenacht, B., Benton, R., Finco, M. V., & Megown, K. (2020). Estimating tree canopy cover using harmonic regression coefficients derived from multitemporal Landsat data. *International Journal of Applied Earth Observation and Geoinformation*, *86*, 101985. https://doi.org/10.1016/j.jag.2019.101985

Eskelson, B. N. I., Temesgen, H., Lemay, V., Barrett, T. M., Crookston, N. L., & Hudak, A. T. (2009). The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. *Scandinavian Journal of Forest Research*, *24*(3), 235–246. https://doi.org/10.1080/02827580902870490

Francini, S., McRoberts, R. E., Giannetti, F., Mencucci, M., Marchetti, M., Scarascia Mugnozza, G., & Chirici, G. (2020). Near-real time forest change detection using PlanetScope imagery. *European Journal of Remote Sensing*, *53*(1), 233–244. https://doi.org/10.1080/22797254.2020.1806734

Gómez, C., White, J. C., & Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *116*, 55–72. https://doi.org/10.1016/j.isprsjprs.2016.03.008

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V. C., Abrudan, I. V., Lieskovsky, J., Munteanu, C., Ostapowicz, K., & Hostert, P. (2014). Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. *Remote Sensing of Environment*, *151*, 72–88. https://doi.org/10.1016/j.rse.2013.04.022

Henderson, E. B., Ohmann, J. L., Gregory, M. J., Roberts, H. M., & Zald, H. (2014). Species distribution modelling for plant communities: Stacked single species or multivariate modelling approaches? *Applied Vegetation Science*, *17*(3), 516–527. https://doi.org/10.1111/avsc.12085

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015a). An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. *Remote Sensing of Environment*, *158*, 220–234. https://doi.org/10.1016/j.rse.2014.11.005

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2015b). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. *Remote Sensing of Environment*, *170*, 121–132. https://doi.org/10.1016/j.rse.2015.09.004

Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., & Hobart, G. W. (2018). Disturbance-Informed Annual Land Cover Classification Maps of Canada's Forested Ecosystems for a 29-Year Landsat Time Series. *Canadian Journal of Remote Sensing*, *44*(1), 67–87. https://doi.org/10.1080/07038992.2018.1437719

Huang, C., Goward, S. N., Masek, J. G., Thomas, N., Zhu, Z., & Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, *114*(1), 183–198. https://doi.org/10.1016/j.rse.2009.08.017

Huang, C., Wylie, B., Yang, L., Homer, C., & Zylstra, G. (2002). Derivation of a tasselled cap transformation based on Landsat 7 at-satellite reflectance. *International Journal of Remote Sensing*, *23*(8), 1741–1748. https://doi.org/10.1080/01431160110106113

Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E., & Falkowski, M. J. (2008). Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment*, *112*(5), 2232–2245. https://doi.org/10.1016/j.rse.2007.10.009

Kennedy, R. E., Ohmann, J., Gregory, M., Roberts, H., Yang, Z., Bell, D. M., Kane, V., Hughes, M. J., Cohen, W. B., Powell, S., Neeti, N., Larrue, T., Hooper, S., Kane, J., Miller, D. L., Perkins, J., Braaten, J., & Seidl, R. (2018). An empirical, integrated forest biomass monitoring system. *Environmental Research Letters*, *13*(2), 025004. https://doi.org/10.1088/1748-9326/aa9d9e

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr – Temporal segmentation algorithms. *Remote Sensing of Environment*, *114*(12), 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

Kennedy, R. E., Yang, Z., Cohen, W. B., Pfaff, E., Braaten, J., & Nelson, P. (2012). Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. *Remote Sensing of Environment*, *122*, 117–133. https://doi.org/10.1016/j.rse.2011.09.024

Kuemmerle, T., Chaskovskyy, O., Knorn, J., Radeloff, V. C., Kruhlov, I., Keeton, W. S., & Hostert, P. (2009). Forest cover change and illegal logging in the Ukrainian Carpathians in the transition period from 1988 to 2007. *Remote Sensing of Environment*, *113*(6), 1194–1207. https://doi.org/10.1016/j.rse.2009.02.006

Kuemmerle, T., Olofsson, P., Chaskovskyy, O., Baumann, M., Ostapowicz, K., Woodcock, C. E., Houghton, R. A., Hostert, P., Keeton, W. S., & Radeloff, V. C. (2011). Post-Soviet farmland abandonment, forest recovery, and carbon sequestration in western Ukraine. *Global Change Biology*, *17*(3), 1335–1349. https://doi.org/10.1111/j.1365-2486.2010.02333.x

Lister, A. J., Andersen, H., Frescino, T., Gatziolis, D., Healey, S., Heath, L. S., Liknes, G. C., McRoberts, R., Moisen, G. G., Nelson, M., Riemann, R., Schleeweis, K., Schroeder, T. A., Westfall, J., & Wilson, B. T. (2020). Use of Remote Sensing Data to Improve the Efficiency of National Forest Inventories: A Case Study from the United States National Forest Inventory. *Forests*, *11*(12), 1364. https://doi.org/10.3390/f11121364

Matasci, G., Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Hobart, G. W., Bolton, D. K., Tompalski, P., & Bater, C. W. (2018). Three decades of forest structural dynamics over Canada's forested ecosystems using Landsat time-series and lidar plots. *Remote Sensing of Environment*, *216*, 697–714. https://doi.org/10.1016/j.rse.2018.07.024

Matsala, M., Bilous, A., Myroniuk, V., Diachuk, P., Burianchuk, M., & Zadorozhniuk, R. (2021). Natural forest regeneration in Chernobyl Exclusion Zone: Predictive mapping and model diagnostics. *Scandinavian Journal of Forest Research*, 1–13. https://doi.org/10.1080/02827581.2021.1890816

Matsala, M., Bilous, A., Myroniuk, V., Holiaka, D., Schepaschenko, D., See, L., & Kraxner, F. (2021). The Return of Nature to the Chernobyl Exclusion Zone: Increases in Forest Cover of 1.5 Times Since the 1986 Disaster. *Forests*, *12*(8), 1024. https://doi.org/10.3390/f12081024

McRoberts, R. E. (2009). Diagnostic tools for nearest neighbors techniques when used with satellite imagery. *Remote Sensing of Environment*, *113*(3), 489–499. https://doi.org/10.1016/j.rse.2008.06.015

McRoberts, R. E. (2012). Estimating forest attribute parameters for small areas using nearest neighbors techniques. *Forest Ecology and Management*, *272*, 3–12. https://doi.org/10.1016/j.foreco.2011.06.039

McRoberts, R. E., Tomppo, E. O., & Næsset, E. (2010). Advances and emerging issues in national forest inventories. *Scandinavian Journal of Forest Research*, *25*(4), 368–381. https://doi.org/10.1080/02827581.2010.496739

Moeur, M., & Stage, A. R. (1995). Most similar neighbor—An improved sampling inference procedure for natural-resource planning. *Forest Science*, *41*(2), 337–359.

Myroniuk, V., Bell, D. M., Gregory, M. J., Vasylyshyn, R., & Bilous, A. (2022). Uncovering forest dynamics using historical forest inventory data and Landsat time series. *Forest Ecology and Management*, *513*, 120184. https://doi.org/10.1016/j.foreco.2022.120184

Myroniuk, V., Bilous, A., Khan, Y., Terentiev, A., Kravets, P., Kovalevskyi, S., & See, L. (2020). Tracking Rates of Forest Disturbance and Associated Carbon Loss in Areas of Illegal Amber Mining in Ukraine Using Landsat Time Series. *Remote Sensing*, *12*(14), 2235. https://doi.org/10.3390/rs12142235

Nguyen, T. H., Jones, S. D., Soto-Berelov, M., Haywood, A., & Hislop, S. (2018a). A spatial and temporal analysis of forest dynamics using Landsat time-series. *Remote Sensing of Environment*, *217*, 461–475. https://doi.org/10.1016/j.rse.2018.08.028

Nguyen, T. H., Jones, S. D., Soto-Berelov, M., Haywood, A., & Hislop, S. (2020). Monitoring aboveground forest biomass dynamics over three decades using Landsat time-series and single-date inventory data. *International Journal of Applied Earth Observation and Geoinformation*, *84*, 101952. https://doi.org/10.1016/j.jag.2019.101952

Nguyen, T. H., Jones, S., Soto-Berelov, M., Haywood, A., & Hislop, S. (2018b). A Comparison of Imputation Approaches for Estimating Forest Biomass Using Landsat Time-Series and Inventory Data. *Remote Sensing*, *10*(11), 1825. https://doi.org/10.3390/rs10111825

Ohmann, J. L., & Gregory, M. J. (2002). Predictive mapping of forest composition and structure with direct gradient analysis and nearest- neighbor imputation in coastal Oregon, U.S.A. *Canadian Journal of Forest Research*, *32*(4), 725–741. https://doi.org/10.1139/x02-011

Ohmann, J. L., Gregory, M. J., & Roberts, H. M. (2014). Scale considerations for integrating forest inventory plot data and satellite image data for regional forest mapping. *Remote Sensing of Environment*, *151*, 3–15. https://doi.org/10.1016/j.rse.2013.08.048

Ohmann, J. L., Gregory, M. J., Roberts, H. M., Cohen, W. B., Kennedy, R. E., & Yang, Z. (2012). Mapping change of older forest with nearest-neighbor imputation and Landsat time-series. *Forest Ecology and Management*, *272*, 13–25. https://doi.org/10.1016/j.foreco.2011.09.021

Potapov, P. V., Turubanova, S. A., Tyukavina, A., Krylov, A. M., McCarty, J. L., Radeloff, V. C., & Hansen, M. C. (2015). Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. *Remote Sensing of Environment*, *159*, 28–43. https://doi.org/10.1016/j.rse.2014.11.027

Rathnayake, Jones, & Soto-Berelov. (2020). Mapping Land Cover Change Over a 25-Year Period (1993–2018) in Sri Lanka Using Landsat Time-Series. *Land*, 9(1), 27. https://doi.org/10.3390/land9010027

Schroeder, T. A., Schleeweis, K. G., Moisen, G. G., Toney, C., Cohen, W. B., Freeman, E. A., Yang, Z., & Huang, C. (2017). Testing a Landsat-based approach for mapping disturbance causality in U.S. forests. *Remote Sensing of Environment*, *195*, 230–243. https://doi.org/10.1016/j.rse.2017.03.033

Senf, C., & Seidl, R. (2021). Mapping the forest disturbance regimes of Europe. *Nature Sustainability*, *4*(1), 63–70. https://doi.org/10.1038/s41893-020-00609-y

Storozhuk, V., & Polley, H. (2017). *Forest Inventory—Status Quo in Ukraine, Experience of Germany, and FAO Recommendations* (Agricultural Policy Report APD/APB/06/2017; p. 52). German-Ukrainian Agricultural Policy Dialogue. https://www.apd-ukraine.de/images/2018/APR/APD_APR_07-2017_Forest_Inventories_ukr.pdf (in Ukrainian)

Tomppo, E. (1990). Designing a satellite image-aided National Forest Survey in Finland. *Proceedings of the SNS/IUFRO Workshop on the Usability of Remote Sensing for Forest Inventory and Planning, 26–28 Feb. 1990, Umeå, Sweden. International Union of Forest Research Organizations, Vienna*, 43–47.

Wilson, B. T., Knight, J. F., & McRoberts, R. E. (2018). Harmonic regression of Landsat time series for modeling attributes from national forest inventory data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *137*, 29–46. https://doi.org/10.1016/j.isprsjprs.2018.01.006

Wilson, B. T., Lister, A. J., & Riemann, R. I. (2012). A nearest-neighbor imputation approach to mapping tree species over large areas using forest inventory plots and moderate resolution raster data. *Forest Ecology and Management*, *271*, 182–198. https://doi.org/10.1016/j.foreco.2012.02.002

Wulder, M. A., Loveland, T. R., Roy, D. P., Crawford, C. J., Masek, J. G., Woodcock, C. E., Allen, R. G., Anderson, M. C., Belward, A. S., Cohen, W. B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J. D., Hostert, P., Hughes, M. J., … Zhu, Z. (2019). Current status of Landsat program, science, and applications. *Remote Sensing of Environment*, *225*, 127–147. https://doi.org/10.1016/j.rse.2019.02.015

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, *122*, 2–10. https://doi.org/10.1016/j.rse.2012.01.010

Ye, S., Rogan, J., Zhu, Z., & Eastman, J. R. (2021). A near-real-time approach for monitoring forest disturbance using Landsat time series: Stochastic continuous change detection. *Remote Sensing of Environment*, *252*, 112167. https://doi.org/10.1016/j.rse.2020.112167

Ye, S., Rogan, J., Zhu, Z., Hawbaker, T. J., Hart, S. J., Andrus, R. A., Meddens, A. J. H., Hicke, J. A., Eastman, J. R., & Kulakowski, D. (2021). Detecting subtle change from dense Landsat time series: Case studies of mountain pine beetle and spruce beetle disturbance. *Remote Sensing of Environment*, 263, 112560. https://doi.org/10.1016/j.rse.2021.112560

Zald, H. S. J., Wulder, M. A., White, J. C., Hilker, T., Hermosilla, T., Hobart, G. W., & Coops, N. C. (2016). Integrating Landsat pixel composites and change metrics with lidar plots to predictively map forest structure and aboveground biomass in Saskatchewan, Canada. *Remote Sensing of Environment*, *176*, 188–201. https://doi.org/10.1016/j.rse.2016.01.015

Zhu, Z., & Woodcock, C. E. (2014). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, *144*, 152–171. https://doi.org/10.1016/j.rse.2014.01.011

Zhu, Z., Zhang, J., Yang, Z., Aljaddani, A. H., Cohen, W. B., Qiu, S., & Zhou, C. (2020). Continuous monitoring of land disturbance based on Landsat time series. *Remote Sensing of Environment*, *238*, 111116. https://doi.org/10.1016/j.rse.2019.03.009

unique land use GmbH
Schnewlinstr. 10
79098 Freiburg, Germany
Tel +49 761 208534 – 0
unique@unique-landuse.de
www.unique-landuse.de