remote sensing

MDPI

Article

Prediction of Site Index and Age Using Time Series of TanDEM-X Phase Heights

Ivan Huuva [1,*], Jörgen Wallerman [1], Johan E. S. Fransson [2] and Henrik J. Persson [1]

[1] Department of Forest Resource Management, Swedish University of Agricultural Sciences, Skogsmarksgränd 17, SE-901 83 Umeå, Sweden; henrik.persson@slu.se (H.J.P.)
[2] Department of Forestry and Wood Technology, Linnaeus University, Georg Lückligs väg 1, SE-351 95 Växjö, Sweden
* Correspondence: ivan.huuva@slu.se

Abstract: Site index and stand age are important variables in forestry. Site index describes the growing potential at a given location, expressed as the height that trees can attain at a given age under favorable growing conditions. It is traditionally used to classify forests in terms of future timber yield potential. Stand age is used for the planning of management activities such as thinning and harvest. SI has previously been predicted using remote sensing, but usually relying on either very short time series or repeated ALS acquisitions. In this study, site index and forest stand age were predicted from time series of interferometric TanDEM-X data spanning seven growth seasons in a hemi-boreal forest in Remningstorp, a test site located in southern Sweden. The goal of the study was to see how satellite-based radar time series could be used to estimate site index and stand age. Compared to previous studies, we used a longer time series and applied a penetration depth correction to the phase heights, thereby avoiding the need for calibration using ancillary field or ALS data. The time series consisted of 30 TanDEM-X strip map scenes acquired between 2011 and 2018. Established height development curves were fitted to the time series of TanDEM-X-based top heights. This enabled simultaneous estimation of both age and site index on 91 field plots with a 10 m radius. The RMSE of predicted SI and age were 6.9 m and 38 years for untreated plots when both SI and age were predicted. When predicting SI and the age was known, the RMSE of the predicted SI was 4.0 m. No significant prediction bias was observed for untreated plots, while underestimation of SI and overestimation of age increased with the intensity of treatment.

Keywords: site index; time series; InSAR; height development curves; growth measurement; forestry; TanDEM-X

Citation: Huuva, I.; Wallerman, J.; Fransson, J.E.S.; Persson, H.J. Prediction of Site Index and Age Using Time Series of TanDEM-X Phase Heights. *Remote Sens.* **2023**, *15*, 4195. https://doi.org/10.3390/rs15174195

Academic Editor: Lars T. Waser

Received: 28 June 2023
Revised: 17 August 2023
Accepted: 18 August 2023
Published: 25 August 2023

1. Introduction

In forests, wood productivity is of interest in commercial forestry to determine economic value and to support the planning of silvicultural treatments. Mapping of forest productivity and age can also be useful in monitoring and modeling forest biomass (carbon stock) and changes in this over time. Forest productivity can be expressed in terms of site index (SI), a variable expressing the expected height of dominant trees at a reference age, given the local conditions. In addition to being a useful tool in economic assessments, forecasting, and planning in the commercial management of forests, large-scale mapping of SI can be used to quantify the effects of environmental changes, such as mean temperature changes or droughts, on the productivity of forests [1]. Such mappings can be used to make predictions about the geographically distinct consequences of climate change.

SI can be determined based on climatic and field conditions such as precipitation, temperature, and classification of soil strata, which is useful when no trees are present on the site. Another way to estimate SI uses age and dominant- or top-height measurements and is generally favored over the previous method due to its practicality, low cost, and higher accuracy. It requires the location to have an established even-aged forest and relies

on a strong correlation between volume growth and height growth [2]. The definitions of top height vary, where some are based on the mean height of dominant trees, and others on the maximum tree height, or the mean height of a certain percentage of the highest trees in an area. In Sweden, the SI estimated from top height describes the site productivity for the dominant species in terms of the achievable height in meters of the largest diameter trees at a specific reference age (A_{SI}). The top height is defined as the mean height of the 100 trees with the largest diameter at breast height per hectare. This definition of top height is sometimes called H100, and is meant to represent the upper height of tree crowns in the forest. Top height has successfully been estimated using different remote sensors. Examples include estimating top height with the maximum airborne laser scanning (ALS) canopy model height in a 10 m × 10 m window, or the maximum height in 500 m^2 plots of an aerial stereo-image-based canopy height model (CHM) [3,4].

While SI is not the most commonly estimated forest variable, it has been successfully predicted, often together with the related variable stand age, using a few different approaches, sensors, and sensor combinations. Commonly stand or tree heights are estimated by some remote sensing techniques and compared to known height–age curves to determine SI and age [5–11]. Véga and Onge [5] used CHMs based on historical aerial photographs and ALS from four time points spanning a period of 58 years to predict SI and age. Their models were estimated by minimizing mean absolute residuals to age-height curves, where the heights were extracted from CHMs calibrated with individual tree growth reconstruction. This method required counting tree rings on the cross-sections of felled trees to derive correction equations between tree heights from manual CMH interpretation and the field reconstructed heights. The procedure resulted in 2.4 m RMSE for SI and seven years RMSE for age predictions on 400 m^2 plots. Kandare et al. [6] used an individual tree crown (ITC) approach for predicting SI in boreal forests using airborne laser scanning (ALS) and hyperspectral data. They estimated the age, height, and diameter at breast height of the dominant trees from ALS and hyperspectral metrics. These were then used in age-height curves to predict SI. When predicting both SI and age, the method by Kandare et al. achieved RMSEs of 4.3 m and 34 years, respectively. When the age from field data was used in the prediction, the RMSE of SI predictions dropped to 1.18 m [6]. Solberg et al. used age-independent equations of top height growth and single tree ALS data to predict SI by matching single dominant trees in repeated ALS measurements six years apart [7]. They estimated SI values very close to field-based values for individual sample trees (bias 0.27 m, RMSE about 2.8 m, as interpreted from a figure). Penner et al. [8] used two successive ALS collections, acquired 13 years apart, to estimate SI with an RMSE of 2.5 m and a bias of 0.3 m on 400 m^2 field plots.

Many of the reported results are good but require access to long time series, as in the case of [5], rely on relatively costly ALS data, usually from several years, or on local calibration of remote sensing data or predicted attributes. Synthetic Aperture Radar (SAR) provides a cost-efficient alternative to ALS and aerial photography that is independent of sunlight and relatively unhindered by clouds and precipitation, thereby providing reliable year-round coverage of large parts of the world from different spaceborne systems. These operate in different parts of the microwave spectrum, called bands, corresponding to different wavelengths. Shorter wavelengths, such as the X and C bands, have significant contributions from the top part of the canopy and are, therefore, well suited for canopy height estimation using single-pass SAR interferometry (InSAR). Sentinel-1, a C band SAR system that provides open access data over large parts of the world, does, however, not have single-pass capability. TanDEM-X (TerraSAR-X add-on for Digital Elevation Measurement) is a two-satellite constellation that captures single-pass interferometric InSAR images at X-band (wavelength 3.1 cm). It provides data over a large part of the world and has proven itself valuable in forest variable retrieval [12,13]. Several studies have used TanDEM-X for the retrieval of forest variables [3,14–17]. Many of these have estimated forest heights from TanDEM-X data [3,14,17–21], and a few have investigated

height development due to deforestation, silvicultural treatments, or growth [22–24], or used phase height development to estimate biomass and volume changes [25–27].

The use of TanDEM-X data for SI prediction has so far been limited to Persson and Fransson [9], Wallerman et al. [10], and Persson and Fransson [11]. In these studies, simple linear models relating TanDEM-X phase heights to ALS percentiles or Lorey's heights (i.e., basal area weighted mean heights) from field data were used as calibration. Wallerman et al. [10] estimated SI when the age was provided, with an RMSE of 18.6% (corresponding to around 6–7 m, as interpreted from a figure) on 314 m^2 plots. They used TanDEM-X image pairs from three growth seasons calibrated using ALS data. Persson and Fransson [11] used four TanDEM-X acquisitions covering three growth seasons, calibrated using ALS data or Lorey's height from field data. They predicted SI with 4.4 m RMSE and age with 17.8 years RMSE on 0.5 ha plots. The need for calibration, however, hampers the scalability of the methods, as it relies on local high-resolution ALS data or field data. Furthermore, the usefulness of calibration data decreases with the time between data collection and the TanDEM-X acquisition date due to forest growth and other changes. Because of this, longer time series may often need calibration data from multiple time points.

In this study, we wanted to use a longer and denser time series of TanDEM-X acquisitions than in the previous studies and simultaneously avoid the use of calibration of the TanDEM-based heights via ancillary remote sensing or field data. Additionally, all remote sensing studies predicting SI that we are aware of use only plots, which appear to be unaffected by silvicultural treatments during the observation period. This study included plots subject to different silvicultural treatments during the study period to assess the potential effects on the predictions.

The remainder of the paper is structured as follows: Section 2 starts with a description of the test site and field data, after which the TanDEM-X data and its processing into TanDEM-X-based top heights are detailed. After this, established height development curves (HDC) and how they are used to calculate SI from field-measured top height and age are described. This is followed by a description of the method by which the SI and age are predicted by fitting an HDC to the time series of TanDEM-X-based top heights and how the results were evaluated. Section 3 presents the results of SI and age predictions. Section 4 contains a discussion of the results, and finally, Section 5 concludes the paper.

2. Materials and Methods

2.1. Test Site and Field Data

The study was conducted in Remningstorp, a forest test site located in southern Sweden (Lat. 58°30′N, Long. 13°40′E), consisting of about 1200 ha of commercially managed hemi-boreal forest. About two-thirds of the forest grows on till, a mixture of glacial debris, with, except in old spruce stands, a field layer of herbs, blueberry (*Vaccinium myrtillus* L.), and narrow-leaf grass (e.g., *Deschampsia flexuosa* (L.) Trin.). The main tree species are Norway spruce (*Picea abies* (L.) H. Karst.), Scots pine (*Pinus sylvestris* L.), and birch (*Betula* spp.). The rest of the forest grows on peatland, dominated by Scots pine. The landscape is mainly flat, with mild slopes, located 120 m to 145 m above sea level.

SI was determined for 91 circular field plots with a 10 m radius in a survey carried out in the fall of 2021. The age and height of two dominant trees per plot were measured, and the dominant species recorded. SI was calculated from the mean age and height for each plot. Using forest treatment records and inspection of biannual aerial orthophotos, the plots were classified into 2 clear-cut plots, 45 thinned plots, 7 pre-commercially thinned plots, and 26 untreated plots. Among the 91 plots, 11 plots were not covered by the available treatment records nor determined clear-cut in the inspection of orthophotos and, therefore, referred to as "undocumented". While clear-cuts were evident in the available orthophotos, thinnings were difficult to detect, and it is likely that a significant portion of these undocumented plots were, in fact, thinned or pre-commercially thinned. Table 1. shows the mean and range of SI and the field-measured variables for each treatment group.

Table 1. Summary statistics on the 10 m radius field plots used in the study.

Treatment	Top Height [m] Min/Mean/Max	Age [Years] Min/Mean/Max	SI Min/Mean/Max	*n*
Untreated	14/25/32	25/52/140	13/34/45	26
Pre-commercially thinned	14/20/28	15/25/50	27/35/44	7
Thinned	12/21/32	20/35/105	16/38/50	45
Clear-cut	25/26/28	60/70/80	30/30/30	2
Undocumented	15/24/29	25/49/96	16/34/40	11

SI values based on a previous field survey in 2014 were also available for 51 of the plots. Since the inherent productive potential of a specific site is not expected to change significantly in seven years, this dataset provided a means to characterize the uncertainty in the reference data. For these 51 plots, the variation in terms of Root Mean Square Deviation (RMSD) and bias, calculated according to Equations (1) and (2), between the 2021 and 2014 surveys of SI was 3.3 m and 2.0 m, respectively.

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N}\left(SI_i^{2021} - SI_i^{2014}\right)^2}{N}} \quad (1)$$

$$bias = \frac{\sum_{i=1}^{N}\left(SI_i^{2021} - SI_i^{2014}\right)}{N} \quad (2)$$

2.2. SAR Data

Thirty TanDEM-X scenes were acquired in a bi-static configuration over Remningstorp between 11 August 2013 and 24 September 2018. The scenes were acquired in strip-map mode and included a vertical transmit/receive (VV) polarization, acquired either as a single polarization or as a single channel from a dual-polarization scene. The bandwidths were 100 MHz or 150 MHz, respectively. A single polarization was chosen to avoid polarization-dependent systematic differences in phase heights, and the VV polarization specifically was chosen because it provided the best temporal coverage of the time period under study. Furthermore, using meteorological records, only acquisitions from dates preceded by a three-day average temperature of above 5 °C were included since freezing temperatures severely affect the observed radar phase heights from vegetation. The height of ambiguity, *HoA*, ranged between 43 m and 100 m, but most scenes were acquired with a *HoA* between 50 m and 65 m. The incidence angles ranged from 19° to 40°.

The data were delivered in the Coregistered Single look Slant range Complex (CoSSC) format. A complex interferogram was computed with 5 × 5 spatial averaging in range and azimuth. The interferogram was flattened with respect to earth curvature, and Goldstein filtered [28]. The flattened phase was unwrapped and converted to phase height by scaling with the wavenumber, and the interferometric coherence was estimated from the flattened interferogram using a coherence window of 3 × 3 pixels. Finally, the scenes were interpolated to a ground resolution of 10 m × 10 m. The estimated coherence was corrected for decreasing signal-to-noise ratio [3,15,17].

The radar signal penetrates significantly into the canopy, and for boreal coniferous forests, this leads to a negative elevation bias of the canopy heights that can sometimes be as large as 10–20 m [3]. In [29], a correction of this bias was proposed, which has successfully been evaluated on TanDEM-X data over temperate and hemi-boreal forests [24,30]. According to [29], the canopy height bias is given by

$$\Delta h = \frac{|HoA|}{2\pi}\tan^{-1}\left(\sqrt{|\gamma|^{-2} - 1}\right), \quad (3)$$

where γ is the volume coherence, and *HoA* is the height of ambiguity. The InSAR phase center heights in the time series were corrected for elevation bias by calculating Δh on pixel level and correcting the height values to produce InSAR-based canopy heights. A thorough derivation of (3) is given in [29]. As the bias correction assumes penetration into an infinite volume, it is not theoretically valid when the signal has significant ground contributions, and as a rule of thumb, the canopy height should be at least twice the bias correction. The correction was applied to all plots, although some field plots with low canopy heights or sparse forests could potentially violate this criterion. The ground contributions to the pixels with the highest phase heights on each plot are generally assumed to be small, and as described in Section 2.3, only the highest InSAR canopy heights on each plot influence the estimated top height in the method applied. However, ground contributions are likely dominant in clear-cut plots after treatment, but these were nevertheless corrected using Equation (3).

Each acquisition was assigned an integer representing its growth period based on the date of acquisition. A growth period was defined to start on 15 June, approximating the halfway point of the actual growth period, and last for one year. The earliest acquisition date was assigned to growth period 0, and the latest acquisitions were assigned to growth period 5.

2.3. Top Height Estimation

For each date, the corrected phase height values of pixels covered or intersected by a polygon defining the field plot region were extracted. In order to estimate the top height from these pixels, the 90th height percentile was calculated for each plot and date. Different percentiles were investigated, and generally, the higher percentiles correlated better with field-measured top heights. The 90th height percentile will hereafter be referred to simply as TanDEM-X top height.

2.4. Site Index

Established HDCs for common Swedish tree species, as developed in [31,32] and summarized in [33], describe the expected top height H_2 at stand age A_2, given a measured top height H_1 at stand age A_1:

$$H_2(H_1,\ A_1, A_2) = \frac{H_1 + d + r(H_1, A_1)}{2 + \frac{4\beta A_2^{b_2}}{H_1 - d + r(H_1, A_1)}}, \tag{4}$$

$$r(H_1, A_1) = \sqrt{(H_1 - d)^2 + 4\beta H_1 {A_1}^{b_2}}, \tag{5}$$

$$d = \beta A_{SI}^{b_2}, \tag{6}$$

where β and b_2 are previously determined tree species-specific fixed parameters, and A_{SI} is the reference age. Equation (4) is commonly used to calculate SI given field measurements of top height and age, as was done with the field data in this study. By setting A_2 to the preferred SI reference age and H_1 and A_1 to the measured height and age, H_2 equals the SI. The HDCs were developed from multiple measurements on sets of field plots in even-aged forests, the predominant forest type in Sweden, and are therefore valid in such forests.

If H_1 is set to a specific SI value instead of a measured height and A_1 is set to the corresponding reference age, H_2 gives the expected top height at any age A_2. For illustration, the resulting HDCs of Norway spruce for a few values of SI are shown in Figure 1.

Figure 1. HDC describing the top height growth of Norway spruce.

2.5. Site Index Estimation

Setting H_1 and A_1 in Equation (4) to SI and the corresponding reference age, respectively, and substituting A_0 + GP (growth period) for A_1, allows us to express TanDEM-X top height H as a function of SI and GP, explicitly

$$H(A_0 + GP, SI) = \frac{SI + d + r(SI, A_{SI})}{2 + \frac{4\beta(A_0+GP)^{b_2}}{SI-d+r(SI,A_{SI})}}. \quad (7)$$

SI and A_0 were determined by applying a weighted non-linear least squares regression of Equation (7) to the time series of TanDEM-X top heights, leaving initial age (age at the time of the first TanDEM-X measurement, A_0) and/or SI as parameters. The function was fit to each field plot using dominant species information from the field data to select the correct fixed parameters, and two different prediction cases were applied; (a) estimates of both SI and A_0 for each plot, and (b) estimates of only SI, assuming that the initial age is known. In case (b), A_0 in the fitting was supplied from the field data. Figure 2 illustrates prediction case (a). In this figure, the process of fitting A_0, can be considered as a translation in time of the time series of TanDEM-X top heights to find the optimal fit, while the fitting of SI corresponds to the choice of optimal curve out of the family defined by Equation (7).

The least squares regression was performed using the *nls* function from the *stats* package of the open-source *R* programming language [34]. By using the *port* algorithm, the solver utilized an implementation of the *nl2sol* algorithm [35]. The SI and A_0 (in prediction case (a)) were initialized to 25 and 75 and constrained to the intervals [4, 60] and [4, 200], respectively. This algorithm was chosen because of the possibility of setting bounds for the parameters. Otherwise, the algorithm tended to diverge or produce implausible parameter values for plots where TanDEM-X phase heights decreased over time. In case the fitting did not converge, it was restarted and initialized using the parameter values obtained in the non-converging fit.

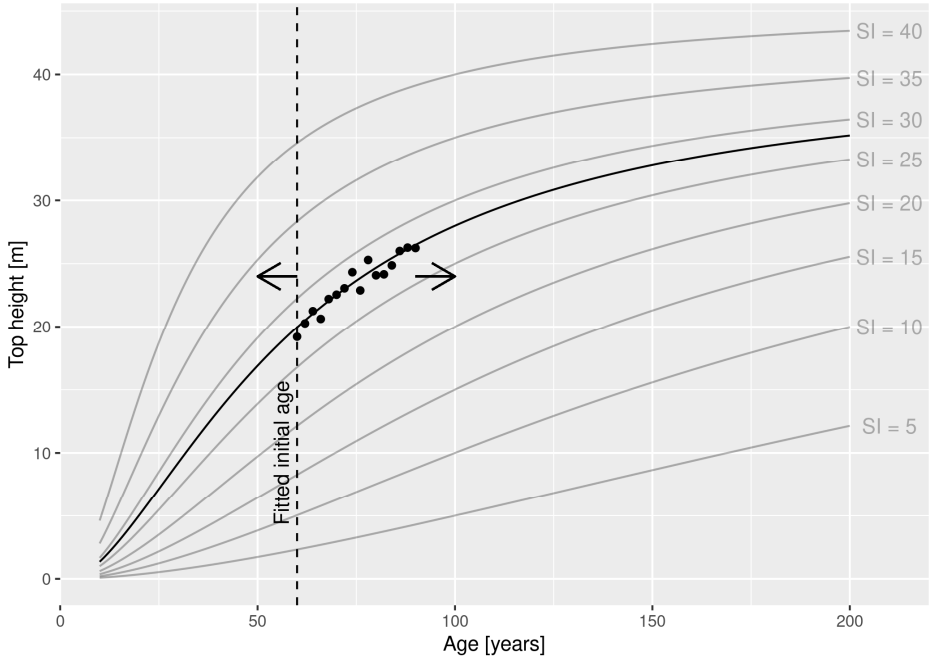

Figure 2. An illustration of the prediction of initial age A_0 and SI by fitting a HDC to a time series of TanDEM-X top heights. The fitting of $A_{0,}$ indicated by a dashed line in the figure, can be thought of as horizontal translation of the time series of data points, while the fitting of SI corresponds to the choice of curve.

The uncertainty of the InSAR phase height is, up to a critical value, roughly inversely proportional to the interferometric baseline [36] and hence proportional to *HoA*. Because of this, scenes with a baseline below, or equally an *HoA* above, some threshold value are often omitted in pursuit of high precision. In order to account for the *HoA*-related uncertainty without sacrificing temporal resolution, each observation was weighted in the fitting procedure with the reciprocal of *HoA*.

The SI and A_0 predicted by parameter estimation were visually inspected via plots of the fitted HDC alongside the TanDEM-X top heights and the HDC expected from the field-data-based SI and age. The quality of predictions of A_0 and SI were evaluated by comparisons with the corresponding field-data-based values, and prediction results were further visually evaluated through plots to investigate possible correlations between prediction errors and SI, stand age, species, or treatment groups. The Root Mean Square Error (RMSE) and bias were calculated for each treatment group k as

$$RMSE_k = \sqrt{\frac{\sum_{i=1}^{i=n_k} (\hat{y}_i - y_i)^2}{n_k}}, \text{ and} \tag{8}$$

$$bias_k = \frac{\sum_{i=1}^{i=n_k} (\hat{y}_i - y_i)}{n_k}, \tag{9}$$

where $\hat{y}_i$ is the ith prediction, y_i the corresponding field data value, and n_k is the number of field plots in group k. Additionally, the coefficient of determination between predictions and reference values was calculated.

3. Results

In both prediction cases, the HDC fitting resulted in convergent solutions for all 91 field plots. For the vast majority of plots, a convergent solution was computed from the first initialization, but for a handful of plots, a re-initialization was required, as described in Section 2.5.

3.1. Predicting Both SI and Stand Age

In prediction case (a), when predicting both SI and A_0, the RMSEs and biases (as defined in Equations (8) and (9)) tended to increase with the intensity (in terms of expected relative biomass reduction) of the treatment. Table 2 shows the evaluation results for this case. For the untreated plots, the predicted SI had an RMSE of 6.9 m. The RMSE of pre-commercially thinned plots was 9.5 m, increasing to 16.1 m for the thinned group. The clear-cut group, however, had a slightly lower RMSE of 13.2 m.

Table 2. Treatment-wise summary statistics of SI and age predictions from HDC fitting to time series of TanDEM-X phase heights. P-c thinned = Pre-commercially thinned.

Treatment	SI RMSE [m]	SI R^2	SI Bias [m]	Age RMSE [Years]	Age R^2	Age Bias [Years]	*n*
Untreated	6.9	0.46	−1.6	38	0.406	9.6	26
P-c thinned	9.5	0.24	−7.8	22	0.321	11.9	7
Thinned	16.1	0.06	−12.6	82	0.015	52.5	45
Clear-cut	13.2	-	−13.2	137	-	136.8	2

The coefficient of determination between predicted and reference SI was 0.46 for the untreated plots and decreased to 0.24 for the pre-commercially thinned plots. It was very low, 0.06, for the thinned plots. R^2 of the clear-cut plots is not reported, as it yields no information with only two observations. A plot of the field measured vs. predicted SI is shown in Figure 3.

Similarly, the magnitude of the biases of predicted SI increased with the relative intensity of the treatment, from a bias of −1.6 m for the untreated group to −7.8 m for the pre-commercially thinned group, −12.6 m for the thinned group, and −13.2 m for the clear-cut group. However, only the biases for the thinned and pre-commercially thinned plots were significant at the 95% confidence level (Table 2 and Figure 4).

The predicted initial age A_0 had an RMSE of 38 years for the untreated group, 22 years for the pre-commercially thinned group, and then increased with the intensity of treatment to 82 for the thinned group and 137 years for the clear-cut plots.

The R^2 between predicted and reference age are similar to those for SI, 0.4 for untreated plots, and decreasing with treatment intensity. The R^2 of the clear-cut plots is not reported, as it yields no information for only two observations.

The bias of predicted initial age increased in magnitude with the intensity of treatment, with a more intense treatment having a larger positive age prediction bias, going from 9.6 years for the untreated group to 137 years for the clear-cut plots. However, only the age prediction biases for thinned and clear-cut plots were significant at a 95% confidence level (Table 2 and Figure 5).

3.2. Predicting SI Assuming Known Age

In prediction case (b), when predicting only SI, using the field-measured age in the HDC fitting, the precision was better than for case (a), with RMSEs between 2.2 m and 5.3 m for all treatment groups. The biases were also smaller for every treatment, all between −2 m and 0.5 m (Table 3 and Figure 6. Moreover, none of the biases were statistically significant at a 95% confidence level. The coefficients of determination, R^2, between true and predicted SI were relatively high, with R^2 values between 0.6 and 0.8. A plot of the field measured vs. predicted SI is shown in Figure 7.

An example of a time series of TanDEM-X top heights for an untreated plot superimposed with the fitted HDC and the reference-data-based HDC is shown in Figure 8. The size of a point representing TanDEM-X top height observation is proportional to its weight (reciprocal of *HoA*) in the regression. In this example, the slope is slightly overestimated, leading to an overestimation of SI and an underestimation of stand age.

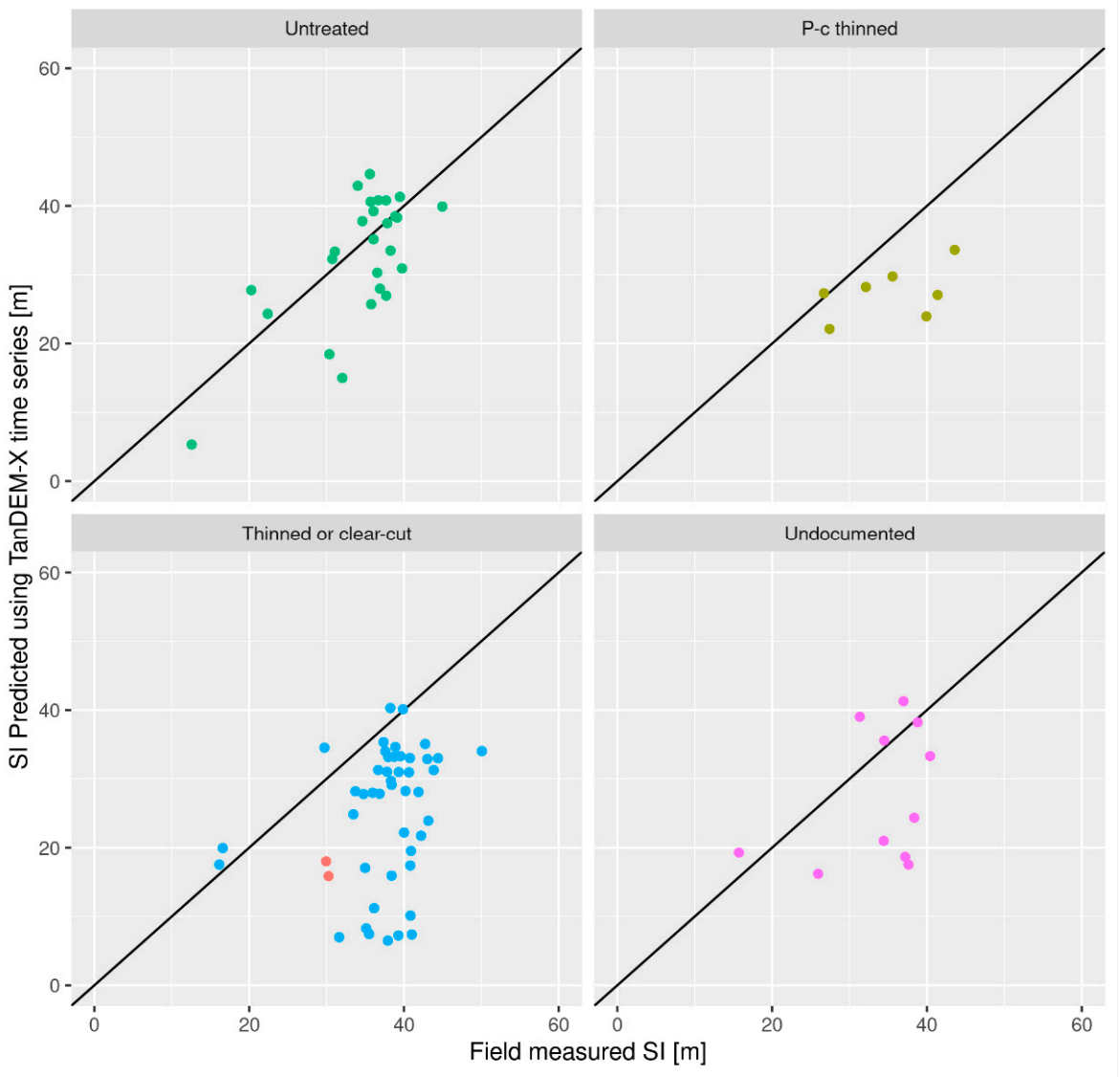

Figure 3. Plot of Predicted vs. field measured SI, obtained via simultaneous prediction of age. Colored by treatment. In the bottom left panel, thinned plots are shown in blue, while clear-cut plots are shown in red.

Table 3. Treatment-wise summary statistics of SI predictions assuming known age.

Treatment	SI RMSE [m]	SI R^2	SI Bias [m]	*n*
Untreated	4.0	0.80	−0.8	26
P-c thinned	5.3	0.63	−1.21	7
Thinned	3.3	0.73	0.47	45
Clear-cut	2.2	—	−1.99	2

3.3. Error Characteristics

A clear correlation between severe underestimations of SI and overestimations of age was observed, as can be deduced from Table 2, and is perhaps even more clear in Figure 9, which shows predicted vs. reference SI, colored by the age prediction error. For both prediction cases, plots of prediction errors against reference age (in case (a)), reference top height and averaged height residuals from HDC fits were inspected to identify additional correlations, but none were found.

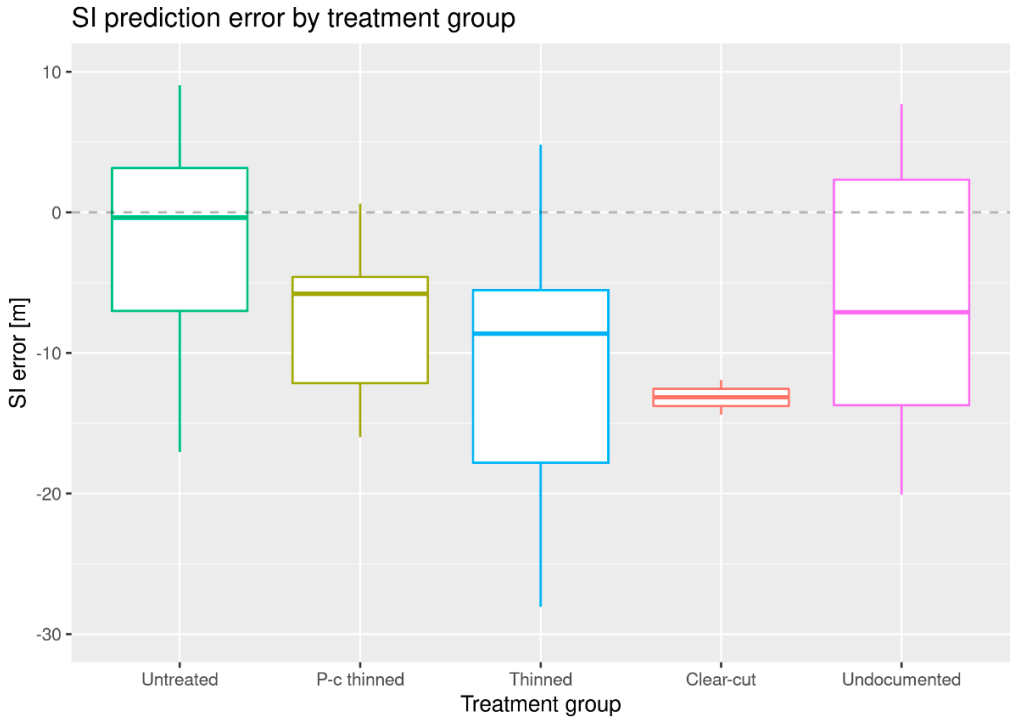

Figure 4. Box plot of SI prediction errors by group, obtained when simultaneously predicting age.

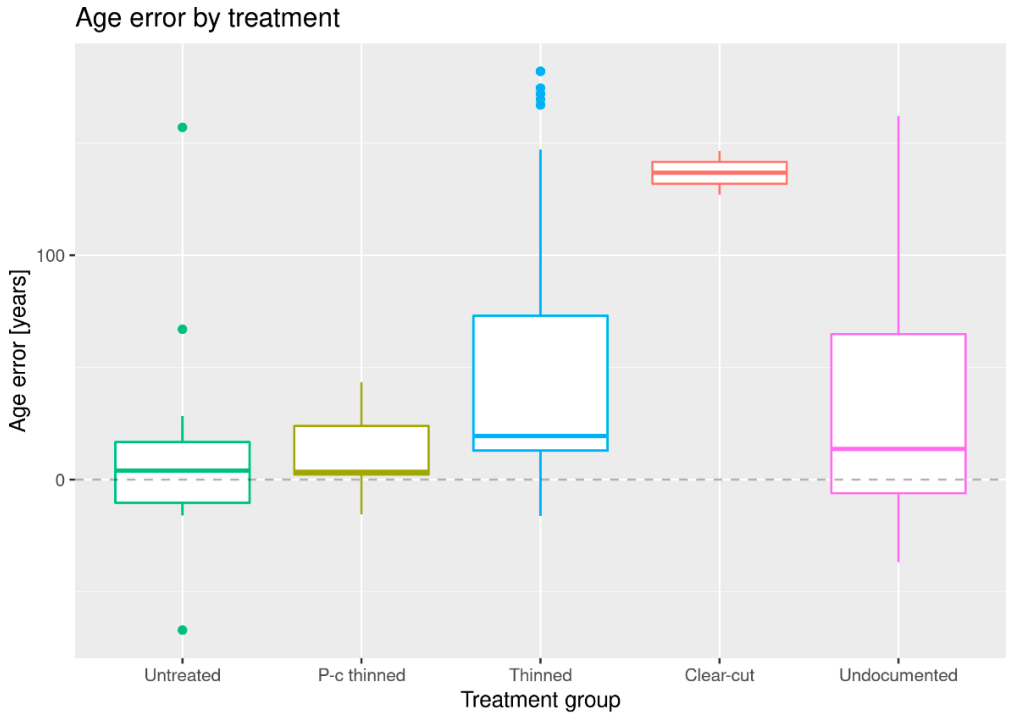

Figure 5. Box plot of age prediction errors by group, obtained when simultaneously predicting SI.

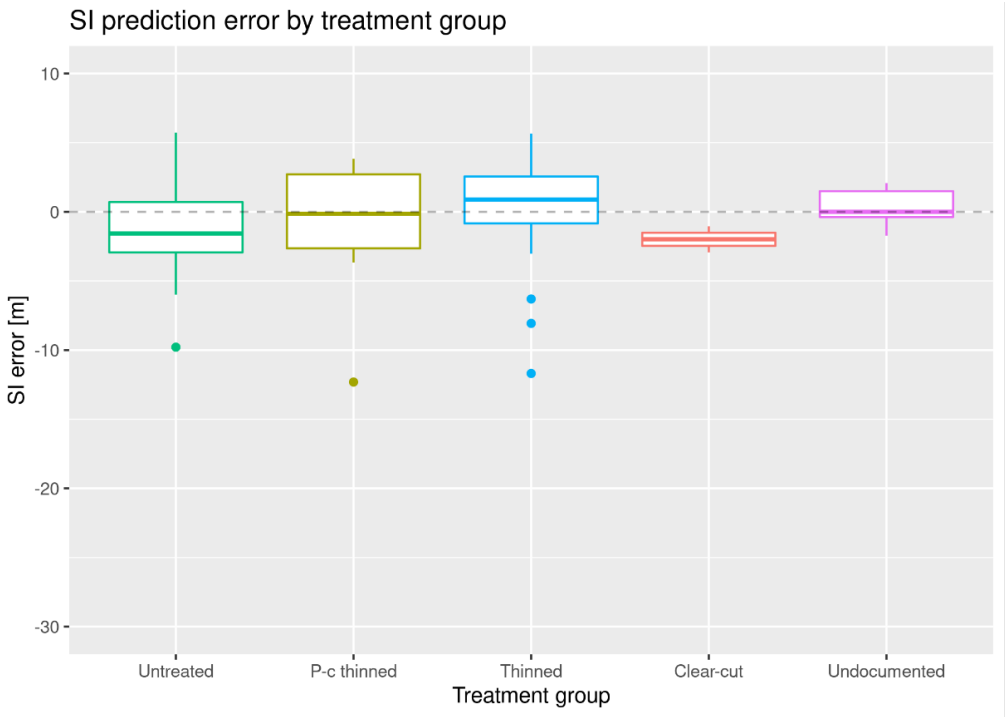

Figure 6. Box plot of SI prediction errors by group, obtained when simultaneously predicting age.

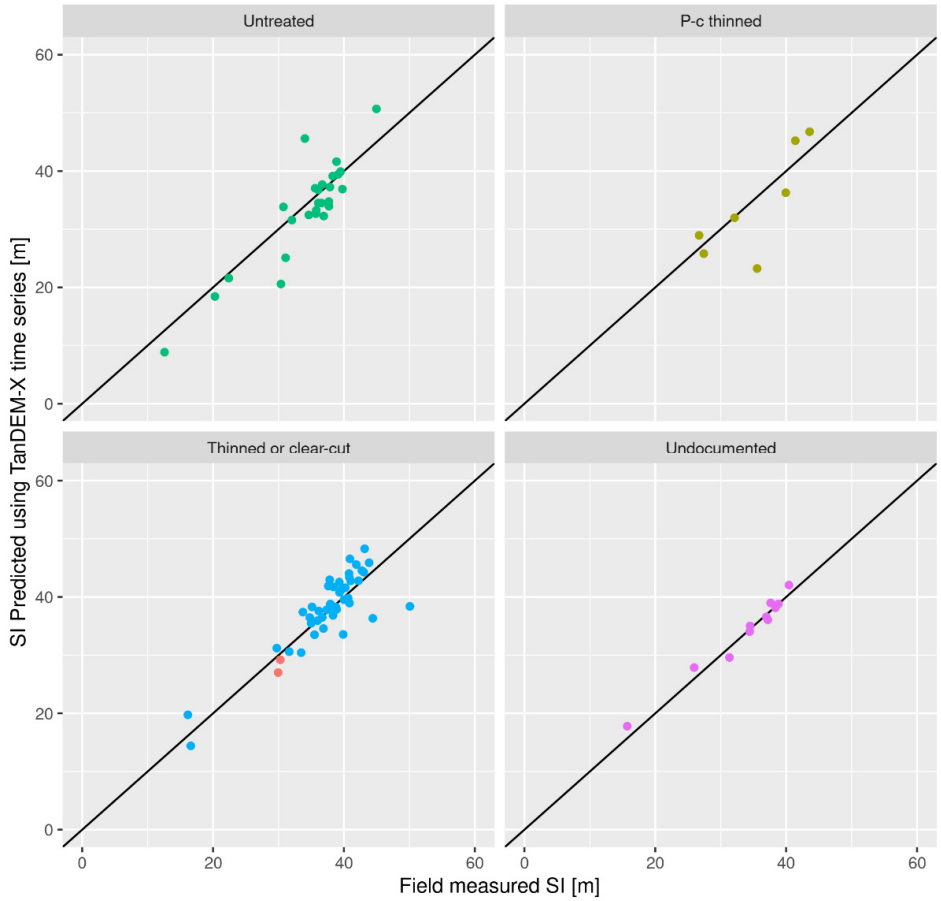

Figure 7. Predicted vs. field measured SI, using field measured ages in the fitting. Colored by treatment. In the bottom left panel, thinned plots are shown in blue, while clear-cut plots are shown in red.

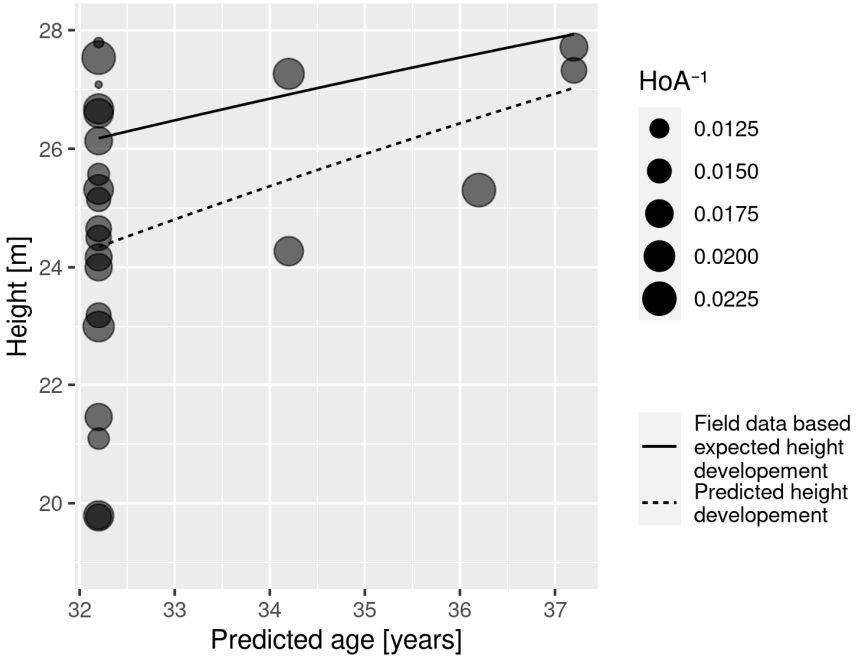

Figure 8. Time series untreated pine dominated plot 169. TanDEM-X top heights superimposed with predicted HDC and reference HDC. Predicted SI: 39.2, field SI: 36.1 predicted age: 32.2 field age: 43.

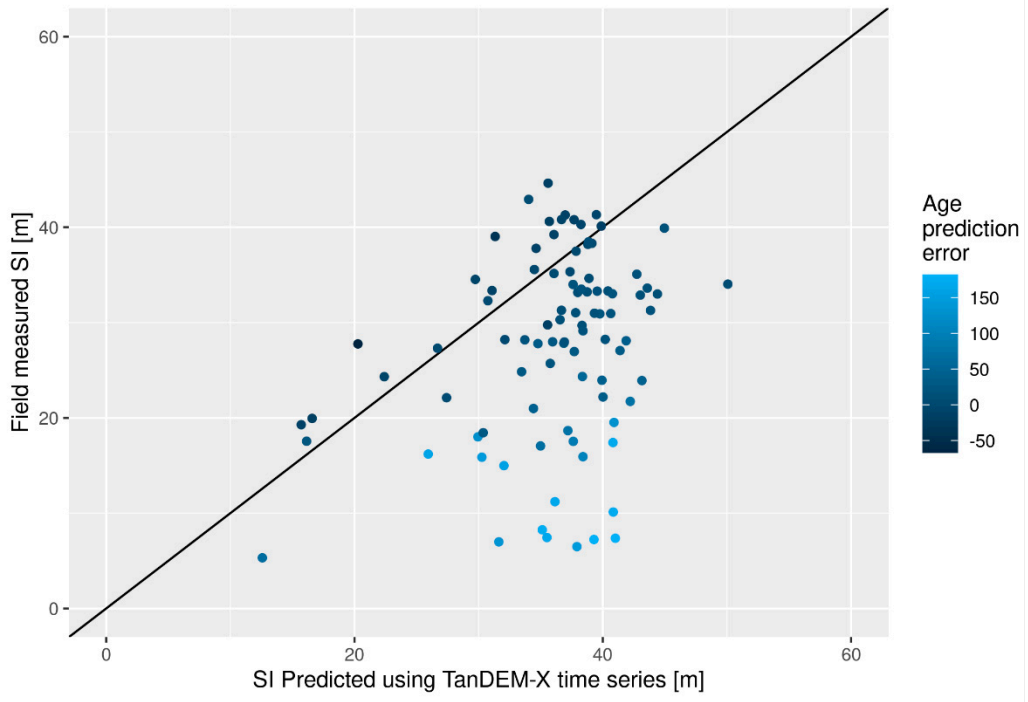

Figure 9. Predicted vs. field measured SI, colored by predicted age.

4. Discussion

The TanDEM-X top height, i.e., the 90th height percentile of bias-corrected phase heights, captured the canopy top height reasonably well. This was evidenced by (1) the relatively small and statistically insignificant prediction biases on untreated plots in case (a), when predicting both SI and age, and (2) the very low RMSE and bias of the predictions of SI in case (b), when age was provided from field data.

The RMSE of 4.0 m for prediction case (b) is significantly lower than the RMSE of 6–7 m in [10], which used ALS-calibrated TanDEM-X heights and similarly predicted only

SI on 10 m radius plots. In [10], the prediction of both SI and age using TanDEM-X data (corresponding to case (a) in this paper) was unsuccessful due to divergent solutions.

The RMSE for untreated plots in prediction case (a), 6.9 m, is larger than some previous studies [5,7–9,11], but they cannot be directly compared, as they differ in SI reference ages and/or the area of evaluation units. For example, this study used 10 m radius plots, while in [11], reporting 4.4 m and 17.8-year RMSEs, the plots were 16 times larger. In other studies, such as [5] (2.4 m RMSE) and [7] (about 2.8 m RMSE), the reported SI predictions differed in species and the reference age (50 years in [5], 40 years in [7]) at which the SI height is defined, due to local functions and practices.

Silvicultural treatments, from pre-commercial thinning to clear-cutting, lead to underestimation of the slope of the HDC, which in turn leads to underestimation of SI and overestimation of age. This tendency increased with the intensity of treatment.

We did not observe systematically larger prediction errors for mature stands, as observed in, for example, [5], where predictions based on measurement periods capturing later stand development stages with lower height growth rates produced more uncertain estimates of both SI and age. As the growth rate decreases with age, the importance of absolute height estimates increases (Figure 2). The absence of increased prediction uncertainty for older untreated plots further indicates that the absolute top height is estimated well by the TanDEM-X top height.

Some of the uncertainty in predictions is explained by edge effects. From inspection of orthophotos, we found that plots with large deviations between predictions and reference values were often located close to, or even across, stand boundaries or roads since the plot locations in the field surveys were distributed in a systematic grid. During the analyses, it was found that for such plots, the measured phase heights could be drastically different depending on look direction, which caused us to use only acquisitions from a descending orbit. This maximized the duration and resolution of the time series. Removal of such edge plots would likely have led to lower RMSEs in the predictions, but they were nevertheless kept in order to better reflect realistic results with a minimum amount of manual intervention. Because of this, it is also reasonable to expect higher precision for larger prediction units, where boundary effects have a smaller impact on the prediction or can be dealt with, for example, by using buffer zones.

Given the small prediction biases observed for untreated plots in case (a) and the apparent overall precision of top height estimates, we expect even longer time series to increase the quality of predictions of both SI and age significantly from the results for prediction case (a) in untreated plots. Longer time series may also mitigate the underestimation of slope resulting from treatments if the influence of such treatments on the TanDEM-X top height is transient in nature. Future studies should investigate the inclusion of multiple polarizations as a way to further increase usable observations.

Further, the prediction of SI and age via weighted non-linear regression readily accommodates the inclusion of other height data sources since top heights based on any source could simply be added to the time series and weighted according to the uncertainty of the source. Alternatively, since the SI prediction quality was shown to be much better assuming known age, the method could also be combined with other data sources by supplying age from field data or predictions from photogrammetric time series, as in [37]. In this study, the dominant species was assumed to be known, but in a practical application, the dominant species could also be predicted based on other data sources.

It should be noted that although the method proposed does not use any ancillary field data or remote sensing data such as ALS for calibration of the TanDEM-X data, it does require an accurate DTM in order to obtain reliable canopy heights from phase heights. This requirement is fulfilled in a rapidly increasing part of the world. Additionally, the HDCs used are developed using data from even-aged stands, and their applicability in other types of forest should be further investigated.

As the method presented does not rely on local calibration and easily accommodates and benefits from additional TanDEM-X scenes that extend or increase the temporal resolu-

tion of the prediction period, it is suited for producing wall-to-wall estimates over large areas of forest.

5. Conclusions

SI, the expected top height at some reference age, and stand age are important variables in forest management and forecasting. This study presented and evaluated a method of predicting SI and age using only time series of TanDEM-X data and a DTM.

The method consists of fitting established HDC to the time series, using the 90th height percentile of canopy penetration corrected phase heights as a surrogate for forest top heights. Predicted SI and age were retrieved as parameter values minimizing the squared top height residuals.

SI and age could be unbiasedly predicted for untreated plots, and the RMSE of predictions is likely to decrease with the length and temporal resolution of the time series available. When the stand age was known, the SI was predicted with an RMSE comparable to that of field-based measurements.

The results for treated plots indicate that the RMSE and bias of predictions increase with the intensity of silvicultural treatments, with a larger relative decrease in stem volume on average leading to a larger underestimation of SI and an overestimation of stand age and a higher RMSE for both variables.

In general, the results demonstrate viability for large-scale wall-to-wall mapping of SI using time series of TanDEM-X data without the need for ancillary data for height calibration. Further studies should investigate the use of multiple polarizations and both orbit directions to increase the length and temporal density of useful time series in an effort to further increase the obtained prediction quality.

Author Contributions: Conceptualization, J.E.S.F., I.H. and H.J.P.; methodology, I.H. and H.J.P.; software, I.H.; validation, I.H. and H.J.P.; formal analysis, I.H.; investigation, I.H.; resources, J.E.S.F. and H.J.P.; data curation, I.H., J.W. and H.J.P.; writing—original draft preparation, I.H.; writing—review and editing, I.H., J.W., H.J.P. and J.E.S.F.; visualization, I.H.; supervision, J.E.S.F., H.J.P. and J.W.; project administration, J.E.S.F. and H.J.P.; funding acquisition, J.E.S.F. and H.J.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the foundation Skogssällskapet, grant numbers 2021-923 and 2022-1023, the Bo Rydin Foundation for Scientific Research, grant number F 12/22, and Stiftelsen Seydlitz MP bolagen, grant number 143. The TanDEM-X images were acquired through the German Aerospace Center (DLR) "3D forest parameter retrieval from TANDEM-X interferometry" project [XTI_VEGE0376]. The field data collection was funded by the Hildur and Sven Wingquist's foundation.

Data Availability Statement: The data used in this study are available upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Roach, W.J.; Simard, S.W.; Defrenne, C.E.; Pickles, B.J.; Lavkulich, L.M.; Ryan, T.L. Tree Diversity, Site Index, and Carbon Storage Decrease With Aridity in Douglas-Fir Forests in Western Canada. *Front. For. Glob. Change* **2021**, *4*, 682076. [CrossRef]
2. Skovsgaard, J.P.; Vanclay, J.K. Forest Site Productivity: A Review of the Evolution of Dendrometric Concepts for Even-Aged Stands. *For. Int. J. For. Res.* **2008**, *81*, 13–31. [CrossRef]
3. Kugler, F.; Schulze, D.; Hajnsek, I.; Pretzsch, H.; Papathanassiou, K.P. TanDEM-X Pol-InSAR Performance for Forest Height Estimation. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 6404–6422. [CrossRef]
4. Stepper, C.; Straub, C.; Pretzsch, H. Assessing Height Changes in a Highly Structured Forest Using Regularly Acquired Aerial Image Data. *For. Int. J. For. Res.* **2015**, *88*, 304–316. [CrossRef]
5. Véga, C.; St-Onge, B. Mapping Site Index and Age by Linking a Time Series of Canopy Height Models with Growth Curves. *For. Ecol. Manag.* **2009**, *257*, 951–959. [CrossRef]
6. Kandare, K.; Ørka, H.O.; Dalponte, M.; Næsset, E.; Gobakken, T. Individual Tree Crown Approach for Predicting Site Index in Boreal Forests Using Airborne Laser Scanning and Hyperspectral Data. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *60*, 72–82. [CrossRef]
7. Solberg, S.; Kvaalen, H.; Puliti, S. Age-Independent Site Index Mapping with Repeated Single-Tree Airborne Laser Scanning. *Scand. J. For. Res.* **2019**, *34*, 763–770. [CrossRef]

8. Penner, M.; Woods, M.; Bilyk, A. Assessing Site Productivity via Remote Sensing—Age-Independent Site Index Estimation in Even-Aged Forests. *Forests* **2023**, *14*, 1541. [CrossRef]
9. Persson, H.J.; Fransson, J.E.S. Analysis of Tree Height Growth with TanDEM-X Data. In Proceedings of the 35th EARSeL Symposium, Stockholm, Sweden, 15–19 June 2015; Volume 1, pp. 1–6.
10. Wallerman, J.; Nyström, K.; Bohlin, J.; Persson, H.J.; Soja, M.J.; Fransson, J.E.S. Estimating Forest Age and Site Productivity Using Time Series of 3D Remote Sensing Data. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 3321–3324.
11. Persson, H.J.; Fransson, J.E.S. Estimating Site Index From Short-Term TanDEM-X Canopy Height Models. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 3598–3606. [CrossRef]
12. Fiedler, H.; Krieger, G.; Zink, M.; Younis, M.; Bachmann, M.; Huber, S.; Hajnsek, I.; Moreira, A. The TanDEM-X Mission: An Overview. In Proceedings of the 2008 International Conference on Radar, Adelaide, Australia, 2–5 September 2008; pp. 60–64.
13. Krieger, G.; Zink, M.; Bachmann, M.; Bräutigam, B.; Schulze, D.; Martone, M.; Rizzoli, P.; Steinbrecher, U.; Walter Antony, J.; De Zan, F.; et al. TanDEM-X: A Radar Interferometer with Two Formation-Flying Satellites. *Acta Astronaut.* **2013**, *89*, 83–98. [CrossRef]
14. Karila, K.; Vastaranta, M.; Karjalainen, M.; Kaasalainen, S. Tandem-X Interferometry in the Prediction of Forest Inventory Attributes in Managed Boreal Forests. *Remote Sens. Environ.* **2015**, *159*, 259–268. [CrossRef]
15. Persson, H.J.; Olsson, H.; Soja, M.J.; Ulander, L.M.H.; Fransson, J.E.S. Experiences from Large-Scale Forest Mapping of Sweden Using TanDEM-X Data. *Remote Sens.* **2017**, *9*, 1253. [CrossRef]
16. Persson, H.J.; Fransson, J.E.S. Comparison between TanDEM-X- and ALS-Based Estimation of Aboveground Biomass and Tree Height in Boreal Forests. *Scand. J. For. Res.* **2017**, *32*, 306–319. [CrossRef]
17. Chen, H.; Cloude, S.R.; Goodenough, D.G. Forest Canopy Height Estimation Using Tandem-X Coherence Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 3177–3188. [CrossRef]
18. Kugler, F.; Hajnsek, I. Forest Characterisation by Means of TerraSAR-X and TanDEM-X (Polarimetric and) Interferometric Data. In Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2011; pp. 2578–2581.
19. Olesk, A.; Praks, J.; Antropov, O.; Zalite, K.; Arumäe, T.; Voormansik, K. Interferometric SAR Coherence Models for Characterization of Hemiboreal Forests Using TanDEM-X Data. *Remote Sens.* **2016**, *8*, 700. [CrossRef]
20. Praks, J.; Antropov, O.; Hallikainen, M.T. LIDAR-Aided SAR Interferometry Studies in Boreal Forest: Scattering Phase Center and Extinction Coefficient at X- and L-Band. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3831–3843. [CrossRef]
21. Schlund, M.; Magdon, P.; Eaton, B.; Aumann, C.; Erasmi, S. Canopy Height Estimation with TanDEM-X in Temperate and Boreal Forests. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *82*, 101904. [CrossRef]
22. Soja, M.J.; Persson, H.J.; Ulander, L.M.H. Modeling and Detection of Deforestation and Forest Growth in Multitemporal TanDEM-X Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 3548–3563. [CrossRef]
23. Schlund, M.; Kukunda, C.B.; Baumann, S.; Wessel, B.; Kiefl, N.; von Poncet, F. Potential of Forest Monitoring with Multi-Temporal TANDEM-X Height Models. In Proceedings of the IGARSS 2020–2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 308–311.
24. Huuva, I.; Persson, H.J.; Wallerman, J.; Fransson, J.E.S. Detectability of Silvicultural Treatments in Time Series of Penetration Depth Corrected Tandem-X Phase Heights. In Proceedings of the IGARSS 2022–2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 5909–5912.
25. Solberg, S.; May, J.; Bogren, W.; Breidenbach, J.; Torp, T.; Gizachew, B. Interferometric SAR DEMs for Forest Change in Uganda 2000–2012. *Remote Sens.* **2018**, *10*, 228. [CrossRef]
26. Solberg, S.; Næsset, E.; Gobakken, T.; Bollandsås, O.-M. Forest Biomass Change Estimated from Height Change in Interferometric SAR Height Models. *Carbon Balance Manag.* **2014**, *9*, 5. [CrossRef]
27. Askne, J.I.H.; Persson, H.J.; Ulander, L.M.H. Biomass Growth from Multi-Temporal TanDEM-X Interferometric Synthetic Aperture Radar Observations of a Boreal Forest Site. *Remote Sens.* **2018**, *10*, 603. [CrossRef]
28. Goldstein, R.M.; Werner, C.L. Radar Interferogram Filtering for Geophysical Applications. *Geophys. Res. Lett.* **1998**, *25*, 4035–4038. [CrossRef]
29. Dall, J. InSAR Elevation Bias Caused by Penetration Into Uniform Volumes. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 2319–2324. [CrossRef]
30. Schlund, M.; Baron, D.; Magdon, P.; Erasmi, S. Canopy Penetration Depth Estimation with TanDEM-X and Its Compensation in Temperate Forests. *ISPRS J. Photogramm. Remote Sens.* **2019**, *147*, 232–241. [CrossRef]
31. Elfving, B.; Kiviste, A. Construction of Site Index Equations for Pinus Sylvestris L. Using Permanent Plot Data in Sweden. *For. Ecol. Manag.* **1997**, *98*, 125–134. [CrossRef]
32. Eriksson, H.; Johansson, U.; Kiviste, A. A Site-index Model for Pure and Mixed Stands of Betula Pendula and Betula Pubescens in Sweden. *Scand. J. For. Res.* **1997**, *12*, 149–156. [CrossRef]
33. Johansson, U.; Ekö, P.M.; Elfving, B.; Johansson, T.; Nilsson, U. *Rön Från Sveriges Lantbruksuniversitet*; Swedish University of Agricultural Sciences: Uppsala, Sweden, 2014.
34. R Core Team. *R: A Language and Environment for Statistical Computing*; R Core Team: Vienna, Austria, 2013.

35. Dennis, J.E.; Gay, D.M.; Welsch, R.E. Algorithm 573: NL2SOL—An Adaptive Nonlinear Least-Squares Algorithm [E4]. *ACM Trans. Math. Softw.* **1981**, *7*, 369–383. [CrossRef]
36. Li, F.K.; Goldstein, R.M. Studies of Multibaseline Spaceborne Interferometric Synthetic Aperture Radars. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 88–97. [CrossRef]
37. Vastaranta, M.; Niemi, M.; Wulder, M.A.; White, J.C.; Nurminen, K.; Litkey, P.; Honkavaara, E.; Holopainen, M.; Hyyppä, J. Forest Stand Age Classification Using Time Series of Photogrammetrically Derived Digital Surface Models. *Scand. J. For. Res.* **2016**, *31*, 194–205. [CrossRef]