Analysis of relationships between parameters of the National Forest Inventory of Finland: case study of mesic forest

Olena Kavats, Dmitriy Khramov and Kateryna Sergieieva*

Foreign Company EOS Ukraine, Dnipro 49055, Ukraine

OK, 0000-0002-0172-7856; DK, 0000-0002-1737-7272; KS, 0000-0001-7345-2209
*Correspondence: ekaterina.sergeeva@eosda.com

Abstract: The use of satellite images and machine learning in addition to *in situ* data in national forest inventories enables covering large areas and significantly reduces costs. However, such combined inventories provide modelled stand properties, the relationships between which are not well understood. An approach to investigating linear and non-linear relationships between forest inventory parameters is proposed. It is applied to a study of the Multi-Source National Forest Inventory (MS-NFI) stand properties for the case of mesic forests. The relationships between MS-NFI parameters and stand reflectance in the visible, red edge, near infrared and shortwave infrared spectral regions were investigated for the Sentinel-2 satellite sensor. Linear models of canopy reflectance as a function of forest stand and elevation properties were developed. These models allowed to assess the comparative influence of MS-NFI parameters on stand reflectance as well as the monthly dynamics of this influence during the season (May–August 2019). Linear relationships between forest inventory parameters were investigated using a correlation matrix. Generalized additive models were used to investigate nonlinear pairwise relationships between forest inventory parameters. The proposed approach can be applied to assess the impact of stand features obtained from conventional ground-based forest inventory on forest canopy reflectance.

Supplementary material: The linear regression models for the Sentinel-2 spectral bands and the generalized additive models for selected features are available at https://doi.org/10.6084/m9.figshare.c.7977465

The use of satellite images to determine forest stand features requires an understanding of how different stand parts reflect electromagnetic radiation at various growth stages and how changes in forest cover affect changes in reflectance.

Vegetation canopy reflectance is a complex phenomenon that is influenced by many factors, including the optical properties of leaves, plant geometry (especially leaf area index (LAI)), leaf angle distribution, soil reflectance and the structure of the vegetation cover (i.e. plant spatial distribution) (Roy 1989; Asner 1998). Forest reflectance at the stand level comprises a mixture of materials with different reflective properties, depending on the properties of single trees (i.e. their species, condition and health) and the forest structure (such as canopy cover), as well as the terrain conditions (e.g. elevation, slope) (Ollinger 2010; Rautiainen *et al.* 2018). These properties are subject to both seasonal variations and long-term changes (Jensen 2007).

The general principles of forest stand reflectance variability have been studied for a long time and are well known (Gates *et al.* 1965), although there is still not enough information about how reflectance changes over time and in space depending on environmental properties/forest stands (Grabska and Socha 2021).

Stand biomass depends on other stand parameters such as density, age, volume, height and basal area (Jagodziński *et al.* 2019). Relationships between stand features have been discovered through years of ground-based research and formulated as allometric equations (Widlowski *et al.* 2003). Similar relationships are widely used in stand biomass models (Repola 2009; Liepiņš *et al.* 2022). An important feature of allometric equations is their locality: the equations vary with climatic conditions, vegetation structure, tree species and growth forms (Kangas *et al.* 2022).

Satellite remote sensing data are actively used in estimating timber stocks and forest biomass (Rodriguez-Veiga *et al.* 2017). Machine-learning models based on multispectral optical, radar and light detection and ranging (LiDAR) data (Ehlers *et al.* 2022) are used for this purpose. It is noted that the ability to use multispectral optical data to estimate above-ground forest biomass is limited due to poor penetration under the forest canopy, cloud cover and saturation (Avitabile *et al.* 2012). Saturation means that once a certain forest biomass

From: Paavola, M., Kotavaara, O., Bogdanov, K., Knobloch, A., Gusat, D. and Joutsenvaara, J. (eds) *Earth Observations and Proximity Sensing Technologies: Enhancing Safety, Sustainability and Efficiency in Mining Operations*. Geological Society, London, Special Publications, **559**,
https://doi.org/10.1144/gslspecpub2023-215

density is reached, the reflectance characteristics (spectral reflectance in a given range or vegetation index) no longer respond to changes in biomass (Rodriguez-Veiga *et al.* 2017).

The use of satellite data in forest inventories tends to increase (Lister *et al.* 2020; Wallner *et al.* 2021; Coops *et al.* 2022; Gschwantner *et al.* 2022). A combination of ground-based measurements, satellite images, digital elevation models and machine-learning models can cover large areas and significantly reduce inventory costs (White *et al.* 2016; Schumacher *et al.* 2020). The problem with such combined inventories is that they provide modelled stand features, adding modelling errors to the errors of traditional inventories. It is not known whether relationships between modelled stand features will be observed similarly to relationships between real stand parameters. The potential and limitations of combined inventories need further investigation (Maltamo *et al.* 2021).

In Finland, medium-resolution satellite images (Landsat 8 OLI, Sentinel-2 MSI) have been used together with map data and digital elevation models (DEMs) for national forest inventories since the 1990s (Tomppo *et al.* 2011). For more than 30 years, extensive research has been conducted to analyse the results and improve the methods of Multi-Source National Forest Inventory (MS-NFI) (Tomppo *et al.* 2014; Mäkisara *et al.* 2019; Mäkisara *et al.* 2022). However, the relationships between variables in the MS-NFI feature space have not been studied in detail.

The objective of the study is to investigate the MS-NFI feature space using one Finnish region Northern Ostrobothnia (Pohjois-Pohjanmaa) and one forest type (mesic forests) as an example. Linear and non-linear relationships between forest inventory parameters and forest canopy reflectance measured by the Sentinel-2 MSI satellite sensor, as well as changes in these relationships over time, will be investigated. This will allow the regularities of the influence of MS-NFI parameters on forest canopy reflectance to be determined and compared with regularities obtained from ground-based measurements.

Materials and methods

Study area and reference data

Located in northern Europe, Finland is one of the most forested countries in the world, with forests covering 75% (22.8 million hectares) of the land area (Peltola *et al.* 2019). Forest conservation is very important in Finland and environmental organizations make great efforts to protect forests and monitor their condition. Information on forest stands is collected regularly, resulting in up-to-date forest datasets that have been created for national forest management and other related purposes. However, 76% of forest habitats are under threat (Kontula and Raunio 2019). This deterioration of ecological conditions has resulted from intensive forest management over past centuries (Watson *et al.* 2018).

The study area is located in the Northern Ostrobothnia region (Fig. 1), with a total area of 316 km^2. The MS-NFI is based on field sampling and remote sensing methods. The methods and results of the Finnish MS-NFI corresponding to 2019 are summarized in Mäkisara *et al.* (2022). To estimate forest parameters, Finland is divided into 19 regions (maakunta). The total number of sampling plots was 74 472. An improved k-nearest neighbours algorithm is used to generate continuous maps of selected forest variables at the local scale for the entire country. The magnitude of the errors of the MS-NFI models is estimated compared to the training sample data, and the estimation results are reported in Mäkisara *et al.* (2022).

Data for 2019 downloaded from the Finnish Forest Research Institute (Metla) website (http://kartta.luke.fi/index-en.html were used. The data are distributed in the ETRS89–TM35FIN (EPSG:3067) coordinate reference system (CRS) and cover all the forest lands of Finland at a 16 m spatial resolution. In total, 45 different variables, such as species composition, growing stock volume and biomass, describe the forest structure of each pixel.

Figure 2 shows the seven largest forest classes by area (according to the MS-NFI 2019, land class 2019 = 1) within the study area, while the other classes are merged into the 'others' category. The total forest area is 204.8 km^2. More than 46% of this area is covered by mesic forests. Mesic forests are quite shady and humid and are dominated by Norway spruce (*Picea abies*), Scots pine (*Pinus sylvestris*) and birch (*Betula pendula, B. pubescens*) (Tonteri *et al.* 1990). Blueberry plants, arctic starflower, twinflower and various mosses predominate at ground level. The study area's growing season for broad-leaved trees starts in the first half of May, whereas the growing season for coniferous trees starts in the third week of April (Böttcher *et al.* 2016). The growing season ends in the early days of November (Böttcher *et al.* 2018).

A raster mask was generated for the mesic forest using MS-NFI data, based on the following condition:

$$\text{Mesic forest mask} = (\text{LC2019} = 1) \wedge (\text{SMC} = 1) \wedge (\text{SFC} = 3) \quad (1)$$

where LC2019 is the land class in 2019, SMC is the site main class and SFC is the site fertility class.

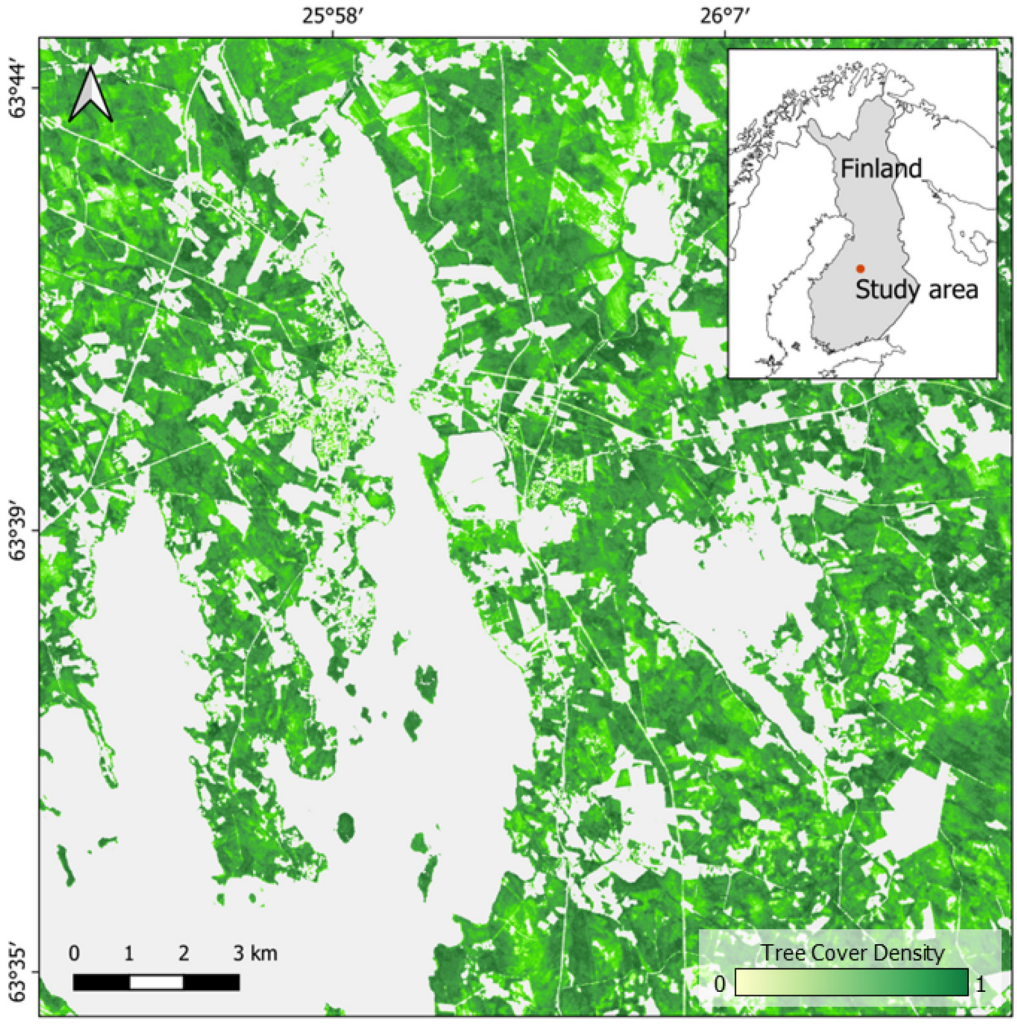

Fig. 1. The study area. Source: Copernicus Tree Cover Density (2018).

In the National Forest Inventory (NFI) of Finland, the SMC indicator divides forest land, poorly productive forest land and unproductive land into mineral soil and peatland site classes. SMC = 1 corresponds to mineral soil. The SFCs are used to group forest by vegetation zone into uniform classes according to their site fertility and wood production capacity. SFC = 3 is mesic forests.

The mesic forest mask was polygonized and a 20 m internal buffer zone was removed to eliminate edge effects (i.e. shadows, other forest classes, etc.) and small forest areas. A 20 m buffer width was chosen to provide pure forest pixels for all bands, especially B8A, B11 and B12, at a 20 m spatial resolution. A stratified spatial sample of 4664 points was collected within the remaining vector mask boundaries.

Satellite data

Satellite observations were made during the growing season (May–October), which was additionally limited by cloud cover. In the study area, predominantly cloudless periods usually occur around May–September. However, there was significant cloud cover in September 2019, which meant that a sufficient number of high-quality images could not be obtained.

To model the forest reflectance, Sentinel-2 Multispectral Instrument (MSI) Level-2A images (tile 35VML) were used, which were acquired from May to August 2019. These images most closely corresponded to the NFI of the same year (http://kartta.luke.fi/index-en.html).

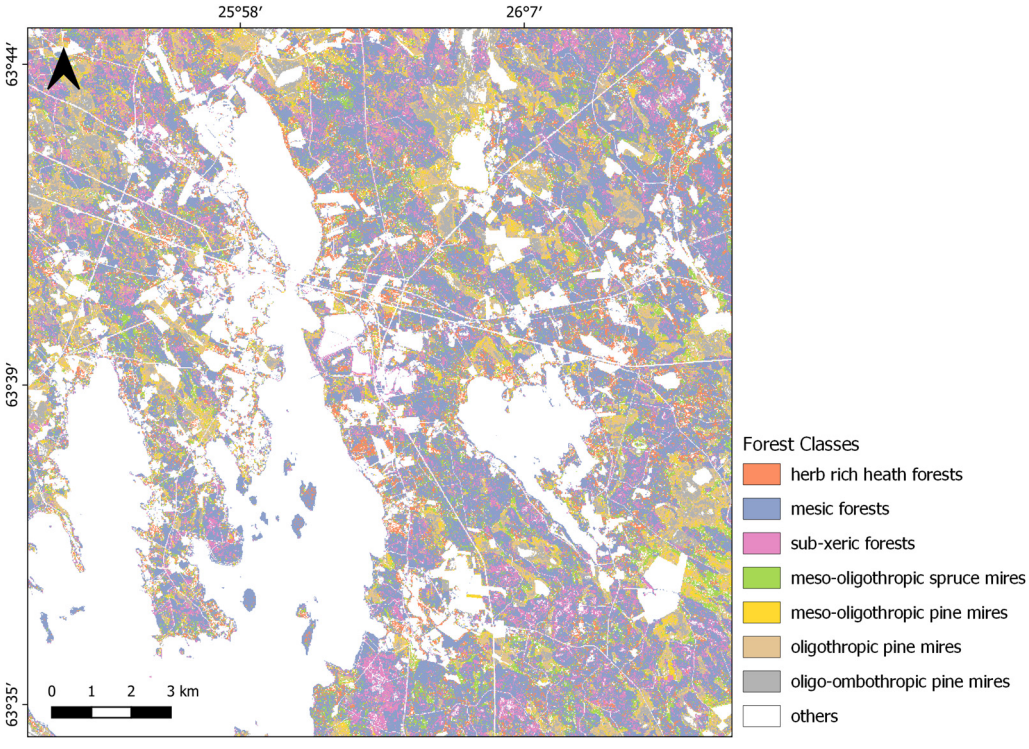

Fig. 2. The main forest classes, according to the Multi-Source National Forest Inventory (MS-NFI) of Finland 2019.

Images with less than 80% cloudy pixel percentage were used for the research. Among them, images with cloud and shadow cover of the study area of less than 40% were selected. Monthly median mosaics were generated to fill the gaps caused by cloudiness and reduce the effects of atmospheric and geometric observation conditions on each image.

Features and responses

Table 1 shows the variables that were used in forest canopy reflectance models (here and below 'feature' and 'variable' are used interchangeably). Variables describing stand properties were obtained from MS-NFI data. Canopy cover referred to the area of the vertical projection of the canopy onto the horizontal plane (without the duplicate counting of overlapping crowns). Mean stand diameter was estimated at a height of 1.3 m. Stand basal area referred to the cross-sectional area of tree trunks per hectare, measured at a height of 1.3 m. In MS-NFI, stand volume was divided into groups based on tree species (Scots pine, Norway spruce, birch and other broad-leaved trees) and wood grade (saw timber, pulpwood, etc.). Tree volume referred to the volume of the trunk from the stump to the top of the tree.

The terrain variables (elevation, slope and aspect) were obtained from the DEM National Land Survey of Finland (https://tiedostopalvelu.maanmittauslaitos.fi/tp/kartta?lang=en). The original DEM at a 2 m resolution was resampled by bilinear interpolation to a spatial resolution of 16 m and converted into CRS EPSG:3067.

Table 1. The variables used in forest canopy reflectance models

Name	Description and units
CC	Canopy cover in 2019 (%)
CC_b	Canopy cover of broad-leaved trees in 2019 (%)
Diam	Mean stand diameter in 2019 (cm)
H	Mean stand height in 2019 (dm)
Age	Stand age in 2019 (years)
BA	Stand basal area in 2019 (m^2 ha^{-1})
V_{pine}	Volume of pine trees in 2019 (m^3 ha^{-1})
V_{spruce}	Volume of spruce trees in 2019 (m^3 ha^{-1})
V_{birch}	Volume of birch trees in 2019 (m^3 ha^{-1})
V_b	Volume of other broad-leaved trees in 2019 (m^3 ha^{-1})
Elev	Elevation (m)
Slope	Slope (rad)
Aspect	Aspect (rad)

The model responses were the reflectance in the B2–B8A, B11 and B12 bands of the Sentinel-2 MSI (Table 2; central wavelength is provided for the Sentinel-2A satellite).

Workflow

The relationships between forest stand and terrain features, and their influence on canopy reflectance, were analysed in three stages:

(1) analysis of linear relationships between features and generating a set of explanatory features with the correlation matrix and variance inflation factor;
(2) development of linear models of forest stand reflectance and assessment of feature importance using linear regression;
(3) analysis of non-linear relationships between features using generalized additive models (GAMs).

To form a set of explanatory features to be used in linear models of forest canopy reflectance, the correlation matrix of Pearson correlation coefficients and variance inflation factor (VIF) were applied sequentially. It is known that the linear interdependence of variables (multicollinearity) introduces errors into model weight (coefficient) estimates. This complicates the analysis of the influence of single explanatory variables on a dependent variable (response). The investigation of variable correlation matrices allows for detection values of close to -1 or $+1$, indicating the pairwise dependence of variables. However, this method does not reveal mutual dependences within variable groups. A more reliable way to detect multicollinearity is to analyse the VIF (James *et al.* 2014). Variables with VIF > 5 were removed.

Linear models of forest canopy reflectance are needed to identify the most influential features.

Table 2. The spectral bands of the Sentinel-2 Multispectral Instrument (MSI)

Band	Name	Central wavelength (nm)
B2	Blue	492.4
B3	Green	559.8
B4	Red	664.6
B5	Red edge 1	704.1
B6	Red edge 2	740.5
B7	Red edge 3	782.8
B8	NIR	832.8
B8A	Narrow NIR	864.7
B11	SWIR 1	1613.7
B12	SWIR 2	2202.4

NIR, near-infrared; SWIR, short-wave infrared.

Multiple linear regression models were created for each spectral band (Table 2) and observation month (May–August). Only statistically significant variables were kept in the models, for which the absolute value of the t-statistic exceeded 2 (i.e. $|\text{t-stat}| > 2$).

Cook's distance was used to reveal outliers with a threshold of $4/n$, where n is the number of observations (control points). Outliers must be distinguished carefully to avoid removing the most interesting observations (Faraway 2014). Therefore, out of the observations that exceeded the Cook's distance threshold, only the most extreme were removed. For this, the outlier candidates were divided into quartiles and only those in the fourth quartile (Q4) were removed from the data.

The selection of the most influential variables simplified the models and, therefore, facilitated their interpretation. Our preliminary studies on influential variable selection showed that after the removal of the insignificant variables, the total number of variables was only reduced by one or two. This did little to facilitate the interpretation of the models. It was therefore decided to retain all significant variables ($|\text{t-stat}| > 2$).

The t-value was used to test the hypothesis that the coefficients of given variables differed significantly from 0 and to estimate the relative importance of the variables in the regression equations (the larger the t-value, the more influential the variable (Kuhn and Johnson 2013)).

GAMs allow for the observation and interpretation of non-linear relationships between stand features (Nath and Ni-Meister 2021). To analyse non-linear relationships, features with moderate or strong correlations are selected (e.g. features with Pearson correlation coefficient (r) in the range of 0.4–0.9). For such features pairwise GAMs are generated, and pairs whose GAMs have a significant increase in the coefficient of determination R^2 compared to the corresponding linear models are considered in detail.

The R packages and the R-Studio IDE were used in this study. The lm() function was used to fit the linear regression models. The following libraries were also used: 'car' to calculate the VIF (Fox and Weisberg 2019), 'mgcv' to generate the GAMs (Wood 2017), 'terra' to process the raster and vector spatial data (Hijmans 2022), 'sf' to process the vector spatial data (Pebesma 2018), 'ggplot2' to draw plots (Wickham 2016), and 'latex2exp' to use LaTeX formulae in plots (Meschiari 2022).

Results

Feature importance

In order to eliminate interdependent variables, the correlation matrix (Fig. 3) was analysed. The closest

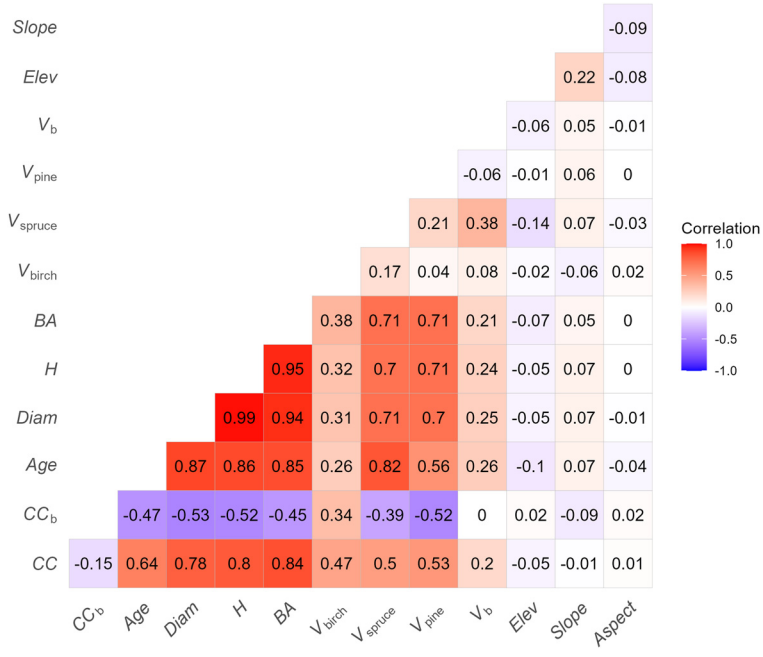

Fig. 3. A correlation diagram of the variables used (see Table 1).

positive relationships were between the variables *Age*, *Diam*, *H* and *BA* (0.85/0.99) (see Table 1 for definition of variables). The variables in this group were characterized by strong positive relationships with *CC* (>0.64), as well V_{spruce} and V_{pine}. All these variables were negatively related to CC_b.

Close relationship between *Diam* and *BA* follows from the definition of these variables (Mäkisara *et al.* 2022). The relationship between *H* and *Diam* is also well known (Mehtätalo 2005). The variables with the strongest correlations were removed (*Diam*, *H* and *BA*) and VIF test calculations were performed. These calculations showed the presence of *Age* and V_{spruce} values in the range of VIF = 5–7. *Age* has been used in many forest reflectance modelling studies (e.g. Grabska and Socha (2021)) and removing it from the list of variables would mean that results could not be compared. Therefore, V_{spruce} was removed from the variables. The final list of variables to be used in linear regression models was: *CC*, CC_b, *Age*, V_{birch}, V_{pine}, V_b, *Elev*, *Slope* and *Aspect*.

Our correlation analysis of the spectral reflectance in each band by month (Fig. 4) showed that the bands were divided into two distinct groups: red edge 2–narrow near-infrared (NIR) bands (B6–B8A) and all other bands (B2–B5, B11 and B12). This phenomenon was already known and was observed in Horler and Ahern (1986). The correlation coefficient values between bands within a group were noticeably higher than those between bands from different groups. This was especially noticeable from June to August (Fig. 4b–d). In May, the intergroup relationships were stronger than those in the summer months (Fig. 4a).

Effect of variables on stand reflectance

The linear regression models for the Sentinel-2 spectral bands are presented in the Supplementary material, Tables S1–S4. The reflectance in all bands was explained well by the variables in Table 1, with the maximum adjusted R-squared (R^2) value being 0.9 for the red edge 1 in July and the minimum R^2 value being 0.58 for the blue band in August (Fig. 5). The models best explain the reflectance in spectral bands B5 (red edge 1), B11 (short-wave infrared (SWIR) 1) and B12 (SWIR 2), where the fraction of explained variance does not fall below $R^2 = 0.82$ (B12 in August). For the B6 (red edge 2), B7 (red edge 3), B8 (NIR) and B8A (narrow NIR) spectral bands, R^2 increased from May to July and decreased to some extent in August. The models best explained the reflectance in the visible range bands in May ($R^2 > 0.84$). The lowest R^2 values are observed for reflectance models in green and red bands from June to August.

The influence of different variables on the reflectance in the different spectral bands (the t-value from Supplementary material, Tables S1–S4) and its change over time are shown in Figure 6.

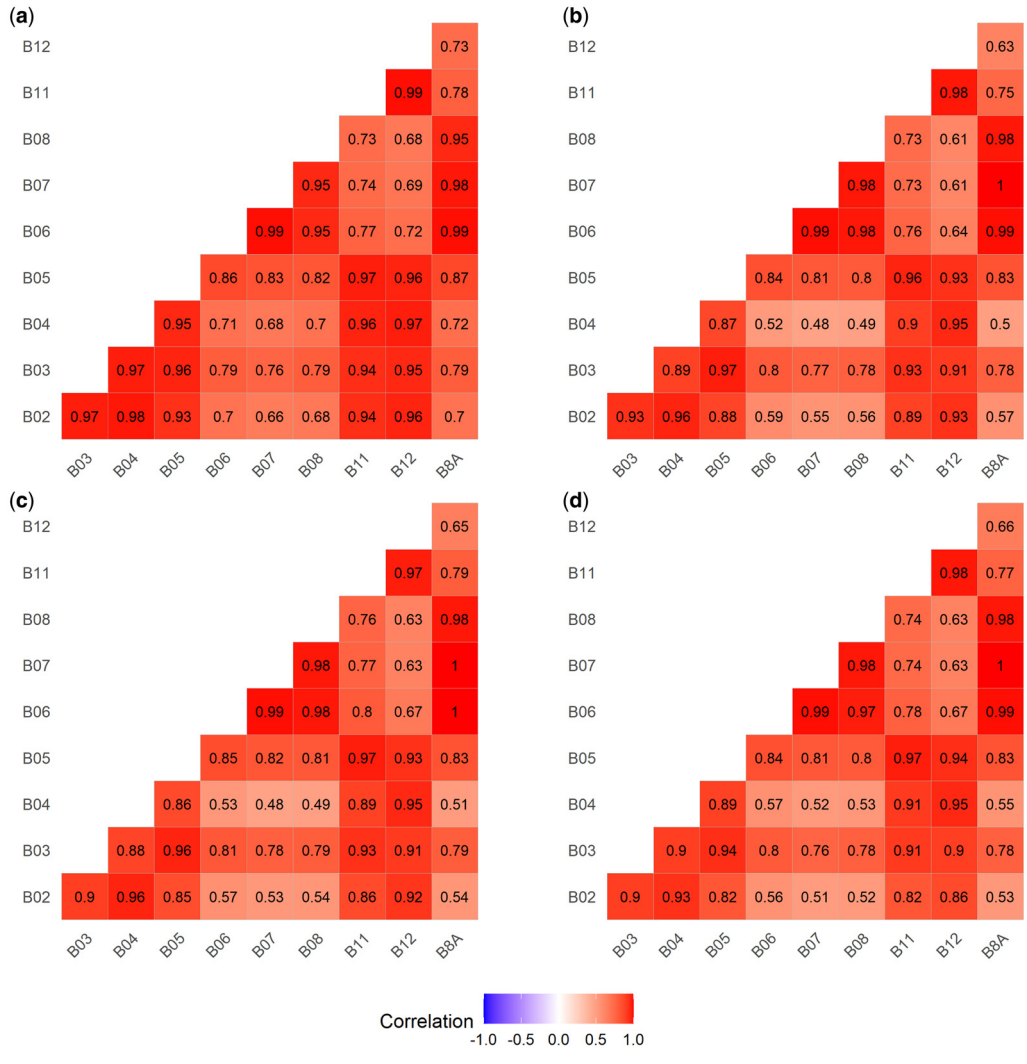

Fig. 4. The correlation of responses (see Table 2) by month: (**a**) May; (**b**) June; (**c**) July; (**d**) August.

Throughout the observation period, an increase in *CC* decreased the reflectance (Fig. 6). *CC* also strongly affected the reflectance in the SWIR and red edge 1 bands. This influence was somewhat weaker in the visible range of the electromagnetic spectrum. The *CC* variable influenced the reflectance in the above-mentioned bands more than the other variables. The influence of *CC* was much weaker in the red edge 2–NIR region, where it was inferior to the influence of CC_b and *Age*.

In the visible range bands, *CC* strongly influenced the reflectance in May. Then, the influence of *CC* decreased over time in the blue band. In the red band, the influence of *CC* decreased from May to July but increased again in August. In the green band, after the May peak there was a decrease in the influence of *CC* in June, an increase in July and another decrease in August. The influence of *CC* on the reflectance in the B5 and B11 bands increased over time, while it decreased in the B12 band. In the red edge 2–NIR region, the influence of *CC* on the reflectance increased from May to July, after which it decreased in August.

A decrease in the reflectance with an increase in stand age was observed throughout the observation period (Fig. 6). The influence of *Age* on the reflectance was most substantial in the red edge and NIR bands, followed by the green, SWIR, blue and red bands.

In the visible, red edge 1 and SWIR bands, the influence of *Age* over the entire observation period

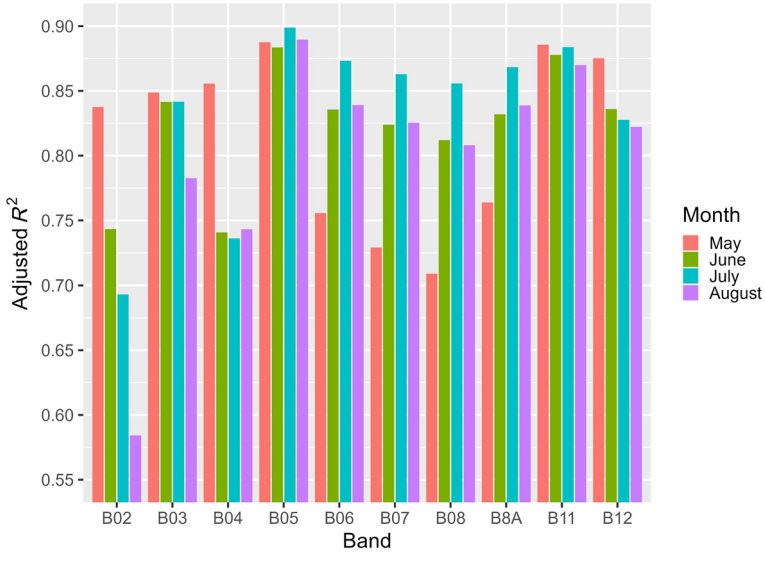

Fig. 5. Values of adjusted R-squared (R^2) linear models of spectral reflectance of Sentinel-2 Multispectral Instrument (MSI) bands.

was exceeded only by that of *CC*. The exceptions were the blue band in August, when the influence of *Age* was inferior to that of *CC* and *Elev*, and the SWIR bands in May, when the influence of *Age* was inferior to that of *CC* and CC_b. In the red edge 2–NIR bands, *Age* was the most important variable in May, and it was also characterized as having significant influence on the reflectance along with CC_b in the other months.

For the red and SWIR 2 bands, the influence of *Age* decreased from May to August. In the other bands, the influence of *Age* reached its maximum in June–July, after which it decreased in August.

The reflectance increased with the canopy cover of broad-leaved trees during the observation period (Fig. 6). The exception was the reflectance in the red band in July, for which a negative relationship with CC_b was observed. CC_b had the strongest effect on the reflectance in the red edge 2–NIR bands and somewhat weaker effects in the red edge 1, SWIR 1 and green bands.

The CC_b variable strongly influenced the reflectance in the visible, red edge 1 and SWIR bands in May (Fig. 6). This was apparently due to the influence of the understory: the higher CC_b, the better the understory could be seen in May when the broad-leaved trees in the study area started coming into leaf.

As a rule, the influence of tree volume on the reflectance was weaker than the influence of *CC*, CC_b and *Age* (Fig. 6). The exception was V_{spruce}, which influenced the reflectance through its relationship with *Age* (Fig. 3). In addition, the influence of the tree volume-related variables in the red edge 2–NIR ranges in May surpassed that of *CC*.

An increase in broad-leaved tree volume (V_{birch} and V_b) resulted in a decrease in the reflectance. The exception was the influence of V_{birch} on the reflectance in the red band in July–August. V_{birch} and V_b strongly influenced the reflectance in the red edge and NIR bands, although their influence on the reflectance in the green and SWIR 1 bands was weaker. Finally, the influence of broad-leaved tree volume on the reflectance was the weakest in the blue and red bands, as well as in the SWIR 2 band.

In contrast to the variables discussed above, elevation and slope had no influence on the reflectance throughout the entire observation period but demonstrated episodic effects. Because the study area was dominated by gently sloping terrain, the topographic variables influenced the reflectance much less than vegetation cover. In our case, *Elev* and *Slope* tended to have less of an influence than the other variables (Fig. 6). The only exception was the reflectance in the blue band in August, which only was exceeded by canopy cover in terms of influence. Elevation had the strongest effect on the reflectance in the visible region and the SWIR and red edge 1 bands, but its influence was much weaker in the red edge 2–NIR region.

Non-linear feature relationships

To analyse non-linear relationships, variables with a moderate to strong relationship between them ($0.4 < r \leq 0.9$) were selected (Fig. 3). Pairwise GAMs were generated for these variables. The pairs whose GAMs have adjusted $R^2 > 0.4$ and whose

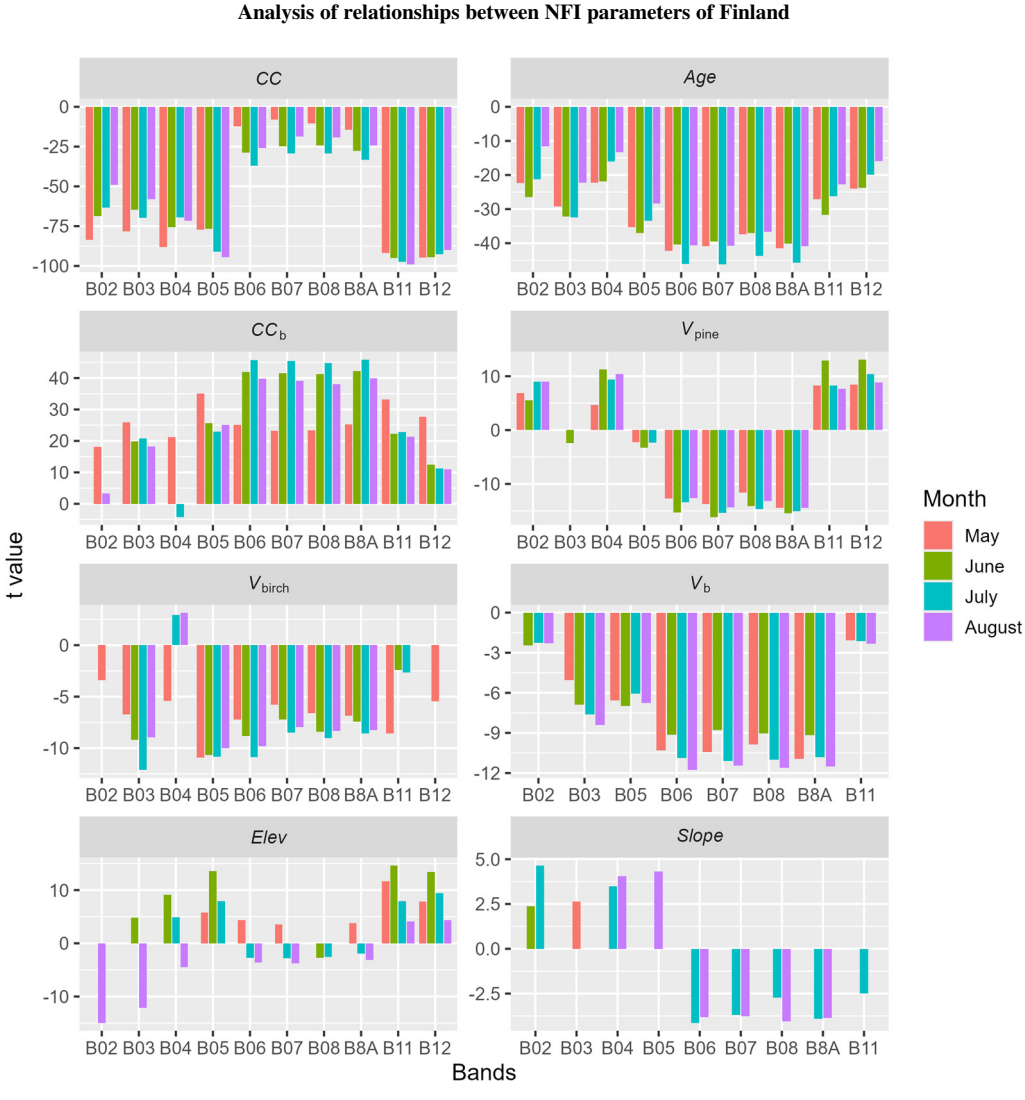

Fig. 6. The influence of different variables on the reflectance in single spectral ranges and its variation over time.

incrementally adjusted R^2 compared to the linear model (for the same variables) was >0.1 were further considered. There were five such pairs (Figs 7 & 8). Figure 7 shows the increase in pine stocks with age (there is a strong linear relationship between spruce stocks and age (Fig. 3), so that using GAM only slightly increases the adjusted R^2). Models relating the canopy cover to stand age and volume of growing stock have a common peculiarity (Fig. 8): the area of rapid growth at small values of age or stock volume is replaced by saturation when *CC* stops changing with increases in the other variables.

The GAM describing the relationship between *Age* and *CC* is shown in Figure 9 (adjusted $R^2 = 0.78$). It should be noted that a closed forest canopy ($CC > 60$) approximately corresponded to over 30–50 years of stand age. The growth and saturation regions on the plot can be divided by both *CC* and *Age*. The plots are divided by *Age*, and the age at which $dCC/dAge = 0.5$ is chosen as the threshold value (Age_{th}). This corresponds to $Age_{th} = 37$ years.

The dependence of the variables on age is considered in more detail. To do this, the data are divided into two classes by age: $Age > Age_{th}$ (class 0) and $Age \leq Age_{th}$ (class 1). The relationships between the variables change as a function of stand age is considered. The corresponding correlation matrices are shown in the Figure 10

A sample consisting of trees aged $Age > Age_{th}$ shows no strong relationships between *CC* and other features (Fig. 10b). The positive relationship

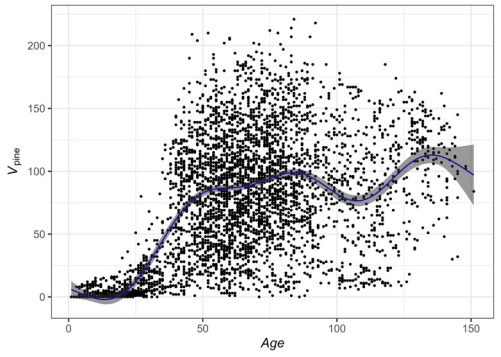

Fig. 7. The generalized additive model (GAM) describing the relationship between *Age* and V_{pine}. The shaded area around the curve represents the standard deviation.

between CC_b and V_{birch} strengthened significantly, while the relationships of CC_b with other traits weakened. The strong relationship between *Age* and V_{spruce} remains, while the relationship between *Age* and V_{pine} has become significantly weaker.

There is a positive relationship between canopy cover *CC* and CC_b (Fig. 10c) for stands whose age satisfies the condition $Age \leq Age_{th}$. The GAM describing the relationship between CC_b and *CC* at $Age > Age_{th}$ (class 0) and $Age \leq Age_{th}$ (class 1) is shown in Figure 11. The positive relationship between *CC* and the other features observed for the whole dataset remains and in most cases is strengthened for the $Age \leq Age_{th}$ stands. The volume of growing stock increases with increasing the age and canopy closure (Fig. 10c), while for stands with $Age > Age_{th}$ such a relationship is only observed for the pairs *Age* and V_{spruce}, as well as CC_b and V_{birch}.

Discussion

The obtained regularities of the influence of modelled stand features on canopy spectral reflectance, and changes in this influence over time, were compared with the known results on the relationships between canopy reflectance and real stand parameters in Finland (e.g. Kuusinen *et al.* 2014; Lukeš *et al.* 2014, 2016; Rautiainen and Lukeš 2015; Hovi *et al.* 2017; Rautiainen *et al.* 2018; Juola *et al.* 2022). Note that the observed regularities (or lack of them) may not only describe real forest

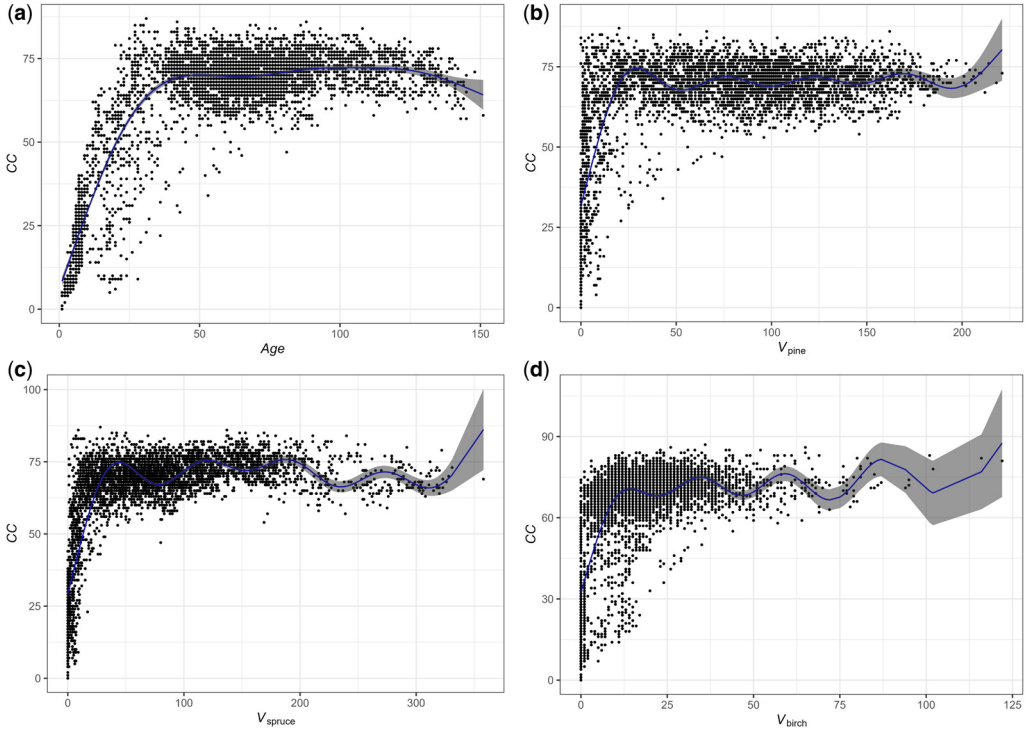

Fig. 8. The generalized additive model (GAM) describing the relationship between: (**a**) *Age* and *CC*; (**b**) V_{pine} and *CC*; (**c**) V_{spruce} and *CC*; (**d**) V_{birch} and *CC*. The shaded areas around the curve represent the standard deviation.

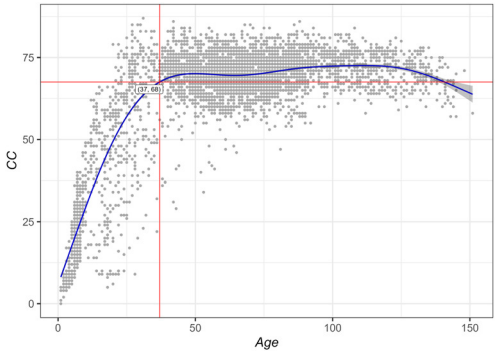

Fig. 9. The generalized additive model (GAM) describing the relationship between *Age* and *CC*. The shaded area around the curve represents the standard deviation.

stand features but also characterize the accuracy of the models used to derive them.

Canopy cover

A negative relationship between reflectance and forest stand canopy cover is observed in all considered spectral ranges during the observation period. This relationship is mentioned in many publications. For example, Lukeš *et al.* (2014) revealed negative relationships between spectral reflectance in the visible range and forest density characteristics, including canopy cover. They also found a weak negative correlation between NIR reflectance and canopy cover (Lukeš *et al.* 2016). A negative relationship between *CC* and reflectance in the visible and in red edge ranges was observed by Hallik *et al.* (2019). A strong negative correlation between canopy cover and reflectance in the visible, SWIR and thermal infrared (TIR) bands was observed by Butera (1986), in which a weaker correlation between canopy cover and NIR reflectance was also determined.

One possible explanation for the decrease in forest stands reflectance with increasing *CC* is the influence of understory. The smaller the canopy cover, the more the understory is seen due to its reflectance, which is higher than that of the predominantly coniferous trees in the overstory. It is known that the contribution of forest litter and understories in sparse conifer stands can explain up to 50% of total forest reflectance (Rautiainen and Lukeš 2015).

Dry grass and bare soil are characterized by higher reflectance in the visible range than turgid vegetation (Jensen 2007). This could explain the strong influence of *CC* on the reflectance in the visible range of the electromagnetic spectrum that was observed in May (Butera 1986). The weaker influence of *CC* on the reflectance in June could be explained by understory greening (Croft *et al.* 2014; Testa *et al.* 2018). This intensified in July, which was confirmed by the growing influence of *CC* on the reflectance in the green band and its decreasing influence on the reflectance in the red band. At the same time, from May to July, the influence of *CC* on the reflectance in the red edge and NIR bands increased. Finally, the influence of *CC* on the reflectance in the red band that was observed in August could have been caused by the increase in the reflectance in the red range associated with chlorophyll disintegration in the leaves. Meanwhile, there was a decrease in the influence of *CC* on the reflectance in the NIR band. Thus, it could be assumed that the observed variations in the influence of *CC* (Fig. 6) mainly followed seasonal changes in understory conditions.

The *CC* saturation effect observed at $CC > 60$–70 (Fig. 9) appears to be related to canopy closure. Lukeš *et al.* (2014, 2016) revealed a weak reflectance sensitivity to canopy cover changes after canopy closure. After full canopy closure, the LAI stops

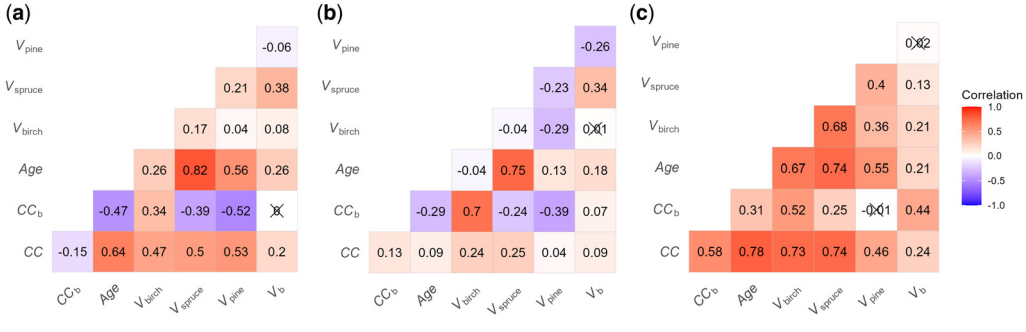

Fig. 10. Correlation matrices of variables: (**a**) without division by *Age*; (**b**) $Age > Age_{th}$; (**c**) $Age \leq Age_{th}$. Insignificant correlation coefficients (*P*-value >0.05) are crossed through.

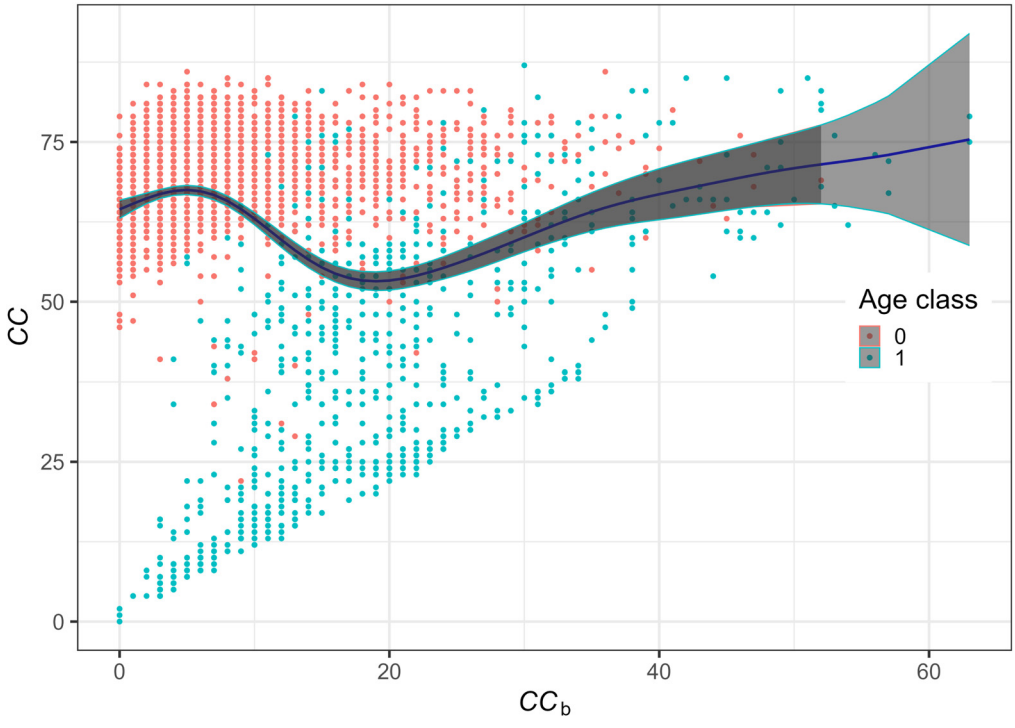

Fig. 11. The generalized additive model (GAM) describing the relationship between CC_b and CC for two classes: $Age > Age_{th}$ (class 0) and $Age \leq Age_{th}$ (class 1). The shaded area around the curve represents the standard deviation.

changing noticeably but the stemwood volume and tree height continue to increase (Lang *et al.* 2016). This explains the non-linear relationship between *CC* and volumes of growing stocks (Fig. 7).

Stand age

Observed decrease in forest canopy reflectance with an increase in stand age has been mentioned by many researchers (Kuusinen *et al.* 2014; Rautiainen *et al.* 2018; Hallik *et al.* 2019). Kuusinen *et al.* (2014) described the spruce and pine albedo in the visible, NIR and SWIR ranges as an exponentially decreasing function of stand age. One of the possible explanations for the decrease in forest reflectance with increasing stand age is the close relationship between stand age and canopy cover (Fig. 3) (Rautiainen *et al.* 2018).

According to Mäkelä (1997), Scots pine stands reached their maximum crown coverage by approximately 20–50 years of age, after which the magnitude and relative proportions of the crown and understory reflectance remained stable. The GAM describing the relationship between *CC* and *Age* (Fig. 9) also demonstrates a saturation effect of *CC* approximately corresponded to over 30–50 years of stand age.

Amiro *et al.* (2006) suggested that the high albedo of young stands could be due to the combined effects of the low LAI and the high proportion of broad-leaved trees in young forests. As conifer stands age, broad-leaved trees are removed, resulting in a decrease in forest albedo. In addition, with the increase in forest age, the proportion of pioneer tree species (mainly broad-leaved species) in the stand gradually decreases, while the share of spruce increases (Linder *et al.* 1997). Thus, forest reflectance can decrease with increasing stand age due to plant care and natural changes in species composition.

According to Hallik *et al.* (2019), the best stand age predictor in summer season was the reflectance in single bands with a central wavelength of 751 or 783 nm. The latter corresponded to the red edge 3 band of the Sentinel-2 MSI (Table 2). Our results showed that the reflectance in the red edge 3 and narrow NIR bands in July was the most closely related to stand age (Fig. 6).

For young trees with $Age \leq Age_{th}$, wood stock volumes increase with age (Fig. 10c). After presumed crown closure (at $Age > Age_{th}$), a strong positive relationship is observed only between age

and V_{spruce}, which is the dominant species in the study area.

The canopy cover of broad-leaved trees

The stand reflectance increased with the proportion of broad-leaved tree cover (Fig. 6) since the reflectance of broad-leaved trees is stronger than that of conifers (Rautiainen *et al.* 2004).

In the red edge 2–NIR bands, the influence of CC_b increased from May to July and decreased in August. In May, the leaves on the trees in the study area were just beginning to appear and the effect of CC_b on the reflectance in the red edge 2–NIR range was the smallest out of entire observation period. Reflectance in NIR bands increases with LAI (Asner 1998). One of the reasons for a decrease in reflectance in the NIR range is the decrease in healthy leaf biomass, which starts in August (Marx and Kleinschmit 2017).

For young and sparse forests, an increase in CC_b results in an increase in *CC* (Fig. 11, class 1). After crown closure, the linear relationship between CC_b and *CC* disappears (Fig. 11, class 0). At the same time, there is a positive relationship between CC_b and V_{birch} as the dominant broad-leaved tree species.

The volume of tree species

Lukeš *et al.* (2016) observed a negative correlation between reflectance in the visible range and total forest biomass. Hallik *et al.* (2019) noted that some forest inventory parameters related to the biomass volume, mainly stem volume, were characterized by weak negative correlations with reflectance near the maximum pigment absorption point (i.e. in the blue and red bands). A stronger negative relationship with reflectance was observed in the green and red edge regions. Similar regularities were observed for V_{spruce} (due to its close relationship with *Age*), V_{birch} and V_b. In the red edge and NIR bands, the reflectance also decreased as V_{pine} increased (V_{pine} influenced the reflectance most strongly in these bands). On the contrary, in the blue and red bands and the SWIR region, an increase in V_{pine} caused an increase in the reflectance.

The negative relationship between reflectance in the NIR range and forest biomass disagrees with the assumption that an increase in biomass increases the reflectance in the NIR range, which is the cornerstone of classical vegetation indices. Nevertheless, a negative relationship between NIR reflectance and biomass has been observed and discussed in many publications, including studies on forests in Finland (Lukeš *et al.* 2016; Rautiainen *et al.* 2018).

The relationship between the reflectance in the red band and V_{birch} was negative in May but positive in July–August. The latter could be explained by the influence of seasonal changes in leaf pigmentation (Jensen 2007).

Young forests with low timber stocks are characterized by low *CC* values and the predominant influence of understory reflectance, which surpasses forest reflectance. The *CC* value increases with timber volume, while reflectance decreases due to the increased influence of the forest canopy. For closed forest canopies, the main factors that affect scattering are the canopy structure and multiple scattering at the crown level (Rautiainen *et al.* 2018). The latter increases the percentage of absorbed radiation with the increase in wood volume and causes a decrease in canopy reflectance.

Elevation and slope

Because the study area was dominated by gently sloping terrain, the topographic variables influenced the reflectance much less than vegetation cover (Karaska *et al.* 1986). In our case, *Elev* and *Slope* tended to have less of an influence than the other variables (Fig. 6). The influence of *Slope* on reflectance is significantly greater than *Elev*, and occurs sporadically.

Note that other factors, in particular soil temperature and moisture, influence stand reflectance through terrain variables. Thus, in the red edge 1 and SWIR bands, a positive relationship was observed between the reflectance and elevation. An increase in SWIR reflectance can be caused by a decrease in leaf moisture content with increasing elevation (Jensen 2007). A positive correlation between red edge 1 reflectance in spring and elevation was also observed (Grabska and Socha 2021).

Conclusions

An approach is proposed to investigate linear and non-linear relationships between forest inventory parameters and satellite sensor reflectance based on multiple linear regression and GAM models. This approach can be applied both to forest stand parameters measured during ground-based inventory and to values of these parameters derived from a joint ground-based and satellite inventory.

The relationships between MS-NFI national forest inventory parameters and stand reflectance in the visible, red edge, NIR and SWIR spectral regions were investigated for the Sentinel-2 MSI satellite sensor.

In the case study of mesic forest in the Northern Ostrobothnia, among the MS-NFI parameters the main factors that influenced the forest reflectance in the Sentinel-2 MSI spectral bands were canopy cover, stand age and the canopy cover of broad-leaved trees. In the visible, SWIR and red edge 1

bands, the influence of canopy cover was dominant, whereas in the other red edge bands and in the NIR region, stand age and the canopy cover of broad-leaved trees influenced the reflectance more strongly.

An increase in canopy cover caused a decrease in the forest reflectance in the considered spectral bands. A possible explanation for this is a reduction in understory lighting. Seasonal changes in the influence of canopy cover on reflectance also follow seasonal changes in understory conditions.

The decrease in stand reflectance with age for stands younger than 30–50 years is due to the close relationship between age and canopy cover. For stands older than 30–50 years with $CC > 60$–70, the relationship of canopy cover with age and other stand parameters disappears, which is presumably caused by forest canopy closure. In all considered spectral bands, positive relationships were found between the reflectance and the canopy cover proportion of broad-leaved trees.

The forest canopy reflectance in the red edge and NIR bands decreased with an increase in the volume of all tree types. A negative relationship was also observed between the reflectance in the visible and SWIR regions and V_b. On the contrary, an increase in V_{pine} caused an increase in the reflectance in the blue and red bands and the SWIR region.

In stands younger than 30–50 years, the volume of trees increases with age. After 30–50 years, the volume of trees of the dominant species (spruce) in the area increases with age. With increasing CC_b, V_{birch} increases.

The area under consideration has a gently sloping landscape and topographic variables influence the reflectance of the forest canopy much more weakly than stand parameters. The research results will help to better understand the effects of forest stand features and terrain conditions on forest reflectance, which is important for application of remote-sensing technologies in forest monitoring. The proposed methodology can be enhanced and used to monitor changes in forests located near mines or other industrial facilities. It can also be used to detect variations in spectral reflectance of plants caused by exposure to potentially toxic elements.

Acknowledgements The authors would like to thank Maria Hänninen, the environmental manager at the Pyhäsalmi Mine, for specifying the locations for measurements and study planning, the University of Oulu Kerttu Saalasti Institute for supporting cooperation at the Pyhäsalmi site, as well as the OPT/NET BV company (opt-net.eu) and GOLDEN-AI platform for the Sentinel-2 data. The authors would also like to thank the European Commission, the European Space Agency and the Copernicus Programme for providing the Sentinel-2 data.

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

Author contributions **OK**: methodology (equal), validation (equal), visualization (equal), writing – original draft (equal), writing – review & editing (equal); **DK**: methodology (equal), validation (equal), visualization (equal), writing – original draft (equal), writing – review & editing (equal); **KS**: methodology (equal), validation (equal), visualization (equal), writing – original draft (equal), writing – review & editing (equal).

Funding This work was funded by the European Union's Horizon 2020 research and innovation programme, under grant agreement number 869398 ('Earth observation and Earth GNSS data acquisition and processing platform for safe, sustainable and cost-efficient mining operations' (GoldenEye)).

Data availability All data generated or analysed during this study are included in this published article (and, if present, its supplementary information files).

References

Amiro, B.D., Orchansky, A.L. *et al*. 2006. The effect of post-fire stand age on the boreal forest energy balance. *Agricultural and Forest Meteorology*, **140**, 41–50, https://doi.org/10.1016/j.agrformet.2006.02.014

Asner, G.P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*, **64**, 234–253, https://doi.org/10.1016/s0034-4257(98)00014-5

Avitabile, V., Baccini, A., Friedl, M.A. and Schmullius, C. 2012. Capabilities and limitations of Landsat and land cover data for aboveground woody biomass estimation of Uganda. *Remote Sensing of Environment*, **117**, 366–380, https://doi.org/10.1016/j.rse.2011.10.012

Böttcher, K., Markkanen, T. *et al*. 2016. Evaluating biosphere model estimates of the start of the vegetation active season in boreal forests by satellite observations. *Remote Sensing*, **8**, 580, https://doi.org/10.3390/rs8070580

Böttcher, K., Rautiainen, K. *et al*. 2018. Proxy indicators for mapping the end of the vegetation active period in boreal forests inferred from satellite-observed soil freeze and ERA-interim reanalysis air temperature. *Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, **86**, 169–185. https://doi.org/10.1007/s41064-018-0059-y

Butera, M. 1986. A correlation and regression analysis of percent canopy closure versus TMS spectral response for selected forest sites in the San Juan National Forest, Colorado. *IEEE Transactions on Geoscience and Remote Sensing*, **GE-24**, 122–129, https://doi.org/10.1109/tgrs.1986.289693

Coops, N.C., Tompalski, P., Goodbody, T.R.H., Achim, A. and Mulverhill, C. 2022. Framework for near real-time

forest inventory using multi source remote sensing data Fassnacht, F. (ed.). *Forestry: An International Journal of Forest Research*, **96**, 1–19, https://doi.org/10.1093/forestry/cpac015

Copernicus Tree Cover Density 2018 (raster 10 m), Europe, 3-yearly, https://doi.org/10.2909/486f77da-d605-423e-93a9-680760ab6791

Croft, H., Chen, J.M. and Zhang, Y. 2014. Temporal disparity in leaf chlorophyll content and leaf area index across a growing season in a temperate deciduous forest. *International Journal of Applied Earth Observation and Geoinformation*, **33**, 312–320, https://doi.org/10.1016/j.jag.2014.06.005

Ehlers, D., Wang, C. *et al*. 2022. Mapping forest aboveground biomass using multisource remotely sensed data. *Remote Sensing*, **14**, 1115, https://doi.org/10.3390/rs14051115

Faraway, J.J. 2014. *Linear Models with R*. 2nd edn. Chapman & Hall/CRC Texts in Statistical Science.

Fox, J. and Weisberg, S. 2019. *An R Companion to Applied Regression*. 3rd edn. Sage Publications.

Gates, D.M., Keegan, H.J., Schleter, J.C. and Weidner, V.R. 1965. Spectral properties of plants. *Applied Optics*, **4**, 11, https://doi.org/10.1364/ao.4.000011

Gschwantner, T., Alberdi, I. *et al*. 2022. Growing stock monitoring by European National Forest Inventories: historical origins, current methods and harmonisation. *Forest Ecology and Management*, **505**, 119 868, https://doi.org/10.1016/j.foreco.2021.119868

Grabska, E. and Socha, J. 2021. Evaluating the effect of stand properties and site conditions on the forest reflectance from Sentinel-2 time series Bosela, M. (ed.). *PLoS ONE*, **16**, e0248459, https://doi.org/10.1371/journal.pone.0248459

Hallik, L., Kuusk, A., Lang, M. and Kuusk, J. 2019. Reflectance properties of hemiboreal mixed forest canopies with focus on red edge and near infrared spectral regions. *Remote Sensing*, **11**, 1717, https://doi.org/10.3390/rs11141717

Hijmans, R.J. 2022. Terra: spatial data analysis. R package version 1.5-21, https://CRAN.R-project.org/package=terra

Horler, D.N.H. and Ahern, F.J. 1986. Forestry information content of Thematic Mapper data. *International Journal of Remote Sensing*, **7**, 405–428, https://doi.org/10.1080/01431168608954695

Hovi, A., Raitio, P. and Rautiainen, M. 2017. A spectral analysis of 25 boreal tree species. *Silva Fennica*, **51**, https://doi.org/10.14214/sf.7753

Jagodziński, A.M., Dyderski, M.K., Gesikiewicz, K. and Horodecki, P. 2019. Effects of stand features on aboveground biomass and biomass conversion and expansion factors based on a Pinus sylvestris L. chronosequence in Western Poland. *European Journal of Forest Research*, **138**, 673–683, https://doi.org/10.1007/s10342-019-01197-z

James, G., Witten, D., Hastie, T. and Tibshirani, R. 2014. *An Introduction to Statistical Learning: With Applications in R*. Springer Texts in Statistics.

Jensen, J.R. 2007. *Remote Sensing of the Environment: An Earth Resource Perspective*. Prentice Hall Series in Geographic Information Science.

Juola, J., Hovi, A. and Rautiainen, M. 2022. A spectral analysis of stem bark for boreal and temperate tree species. *Ecology and Evolution*, **12**, https://doi.org/10.1002/ece3.8718

Kangas, A., Pitkänen, T.P., Mehtätalo, L. and Heikkinen, J. 2022. Mixed linear and non-linear tree volume models with regional parameters to main tree species in Finland. *Forestry: An International Journal of Forest Research*, **96**, 188–206, https://doi.org/10.1093/forestry/cpac038

Karaska, M.A., Walsh, S.J. and Butler, D.R. 1986. Impact of environmental variables on spectral signatures acquired by the LANDSAT thematic mapper. *International Journal of Remote Sensing*, **7**, 1653–1667, https://doi.org/10.1080/01431168608948959

Kontula, T. and Raunio, A. (eds) 2019. *Threatened Habitat Types in Finland 2018 – Red List of Habitats Results and Basis for Assessment*. Finnish Environment Institute and Ministry of the Environment.

Kuhn, M. and Johnson, K. 2013. *Applied Predictive Modeling*. SpringerLink: Bücher, https://doi.org/10.1007/978-1-4614-6849-3

Kuusinen, N., Tomppo, E., Shuai, Y. and Berninger, F. 2014. Effects of forest age on albedo in boreal forests estimated from MODIS and Landsat albedo retrievals. *Remote Sensing of Environment*, **145**, 145–153, https://doi.org/10.1016/j.rse.2014.02.005

Lang, M., Lilleleht, A. *et al*. 2016. Estimation of aboveground biomass in forest stands from regression on their basal area and height. *Forestry Studies*, **64**, 70–92, https://doi.org/10.1515/fsmu-2016-0005

Liepinš, J., Lazdinš, A., Kaleja, S. and Liepinš, K. 2022. Species composition affects the accuracy of stand-level biomass models in Hemiboreal forests. *Land*, **11**, 1108, https://doi.org/10.3390/land11071108

Linder, P., Elfving, B. and Zackrisson, O. 1997. Stand structure and successional trends in virgin boreal forest reserves in Sweden. *Forest Ecology and Management*, **98**, 17–33, https://doi.org/10.1016/s0378-1127(97)00076-5

Lister, A.J., Andersen, H. *et al*. 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**, 1364, https://doi.org/10.3390/f11121364

Lukeš, P., Rautiainen, M., Manninen, T., Stenberg, P. and Mõttus, M. 2014. Geographical gradients in boreal forest albedo and structure in Finland. *Remote Sensing of Environment*, **152**, 526–535, https://doi.org/10.1016/j.rse.2014.06.023

Lukeš, P., Stenberg, P., Mõttus, M., Manninen, T. and Rautiainen, M. 2016. Multidecadal analysis of forest growth and albedo in boreal Finland. *International Journal of Applied Earth Observation and Geoinformation*, **52**, 296–305, https://doi.org/10.1016/j.jag.2016.07.001

Mäkelä, A. 1997. A carbon balance model of growth and self-pruning in trees based on structural relationships. *Forest Science*, **43**, 7–24, https://doi.org/10.1093/forestscience/43.1.7

Mäkisara, K., Katila, M. and Peräsaari, J. 2019. *The Multi-Source National Forest Inventory of Finland – Methods and Results 2015*. Natural Resources Institute Finland (Luke).

Mäkisara, K., Katila, M. and Peräsaari, J. 2022. *The Multi-Source National Forest Inventory of Finland – Methods*

and Results 2017 and 2019. Natural Resources Institute Finland (Luke).

Maltamo, M., Packalen, P. and Kangas, A. 2021. From comprehensive field inventories to remotely sensed wall-to-wall stand attribute data – a brief history of management inventories in the Nordic countries. *Canadian Journal of Forest Research*, **51**, 257–266, https://doi.org/10.1139/cjfr-2020-0322

Marx, A. and Kleinschmit, B. 2017. Sensitivity analysis of RapidEye spectral bands and derived vegetation indices for insect defoliation detection in pure Scots pine stands. *iForest – Biogeosciences and Forestry*, **10**, 659–668, https://doi.org/10.3832/ifor1727-010

Mehtätalo, L. 2005. Height-diameter models for Scots pine and birch in Finland. *Silva Fennica*, **39**, https://doi.org/10.14214/sf.395

Meschiari, S. 2022. latex2exp: Use LaTeX Expressions in Plots, https://CRAN.R-project.org/package=latex2exp

Nath, B. and Ni-Meister, W. 2021. The interplay between canopy structure and topography and its impacts on seasonal variations in surface reflectance patterns in the Boreal Region of Alaska – implications for surface radiation budget. *Remote Sensing*, **13**, 3108, https://doi.org/10.3390/rs13163108

Ollinger, S.V. 2010. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist*, **189**, 375–394, https://doi.org/10.1111/j.1469-8137.2010.03536.x

Pebesma, E. 2018. Simple features for R: standardized support for spatial vector data. *The R Journal*, **10**, 439–446, https://doi.org/10.32614/RJ-2018-009

Peltola, A., Ihalainen, A. *et al*. 2019. *Suomen Metsätilastot: Finnish Forest Statistics 2019*. Natural Resources Institute Finland (Luke).

Rautiainen, M. and Lukeš, P. 2015. Spectral contribution of understory to forest reflectance in a boreal site: an analysis of EO-1 Hyperion data. *Remote Sensing of Environment*, **171**, 98–104, https://doi.org/10.1016/j.rse.2015.10.009

Rautiainen, M., Stenberg, P., Nilson, T. and Kuusk, A. 2004. The effect of crown shape on the reflectance of coniferous stands. *Remote Sensing of Environment*, **89**, 41–52, https://doi.org/10.1016/j.rse.2003.10.001

Rautiainen, M., Lukeš, P., Homolová, L., Hovi, A., Pisek, J. and Mõttus, M. 2018. Spectral properties of coniferous forests: a review of in situ and laboratory measurements. *Remote Sensing*, **10**, 207, https://doi.org/10.3390/rs10020207

Repola, J. 2009. Biomass equations for Scots pine and Norway spruce in Finland. *Silva Fennica*, **43**, https://doi.org/10.14214/sf.184

Rodriguez-Veiga, P., Wheeler, J., Louis, V., Tansey, K. and Balzter, H. 2017. Quantifying forest biomass carbon stocks from space. *Current Forestry Reports*, **3**, 1–18, https://doi.org/10.1007/s40725-017-0052-5

Roy, P.S. 1989. Spectral reflectance characteristics of vegetation and their use in estimating productive potential. *Proceedings/Indian Academy of Sciences*, **99**, 59–81, https://doi.org/10.1007/bf03053419

Schumacher, J., Hauglin, M., Astrup, R. and Breidenbach, J. 2020. Mapping forest age using National Forest Inventory, airborne laser scanning, and Sentinel-2 data. *Forest Ecosystems*, **7**, https://doi.org/10.1186/s40663-020-00274-9

Testa, S., Soudani, K., Boschetti, L. and Mondino, E.B. 2018. MODIS-derived EVI, NDVI and WDRVI time series to estimate phenological metrics in French deciduous forests. *International Journal of Applied Earth Observation and Geoinformation*, **64**, 132–144, https://doi.org/10.1016/j.jag.2017.08.006

Tomppo, E., Heikkinen, J. *et al*. 2011. *Designing and Conducting a Forest Inventory - Case: 9th National Forest Inventory of Finland*. Springer, https://doi.org/10.1007/978-94-007-1652-0

Tomppo, E., Katila, M., Mäkisara, K. and Peräsaari, J. 2014. *The Multi-Source National Forest Inventory of Finland – Methods and Results 2011*. Natural Resources Institute Finland (Luke).

Tonteri, T., Hotanen, J.-P. and Kuusipalo, J. 1990. The Finnish forest site type approach: ordination and classification studies of mesic forest sites in southern Finland. *Vegetatio*, **87**, 85–98, https://doi.org/10.1007/bf00045658

Wallner, A., Friedrich, S. *et al*. 2021. A remote sensing-guided forest inventory concept using multispectral 3D and height information from {ZiYuan}-3 satellite data. *Forestry: An International Journal of Forest Research*, **95**, 331–346, https://doi.org/10.1093/forestry/cpab055

Watson, J.E.M., Evans, T. *et al*. 2018. The exceptional value of intact forest ecosystems. *Nature Ecology & Evolution*, **2**, 599–610, https://doi.org/10.1038/s41559-018-0490-x

White, J.C., Coops, N.C., Wulder, M.A., Vastaranta, M., Hilker, T. and Tompalski, P. 2016. Remote sensing technologies for enhancing forest inventories: a review. *Canadian Journal of Remote Sensing*, **42**, 619–641, https://doi.org/10.1080/07038992.2016.1207484

Widlowski, J.-L., Verstraete, M., Pinty, B. and Gobron, N. 2003. *Allometric Relationships of Selected European Tree Species*. European Commission Joint Research Centre/Institute for Environment and Sustainability.

Wickham, H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer.

Wood, S.N. 2017. *Generalized Additive Models: An Introduction with R*, 2nd edn. Chapman and Hall/CRC Press.