Luke
NATURAL RESOURCES INSTITUTE FINLAND

Jukuri, open repository of the Natural Resources Institute Finland (Luke)

This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.

Author(s): Timo P. Pitkänen, Andras Balazs & Sakari Tuominen

Title: Automatized Sentinel-2 mosaicking for large area forest mapping

Year: 2024

Version: Published version

Copyright: The Author(s) 2024

Rights: CC BY 4.0

Rights url: http://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Pitkänen, T. P., Balazs, A., & Tuominen, S. (2024). Automatized Sentinel-2 mosaicking for large area forest mapping. International Journal of Applied Earth Observation and Geoinformation, 127, 103659. https://doi.org/10.1016/j.jag.2024.103659

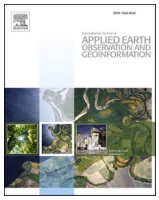
All material supplied via *Jukuri* is protected by copyright and other intellectual property rights. Duplication or sale, in electronic or print form, of any part of the repository collections is prohibited. Making electronic or print copies of the material is permitted only for your own personal use or for educational purposes. For other purposes, this article may be used in accordance with the publisher's terms. There may be differences between this version and the publisher's version. You are advised to cite the publisher's version.

International Journal of Applied Earth Observation and Geoinformation 127 (2024) 103659

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Automatized Sentinel-2 mosaicking for large area forest mapping

Timo P. Pitkänen *, Andras Balazs, Sakari Tuominen

Natural Resources Institute Finland (Luke), Latokartanonkaari 9 00790, Helsinki, Finland

ARTICLE INFO

Keywords:
Sentinel-2
Image mosaicking
Forest inventory
Feature prediction
Google Earth Engine

ABSTRACT

Creating maps of forest inventory variables is commonly taking advantage of satellite images, which are mosaicked together for gaining larger coverage. Recently, mosaicking has increasingly shifted towards user friendly cloud-based online environments such as Google Earth Engine (GEE), which are equipped with huge image repositories and extensive processing capabilities. This enables the easy transferability of workflows into new image sets and diversifies the range of methodological options for mosaicking. The quality control of the output mosaic, ensuring that the reflectance values are representative to the targeted land cover, is however primarily based on certain assumptions or pre-set rules which may not always produce an optimal result. Our study focuses on assessing and comparing the performance of three different mosaicking algorithms for predicting forest inventory variables, based on an extensive set of field data on the main site type, fertility class, and volume and biomass of growing stock. One of the compared mosaics derives from manual image selection, thus enabling rigorous visual quality control, and two others are resting on GEE-assisted automatized methods which include applying a percentile-based statistic over all the input reflectance values and selecting the best pixels using predefined quality indicators. The results indicate that the manual and the percentile-based mosaics are generally providing the best and relatively equal performance levels. Compared to them, the quality-based mosaic has slightly lower accuracy particularly when predicting continuous variables (*i.e.*, the volume and biomass of growing stock) and it suffers from minor image defects. For the total volume of growing stock, for example, the RMS errors are 56.22 % for the manual, 56.33 % for the percentile-based, and 59.47 % for the quality-based mosaics, respectively. These results indicate that from the perspective of large area forest mapping, automatically generated mosaics may provide approximately similar accuracy as compared to manually controlled workflow at a fraction of the workload.

1. Introduction

In forest inventory and monitoring, images from medium-resolution optical satellites, such as Landsat-8 and Sentinel-2, are widely applied for modelling and mapping spatially scattered field observations into larger coverages (Kangas et al., 2017, McRoberts and Tomppo, 2007, Wulder, 1998). In this context, a model is built to predict the targeted inventory variable based on the reflectance values, which allows producing thematic maps and aggregating them further into, for example, stand-wise mean values (Ohmann et al., 2014, Mäkisara et al., 2022). As the mapped area may be large, the coverage of a single scene is often insufficient and combining information from several images is needed. In this case, the challenge is to produce a consistent result over the whole area of interest, given that reflectance characteristics of similar land cover types may vary, particularly if the input images are acquired on several dates.

These reflectance differences result from multiple reasons. One principal characteristic of passive optical sensors is their sensitivity to prevailing weather conditions such as clouds, atmospheric water vapor or haze at the time of image acquisition, as well as to illumination differences depending on the solar angle, soil moisture, terrain topography, and existence of shadows. Additionally, reflectance values derive normally from multiple land cover components within the pixel area, and they may not remain stable over longer periods of time. The primary challenge for prediction and mapping applications depends on the study region; for example, the tropics suffer from persistent cloud coverage which limits the usable coverage of each image, mountainous areas are requiring specific attention to correcting topography-derived spatial and spectral distortions, and areas with strong seasonal phenological variation during the growing season (frondescence and seasonal yellowing of leafy vegetation due to drought or autumn) pose significant restrictions for the annual span of suitable images (Kalinaki et al., 2023, Santini and

* Corresponding author.
E-mail addresses: timo.p.pitkanen@luke.fi (T.P. Pitkänen), andras.balazs@luke.fi (A. Balazs), sakari.tuominen@luke.fi (S. Tuominen).

https://doi.org/10.1016/j.jag.2024.103659
Received 19 October 2023; Received in revised form 4 January 2024; Accepted 13 January 2024
Available online 22 January 2024
1569-8432/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Palombo, 2022, Zhang et al., 2023).

While major part of the reflectance variation between multiple images can be minimized by selecting and carefully preprocessing the most optimal images, this may not remove all the differences derived from their dissimilar acquisition conditions. One option for reducing prediction errors among multiple image frames to process each image scene separately, or as a batch of scenes derived from the same satellite overpass, which removes temporal variation and promotes constant atmospheric conditions (*e.g.*, Mäkisara et al., 2022, Tomppo et al., 2008). Prediction is then performed separately for each image or batch and the resulting pixel level estimates are mosaicked into a larger map. This approach may, however, become quickly unfeasibly tedious if the number of images is large, or they have numerous clouds or other defects, which finally generates a complex patchwork of individual scenes (Helmer and Ruefenacht, 2005). Furthermore, field observations used for variable prediction can only be selected within the focused scene or overpass, which complicates the prediction process and reduces the areal representativeness of the modelling data.

Alternatively, the single scenes may be mosaicked first, referring to spatial juxtaposition of initial images and merging their overlapping areas, which allows variable prediction over the whole area at once (Vaudour et al., 2021). If the distinct image scenes are expected to be spectrally representative and having consistent geometries, mosaicking operation is principally a matter of balancing their radiometric differences, determining seam lines between the images, and deciding a strategy how to handle the pixel values in the overlapping areas (Li et al., 2019, Yu et al., 2017).

For prediction applications based on reflectance values, however, the radiometric consistency of the final product is essential. Major part of this is the reliable detection of clouds and their shadows, followed by substituting their areas from other images, where the goal of mosaicking is to improve the spatial continuity rather than to enlarge the image area (Li et al., 2019). As the manual delineation of these gaps is slow and laborious, various automatic cloud detection methods have been developed. These range from simple thresholding approaches to complicated schemes, where clouds (and their shadows) may be detected by their reflectance properties, shape, texture, and solar illumination direction (Li et al., 2020, Segal-Rozenhaimer et al., 2020). While automatic methods provide an easy way to produce predominantly good results, detection accuracy may be decreased, for example, by detecting high albedo surfaces as clouds, or failing to properly recognize thinner clouds (Segal-Rozenhaimer et al., 2020, Shen et al., 2014, Skakun et al., 2021).

Recently, a range of web-based platforms have also became available to facilitate the production of seamless and cloudless mosaics via an online environment. Perhaps the most widely used of them is Google Earth Engine (GEE; https://code.earthengine.google.com/) which offers a wide range of satellite products, processing tools, support for coding and potential for the collaborative use of analysis resources, thus enabling the user to self-build the desired mosaicking workflows (Chen and Zhao, 2022; Tamiminia et al., 2020). A mosaic may be built using the best pixels by predefined quality criteria, which are often connected to quality bands, band-wise reflection ratios, or indices, such as the normalized difference vegetation index (DeLancey et al., 2019, Helmer and Ruefenacht, 2005, Schmitt et al., 2019). Another common approach is to construct a mosaic using the pixel- and band-wise median (or other percentile) values over the image stack (*e.g.*, Ghorbanian et al., 2020, Shafeian et al., 2021), expecting that the majority of the focused defects are concentrated on the tails of the reflectance distribution. Instead of median, also medoid selection may be applied (*i.e.*, the middlemost single pixel in the spectral space), which preserves the spectral profile of each pixel (Simonetti et al., 2021).

Using automatized methods for detecting defective pixels and constructing seamless mosaics may considerably speed up the creation of forest inventory maps. The output however is prone to errors if the applied threshold or quality rules cannot ensure the consistent quality of the resulting mosaic. In northern Europe, for example, suitable satellite data is limited due to short phenologically optimal season combined with frequent cloudiness, which reduces the quantity and quality of the suitable input images with further implications, *e.g.*, to spectral median values (Castaldi et al., 2023, Puliti et al., 2021). Particularly, the reliable detection of thin and hazy clouds is difficult, which are partially transparent but significantly changing the initial ground reflectance (Shen et al., 2014). Additionally, if the adjacent pixels are selected from different input images which may be temporally distant, the homogeneity of the output may be compromised (Castaldi et al., 2023, Simonetti et al., 2021). Effects of these factors may be difficult to notice or correct within an automatized processing chain, but eventually they may lead to higher prediction error and spatially heterogeneous results.

The objective of our study was to compare the quality of three satellite image mosaics and prediction errors of forest inventory variables, modelled by their reflectance values. One of the mosaics had been constructed using manually selected and cloud-masked images, considered as representing a higher degree of quality control, and the two others by fully automatized processing chains. The automatic mosaics were produced using the GEE service, one based on selecting the best pixels according to certain quality criteria and the other on pixels' band-wise percentile values over all the contributing images. The mosaics were then used to predict selected discrete and continuous forest inventory variables, based on field plots derived from the national forest inventory. Additionally, the prediction results were inspected in detail on smaller subareas.

2. Study area

The study area (Fig. 1) locates in Finland and covers approximately 75,000 km^2. It belongs to the boreal forest zone, ranging between the south and mid-boreal ecoregions. The delineation was based on the borders of a set of Sentinel-2 tiles (35VLJ, 35VLK, 35VMJ, 35VMK, 35VML, 35VNK and 35VNL), located on sparsely inhabited regions and guaranteeing large forest coverage. According to the CORINE Land Cover 2018 data (CLC2018; EEA, 2017), 63.8 % of the study area is covered by forests, 21.9 % by water bodies, 6.2 % by agricultural areas, 5.7 % by wetlands, and 2.5 % by built-up areas, respectively. Most of the forests in the study area are managed for wood production, dominated by Scots pine (Pinus sylvestris), Norwegian spruce (Picea abies) and birches (Betula pendula and B. pubescens). In terms of topography, the area is generally quite flat and located between 50 and 300 m above sea level, characterized by gently undulating hills and numerous lakes.

3. Materials and methods

3.1. Methodological overview

The general outline of the study is shown in Fig. 2 and presented in more detail in chapters 3.2–3.5. For manual mosaicking, selected input Sentinel-2 Level-2A images were downloaded to a local server, carefully masked to exclude clouds and their shadows, and radiometrically normalized using global (*i.e.*, concerning the overall relationships between the images) regression-based models. Two automatically generated mosaics were fully generated within the GEE service, using certain preconditions for selecting the most suitable images and excluding defective pixels, followed by mosaicking either by percentile values or quality-based rules. Then, the reflectance values of each mosaic were used to model and predict a range of forest inventory variables, derived from plot-wise measurements of the Finnish National Forest Inventory. The rationale of modelling was to generate a non-parametric estimator of the reference measurements based on Sentinel-2 reflectance values, expecting that the representativeness and homogeneity of the applied mosaic contributes to this correlation. Modelling was based on the k-nearest neighbor approach with leave-one-out cross-validation. Finally, prediction accuracies of the different mosaics were compared and

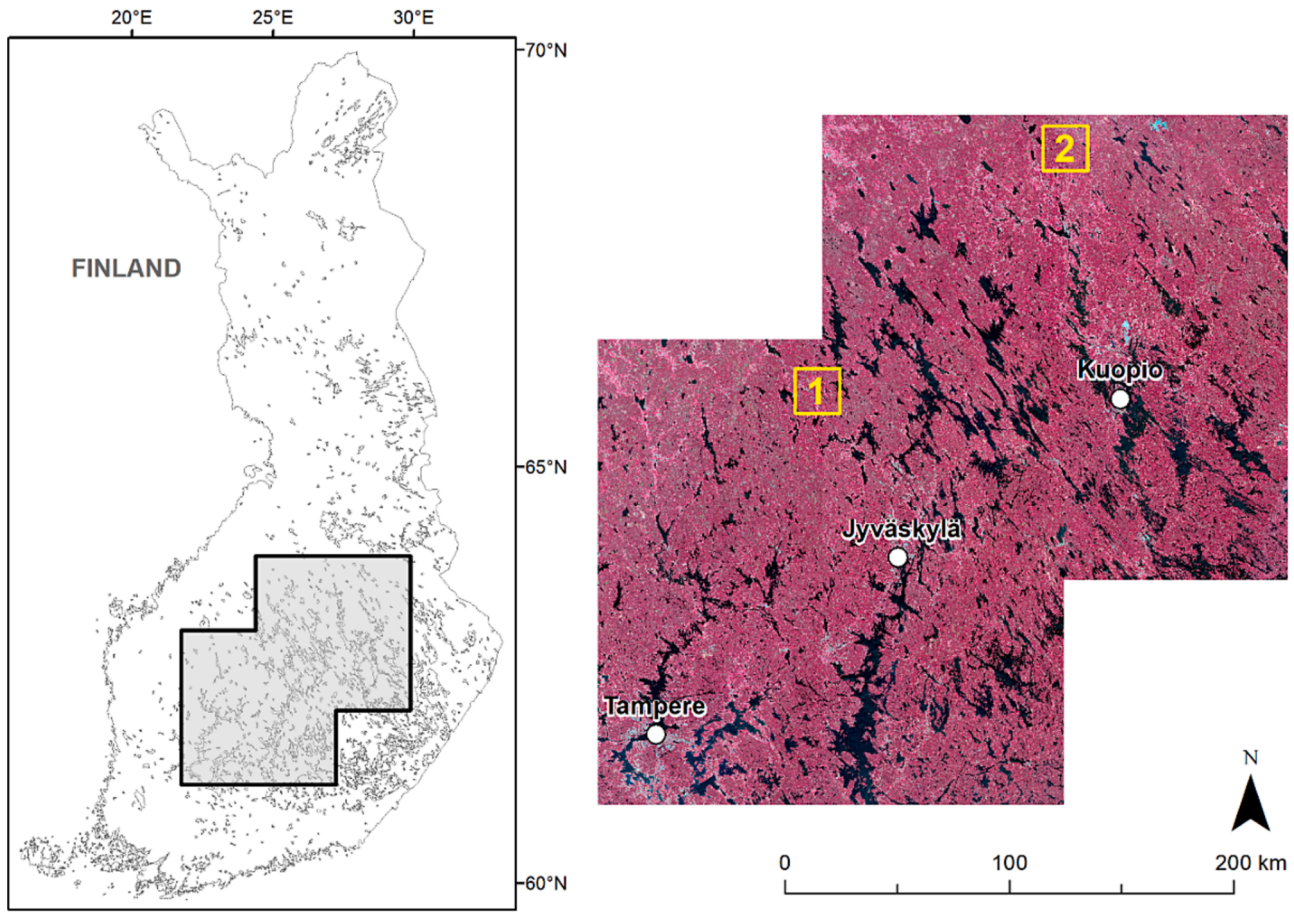


Fig. 1. Study area with general view on the left and Sentinel-2 based false color composite on the right. Yellow squares indicate the two 20 x 20 km subareas used for detailed evaluation of results. Contains National Land Survey of Finland general map data, 2022. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

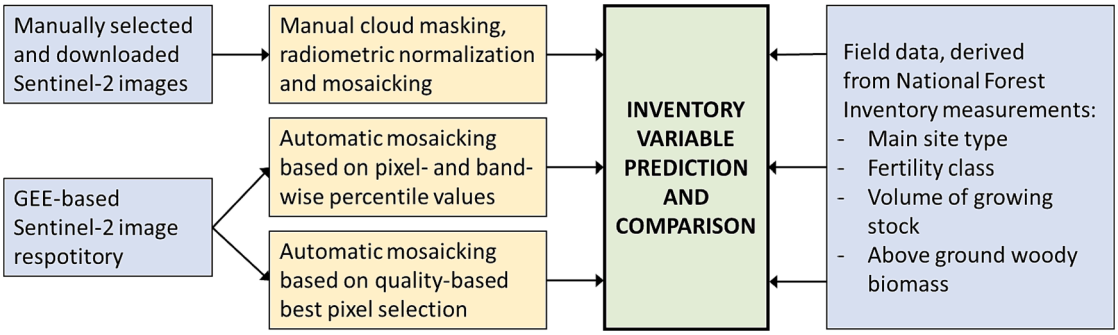


Fig. 2. General outline of the study.

additional tests on the predicted values' spatial heterogeneity were performed.

3.2. Image mosaicking

All the three mosaics were created nominally for summer 2020 with +/-1 year temporal flexibility, which was needed to overcome the challenges of a short yearly span for usable cloudless images. Only bands initially in 10/20 m resolution were preserved in the final mosaics, and 20 m bands were resampled to 10 m pixel size using the nearest neighbor method. Topographic correction was not applied to the mosaics as it was not expected significantly altering the comparison results, given that the study area was generally flat, and the yearly temporal window was narrow, thus minimizing the differences in the sun elevation.

3.2.1. Manual mosaic

Sentinel-2 Level-2A image tiles for the manual mosaic were downloaded from the Copernicus Open Access Hub, selecting the clearest and least cloudy images between early June and late August 2019–2021. Their remaining clouds, shadows and other defects were visually delineated and excluded from the mosaic. Processing of the images started from the middlemost tile, and following images were added based on the largest cloudless overlaps until the whole study area was covered. Regardless of using atmospherically corrected images, reflectance values had minor variations between the images. They were corrected band-wise based on the value pairs of the overlapping pixels, and

standard deviations in 3 × 3 windows. Only pixels on forest (CLC2018 Level 2 class 3.1) or wetland (class 4.1) areas were used, *i.e.*, targeting on natural vegetation likely to have the smallest fluctuations between the years. Corrected values were calculated using a linear regression model, which was weighted by the inverse of standard deviations (of the image to be added) to emphasize the most homogeneous areas.

3.2.2. Percentile-based automatic mosaic

The initial set of Sentinel-2 Level-2A images were selected in GEE according to their dates (between June 15 and August 15 in the years 2019–2021) and overall cloud coverage (up to 20 %). Early June and late August images were excluded to avoid potentially unrepresentative signals from the early and late growing season, as no further manual control could be used to ensure the quality. Then, based on S2_CLOUD_PROBABILITY layers (Skakun et al., 2022, Zupanc, 2017), potential remaining clouds above the threshold of 30 % were masked out. The selected threshold value was relatively low and expected to generate some commission errors (*i.e.*, masking out non-cloudy pixels), but considered to be effective also for most semi-transparent clouds. Finally, all the unmasked image areas were used to construct the mosaic, consisting of band-wise 40th percentile reflectance values for each pixel. As compared to the more common median-based compilation strategy (*e.g.*, Kollert et al., 2021, Shafeian et al., 2021), the 40th percentile appeared to produce a better outcome with less remaining cloud fragments or haze, while still avoiding low-reflectance areas derived from cloud shadows.

3.2.3. Quality-based automatic mosaic

The initial set of images for the quality-based mosaic was selected and masked similarly to the percentile-based approach with the exception that only images from the Sentinel-2A satellite were used to avoid potential differences between the A/B satellites (Lamquin et al., 2019). Then, two supplementary quality indicators were calculated. First, possibly remaining cloud fragments were detected by selecting pixels with cloud probability of 10–30 % and simultaneously having the visible bands' (B2-4) reflectance sum over 50 % higher than the previously calculated 40th percentile value. These were expected to indicate potential for translucent clouds (Baetens et al., 2019, Schmitt et al., 2019) while avoiding naturally bright land cover elements. Then, cloud shadows were detected as having the NIR (B8A) and SWIR (B11) bands' reflectance sum of over 500 (to exclude water areas) but below 75 % as compared to the 40th percentile value (Chandra et al., 2020, Zhu and Helmer, 2018). The locations of clouds and shadows were not matched for making the different detection methods working independently.

These two indicators were used for constructing image-wise quality layers composed of values 0 (the existence of both defects; an unlikely but possible alternative), 1 (one defect found), and 2 (no detected defects). The quality layers were further processed via morphological opening and closing similarly to Schmitt et al. (2019), using *focalMax* function first to remove single or small groups of low-quality pixels, and *focalMin* then to extend the remaining defective pixels to their initial extents. Finally, the actual mosaicking was conducted using the *qualityMosaic* function, starting from the least cloudy image and using the quality layers to select the best pixels.

3.3. Field data

Field measurements from the sample plots of 13th Finnish national forest inventory (NFI13), collected within a similar span to the image mosaics, were used as test data. The NFI13 sampling is based on a systematic grid and on each plot, a number of tree-, stand- and plot-level variables are recorded. In the study area there were 6,358 measured plots located on forestry land (*i.e.*, areas dedicated to forestry and including wet peatlands although not depending on if any trees are currently growing), and we focused on the following variables:

- **main site type** which classifies the forestry land into mineral land, spruce/hardwood-dominated swamps, pine bogs and open mires;
- **fertility class** with seven-step classification, ranging from herb-rich mineral sites / eutrophic mires into barren or rocky mineral sites / ombrotrophic mires;
- **volume of growing stock** according to Laasasenaho (1982) equations, calculated as total growing stock and separately for Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies (L.) Karst.*) and deciduous trees consisting primarily of downy birch (*Betula pubescens*) and silver birch (*Betula pendula*); and
- **above ground woody biomass (AGB)** according to Repola (2008, 2009) equations, calculated separately for Scots pine, Norway spruce, and deciduous trees.

Since most of the forests in the study area were managed for timber production, unrepresentative plots (*i.e.*, those cut between the satellite image and field measurement dates) were removed to exclude their potential bias in the results. This removal was targeted on plots with recent forest use notifications, indicating the forest owners' plans for management activities and required to be submitted by the Forest Act. As the cutting intensities vary and the eligible span of committing the plan is long, plots with actual and observable changes needed to be further filtered out. For this purpose, Pearson's correlations were first calculated between a large range of NFI-measured variables and spectral values (single bands + NDVI) of mosaic pixels, which were found to be the highest between the band B5 (705 nm, red edge) and NFI-measured total basal areas. Then, separately for each mosaic, species-wise linear regression models were estimated, and their residual values were calculated. Finally, starting from the highest residuals, plots where the field data didn't visually appear to correspond to the satellite image were removed. This process was continued until reaching a set of ten consecutive plots with none found to be discarded, resulting in 6,299 approved NFI plots shared by all the three mosaics. While potentially missing some changed plots, this procedure was regarded as sufficient for the purpose of comparing the mosaics.

3.4. Prediction of forest variables

Forest variables were predicted using the k-nearest neighbor method (k-NN) combined with a genetic algorithm, implemented in R statistical software (R Core Team, 2022). In addition to mosaics' band-wise pixel values, other features were calculated including all the possible band ratios (*e.g.*, B2/B3; excluding reciprocal band combinations) as well as NDVI-like normalized difference indices (NDI; exemplified in Eq. (1). These additional features are known to be helpful for vegetation monitoring applications (*e.g.*, Korhonen et al., 2017, Verrelst et al., 2015). The total number of satellite image features was 100.

$$NDI_{B3,B2} = \frac{(B3 - B2)}{(B3 + B2)} \tag{1}$$

In the k-NN method, Euclidean distances between sample plots are calculated in the feature space with n dimensions, equaling the number of features. Distances of nearest neighbors were inversely weighed as shown in Eq. (2)–(3) to diminish bias (Altman, 1992). Several options for k and g were tested for each prediction run.

$$\hat{y} = \sum_{i=1}^{k} w_i y_i \tag{2}$$

where

$\hat{y}$ = estimate for variable y,
k = number of nearest neighbors,
y_i = field observation of variable y of the i^{th} nearest neighbor,
w_i = weight of the i^{th} nearest neighbor, defined as (3):

$$w_i = \frac{1}{d_i^g} \Big/ \sum_{i=1}^{k} \frac{1}{d_i^g} \tag{3}$$

where

d_i is the Euclidean distance of the i^{th} nearest neighbor and g distance weighting parameter.

The genetic algorithm, applied using the R package *genalg* (Willighagen and Ballings, 2022), was used for feature selection to reduce the dimensionality of the feature space and select a close to optimal subset of features. The best combination of features based on an evaluation value (EV). The EV for continuous variables is shown in Eq. (4), and for discrete variables it was the proportion of correctly classified instances.

$$EV_{cont} = \sum_{j=1}^{J} w_j * \left(RMSE_j + \left|bias_j\right|\right) \tag{4}$$

where

w_j = weight of RMSE of variable j,

$RMSE_j$ = root mean square error of variable j,

$bias_j$ = bias of variable j.

In all prediction runs $k = 6$ was used and the value of g varied between 0.0 and 1.2. Weights applied to continuous variables were 0.4, 0.0, 0.0, 0.0, 0.2, 0.2 and 0.2 for species-wise volumes of growing stocks and AGBs, respectively. As correlation is high between the volume of growing stock and biomass, one set of tree species' variables was sufficient for training, and using zero weights made it possible to predict all continuous variables simultaneously. Discrete variables were not weighed. The initial sets of randomly selected features were fed into the algorithm which calculated the performances of the feature sets using leave-one-out cross-validation. The initial feature sets were evolving to maximize prediction performance over a preset number of iterations.

Prior to the training and prediction process, plots were divided into peatland (n = 1515) and mineral soil (n = 4784) strata based on the Topographic Database, produced by the National Land Survey of Finland. Models of these two strata were trained separately due to their differences in soil moisture and consequent reflectance properties as well as in forest characteristics (Katila & Tomppo, 2001). Model training was carried out separately for continuous (growing stock volumes and biomasses) and discrete (main site and fertility class) variables, totaling altogether 12 models for the three mosaic products. Further, wall-to-wall prediction maps for two selected subareas (see Fig. 1) were produced by applying the respective models.

3.5. Evaluation of the results

The mosaic-specific errors in forest variable prediction were assessed using leave-one-out cross-validation, expecting that pixel values' differences between the mosaics were primarily connected to their selection or calculation strategy. The outputs of two separately predicted strata, *i.e.*, peatlands and mineral soils, were combined prior to evaluating the results. Results were compared based on metrics commonly applied for evaluating remote sensing based predictions (see *e.g.*, Castaldi et al., 2023, Chen and Zhao, 2022, Li et al., 2023). For discrete variables, these included overall accuracy (the proportion of correct classifications over all the cases), Cohen's kappa value, precision, recall, and F_1 score. Kappa values κ were defined as (eq. (5)):

$$\kappa = \frac{p_o - p_e}{1 - p_e} \tag{5}$$

where p_o indicates the proportion of correct predictions over all the distinct classes and p_e denotes successful predictions by chance given their marginal probabilities. Precision, recall, and F_1 score are presented in eq. (6)–(8):

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{8}$$

where *TP* are true positive, *FP* false positive and *FN* false negative values. Continuous variables were assessed based on prediction bias, root mean square error and R^2 coefficient (eq. (9):

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{9}$$

where n is the number of observations, y are their field-based reference values which have a mean value of y, and $\hat{y}$ the corresponding values predicted by the model, respectively.

In addition, the results were assessed in terms of spatial characteristics and heterogeneity of the predicted variables. The principal motivation of this was to investigate whether the percentile-based automatic mosaic (*i.e.*, based on statistics rather than single input pixels) contrasted with the other two, and if the images selected for the manual and the quality-based mosaics induced differences for the results. The underlying assumption was that while the predicted values vary in space, the nature of this variation is gradual due to strong spatial autocorrelation. Therefore, a good model built with representative reflectance values should be expected to provide relatively similar predictions for adjacent pixels.

Assessment was based on predicting wall-to-wall maps on two 20 × 20 km subareas (see Fig. 1). These areas were selected so that in addition to high forest proportion, they had several input images in both manual and quality-based mosaics (Fig. 3). Other land uses than forestry (*i.e.*, agriculture, infrastructure, urban areas, and water bodies) were excluded from the maps. Qualitatively, the prediction maps were evaluated visually targeted on spotting any major differences in their outlook, and detecting if transition lines between the images were visible in the predictions due to spectral inconsistencies of the pixel values. And quantitatively, all the maps were analyzed using a 3 × 3 pixel moving window to calculate statistics on variation of the predicted values in the local neighborhood, which for continuous variables was the standard deviation of the measurement unit and for discrete variables the number of distinct classes. These were further aggregated into mean values over all the windows which included at least two predicted pixels located on forestry land.

4. Results

4.1. Prediction performance

For discrete variables, prediction differences between the mosaics are generally small (Table 1). In terms of the main site type, the quality-based mosaic is the best with 84.22 % overall accuracy and 0.6240 kappa with a slight margin compared to the two other mosaics. Differences in the two most common main site types (1 = mineral land and 3 = pine bogs) between the mosaics are almost negligible, but the quality mosaic performs somewhat better with the less common classes (2 = spruce/hardwood-dominated swamps and 4 = open mires). For spruce/hardwood-dominated swamps (2), both manual and percentile mosaics predict them as mineral land more often than the quality mosaic, and for open mires (4), the quality mosaic is generally better for finding the correct class with fewer mispredictions.

With respect to fertility class, however, the quality mosaic has the lowest overall accuracy of 55.71 % and kappa value of 0.3489, whereas the percentile mosaic ranks best with 57.96 % overall accuracy and

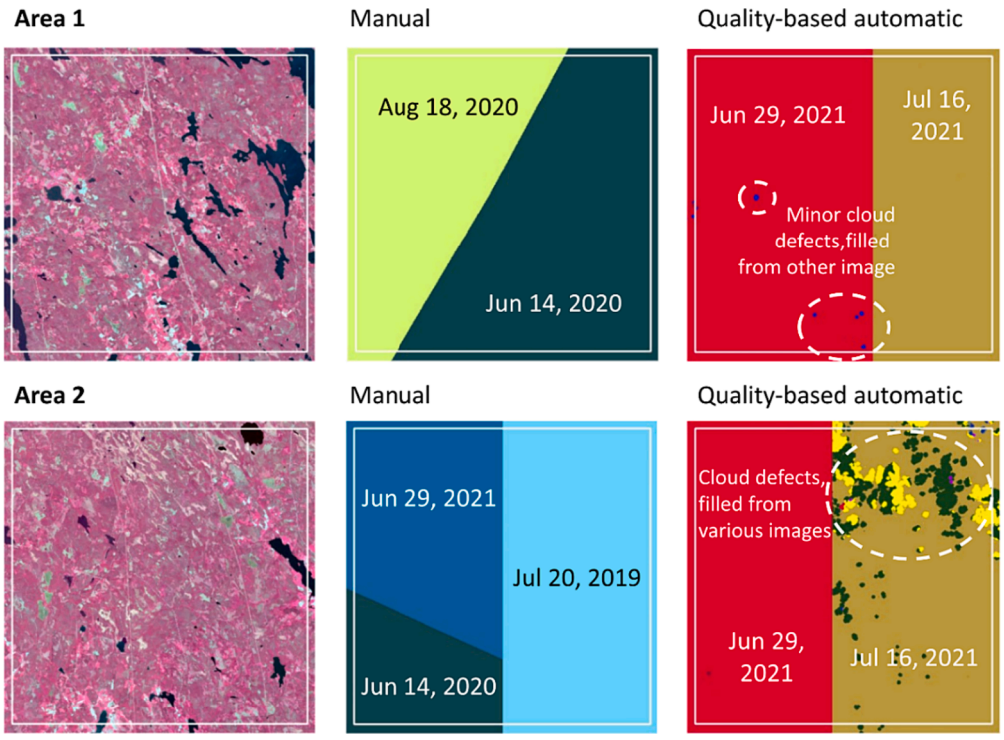


Fig. 3. Subareas selected for detailed assessment of variable prediction, including the areas and dates of the initial manual and quality-based automatic mosaics.

Table 1
Prediction accuracies of discrete variables, including class-wise predictions (N) compared with reference data (ref), precisions, recall values, and F1 scores, as well as overall accuracies and kappa values.

MAIN SITE TYPE													
Class	**N (ref)**	**MANUAL**				**PERCENTILE**				**QUALITY**			
		N	**Precision**	**Recall**	**F1**	**N**	**Precision**	**Recall**	**F1**	**N**	**Precision**	**Recall**	**F1**
1	4486	4862	0.8801	0.9539	0.9155	4857	0.8831	0.9561	0.9181	4840	0.8849	0.9547	0.9185
2	719	390	0.5282	0.2865	0.3715	373	0.5523	0.2865	0.3773	382	0.5759	0.3060	0.3996
3	1006	1005	0.7552	0.7545	0.7548	1016	0.7559	0.7634	0.7596	1029	0.7415	0.7584	0.7499
4	88	42	0.7381	0.3523	0.4769	53	0.7547	0.4545	0.5674	48	0.8125	0.4432	0.5735
FERTILITY CLASS													
1	223	47	0.2766	0.0583	0.0963	51	0.2941	0.0673	0.1095	65	0.3077	0.0897	0.1389
2	1461	1262	0.5872	0.5072	0.5443	1229	0.5989	0.5038	0.5472	1159	0.5695	0.4517	0.5038
3	2749	3143	0.5972	0.6828	0.6371	3131	0.6075	0.6919	0.6469	3228	0.5805	0.6817	0.6271
4	1397	1401	0.5261	0.5276	0.5268	1440	0.5285	0.5447	0.5365	1382	0.5094	0.5039	0.5067
5	410	403	0.5434	0.5341	0.5387	389	0.5656	0.5366	0.5507	420	0.5524	0.5659	0.5590
6	50	43	0.4186	0.3600	0.3871	59	0.2881	0.3400	0.3119	45	0.4222	0.3800	0.4000
7	9	0	–	–	–	0	–	–	–	0	–	–	–

Overall accuracies / main site type: manual 83.74 %, percentile 84.19 %, quality 84.22 %.
Kappa values / main site type: manual 0.6108, percentile 0.6219, quality 0.6240.
Overall accuracies / fertility class: manual 57.23 %, percentile 57.96 %, quality 55.71 %.
Kappa values / fertility class: manual 0.3729, percentile 0.3843, quality 0.3489.

0.3843 kappa value, and the manual mosaic is close to similar figures. Contrary to the main site type, the quality mosaic struggles with detecting the most frequent classes (2–4) while predicting the most fertile and infertile locations better. This tendency seems be driven by the quality mosaic's generally larger proportion of mispredictions to the adjacent classes. All the mosaics neglect the least fertile sites (class 7; n = 9), deriving mainly from the small number of instances combined with their possibly minor spectral differences to class 6 and, therefore, inability of the k-NN methoddiscriminating them properly.

With respect to the prediction performance of continuous variables, the differences of the mosaics are somewhat more discernible (Table 2). In this comparison, the overall performance of the quality mosaic is distinctly the worst: it has the highest RMS errors and lowest R2 scores for all the variables of interest, and it never achieves the lowest bias among the three mosaics. In particular, the quality mosaic underestimates the highest volumes and biomasses more often than the manual or percentile mosaics. Compared with that, the statistics are much closer between the manual and the percentile mosaics. While the manual mosaic almost invariably has the least bias, in terms of the RMSE and R2 figures the percentile mosaic is generally better or at least equal. For some variables, the percentile mosaic has the highest bias of all three options but simultaneously the lowest RMSE, indicating a relatively low tendency for gross errors.

Table 2
Prediction accuracies of continuous variables assessed using bias (%), RMSE (%) and R2 score.

VOLUME OF GROWING STOCK										
Species	**Mean value (m^3 / ha)**	**MANUAL**			**PERCENTILE**			**QUALITY**		
		Bias%	**RMSE%**	**R2**	**Bias%**	**RMSE%**	**R2**	**Bias%**	**RMSE%**	**R2**
Scots pine	60.681	2.690	101.48	0.347	5.768	100.71	0.357	4.036	106.65	0.279
Norway spruce	56.230	−1.886	110.18	0.572	−5.059	105.28	0.610	−3.037	116.97	0.518
Deciduous	28.485	−1.887	131.43	0.390	−2.296	131.48	0.390	−2.255	140.20	0.306
Total	145.396	0.024	56.22	0.525	0.001	56.33	0.523	0.068	59.47	0.469
BIOMASS OF GROWING STOCK										
Scots pine	3081.418	2.676	98.49	0.359	5.668	97.61	0.370	4.007	103.50	0.292
Norway spruce	3205.597	−1.841	102.20	0.580	−4.763	97.71	0.616	−3.026	109.31	0.519
Deciduous	1790.061	−2.003	132.04	0.384	−2.380	131.93	0.385	−2.470	140.54	0.302

4.2. Predictions' spatial characteristics

Thematic maps of predicted values from two subareas are exemplified in Figs. 4 and 5 by using the growing stock volume of Scots pine and fertility classification. By visual evaluation, it is obvious that variation of Scots pine volume between the adjacent pixels is relatively high, and finding any texture-related differences between the mosaics is difficult. The high degree of heterogeneity is also reflected by the approximately 100 % volumetric RMS error in the prediction statistics. Some of the largest differences in the predicted values are derived from recent clear-cuts, which have occurred between the dates of the input images. Borderlines of the image dates however do not appear to make visible differences for the predicted volumes in either manual or quality-based mosaics.

Regarding the fertility classification, the three mosaics have similar overall characteristics and the patch structure follows relatively similar pattern. With respect to image borders, however, one important difference is a large patch of class 3 in the middle of the quality mosaic visualized in Fig. 5, which is not present in the two other mosaics and ends at the nearest borderlines. The same area is associated with relatively low Scots pine volumes. This derives from a partial cloud shadow which has been included in the quality mosaic to fill a cloudy patch of another image. The automatic quality evaluation has failed in detecting any defects with these pixels, but the reflectance values have transformed enough to alter the resulting predictions.

The quantitative evaluation of the thematic maps indicates that the variation in local neighborhoods is invariably the highest with the quality mosaic (Table 3). This tendency is similar both with discrete and continuous variables. With respect to the manual and the percentile mosaics, variation in the number of discrete classes is virtually equal. For continuous variables, the manual mosaic has the invariably lowest standard deviation, and the percentile mosaic ranks between the two other mosaics. However, the differences between the three mosaics with any of the inventory variables are not drastic.

5. Discussion

The results of our paper indicate that from the perspective of predicting forest inventory variables, automatized mosaicking methods may reach approximately similar performance as compared to the manually constructed mosaic. From a simplistic point of view, mosaicking is a straightforward task where the most representative pixel values are selected or calculated over a set of properly georeferenced images. There are however a vast number of different mosaicking

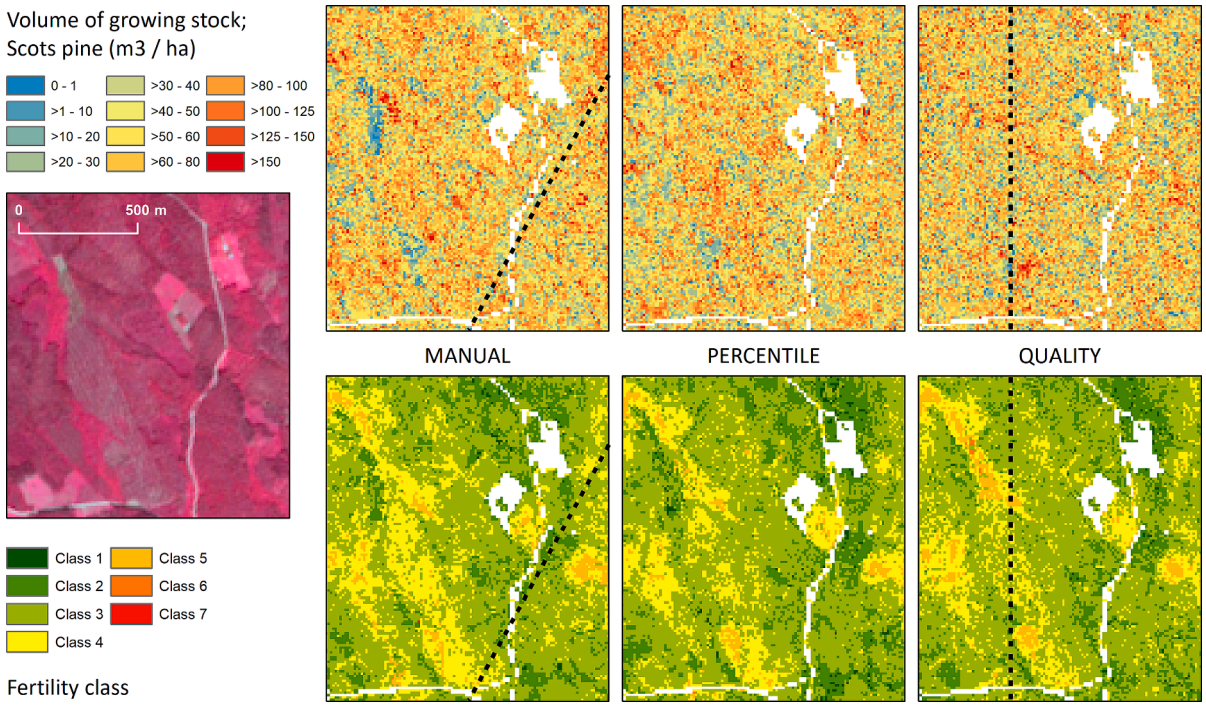


Fig. 4. Predictions of Scots pine growing stock and fertility classes on an approximately 1.2 × 1.4 km square within subarea 1. The dashed lines indicate borderlines between the initial input images from different dates. White pixels are non-forested areas.

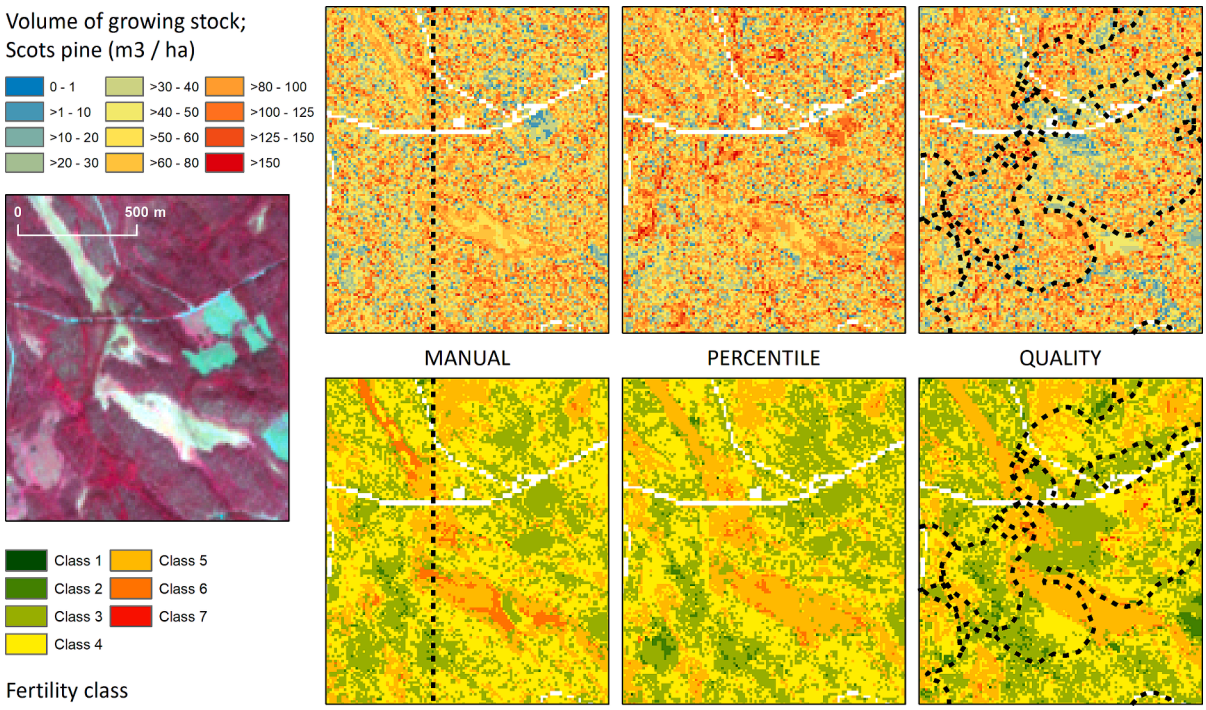


Fig. 5. Predictions of Scots pine growing stock and fertility classes on an approximately 1.2 × 1.4 km square within subarea 2. The dashed lines indicate borderlines between the initial input images from different dates. White pixels are non-forested areas.

Table 3
Number of classes (nclass) or standard deviation (std) of the predicted inventory variables, calculated as the mean value over all the 3 × 3 windows.

Statistic	MANUAL	PERCENTILE	QUALITY
Main site type (nclass)	1.511	1.508	1.526
Fertility class (nclass)	2.154	2.152	2.243
Volume of Scots pine (std of m^3 / ha)	26.671	27.096	27.882
Volume of Norway spruce (std of m^3 / ha)	31.989	33.277	34.328
Volume of deciduous trees (std of m^3 / ha)	16.345	16.360	16.854
Total volume (std of m^3 / ha)	41.593	42.961	43.526
Biomass of Scots pine (std of kg / ha)	1329.19	1348.97	1388.04
Biomass of Norway spruce (std of kg / ha)	1712.37	1778.72	1836.53
Biomass of deciduous trees (std of kg / ha)	1027.98	1029.61	1060.03

methods which have been developed principally for the same task, but from different perspectives and often focusing on a certain range of prevailing conditions. Therefore, not all of them are universally applicable if, for example, the supply of suitable images is very limited, or the output needs to pass strict quality rules. While a manually supervised, expert opinion based mosaicking approach may be regarded as retaining the highest level of quality control and enabling better potential for detecting and fixing radiometric inconsistencies, its applicability is limited due to laboriousness. Particularly applications requiring extensive image coverages combined with high temporal frequency are benefiting from automatized workflows with low level of manual intervention, as long as the quality of the results is satisfactory.

The prediction errors of our study are approximately comparable to the recent operational Finnish multi-source national forest inventory (MS-NFI; Mäkisara et al., 2022) and as reported in the study by Tuominen et al., (2017; the growing stock volumes derived from 2D satellite images), based on satellite images and k-NN prediction in similar conditions. Regarding to the two automatically generated mosaics, particularly the percentile-based predictions indicate relatively good performance without major weaknesses. Its consistency in prediction accuracy derives assumedly from being based on the reflectance distribution over all the input images, which reduces the potential for including outlier values. Processing each pixel separately may, however, diminish the homogeneity of the results, as indicated by the slightly higher neighborhood variation with continuous predicted values as compared to the manual mosaic. To reduce this tendency, the consistency of the input images should be promoted by selecting a limited number of representative images and exclude at least the easily detectable clouds (Kempeneers and Soille, 2017). Additionally, while median (or medoid) values are often preferred, selecting of the applied percentile value is not always a trivial task. For example, Corbane et al. (2020) used 25th percentile for constructing Sentinel-2 mosaics on a global scale, while Castaldi et al. (2023) applied a 90th percentile threshold with intention of finding the driest pixels. In our case, trials suggested selecting the 40th percentile which was just below the threshold where the thinnest clouds started to appear, but which helped to avoid partial shadows.

With respect to the quality-based mosaic, most of the mosaic area consisted of images from only two dates (June 29, 2021, and July 16, 2021). Good performance for predicting the main site type may derive from a narrower time window as compared to the other two mosaics, provided that each class consists of a range of different forest types and their reliable detection requires phenological consistency. For fertility classes, reflectance-based prediction is even more complicated and depends, for example, on the dominant tree species and the existence of brighter pixels derived from sunlit parts of the tree crowns (Mõttus and Takala, 2014). Difficulties of the quality mosaic to predict the exact fertility class and the tendency of underestimating the largest continuous values are however suggesting that the spectral contrast may not be optimal on the selected dates. In this case, the reflectance values will not provide optimal performance for k-NN based prediction, and this may also explain the observed higher neighborhood variation.

The lower performance of the quality mosaic doesn't indicate that the applied method is categorically inferior but setting the optimal threshold values often requires extensive testing, and the final product may still contain flaws such as haziness which cannot be solved by this

approach (Li et al., 2021, Schmitt et al., 2019). Another option would be to use a single index as a quality indicator, such as the highest NDVI value for emphasizing the maximum photosynthetic activity of vegetation or the key phenology window (Li et al., 2023, Thomas et al., 2021). This approach will similarly retain the initial pixel values and their band ratios but will not emphasize selecting larger continuous areas from single images, as was one of the aims in our mosaicking strategy. Altogether, identifying the highest quality pixels is subject to both the characteristics of the study area and to the targets of the study, and therefore the most applicable detecting strategy will vary as well.

The biggest challenge in automatized mosaicking is not related to detecting thick clouds, but rather extending reliable detections to their furthest margins, thinner clouds, cloud shadows, and general haziness. One of the drawbacks with using Sentinel-2 images is the lack of a thermal infrared band, which would allow distinguishing clouds from the ground using temperature differences and prevent bright targets from being misinterpreted as clouds (Franz et al., 2018). Cloud shadows, in turn, may be very obscure over a complex landscape and mix easily with other dark surfaces such as lakes (Shephard et al., 2020). Principally, the direction of cast shadow is always known and the distance may be attempted to be detected by matching the bright and the dark targets or using more advanced approaches such as parallax between the different bands (Franz et al., 2018, Schmitt et al., 2019). Generally, it may be advisable to favor relatively simple mosaicking methods rather than highly complicated ones, as for larger extents the simpler ones are likely to increase the robustness of the output results.

6. Conclusions

The inventory of forest resources is generally relying on field measurements, but satellite images enable extending the scattered observations into larger wall-to-wall coverages. This often requires mosaicking images from different dates and our results emphasize taking advantage of automatized mosaicking methods, given that differences in the prediction accuracies as compared to manual image selection are relatively small. In our study, particularly the performance of the percentile mosaic was found relatively good although the applied threshold needs to be carefully confirmed. Further, while a set of expert-based quality rules may similarly produce a high-quality image mosaic, attention should be paid to adjusting these rules to select the clearest and the most representative pixels over the whole study area.

CRediT authorship contribution statement

Timo P. Pitkänen: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Andras Balazs:** Writing – review & editing, Writing – original draft, Validation, Data curation. **Sakari Tuominen:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the Finnish Ministry of Agriculture and Forestry, Catch the carbon research and innovation program (MaaTi-project, PI R. Mäkipää) and from the European Union's Horizon Europe research and innovation program, under Grant Agreement 101091010 for a Soil Mission project Benchmarks. The study has been made with affiliation to the UNITE Research Flagship funded by the Research Council of Finland (decision no 337655).

References

Altman, N.S., 1992. An introduction to kernel and nearest-neighbor nonparametric regression. Am. Stat. 46, 175–185. https://doi.org/10.1080/00031305.1992.10475879.

Baetens, L., Desjardings, C., Hagolle, O., 2019. Validation of Copernicus Sentinel-2 cloud masks obtained from MAJA, Sen2Cor, and FMask processors using reference cloud masks generated with a supervised active learning procedure. Remote Sens. 11, 433. https://doi.org/10.3390/rs11040433.

Castaldi, F., Koparan, M.H., Wetterlind, J., Žydelis, R., Vinci, I., Savaş, A.Ö., Kıvrak, C., Tunçay, T., Volungevičius, J., Obber, S., Ragazzi, F., Malo, D., Vaudour, E., 2023. Assessing the capability of Sentinel-2 time-series to estimate soil organic carbon and clay content at local scale in croplands. ISPRS J. Photogramm. Remote Sens. 199, 40–60. https://doi.org/10.1016/j.isprsjprs.2023.03.016.

Chandra, D.S., Phinn, S., Scarth, P., 2020. Cloud and cloud shadow masking for Sentinel-2 using multitemporal images in global area. Int. J. Remote Sens. 41, 2877–2904. https://doi.org/10.1080/01431161.2019.1697006.

Chen, Z., Zhao, S., 2022. Automatic monitoring of surface water dynamics using Sentinel-1 and Sentinel-2 data with Google Earth Engine. Int. J. Appl. Earth Obs. Geoinf. 113, 103010 https://doi.org/10.1016/j.jag.2022.103010.

Corbane, C., Politis, B., Kempeneers, P., Simonetti, D., Soille, P., Burger, A., Pesaresi, M., Sabo, F., Syrris, V., Kemper, T., 2020. A global cloud free pixel- based image composite from Sentinel-2 data. Data Brief 31, 105737. https://doi.org/10.1016/j.dib.2020.105737.

DeLancey, E.R., Kariyeva, J., Bried, J.T., Hird, J.N., 2019. Large-scale probabilistic identification of boreal peatlands using Google Earth Engine, open-access satellite data, and machine learning. PLoS One 14, e0218165.

EEA, 2017. CLC2018 Technical Guidelines, 25 October 2017. https://land.copernicus.eu/en/technical-library/clc-2018-technical-guidelines/@@download/file (accessed 2 October 2023).

Franz, D., Haß, E., Uhl, A., Stoffels, J., Hill, J., 2018. Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. Remote Sens. Environ. 215, 471–481. https://doi.org/10.1016/j.rse.2018.04.046.

Ghorbanian, A., Kakooei, M., Amani, M., Mahdavi, S., Mohammadzadeh, A., Hasanlou, M., 2020. Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples. ISPRS J. Photogramm. Remote Sens. 167, 276–288. https://doi.org/10.1016/j.isprsjprs.2020.07.013.

Helmer, E.H., Ruefenacth, B., 2005. Cloud-Free Satellite Image Mosaics with Regression Trees and Histogram Matching. Photogramm. Eng. Remote Sens. 71, 1079–1089. 10.14358/PERS.71.9.1079.

Kalinaki, K., Malik, O.A., Lai, D.T.C., 2023. FCD-AttResU-Net: An improved forest change detection in Sentinel-2 satellite images using attention residual U-Net. Int. J. Appl. Earth Obs. Geoinf. 122, 103453 https://doi.org/10.1016/j.jag.2023.103453.

Kangas, A., Astrup, R., Breidenbach, J., Fridman, J., Gobakken, T., Korhonen, K.T., Maltamo, M., Nilsson, M., Nord-Larsen, T., Næsset, E., Olsson, H., 2017. Remote sensing and forest inventories in Nordic countries – roadmap for the future. Scand. J. For. Res. 33, 397–412. https://doi.org/10.1080/02827581.2017.1416666.

Katila, M., Tomppo, E., 2001. Selecting estimation parameters for the Finnish multisource National Forest Inventory. Remote Sens. Environ. 76, 16–32. https://doi.org/10.1016/S0034-4257(00)00188-7.

Kempeneers, P., Soille, P., 2017. Optimizing Sentinel-2 image selection in a Big Data context. Big Earth Data 1, 145–158. https://doi.org/10.1080/20964471.2017.1407489.

Kollert, A., Bremer, M., Löw, M., Rutzinger, M., 2021. Exploring the potential of land surface phenology and seasonal cloud free composites of one year of Sentinel-2 imagery for tree species mapping in a mountainous region. Int. J. Appl. Earth Obs. Geoinf. 94, 102208 https://doi.org/10.1016/j.jag.2020.102208.

Korhonen, L., Hadi, Packalen, P., Rautiainen, M., 2017. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. Remote Sens. Environ 195, 259–274. 10.1016/j.rse.2017.03.021.

Laasasenaho, J., 1982. Taper curve and volume functions for pine, spruce and birch. Commun Inst For Fenn, 108. http://urn.fi/URN:ISBN:951-40-0589-9.

Lamquin, N., Woolliams, E., Bruniquel, V., Gascon, F., Gorroño, J., Govaerts, Y., Leroy, V., Lonjou, V., Alhammoud, B., Barsi, J.A., Czapla-Myers, J.S., McCorkel, J., Helder, D., Lafrance, B., Clerc, S., Holben, B.N., 2019. An inter-comparison exercise of Sentinel-2 radiometric validations assessed by independent expert groups. Remote Sens. Environ. 233, 111369 https://doi.org/10.1016/j.rse.2019.111369.

Li, X., Feng, R., Guan, X., Shen, H., Zhang, L., 2019. Remote sensing image mosaicking: Achievements and challenges. IEEE Geosci. Remote Sens. Mag. 7, 8–22. https://doi.org/10.1109/MGRS.2019.2921780.

Li, S., He, S., Yue, T., Du, Z., Zhao, N., Zhao, Y., Jiao, Y., Fan, B., Liu, Y., Xu, Z., Yang, Y., 2023. Mapping bamboo forest and expansion intensity in China by coupling vegetation phenology and C-band SAR with Sentinel-1 and Sentinel-2 images. Int. J. Appl. Earth Obs. Geoinf. 121, 103384 https://doi.org/10.1016/j.jag.2023.103384.

Li, J., Knapp, D.E., Lyons, M., Roelfsema, C., Phinn, S., Schill, S.R., Asner, G.P., 2021. Automated global shallow water bathymetry mapping using Google Earth Engine. Remote Sens. 13, 1469. https://doi.org/10.3390/rs13081469.

Li, X., Li, Z., Feng, R., Luo, S., Zhang, C., Jiang, M., Shen, H., 2020. Generating high-quality and high-resolution seamless satellite imagery for large-scale urban regions. Remote Sens. 12, 81. https://doi.org/10.3390/rs12010081.

Mäkisara, K., Katila, M., Peräsaari, J., 2022. The Multi-Source National Forest Inventory of Finland — methods and results 2017 and 2019. Natural resources and bioeconomy studies 90/2022. Natural Resources Institute Finland, Helsinki.

McRoberts, R.E., Tomppo, E.O., 2007. Remote sensing support for national forest inventories. Remote Sens. Environ. 110, 412–419. https://doi.org/10.1016/j.rse.2006.09.034.

Mõttus, M., Takala, T., 2014. A forestry GIS-based study on evaluating the potential of imaging spectroscopy in mapping forest land fertility. Int. J. Appl. Earth Obs. Geoinf. 33, 302–311. https://doi.org/10.1016/j.jag.2014.06.010.

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 Sens. Environ. 151, 3–15. https://doi.org/10.1016/j.rse.2013.08.048.

Puliti, S., Breidenbach, J., Schumacher, J., Hauglin, M., Klingenberg, T.F., Astrup, R., 2021. Above-ground biomass change estimation using national forest inventory data with Sentinel-2 and Landsat. Remote Sens. Environ. 265, 112644 https://doi.org/10.1016/j.rse.2021.112644.

R Core Team, 2022. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (accessed 2 October 2023).

Repola, J., 2008. Biomass equations for birch in Finland. Silva Fenn 42, 605–624. https://doi.org/10.14214/sf.236.

Repola, J., 2009. Biomass equations for Scots pine and Norway spruce in Finland. Silva Fenn 43, 625–647. https://doi.org/10.14214/sf.184.

Santini, F., Palombo, A., 2022. Impact of Topographic Correction on PRISMA Sentinel 2 and Landsat 8 Images. Remote Sens. 14, 3903. https://doi.org/10.3390/rs14163903.

Schmitt, M., Hughes, L.H., Qiu, C., Zhu, X.X., 2019. Aggregating cloud-free Sentinel-2 images with Google Earth Engine. ISPRS Ann. Photogramm. Remote Sens. Spat. Inform. Sci., Volume IV-2/W7. 10.5194/isprs-annals-IV-2-W7-145-2019.

Segal-Rozenhaimer, M., Li, A., Das, K., Chirayath, V., 2020. Cloud detection algorithm for multi-modal satellite imagery using convolutional neural-networks (CNN). Remote Sens. Environ. 237, 111446 https://doi.org/10.1016/j.rse.2019.111446.

Shafeian, E., Fassnacht, F.E., Latifi, H., 2021. Mapping fractional woody cover in an extensive semi-arid woodland area at different spatial grains with Sentinel-2 and very high-resolution data. Int. J. Appl. Earth Obs. Geoinf. 105, 102621 https://doi.org/10.1016/j.jag.2021.102621.

Shen, H., Li, H., Qian, Y., Zhang, L., Yuan, Q., 2014. An effective thin cloud removal procedure for visible remote sensing images. ISPRS J. Photogramm. Remote Sens. 96, 224–235. https://doi.org/10.1016/j.isprsjprs.2014.06.011.

Shephard, J.D., Schindler, J., Dymond, J.R., 2020. Automated mosaicking of Sentinel-2 satellite imagery. Remote Sens. 12, 3680. https://doi.org/10.3390/rs12223680.

Simonetti, D., Pimple, U., Langner, A., Marelli, A., 2021. Pan-tropical Sentinel-2 cloud-free annual composite datasets. Data Brief 39, 107488. https://doi.org/10.1016/j.dib.2021.107488.

Skakun, S., et al., 2022. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. Remote Sens. Environ. 274, 112990 https://doi.org/10.1016/j.rse.2022.112990.

Skakun, S., Vermote, E.F., Artigas, A.E.S., Rountree, W.H., Roger, J.-C., 2021. An experimental sky-image-derived cloud validation dataset for Sentinel-2 and Landsat 8 satellites over NASA GSFC. Int. J. Appl. Earth Obs. Geoinf. 95, 102253 https://doi.org/10.1016/j.jag.2020.102253.

Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S., Brisco, B., 2020. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. ISPRS J. Photogramm. Remote Sens. 164, 152–170. https://doi.org/10.1016/j.isprsjprs.2020.04.001.

Thomas, N., Baltezar, P., Lagomasino, D., Stovall, A., Iqbal, Z., Fatoyinbo, L., 2021. Trees outside forests are an underestimated resource in a country with low forest cover. Sci. Rep. 11, 7919. https://doi.org/10.1038/s41598-021-86944-2.

Tomppo, E., Olsson, H., Ståhl, G., Nilsson, M., Hagner, O., Katila, M., 2008. Combining national forest inventory field plots and remote sensing data for forest databases. Remote Sens. Environ. 112, 1982–1999. https://doi.org/10.1016/j.rse.2007.03.032.

Tuominen, S., Pitkänen, T., Balazs, A., Kangas, A., 2017. Improving Finnish Multi-Source National Forest Inventory by 3D aerial imaging. Silva Fenn. 51, 7743. 10.14214/sf.7743.

Vaudour, E., Gomez, C., Lagacherie, P., Loiseau, T., Baghdadi, N., Urbina-Salazar, D., Loubet, B., Arrouays, D., 2021. Temporal mosaicking approaches of Sentinel-2 images for extending topsoil organic carbon content mapping in croplands. Int. J. Appl. Earth Obs. Geoinf. 96, 102277 https://doi.org/10.1016/j.jag.2020.102277.

Verrelst, J., Rivera, J.B., Veroustraete, F., Muñoz-Mari, J., Clevers, J.G.P.W., Camps-Valls, G., Moreno, J., 2015. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods – A comparison. ISPRS J. Photogramm. Remote Sens. 108, 260–272. https://doi.org/10.1016/j.isprsjprs.2015.04.013.

Willighagen, E., Ballings, M., 2022. genalg: R Based Genetic Algorithm. https://cran.r-project.org/web/packages/genalg/index.html (accessed 2 October 2023).

Wulder, M., 1998. Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. Prog. Phys. Geog. Earth Environ. 22, 449–476. https://doi.org/10.1177/030913339802200402.

Yu, L., Zhang, Y., Sun, M., Zhou, X., Liu, C., 2017. An auto-adapting global-to-local color balancing method for optical imagery mosaic. ISPRS J. Photogramm. Remote Sens. 132, 1–19. https://doi.org/10.1016/j.isprsjprs.2017.08.002.

Zhang, C., Xiao, X., Zhao, L., Qin, Y., Doughty, R., Wang, X., Dong, J., Yang, X., 2023. Mapping Eucalyptus plantation in Guangxi, China by using knowledge-based algorithms and PALSAR-2, Sentinel-2, and Landsat images in 2020. Int. J. Appl. Earth Obs. Geoinf. 120, 103348 https://doi.org/10.1016/j.jag.2023.103348.

Zhu, X., Helmer, E.H., 2018. An automatic method for screening clouds and cloud shadows in optical satellite image time series in cloudy regions. Remote Sens. Environ. 214, 135–153. https://doi.org/10.1016/j.rse.2018.05.024.

Zupanc, A., 2017. Improving Cloud Detection with Machine Learning. Sentinel Hub Blog. https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learning-c09dc5d7cf13 (accessed 2 October 2023).