Combining TanDEM-X and Sentinel-2 for large-area species-wise prediction of forest biomass and volume

Henrik J. Persson [*], Jonas Jonzén, Mats Nilsson

Department of Forest Resource Management, Swedish University of Agricultural Sciences (SLU), 901 83 Umeå, Sweden

ARTICLE INFO

ABSTRACT

In this study, data from the satellite sensors TanDEM-X and Sentinel-2 were combined with national field inventory data to predict forest above-ground biomass (AGB) and stem volume (VOL) over a large area in Sweden. The data sources were evaluated both separately and in combination. The study area covers approximately 20,000,000 ha and corresponds to about 70% of the Swedish forest area. The study area was divided into tiles of 2.5×2.5 km^2, which were processed sequentially. The field plots were inventoried on 7 m and 10 m circular plots by the Swedish National Forest Inventory, and plot AGB and VOL at the year of the satellite data were estimated based on a 10-year period of field data. The AGB and VOL were modelled using the k nearest neighbor (kNN) algorithm, with k = 5 neighbors. The combined use of two data sources with different scene extents enabled the generation of seamless AGB and VOL maps. Moreover, the kNN algorithm provided the VOL divided per tree species, which was used for classification of the dominant tree species at stand-level. The overall accuracy for the dominant tree species classification was 77%. The predicted AGB and VOL rasters were evaluated using 549 field inventoried forest stands distributed over Sweden. The RMSE for the predictions based on both data sources were 31.4 t/ha (29.1%) for AGB, and 59.0 m^3/ha (30.2%) for VOL. By estimating and removing the variance due to sampling (the stand values were estimated from sample plots), the RMSE was improved to 18.0 t/ha (16.6%). The evaluated approach of using kNN was suitable for estimating forest variables from a combination of different satellite sensors, provided sufficient field reference data are available. The TanDEM-X data were most important for the AGB and VOL predictions, while Sentinel-2 data were essential to map the tree species.

1. Introduction

Remote sensing (RS) has proved invaluable for many sectors that rely on forest, since it enables predictions of variables with complete coverage in terms of raster maps. Yet, practical challenges hampers the continuous generation of such maps, especially at large scales, e.g., due to temporally harmonizing reference data (mostly inventoried in field) with the RS data, pre- and post-processing of the RS data, and mosaicking and filtering.

In Sweden, wall-to-wall raster maps with predictions of common forest variables (e.g., above-ground biomass (AGB), stem volume (VOL), tree height, and basal area) have been produced for the entire country using laser scanning data and field data from the Swedish National Forest Inventory (NFI) (Nilsson et al., 2017). In the past, such maps were produced on a five years cycle (starting in 2000), based on Landsat or SPOT imagery. Swedish authorities and users in the forest industry have

now got used to free high quality forest variable maps, but new laser scanning campaigns will only be repeated on five to seven years cycles (the time required to scan the entire country). Since most of the Swedish forests are managed, there is a strong demand for maps updated more frequently. It is also preferred to acquire all data within a short period to avoid temporal mismatches. An ongoing project aims at combining Sentinel-2 data with tree heights obtained from photogrammetric image matching, in order to generate forest maps more frequently. Sentinel-2 data are acquired from Sweden every three to five days, but the time for complete coverage depends greatly on the weather conditions. Photogrammetric height data are acquired nationally at two- to four years cycles. In this study, the photogrammetric heights were replaced with synthetic aperture radar (SAR) TanDEM-X data from a previous project. TanDEM-X has shown good potential for mapping of forest AGB and VOL, and has the advantage of working in all weather conditions (Karila et al., 2015; Persson et al., 2017; Persson and Fransson, 2017;

Rahlf et al., 2014; Soja et al., 2017). TanDEM-X acquires data at an 11 days repeat-cycle, but for complete coverage of the entire Sweden, a period of about 4 months was nevertheless required in this project (October 2015–January 2016) (Persson et al., 2017). For efficient co-use of resources, this study is based on the frameworks developed within the two projects.

Past studies that combined data from different satellite sensors for forest mapping have also aimed at taking advantage of complementing sensor techniques, such as SAR and spectral sensors (Alan et al., 2017; Chang and Shoshany, 2016; Chen et al., 2018; Chirici et al., 2020; Debastiani et al., 2019; Mendes et al., 2019). Several studies combined Sentinel-1 and Sentinel-2 or Landsat, due to its free availability. However, these studies have been based on SAR backscatter which suffer from saturation of volumetric predictions, as does spectral sensors (Fagan and DeFries, 2009; Imhoff, 1995; Lu, 2006). They have also aimed at predicting the total AGB or VOL, and not tree species. The limitation of saturation in the SAR data is largely overcome by using the single-pass interferometry available with TanDEM-X, since it provides observations related to both forest height and density (Askne et al., 2019, 2018). The saturation in the optical data is less critical, since the main contribution from the spectral wavelengths is to improve species separation.

The objective of this study was to generate and evaluate large-area maps of forest above-ground biomass and stem-volume, predicted from the combination of TanDEM-X, Sentinel-2, and field reference data using the k nearest neighbor algorithm. An additional objective was to investigate the possibly added value of separating the predictions to individual tree species (pine, spruce and birch), and to assess the tree species-classification accuracy with this approach. Furthermore, the effects of using a recently developed error model was assessed, by compensating for sampling errors in the field references.

2. Material and methods

2.1. Background

This paper is based on the work from two other applied projects, which therefore have set boundary conditions for some of the practicalities, such as the study area (the map is not covering the entire Sweden) and the investigated methods. The study area was selected in the other projects due to the availability of high-resolution airborne photogrammetric data at 2-year cycles. Thus, the replacement of airborne data with TanDEM-X data for this study was limited by the extent of that project. The study area covers approximately 20,000,000 ha (Fig. 1) and corresponds to about 70% of the Swedish forest area

2.2. Satellite data

The SAR data from TanDEM-X were acquired between October 2015 and January 2016. In total, 518 scenes for entire Sweden were processed within a previous project (Persson et al., 2017), but only 169 scenes were used in this paper, since these were overlapping the study area. The height-of-ambiguity (HOA) is a parameter determining the sensitivity of the InSAR measurements to height changes (Krieger and De Zan, 2014; Papathanassiou and Cloude, 2001; Soja and Ulander, 2013). For all images, the HOA was between 46 m and 68 m, resulting in good forest height sensitivity and a low probability of phase ambiguities. The images were acquired during late fall and winter conditions, with temperatures ranging from −32 °C to +11 °C, and they were acquired in strip-map mode, HH polarization, and with the single-look complex resolution 2.5 m in slant range and 3.3 m in azimuth. The incidence angles at the scene centres were all between 38° and 45°. A digital terrain model (DTM) derived from the national laser scanning data was provided by Lantmäteriet (The Swedish National Land Survey) at a 2 m
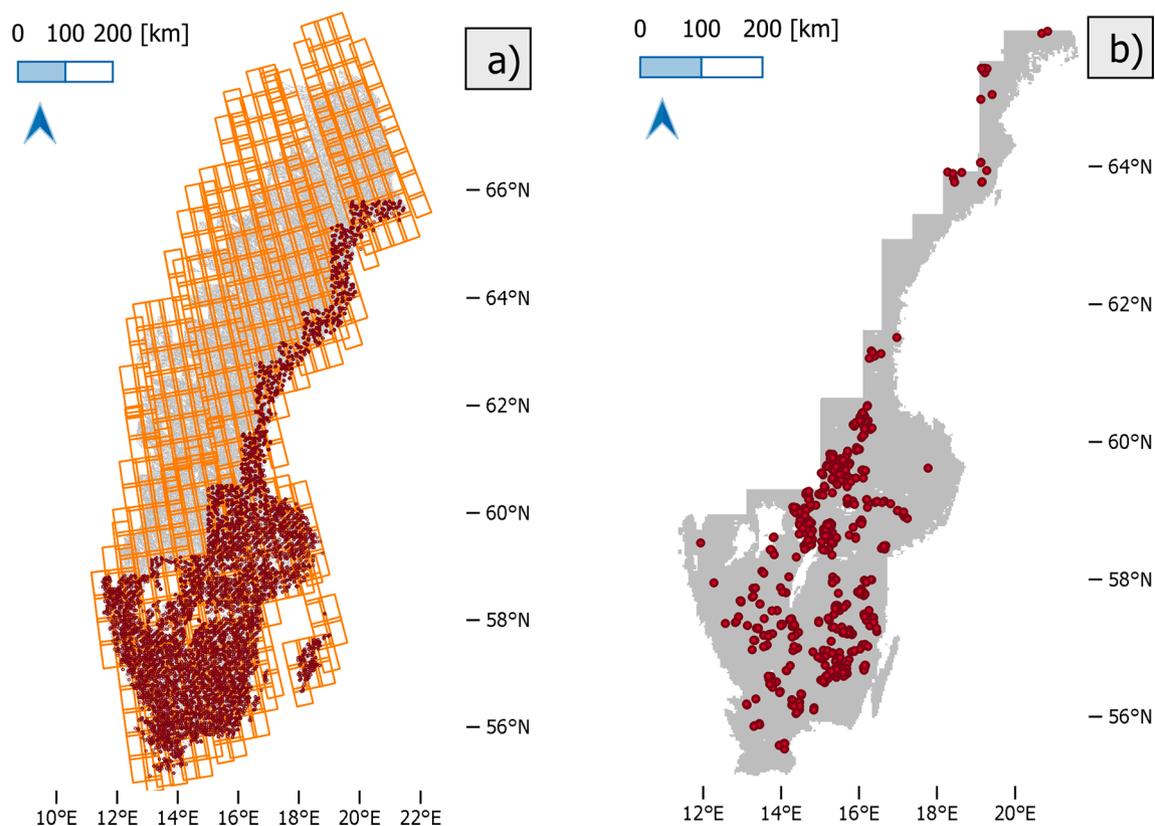

Fig. 1. (a) The extent of all 518 TanDEM-X scenes (outlined in orange) and the locations of all NFI plots (in red) for the investigated area. **(b)** Extent in gray of the study area and the locations (in red) of the 549 evaluation stands. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

grid cell resolution. The DTM was used as ground reference during interferometric processing.

The TanDEM-X data were delivered in the Coregistered Single look Slant range Complex (CoSSC) format. Following a common InSAR processing approach, explained extensively in Persson et al. (2017), a complex interferogram ($\widetilde{\gamma}$) was computed using:

$$\widetilde{\gamma} = \frac{\langle s_1 s_2^* e^{-i\phi_0} \rangle}{\sqrt{\langle |s_1|^2 \rangle \langle |s_2|^2 \rangle}} \qquad (1)$$

where s_1 and s_2 are the two interferometric images, * is the complex conjugate operator, ϕ_0 is the interferometric phase due to topographic variations (simulated from the DTM), and $\langle \cdot \rangle$ denotes spatial averaging using a window of 5×5 in range and azimuth, respectively.

The complex interferogram was flattened with respect to the Earth curvature, and then filtered using the method by Richard M. Goldstein and Werner (1998). The flattened phase was unwrapped and converted to the phase height by scaling with the wavenumber k_z (Persson et al., 2020b, 2017; Persson and Fransson, 2017), where

$$k_z = \frac{2\pi B_\perp}{\lambda R sin(\theta)} \qquad (2)$$

and $B_\perp$ is the perpendicular baseline, λ is the wavelength 3.1 cm, R is the average range to the satellites, and θ is the angle of incidence. The interferometric coherence γ was estimated from the flattened interferogram, using a coherence window of 3×3 pixels:

$$\gamma = \frac{\left| \sum s_1 s_2^* \right|}{\sqrt{\sum |s_1|^2 \sum |s_2|^2}} \qquad (3)$$

The estimated coherence was then corrected for decreasing signal-to-noise ratio as described in (Chen et al., 2016; Kugler et al., 2014; Persson et al., 2017).

The backscatter σ^0 was computed from the multilooked intensity images (β^0), based on the CoSSC data. The calibration gain (provided in the metadata for each acquisition) was subtracted and then a radiometric normalization based on the pixel-area method was applied, as described in (Frey et al., 2013; Small, 2011). The backscatter for each scene was computed as the arithmetic mean of the normalized backscatter image from respective satellite.

The spectral Sentinel-2 images were downloaded from the image portal 'Saccess', which is provided by the Swedish National Land Survey. The images were acquired with both the Sentinel-2A and Sentinel-2B satellites, and made available in the portal as level 1C only (top-of-atmosphere reflectance). The images were acquired mostly in 2017, but due to clouds, the period was extended to 2015 and 2016 as well to obtain a full cloud free mosaic over the study area. About 88% of the study area was covered with only four Sentinel-2 images (the full 290 × 290 km^2), but eight additional images were required to reach 100% coverage. All 12 images were manually selected, and resampled to 12.5 × 12.5 m^2 resolution using the nearest neighbor algorithm, in order to match the other data used. All 10 m and 20 m bands were downloaded, while the 60 m bands were considered to have too low resolution and were hence skipped. Additional pre-processing of the Sentinel-2 images was found to be unnecessary.

The pixel values corresponding to the NFI plots were extracted from the TanDEM-X and Sentinel-2 images, using the average from a circle with 30 m radius for the TanDEM-X metrics (PH, COH, S0). Experience from previous studies (Persson et al., 2019, 2017) showed that the extraction of SAR data for areas of such sizes decreased the variability and improved the estimation accuracy, in terms of lower standard error, compared to using only a few pixels overlapping the small field plots. The bilinear average of the pixels matching the field plots were used for the Sentinel-2 images.

2.3. Field reference data

Sweden has a land area of 41 million ha and about 23.6 million ha is productive forest (SLU, 2019). Sweden is located mainly in the boreal forest region, though the southernmost parts are within the hemi-boreal and nemoral regions. The forest is dominated by Norway spruce (*Picea abies* (L.) H. Karst.), Scots pine (*Pinus sylvestris* L.) and birch (*Betula* spp.), where pine and spruce constitutes about 80% of the growing stock and birch 12%. In this study, all deciduous forest types were approximated as birch forest, since birch constitute 86% of the deciduous forest volume in Sweden. Yet, in the hemi-boreal and nemoral regions there are also forests dominated by beech (*Fagus sylvatica*) and oak (*Quercus robur*). The current total stem volume of growing stock is 3,533 million m^3 and the total AGB is more than 2,660 million tons (t). For the area used in this study, the average AGB and VOL is 91.9 t/ha and 165 m^3/ha, respectively.

The Swedish NFI is collecting data on about 11,000 circular plots annually, where about 1/3 are temporary plots, and 2/3 are permanent plots (Fridman et al., 2014). The latter are revisited every 5 years. The temporary plots have a radius of 7 m and the permanent plots 10 m. The plots are distributed in clusters (Fig. 1a), but the spatial separation of plots within each cluster is always >200 m, which means the plots can be regarded as spatially uncorrelated. All trees are registered for species, diameter at breast height (DBH), and the tree heights are measured on a sub-sample of trees. Refined variables, i.e., the AGB and VOL, were estimated tree-wise using established formulae (Brandel, 1990; Marklund, 1988, 1987; Näslund, 1947) and then aggregated to plot-level values. Statistics for the field data are presented in Table 1.

The entire forest map was processed tile-wise, but the number of available field plots for a single tile (2.5 × 2.5 km^2) in one year was too few to attain robust predictions. Therefore, plots from a longer time period (forecasted or backdated to 2016 to match the satellite data) and outside the tile (but within the same Sentinel-2 satellite scene) were used. Earlier projects (Persson et al., 2020b; Reese et al., 2003; Tomppo et al., 2008) have shown that about 500–1000 plots are suitable, to sufficiently cover the range of available volumes, which is necessary for accurate estimations using the kNN method. The search radius (from the tile center) and temporal period were therefore successively increased in an iterative process until a sufficient number of plots were included. The time range was on average −4 years (max 10 years) and 2.5 year standard deviation, while the spatial radius was on average 70 km (max 110 km in northern Sweden due to the narrow shape) and 32 km standard deviation. To filter field plots that had been thinned or clear-cut between the field inventory and the satellite data acquisition, a linear regression model was created as $AGB = k \cdot Band11$, where k denotes the slope coefficient. Then, all plots with model residuals exceeding 2.5 standard deviations were removed, which resulted in approximately 800 available field plots for each tile.

For evaluation, stand-level data from the company Sveaskog were used. The company is owned by the Swedish state and is the largest forest owner in Sweden (it owns about 13% of Sweden's total forests). In 2017, they carried out an extensive field inventory of about 2% of their entire forest assortment, covering 2400 forest stands. The purpose was to obtain an accurate estimate of their entire forest holding, reference data for prediction modeling, and evaluation data for their previous

Table 1
Properties for field datasets (for the evaluation also divided per tree-species).

Dataset	Mean (t/ha)	Sd (t/ha)	Min (t/ha)	Max (t/ha)	N
NFI (training)	91.9	74.3	0	598.1	25,520
Sveaskog (evaluation)	108.1	48.3	2.75	289.3	549
Sveaskog (pine)	44.0	34.4	0	167.5	549
Sveaskog (spruce)	50.3	49.3	0	270.9	549
Sveaskog (birch)	10.0	15.2	0	142.1	549

forest management. Out of these, 928 stands were covering the area investigated in this paper. Due to the time difference between the satellite and field datasets, clear-cut stands and stands largely overlapping with non-forest land features (e.g., roads and mires) were removed. This resulted in 549 forest stands (Fig. 1b) used for the evaluation (Tables 1 and 2). The inventoried stands were sampled with circular plots of dynamic radius (from 5 to 10 m), using a probability sample approach with a systematic grid, but with random reference location. The radius was fixed for all plots within a stand, but adjusted between stands, to obtain on average 15–20 trees per plot. This approach led to an average of 7 plots per stand (5–12, depending on the stand size and shape). The separate plot averages were used to compute the within-stand variance between the plots. Trees with a DBH > 4 cm were measured and the subjectively estimated tree-species composition was registered. The plot-level AGB and VOL was computed using established equations valid for the region (the same approach as for the NFI plots in the first dataset). Stand-level estimates were computed as plot averages for each stand, and the dominant tree species was determined. The stand size properties are presented in Table 2, and the distribution of stand size is presented in Fig. 2.

2.4. Modelling and prediction

The modelling was based on the kNN algorithm, which was implemented in the software R using the package yaImpute (Crookston and Finley, 2007). It requires two parameters: *method* and *k*. yaImpute provides an implementation of kNN regression where the "randomForest" method was used for computing the distance and finding neighbors. It assumes the distances are following a beta distribution and the distance measure is one minus the proportion of trees where a target observation is in the same terminal node as a reference observation (Crookston and Finley, 2007). Different values for the number of neighbors (k) have been tested in a previous Swedish study, and the results for k = 1, 3, or 5 neighbors were reported in Persson et al. (2020a). It was found, that lower values on k allowed a wider range of predicted values, but also a larger variance at the pixel level. A larger k caused predictions with a smaller variety between nearby pixels, hence giving a map that appeared smoother. The parameter k = 5 was selected for this study, since this caused the smallest bias and with a low RMSE in Persson et al. (2020a). Unbiased predictions are particularly important for products where aggregation over larger regions may be used, since a bias would otherwise cause the error to accumulate with increased areas.

Previous experience from generating national forest maps from optical satellite data and kNN (Reese et al., 2003, 2002; Tomppo et al., 2008) were used as a basis for the selection of relevant bands and radar metrics, and then further empirical testing were carried out at the current study area. We found that the Sentinel-2 bands 4 (red) and 8 (NIR) with 10 m resolution, and the 20 m bands 5 (vegetation red edge), 7 (vegetation red edge), 8a (narrow NIR), 11 (SWIR), and 12 (SWIR) were meaningful to include and contributed to the overall performance. From the available radar metrics, the inclusion of only the interferometric phase height (PH) gave the best results. A project requirement was that the same explanatory variables should be used for all tiles, which facilitates the statistical assessments. The kNN model was used to predict values for all pixels (12.5 × 12.5 m^2) within each tile. The predicted variables were AGB, and VOL divided per tree-species (pine, spruce and birch). The proportion of different tree-species was assumed the same for both AGB and VOL. Negative AGB or PH < 2 m were removed.

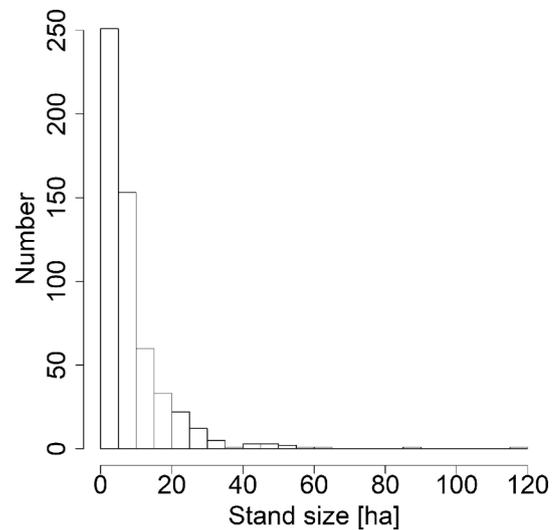To determine the accuracy of dominant tree species classification, a

Fig. 2. Distribution of stand size in hectare for the stands used for evaluation.

confusion matrix was computed, where the dominant predicted tree-species was compared with the dominant tree species from the field. The dominant species was defined as the most common of pine, spruce, and deciduous trees in terms of volume.

2.5. Uncertainty evaluation

The predictions were evaluated with stand level inventory data from Sveaskog, using root mean square error (RMSE) and bias defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{Y}_i - Y_i\right)^2} \tag{4}$$

$$Bias = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{Y}_i - Y_i\right) \tag{5}$$

with Y being the reference, and $\widehat{Y}$ the prediction for stand i, and n denoting the number of stands.

The error at stand-level was furthermore estimated for AGB with a linear error model (Persson and Ståhl, 2020), which enabled comparisons with other studies using other datasets. The error model was defined as in Persson and Ståhl (2020):

$$T_{RS} = \lambda_0 + \lambda_1 \cdot T + \varepsilon \tag{6}$$

where T_{RS} denotes the RS-based estimate, T is the true AGB value, λ_0 is a systematic displacement, and λ_1 makes the systematic error change across the range of true values, and ε denotes the random errors, quantified by their standard deviation σ_ε. There were two main purposes for using the error model. First, to illustrate how the error varies with varying reference value, instead of reporting a single value as RMSE. Second, the error model (and underlying estimators) provided a framework to compute estimates of the errors in the stand-level field data, which were propagated to the reported ordinary RMSE. It therefore allowed to also report the corrected RMSE*, where the variance contribution due to random errors in the field data was removed (which was assumed to appear due to the use of a sample instead of a completely inventoried reference).

3. Results

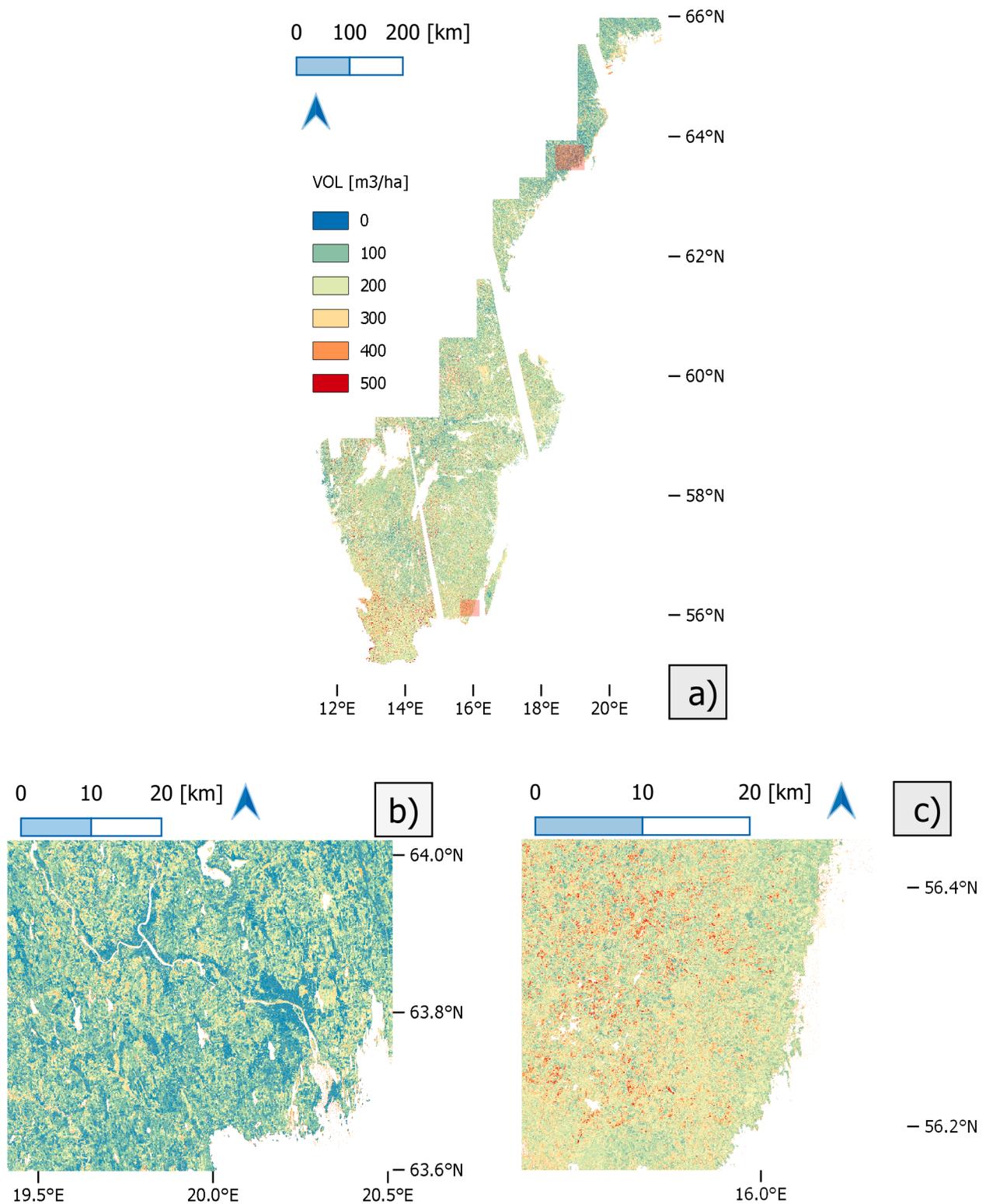The predicted AGB and VOL maps were mosaicked, and they covered the area illustrated in Fig. 3a.

Table 2
Stand size properties for the stands used for evaluation.

Dataset	Mean (ha)	Sd (ha)	Min (ha)	Max (ha)
Sveaskog (evaluation)	8.68	10.4	0.370	116

Fig. 3. a) The stem volume (VOL) predicted from Sentinel-2 and TanDEM-X and with the extents of b) and c) overlaid in red. **b)** zoomed-in extent from northeastern Sweden. **c)** zoomed-in extent from southeastern Sweden. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.1. Plot-level accuracy training data

Since the processing of the single tiles were using plot references from a surrounding, extending outside the current tile, the same plots were used many times but for different tiles. Therefore, both the statistics for the accumulated NFI references and the corresponding

estimated values with kNN are presented in Table 3. It also shows the RMSE, computed from the estimated-vs-NFI references for all the plots. Due to the large number of tiles (29,932) and reference plots (N = 25,520), the total number of plots became too large (N > 20 millions) to derive the leave-one-out accuracy, due to the processing time. The average VOL for the NFI plots used for the Sentinel-2 estimates was

Table 3

The statistics for the NFI plots (for TanDEM-X and Sentinel-2 + TanDEM-X), and also RMSE for the estimations, when accumulated over all tiles. The number of plots N is higher for S2, since there were orbits with missing data in the TDM images.

Datasets (training)	Mean (m³/ha)	Sd (m³/ha)	Min (m³/ha)	Max (m³/ha)	N	RMSE	RMSE%
NFI	221.4	136.3	0	1050.9	20,985,934	–	–
Estimated S2	188.0	126.3	0	1050.9	23,667,925	143.1	68.0
Estimated TDM	215.6.	136.4	0	1050.9	20,985,934	149.5	67.6
Estimated S2 + TDM	201.9	125.8	0	1050.9	20,985,934	131.3	59.3

lower (210.4 m³/ha) than those that used TanDEM-X (221.4 m³/ha).

3.2. Stand-level accuracy evaluation data

The accuracy for AGB and VOL at stand-level are presented in Table 4. The relative error values for AGB and VOL were similar, and the biases were small but not negligible. The bias constituted approximately 8% (VOL) to 10% (AGB) of the RMSE (where MSE is bias^2 + variance). For higher reference values, the underestimation increased, while for lower values, the overestimation increased (Fig. 4a and 4b). The correlation between the predicted AGB and VOL was 0.98, and this indicates that the prediction performance for any of these variables are indicative for the other as well.

The under- and overestimation visible in the scatter plots can be parametrized with the error model, which parameters for AGB are listed in Table 5, and illustrated in Fig. 5. The positive λ_0^* parameter indicates the overestimation at low values, and with $\lambda_1^* < 1$ the overestimation decreases with increasing reference values. Approximately unbiased predictions were obtained for values close to the sample mean for the NFI dataset used as training (Table 1). The RMSE* values that were corrected for the uncertainty in the reference data (due to using a sample) are also listed. The corrected RMSE decreased to 18.0 t/ha for AGB, which reduced the relative accuracy to 16.6%. This means that about 43% of the RMSE computed with the traditional estimator (Eq. 4) appeared due to random errors in the sampling of the stands. The bias is not affected by the correction in absolute terms, but with the significantly lower corrected RMSE*, the relative contribution due to bias* increased to 18%.

3.3. Influence of tree species

The accuracy of the predictions were also assessed for the species-wise AGB, predicted from the combination of TanDEM-X and Sentinel-2 (Table 6). The AGB for respective species-fraction for each stand (hence n = 549 for each species) was plotted against the corresponding reference in Fig. 6. The RMSE in absolute terms was similar (pine) or better (spruce and birch) compared to the accuracies for the overall AGB (in Table 4). However, since the prediction accuracy largely depends on the total AGB, comparisons with fraction-wise AGB averages has little meaning. The bias was small for pine and spruce, while evident for birch (Table 6). The same trend for under- and overestimation was observed when inspecting the individual tree species compared to the overall trend.

A species-classification at the stand-level was derived by classifying the dominant tree species by assigning the majority species in terms of
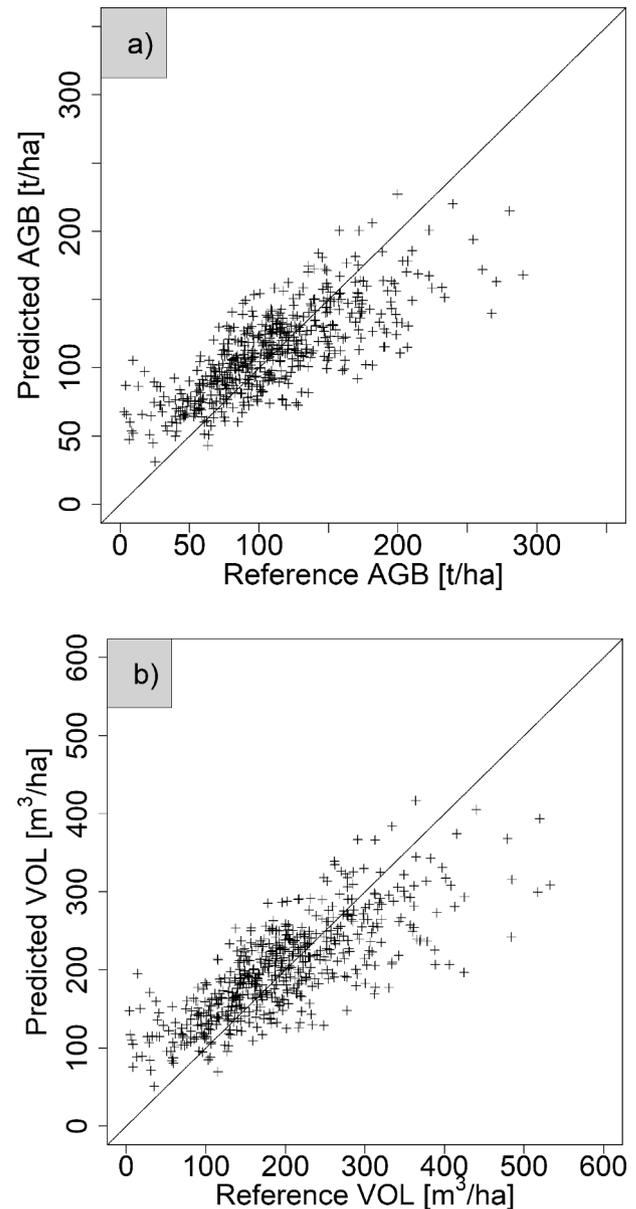

Fig. 4. Scatter plots of (a) AGB and (b) VOL, predicted with kNN vs. field references at stand-level.

AGB as the dominant predicted tree-species. These were compared with the dominant species in the field reference and a confusion matrix was calculated (Table 7) in addition to the user's, producer's, and overall accuracy, and kappa statistics. The classification clearly captured coniferous vs. deciduous forest. The largest confusion appeared between pine and spruce forest, where 24% of the pine-dominated stands were classified as spruce, and on the other hand, 22% of the spruce stands were classified as pine. However, the stands were not consisting of a single tree-species. The overall accuracy was 77% and the Kappa value

Table 4

Stand-level accuracy for kNN predictions based on the combination of Sentinel-2 (S2) and TanDEM-X (TDM), or each data source separately.

Dataset	Variable	RMSE	Bias
S2 + TDM	AGB	31.4 t/ha (29.1%)	3.17 t/ha (2.93%)
S2 + TDM	VOL	59.0 m³/ha (30.2%)	4.71 m³/ha (2.41%)
S2	AGB	39.2 t/ha (36.3%)	5.01 t/ha (4.64%)
S2	VOL	74.2 m³/ha (37.9%)	8.36 m³/ha (4.27%)
TDM	AGB	34.0 t/ha (31.5%)	9.17 t/ha (8.49%)
TDM	VOL	62.4 m³/ha (32.0%)	15.1 m³/ha (7.75%)

Table 5
Error model results.

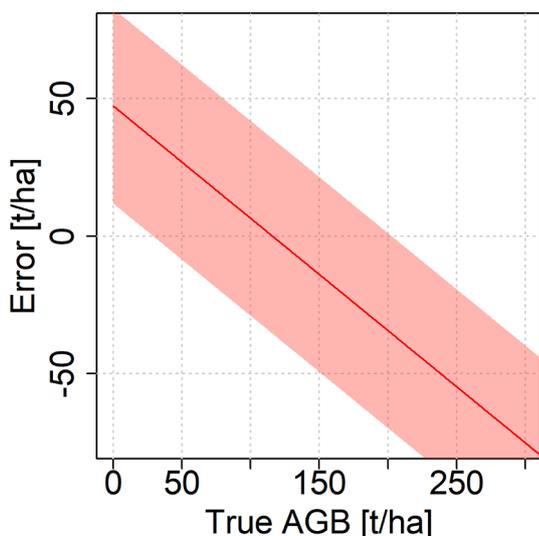| Parameter	Value
λ_0^*	47.3
λ_1^*	0.592
$\sigma_\varepsilon^{2^*}$ (t^2/ha^2)	314
RMSE* (t/ha)	18.0 (16.6%)
Bias* (t/ha)	3.17 (2.93%)

Fig. 5. Visualization of the error structure for the stand-level evaluation. The solid line is due to the residuals, computed from λ_0^* and λ_1^*, and the band width represents the random errors as $\pm 2\sigma_\varepsilon^*$.

Table 6
Accuracy per tree species for evaluation stands. Since both the total AGB and the fractions for each tree-species is available for the stands, it is possible to derive separate RMSEs, and hence n = 549 for all species.

Model	RMSE (t/ha)	Bias (t/ha)	n
Spruce AGB	25.1	0.572	549
Pine AGB	30.7	0.610	549
Birch AGB	18.4	5.49	549

57% (Table 7), which might be perceived as moderate results. The spectral values were extracted from a single image for each location, and previous studies (Grabska et al., 2019; Persson, 2018) have shown that using several spectral images improves such classifications.

3.4. Influence of location and size

The error in terms of the residuals (predicted – reference) was investigated with respect to location and stand-size (Fig. 7), for the predictions from the combination of TanDEM-X and Sentinel-2. The residuals in terms of absolute values did not show any trend with respect to northing or easting. However, since AGB tend to be higher in southern Sweden, the expectation was to obtain larger residuals further south and closer to the coasts (generally further east).

The residuals were approximately normally distributed with respect to stand size, but the distribution was negatively skewed, with a median of 8.56 t/ha.

4. Discussion

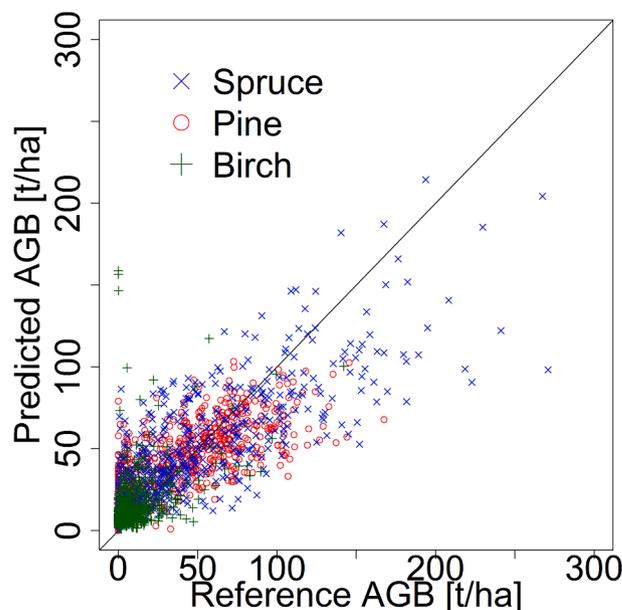This study used kNN to predict AGB and VOL from TanDEM-X,

Fig. 6. Visualization of the predicted vs. reference AGB fractions divided on tree-species.

Table 7
Confusion matrix for the dominant tree species.

Prediction	Reference			
	Birch	Pine	Spruce	Users accuracy
Birch	15	9	12	0.417
Pine	1	240	61	0.795
Spruce	3	77	207	0.721
Producers accuracy	0.789	0.736	0.739	Overall 0.769

Unweighted Kappa statistic: 0.570.

Sentinel-2 and field plots. The contributions from each data source separately and in combination were evaluated (Tables 3 and 4). The RMSE at stand-level (Table 4) for the AGB and VOL predictions were lower when only TanDEM-X (~32% RMSE) was used as data source, compared with the predictions from only Sentinel-2 (~37%). Such differences could not be seen clearly at the plot-level (Table 3, training data), where the accuracies for the separate data sources were both about 68%, but the use of their combination still showed a clear improvement, reaching an RMSE of 59%. The predictions from the combination of both TanDEM-X and Sentinel-2 were significantly better at the stand-level, with ~29% RMSE, which indicates that both data sources contribute with non-overlapping information, useful for mapping of AGB and VOL. The authors do not know about other studies where TanDEM-X and Sentinel-2 have been combined for mapping of AGB and VOL. However, Wittke et al. (2019) compared Sentinel-2 with other sensors for the prediction of forest variables in boreal forest. They reported that Sentinel-2 had a similar accuracy as predictions of AGB and VOL based on TerraSAR-X, which in another study using the same test site, showed lower accuracy than predictions based on TanDEM-X. In Persson et al. (2013), they investigated how the inclusion of heights from different image-matched stereo data combined with spectral SPOT-5 images could be used for AGB estimation. The best combination improved the results based only on spectral data by lowering the RMSE from 33% to 22%. Other studies combining optical satellite images and some height data source were mainly relying on airborne laser (Badreldin and Sanchez-Azofeifa, 2015; Fernández-Landa et al., 2018) or image matched heights from drones (Puliti et al., 2018). Other combinations did not add any height data, but only two-dimensional metrics, e.g., backscatter (Alan et al., 2017; Chang and Shoshany,
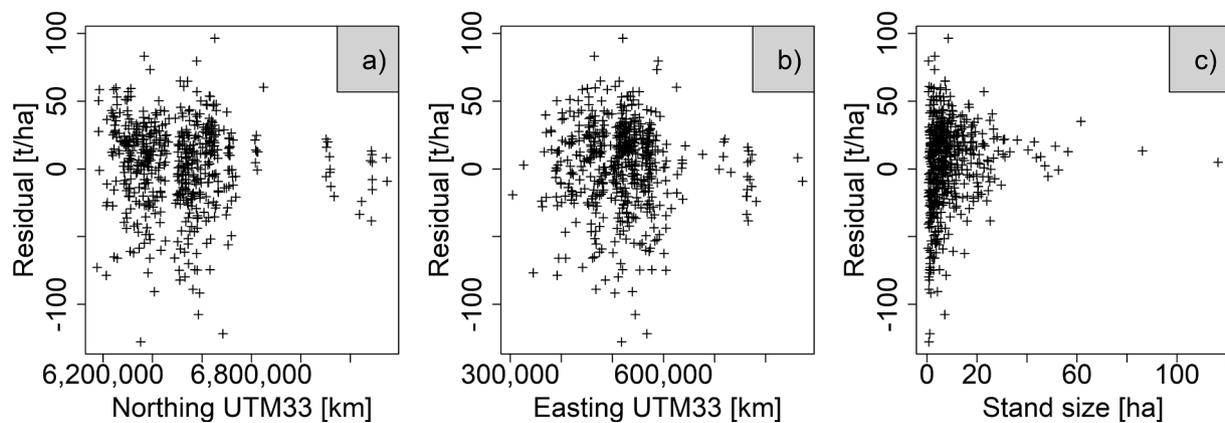
Fig. 7. Visualization of the error vs. northing, easting and stand size.

2016; Debastiani et al., 2019).

The RMSEs of our study were in the upper range in absolute terms (31.4 t/ha, 59.0 m^3/ha) compared to those presented in Persson et al. (2017), 27–30 t/ha and 52–65 m^3/ha, where a regression model was used for the entire country. However, Persson et al. (2017) used a different reference dataset for the evaluation, and different samples of NFI plots were used for the predictions. In relative terms, the accuracy of 29–30% RMSE (Table 4) at stand-level was similar or lower than most large-area studies based on SAR backscatter for the boreal region (Fransson and Israelsson, 1999; Santoro et al., 2011, 2006, 2002). It was higher though, compared to the previous study by Persson et al. (2017, 21–25%) which was based on TanDEM-X data, but employing linear regression instead of kNN as estimation method.

Studies reporting lower RMSE have generally used multitemporal data and/or reported the accuracies from local regions. This tendency has been noted with the dataset used in this study as well, where its accuracy at a local test site was reported in Persson et al. (2020a). The RMSE was about 77–83 m^3/ha (24–26%), compared to the 59 m^3/ha (30%) in the current study where a large part of the country was used. Hence, the relative RMSE can be misleading, due to the different sample means. In addition, reference data from test sites are often collected with greater care (and sometimes without using a sample – fully inventoried stands), than in large-scale operational inventories. The evaluation stands in our study were collected within such a large-scale inventory and may therefore contain more errors than the evaluation data used in Persson et al. (2020a). The stands were reasonably distributed over the entire study area, which decreases the risk of large deviations from the prediction accuracies if used locally.

When the uncertainties in this study were removed (related to using sampled stands as reference), the accuracy was evidently improved, to 16.6% (18.0 t/ha). That difference supports the previous statement that the reported accuracy really depends on the quality, size and number of reference data. Additional support for this premise can be noticed by comparing the error parameter value σ_ε^{2*}. It was 314 t^2/ha^2 in this study, and 114 t^2/ha^2 at a test site in northern Sweden, reported in (Persson and Ståhl, 2020) and based on estimates made from TanDEM-X data only. The λ_1^* parameter (0.592, Table 5) deviated considerably from one, which showed a weakness with the current approach, since it should be close to one to avoid varying under- or overestimation.

Reported accuracies depend largely on the empirical data used, and the size of the evaluation units. In Tomppo et al. (2008), they reported the relative RMSE vs. stand size, when using the kNN method with spectral data and field plots, and when the areas decreased below 10 ha (which was approximately the mean evaluation stand size in this study, Table 2), the relative RMSE reached 40–80%. In our study, no such changes in RMSE were noticed with varying stand size (Fig. 7c). The residuals appeared approximately normally distributed, but positively

biased (3 t/ha, described in Table 4). The residuals were similar regardless of northing, and easting. This suggests that the evaluated approach is robust, likely explained by the predictions being carried out locally and tile-wise.

The assessed approach based on kNN is generally taking advantage of using reference data only from the "local" surrounding (region), which therefore should make the predictions more valid at the local level, although the overall accuracy in terms of RMSE appears moderate. Yet, a large advantage of using kNN over parametric models was that it allowed to easily (and smoothly) combine the TanDEM-X and Sentinel-2 data in a way that caused the predicted tiles to not show any seamlines. The Sentinel-2 images extended over many TanDEM-X scenes and hence balanced this problem. Seamlines were obvious in the kNN raster based only on TanDEM-X, and they were also noticed between different satellite scenes in Persson et al. (2017), particularly for scenes acquired in different weather conditions. The kNN approach therefore facilitates combining different data sources, which implies that when the data sources have their extents at different locations, it contributes to a smoother mosaic.

Species-wise information was possible to derive due to the inclusion of Sentinel-2 images, and although not all the spectral data may be possible to acquire within a single year, the tree species do not change rapidly, and hence a higher tolerance can be accepted for a temporal discrepancy between the spectral data and the other data sources. The species-wise fractions of AGB or VOL were used to classify the dominant tree-species at stand-level, and the overall accuracy of 77% (Table 7) was similar as reported by Grabska et al. (2019) and Persson et al. (2018). Persson et al. (2018) found, that by using a single Sentinel-2 image, classification accuracies of about 72–79% could be reached. By extending the dataset from one to four images, these values were improved to 86%. The Kappa value of 0.57 in our study indicated a moderate reliability of the classification, and was slightly lower than those reported by Persson et al. (2018).

The main application in Sweden for this type of map is to cover the time gaps in-between the national laser scannings, which currently repeat about every seven years. This time-span, however, describes the interval before the same area is covered again. The scanning of Sweden is carried out continuously. Corresponding laser based products provide stand-estimates with an accuracy of about 20% RMSE (Nilsson et al., 2017). Therefore, some users may use the kNN prediction rasters proposed in this study directly, while others may prefer to forecast a laser based base product, and only use the kNN product to filter out large changes or to estimate the growth rate for a forecasting. In the past, kNN based maps were generated every five years and used by many Swedish users, since there were no other alternatives at the national level. With the proposed method, the map gives a national snapshot valid for a single year, in contrast to laser based maps where different parts of the country are acquired in different years. Many countries globally do not

have operational laser scanning programs, although most countries operate national forest inventories. Therefore, the tested kNN approach is likely even more valuable globally. The approach is expected to transfer well, since the required training data come from a limited, rather local, surrounding of each tile.

To contrast the positive aspects discussed earlier, the kNN method is however rather slow to apply, requires much field data, and the predicted values cannot be extrapolated. Depending on the preconditions, these may crucially affect the best choice of method. The approach was yet convenient for combining uncalibrated data sources (TanDEM-X with respect to weather conditions, especially temperature, and Sentinel-2 reflectance levels with respect to the other scenes) for prediction of forest variables.

5. Conclusions

This study demonstrated the use of the kNN method with k = 5 neighbors, applied to data from TanDEM-X, Sentinel-2, and their combination, together with field reference data, to predict above-ground biomass and stem volume of a large area in Sweden. The kNN method was an efficient and convenient way of combining the data sources in order to generate a seamless mosaic of predictions with sufficient accuracy. It did furthermore provide species-wise fractions, which were used to classify the dominant tree species at stand-level. The overall accuracy of 77% was similar to what has been shown in previous studies, and since the prediction accuracy of above-ground biomass and stem volume was not negatively affected, this approach may be in favor over other methods with similar accuracy, due to its benefits.

The prediction accuracy was similar to previous large-area products based on the assessed data sources, and similar or better than large-area products based on SAR backscatter or only spectral data. The use of a linear error model enabled the use of estimators that correct for the sampling errors in the field references. This improved the accuracy of above-ground biomass in terms of RMSE with 43%, which was reduced from 31.4 t/ha to 18.0 t/ha. The error model moreover enabled the separation of bias and variance, where the influence of bias was generally limited.

We conclude that when sufficient amounts of field data are available, the kNN algorithm provides a convenient way of deriving above-ground biomass and stem volume, and obtain a solid source for the dominating tree species. The use of interferometric phase height from TanDEM-X provided the basis for accurate volumetric predictions, while the spectral reflectance from Sentinel-2 was key to the predictions of tree-species.

CRediT authorship contribution statement

Henrik J. Persson: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization, Project administration, Funding acquisition. **Jonas Jonzén:** Data curation, Software, Resources, Writing - review & editing. **Mats Nilsson:** Methodology, Resources, Software, Writing - review & editing, Project administration, Funding acquisition.

Declaration of Competing Interest

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

References

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

Askne, J.I.H., Persson, H.J., Ulander, L.M.H., 2019. On the Sensitivity of TanDEM-X-Observations to Boreal Forest Structure. Remote Sens. 11, 1–22.

Askne, J.I.H.H., Persson, H.J., Ulander, L.M.H., 2018. Biomass Growth from Multi-Temporal TanDEM-X Interferometric Synthetic Aperture Radar Observations of a Boreal Forest Site. Remote Sens. 10, 18. https://doi.org/10.3390/rs10040603.

Badreldin, N., Sanchez-Azofeifa, A., 2015. Estimating forest biomass dynamics by integrating multi-temporal Landsat satellite images with ground and airborne LiDAR data in the Coal Valley Mine, Alberta, Canada. Remote Sens. 7, 2832–2849. https://doi.org/10.3390/rs70302832.

Brandel, G., 1990. Volymfunktioner för enskilda Träd : tall, gran och björk = Volume functions for individual trees: Scots pine (Pinus sylvestris), Norway spruce (Picea abies) and birch (Betula pendula & Betula pubescens). Swedish University of Agricultural Sciences, Garpenberg.

Chang, J., Shoshany, M., 2016. Mediterranean Shrublands Biomass Estimation Using Sentinel-1 and Sentinel-2. In: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, pp. 5300–5303. https://doi.org/10.1109/IGARSS.2016.7730380.

Chen, H., Cloude, S.R., Goodenough, D.G., 2016. Forest Canopy Height Estimation Using Tandem-X Coherence Data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 9, 3177–3188. https://doi.org/10.1109/JSTARS.2016.2582722.

Chen, L., Ren, C., Zhang, B., Wang, Z., Xi, Y., 2018. Estimation of Forest Above-Ground Biomass by Geographically Weighted Regression and Machine Learning with Sentinel Imagery. Forests 9, 1–20. https://doi.org/10.3390/f9100582.

Chirici, G., Giannetti, F., McRoberts, R.E., Travaglini, D., Pecchi, M., Maselli, F., Chiesi, M., Corona, P., 2020. Wall-to-wall spatial prediction of growing stock volume based on Italian National Forest Inventory plots and remotely sensed data. Int. J. Appl. Earth Obs. Geoinf. 84, 101959 https://doi.org/10.1016/j.jag.2019.101959.

Crookston, N.L., Finley, A.O., 2007a. yaImpute: An R Package for kNN Imputation. J. Stat. Softw. 23.

Debastiani, Aline Bernarda, Sanquetta, Carlos Roberto, Paula, A., Corte, D., Rex, Franciel Eduardo, Pinto, N.S., Debastiani, A.B., Sanquetta, C.R., Corte, A.P.D., Rex, F.E., Pinto, A.N., 2019. Evaluating SAR-optical sensor fusion for aboveground biomass estimation in a Brazilian tropical forest. Ann. For. Res. 62, 109–122. https://doi.org/10.15287/afr.2018.1267.

Fagan, M.E., DeFries, R.S., 2009. Measurement and Monitoring of the World's Forests: A Review and Summary of Technical Capability, 2009-2015., The World's Forests: Design and Implementation of Effective Measurement and Monitoring.

Fernández-Landa, A., Fernández-Moya, J., Tomé, J.L., Algeet-Abarquero, N., Guillén-Climent, M.L., Vallejo, R., Sandoval, V., Marchamalo, M., 2018. High resolution forest inventory of pure and mixed stands at regional level combining National Forest Inventory field plots, Landsat, and low density lidar. Int. J. Remote Sens. 39, 4830–4844. https://doi.org/10.1080/01431161.2018.1430406.

Fransson, J.E.S., Israelsson, H., 1999. Estimation of stem volume in boreal forests using ERS-1 C- and JERS-1 L-band SAR data. Int. J. Remote Sens. 20, 123–137. https://doi.org/10.1080/014311699213640.

Frey, O., Santoro, M., Werner, C.L., Wegmuller, U., 2013. DEM-Based SAR Pixel-Area Estimation for Enhanced Geocoding Refinement and Radiometric Normalization. IEEE Geosci. Remote Sens. Lett. 10, 48–52. https://doi.org/10.1109/LGRS.2012.2192093.

Fridman, J., Holm, S., Nilsson, M., Nilsson, P., 2014. Adapting National Forest Inventories to changing requirements–the case of the Swedish National Forest Inventory at the turn of the 20th century. Silva Fenn. 48, 1–29.

Grabska, E., Hostert, P., Pflugmacher, D., Ostapowicz, K., 2019. Forest stand species mapping using the sentinel-2 time series. Remote Sens. 11, 1–24. https://doi.org/10.3390/rs11101197.

Imhoff, M.L., 1995. Radar backscatter and biomass saturation: Ramifications for global biomass inventory. IEEE Trans. Geosci. Remote Sens. 33, 511–518. https://doi.org/10.1109/36.377953.

Karila, K., Vastaranta, M., Karjalainen, M., Kaasalainen, S., 2015. Tandem-X interferometry in the prediction of forest inventory attributes in managed boreal forests. Remote Sens. Environ. 159, 259–268. https://doi.org/10.1016/j.rse.2014.12.012.

Krieger, G., De Zan, F., 2014. Relativistic effects in bistatic synthetic aperture radar. IEEE Trans. Geosci. Remote Sens. 52, 1480–1488. https://doi.org/10.1109/TGRS.2013.2251640.

Kugler, F., Schulze, D., Hajnsek, I., Pretzsch, H., Papathanassiou, K.P., 2014. TanDEM-X Pol-InSAR performance for forest height estimation. IEEE Trans. Geosci. Remote Sens. 52, 6404–6422. https://doi.org/10.1109/TGRS.2013.2296533.

Lu, D., 2006. The potential and challenge of remote sensing-based biomass estimation. Int. J. Remote Sens. 27, 1297–1328. https://doi.org/10.1080/01431160500486732.

Marklund, L.G., 1988. Biomassafunktioner för tall, gran och björk i Sverige. Umeå, Sweden.

Marklund, L.G., 1987. Biomass functions for Norway spruce (Picea abies (L.) Karst.). Sweden. Swedish University of Agricultural Sciences, Umeå, Sweden.

Mendes, F. de S., Baron, D., Gerold, G., Liesenberg, V., Erasmi, S., 2019. Optical and SAR remote sensing synergism for mapping vegetation types in the endangered Cerrado/Amazon ecotone of Nova Mutum-Mato Grosso. Remote Sens. 11. 10.3390/rs11101161.

Näslund, M., 1947. Funktioner och tabeller för kubering av stående träd. Meddelanden från statens skogsforskningsinstitut, Stockholm.

Nilsson, M., Nordkvist, K., Jonzén, J., Lindgren, N., Axensten, P., Wallerman, J., Egberth, M., Larsson, S., Nilsson, L., Eriksson, J., Olsson, H., 2017. A nationwide forest attribute map of Sweden derived using airborne laser scanning data and field data from the national forest inventory. Remote Sens. Environ. 194, 447–454. https://doi.org/10.1016/j.rse.2016.10.022.

Papathanassiou, K.P., Cloude, S.R., 2001. Single-baseline polarimetric SAR interferometry. IEEE Trans. Geosci. Remote Sens. 39, 2352–2363. https://doi.org/10.1109/36.964971.

Persson, H., Wallerman, J., Olsson, H., Fransson, J.E.S., 2013. Estimating forest biomass and height using optical stereo satellite data and a DTM from laser scanning data. Can. J. Remote Sens. 39, 251–262. https://doi.org/10.5589/m13-032.

Persson, H.J., Fransson, J.E.S., 2017. Comparison between TanDEM-X and ALS based estimation of above ground biomass and tree height in boreal forests. Scand. J. For. Res. 32, 306–319. https://doi.org/10.1080/02827581.2016.1220618.

Persson, H.J., Fransson, J.E.S., Jonzén, J., Nilsson, M., 2020. Combining Tandem-x, Sentinel-2 and field data for prediction of species-wise stem volumes (Submitted). In: IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium. IEEE, p. 4.

Persson, H.J., Olsson, H., Soja, M.J., Ulander, L.M.H., Fransson, J.E.S., 2017. Experiences from large-scale forest mapping of Sweden using TanDEM-X data. Remote Sens. 9, 1253–1279. https://doi.org/10.3390/rs9121253.

Persson, H.J., Soja, M.J., Fransson, J.E.S., H. Ulander, L.M., 2019. Using the Two-Level Model with Tandem-X for Large-Scale Forest Mapping, in: IGARSS 2019 - IEEE International Geoscience and Remote Sensing Symposium. pp. 4484–4487. 10.1109/IGARSS.2019.8899886.

Persson, H.J., Soja, M.J., Fransson, J.E.S., Ulander, L.M.H., 2020b. National biomass mapping using the Two-Level model. Submitted 9.

Persson, H.J., Ståhl, G., 2020. Characterizing uncertainty in forest remote sensing studies at plot and stand level. Remote Sens. 12, 1–21. https://doi.org/10.3390/rs12030505.

Persson, M., 2018. Tree species classification using multi-temporal Sentinel-2 data. Swedish University of Agricultural Sciences.

Persson, M., Lindberg, E., Reese, H., 2018. Tree species classification with multi-temporal Sentinel-2 data. Remote Sens. 10, 1–17. https://doi.org/10.3390/rs10111794.

Puliti, S., Saarela, S., Gobakken, T., Ståhl, G., Næsset, E., 2018. Combining UAV and Sentinel-2 auxiliary data for forest growing stock volume estimation through hierarchical model-based inference. Remote Sens. Environ. 204, 485–497. https://doi.org/10.1016/j.rse.2017.10.007.

Rahlf, J., Breidenbach, J., Solberg, S., Næsset, E., Astrup, R., 2014. Comparison of four types of 3D data for timber volume estimation. Remote Sens. Environ. 155, 325–333. https://doi.org/10.1016/j.rse.2014.08.036.

Reese, H., Nilsson, M., Pahlén, T., 2003. Countrywide estimates of forest variables using satellite data and field data from the national forest inventory. AMBIO A J. 32, 542–548. https://doi.org/10.1639/0044-7447(2003)032.

Reese, H., Nilsson, M., Sandström, P., Olsson, H., 2002. Applications using estimates of forest parameters derived from satellite and forest inventory data. Comput. Electron. Agric. 37, 37–55. https://doi.org/10.1016/S0168-1699(02)00118-7.

Goldstein, Richard M., Werner, C.L., 1998. Radar interferogram filtering for geophysical applications. Geophys. Res. Lett. 25, 4035–4038.

Santoro, M., Askne, J., Smith, G., Fransson, J.E., 2002. Stem volume retrieval in boreal forests from ERS-1/2 interferometry. Remote Sens. Environ. 81, 19–35. https://doi.org/10.1016/S0034-4257(01)00329-7.

Santoro, M., Beer, C., Cartus, O., Schmullius, C., Shvidenko, A., McCallum, I., Wegmüller, U., Wiesmann, A., 2011. Retrieval of growing stock volume in boreal forest using hyper-temporal series of Envisat ASAR ScanSAR backscatter measurements. Remote Sens. Environ. 115, 490–507. https://doi.org/10.1016/j.rse.2010.09.018.

Santoro, M., Eriksson, L., Askne, J., Schmullius, C., 2006. Assessment of stand-wise stem volume retrieval in boreal forest from JERS-1 L-band SAR backscatter. Int. J. Remote Sens. 27, 3425–3454. https://doi.org/10.1080/01431160600646037.

SLU, 2019. Skogsdata 2019. Umeå, Sweden.

Small, D., 2011. Flattening gamma: Radiometric terrain correction for SAR imagery. IEEE Trans. Geosci. Remote Sens. 49, 3081–3093. https://doi.org/10.1109/TGRS.2011.2120616.

Soja, M.J., Askne, J.I.H., Ulander, L.M.H., 2017. Estimation of Boreal Forest Properties from TanDEM-X Data using Inversion of the Interferometric Water Cloud Model. IEEE Geosci. Remote Sens. Lett. 14, 1–5. https://doi.org/10.1109/LGRS.2017.2691355.

Soja, M.J., Ulander, L.M.H., 2013. Digital canopy model estimation from TanDEM-X interferometry using high-resolution lidar DEM. Int. Geosci. Remote Sens. Symp. 165–168 https://doi.org/10.1109/IGARSS.2013.6721117.

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.

Wittke, S., Yu, X., Karjalainen, M., Hyyppä, J., Puttonen, E., 2019. Comparison of two-dimensional multitemporal Sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest. Int. J. Appl. Earth Obs. Geoinf. 76, 167–178. https://doi.org/10.1016/j.jag.2018.11.009.