Estimation of Forest Biomass from Two-Level Model Inversion of Single-Pass InSAR Data

Maciej Jerzy Soja, Henrik Persson, and Lars M. H. Ulander

Abstract

A model for above-ground biomass estimation from single-pass interferometric synthetic-aperture radar data is presented. Forest height and canopy density estimates $\Delta h$ and $\eta_0$, obtained from two-level model inversion, are used in a power function with slope $K$ and exponents $\alpha$ and $\beta$. The model is compared to a linear, zero-intercept model, scaling the interferometric height to biomass.

Eighteen bistatic, VV-polarized TanDEM-X (TDM) acquisitions made over two test sites in the summers of 2011, 2012, and 2013 are used. Remningstorp is a hemi-boreal forest in southern Sweden, with flat topography and 32 circular plots (area: 0.5 ha, biomass: 42–242 t/ha, height: 14–32 m). Krycklan is a boreal forest in northern Sweden, 720 km north-north-east from Remningstorp, with significant topography and 31 stands (area: 2.4–26.3 ha, biomass: 23–183 t/ha, height: 7–21 m). For all acquisitions, the nominal incidence angle is 41° and the height-of-ambiguity is in the interval 32–63 m. High-resolution digital terrain model has been used for ground correction during InSAR processing.

The proposed model explains 65–89% of the variance observed in the data, with a residual root-mean-square error (RMSE) 12–19% (median: 15%). If model training and validation are carried out on different acquisitions or between test sites, the prediction RMSE increases (12–80%, median: 30%). With $\alpha$ fixed and $\beta$ a site-dependent constant, the prediction RMSE is lower (12–56%, median: 17%), while the residual RMSE is similar (12–29%, median: 16%). The linear, zero-intercept model shows similar residual and prediction performance for the Krycklan data, whereas for the Remningstorp data and across-site retrieval, the performance is poorer.

Index Terms

above-ground biomass (AGB), forest height, canopy density, interferometric model, two-level model (TLM), interferometric synthetic-aperture radar (InSAR), TanDEM-X

I. INTRODUCTION

Forest are important natural resources because of their economic value and their crucial role in the local and global ecosystems [1]. Efficient and sustainable management procedures are required to maintain healthy and productive forests.
One of the key elements in forest management is to have reliable information for short- and long-term planning. The above-ground dry biomass, here shortly called biomass or the AGB, is especially important for carbon cycle studies, while other parameters such as forest height and canopy density can both aid biomass estimation and provide additional information on the forests. However, current methods of collecting forest information are expensive, and therefore more cost-effective methods need to be developed. Remote sensing in combination with field inventories has the potential to meet these requirements, and provide frequent and high-resolution mapping of forest variables. Large-scale mapping is also needed for natural disaster management, so that the damages caused by, e.g., storms, can be minimized.

Aerial photography has traditionally been used for forest mapping [2], [3]. This technique has the advantage of being relatively easy to implement and interpret, but it requires cloud-free acquisitions and good flying weather. Moreover, it is less efficient on large scale and whenever frequent updates are needed. Spaceborne photography is more efficient in terms of coverage and acquisition rate, but it has lower resolution. More advanced optical techniques, such as photogrammetry [4], [5], can provide additional information on forests, but the even stricter requirements on the acquired data make their use more difficult on an operational scale.

In recent years, airborne lidar scanning (ALS) has become popular. The technique uses laser pulses transmitted downwards from an airborne platform, which are used to sample height at high vertical and horizontal resolutions [4], [6]–[10]. Due to the high resolutions and the penetration of laser pulses through canopy gaps, ALS can provide information on both horizontal and vertical forest structure, and many important forest parameters can be derived from the data. ALS is today considered the most accurate remote sensing technique in forestry [10]. However, the technique is relatively expensive, and thus inefficient for frequent and large scale mapping. Spaceborne lidar, on the other hand, has yet unresolved resolution, coverage, and technology limitations.

Synthetic-aperture radar (SAR) is an active remote sensing technique in which radio- or microwave-frequency pulses are used to probe the environment. Spaceborne SAR sensors can provide weather- and daylight-independent imagery of the Earth with resolutions down to a couple of meters. Through the choice of the center frequency, SAR systems can be optimized to fit different needs [11]. In forestry, low frequency bands, like the VHF-band (30–300 MHz) and the lower UHF-band (300–1000 MHz, according to the IEEE standard) are more suitable for imaging of tree trunks and ground surface, while the high frequency bands, like the X-band (8–12 GHz), are more suitable for the imaging of tree canopies. SAR is one of the most promising tools for forest remote sensing and many past and ongoing studies are dedicated to the retrieval of forest parameters from SAR data [12].

The TanDEM-X system consists of two, almost identical X-band SAR satellites flying in a tight tandem formation, at a distance of a few hundred meters during the operational phase. Using the principles of interferometric SAR (InSAR), small phase differences between the two acquired SAR images are used to measure the position of the
scattering center [13], i.e. to create a digital elevation model (DEM). With the tight tandem formation and bistatic-mode acquisitions of TanDEM-X, the temporal changes between the two SAR acquisitions are minimal, and the acquired height measurements are very precise.

The acquired DEM can be corrected for ground topography if a high-resolution digital terrain model (DTM) is available, and a map of the scattering center elevation above ground can be obtained. In Sweden, there is a national, lidar-scanned DTM with a grid posting of $2\,\text{m} \times 2\,\text{m}$ and a height accuracy better than $0.5\,\text{m}$ [14]. Similar DTMs exist or are being created in many other countries. Since the changes of the ground surface are very slow in most forested regions, only one lidar scanning is required to obtain a high-resolution DTM, and after that mapping of forest canopy can be done with the TanDEM-X system.

The exact position of the scattering center above ground in forests is related to the structure of the forest and it depends on forest properties such as forest height and canopy density. In several studies, this relation has been investigated. In [15], [16], random volume over ground (RVoG) model inversion has been applied to estimate forest height from single-pass X-band InSAR data. In [17], a linear relation between biomass and the measured elevation of the scattering center above ground has been observed. This study has been based on an approach developed earlier for the SRTM X-SAR data [18]. In [19], biomass estimates have been obtained from ground-corrected TDM interferograms using the inversion of the interferometric water cloud model (IWCM), which includes an allometric relation between forest height and biomass, as well as temporal decorrelation. Both interferometric coherence and phase, and backscatter intensity data have been used in the inversion process. In [20], a multiple regression approach using interferometric height, coherence, and their transformed versions has been used to estimate biomass, separately for two test sites in Sweden.

It has been shown in [21], [22] that the inversion of a two-level model (TLM) can provide estimates of forest height and canopy density in a hemi-boreal forest in Sweden. In this study, the inverted parameters will be used to estimate biomass. Data from two boreal test sites in Sweden, separated by 720 km, will be used. The new model will be evaluated both for its explanatory and predictive values.

II. Method

In this section, the basic models used in this study will be described. First, it will be shown how biomass can be estimated from forest height and canopy density. Next, it will be shown how forest height and canopy density can be estimated from InSAR data. Thereafter, the results from the first two sections will be used together, and a new model for biomass estimation from InSAR data will be presented. Finally, the evaluation method used in this study will be described.
Fig. 1. Geometrical visualization of a forest plot.

A. AGB from Forest Height and Canopy Density

Above-ground biomass (AGB) is defined as the total dry mass of all above-ground forest, most commonly measured in terms of biomass density, i.e., as mass per area unit. As a large part of the AGB is confined to the stem (around 3/4 for spruce and pine in Sweden, according to [23]), a simple geometrical argument [24], [25] suggests that the AGB is a function of forest height, basal area, a taper factor accounting for the non-cylindrical trunk shape, the oven-dry wood density, and an expansion factor for the conversion of stem biomass to total aboveground biomass. After merging the last three factors into a single, forest type-dependent constant \( C \), the AGB can be estimated from:

\[
\hat{AGB} = C \cdot h \cdot \frac{A_{st}}{A_0},
\]

where \( A_{st} \) is the total basal area for all trees, \( A_0 \) is the ground area of the plot, see Figure 1, and \( h \) is the basal area-weighted forest height.

In field inventories, the total basal area is estimated from stem diameter measurements. The measurement of forest height is more time-consuming, and many allometric equations for biomass computation require only stem diameter measurements [25]–[27].

In remote sensing, there are several techniques for forest height estimation, including lidar scanning [7], [28], [29], polarimetric SAR interferometry [30]–[32], photogrammetry [4], [5], and radargrammetry [33], [34], but the estimation of the basal area is difficult, due to canopy closure, shadowing, and too low resolution. On the other hand, the size of tree crowns can be estimated from aerial photography [35]–[37], lidar [9], [38], [39], or SAR interferometry [21], [22]. Since several studies show a reasonable correlation between canopy diameter and stem diameter for many tree species [40]–[42], the total crown area \( A_{cr} \) will in the following be used as a predictor of
the total basal area:

\[ \text{AGB} = C' \cdot h \cdot \eta_{cr}, \]  

(2)

where \(C'\) is a forest type-dependent constant, which is the product of \(C\) and the ratio between the total basal area and the total canopy area, and

\[ \eta_{cr} = \frac{A_{cr}}{A_0} \]  

(3)

is the fractional canopy coverage, which is a measure of canopy density.

The main purpose of the argument above is to show that a multiplicative model is appropriate for biomass estimation from forest height and canopy density. However, this argument is based on several simplifying assumptions regarding the shape of the trees, their intrinsic wood properties, their spatial distribution, tree parameter distribution within a plot, etc. In reality, the dependence of the AGB on the two forest parameters is expected to be more complicated. For instance, there will be a residual dependence of \(C'\) on height and basal area, which may affect the dependence of the AGB estimate on \(h\) and \(\eta_{cr}\) in (2). Exponents \(\alpha\) and \(\beta\) are therefore introduced to create an improved model, based on the experience from field inventories [25], [26]:

\[ \text{AGB} = C'' \cdot h^\alpha \cdot \eta_{cr}^\beta, \]  

(4)

where \(C''\) is a new, forest type-dependent constant.

B. Forest Height and Canopy Density from InSAR

In synthetic-aperture radar interferometry [13], the complex correlation coefficient is the main observable and it is defined as:

\[ \tilde{\gamma} = \frac{\mathbb{E}[s_1 s_2^*]}{\sqrt{\mathbb{E}[|s_1|^2] \mathbb{E}[|s_2|^2]}}, \]  

(5)

where \(s_1\) and \(s_2\) are the two interferometric images, \(^*\) is the complex conjugate operator, and \(\mathbb{E}[\bullet]\) is the expectation value operator.

Coherence is the magnitude of the complex correlation coefficient and it is a measure of similarity between two images. The phase of the correlation coefficient carries information about the vertical distribution of the scatterers. In applications, the complex correlation coefficient is estimated from a finite number of samples, and the interferometric
phase is affected by noise. The total noise level will increase with decreasing number of independent samples and/or decreasing coherence \[43\].

The loss of coherence (decorrelation) can be caused by up to four different effects: temporal changes in the scene, geometric differences between the two images, thermal noise, and system imperfections \[44\], \[45\].

Volume decorrelation is a geometric effect caused by the distribution of scatterers in the vertical direction \(z\). It can be modeled from the vertical backscattering profile \(\sigma(z)\) using \[46\], \[47\]:

\[
\tilde{\gamma}_{vol} = \frac{\int_{-\infty}^{\infty} \sigma(z) e^{ik_z z} dz}{\int_{-\infty}^{\infty} \sigma(z) dz},
\]

with \(k_z\) being the vertical wavenumber, which for a bistatic acquisition geometry is:

\[
k_z = \frac{2\pi}{\text{HOA}} = \frac{2\pi B_{\perp}}{\lambda R \sin \theta},
\]

where HOA is the height-of-ambiguity, \(B_{\perp}\) is the perpendicular baseline, \(\lambda\) is the wavelength, \(R\) is the average range, and \(\theta\) is the average angle of incidence. HOA is the height corresponding to a \(2\pi\)-phase shift in the interferogram, and it is the maximal height difference, which can be unambiguously resolved by the interferometric system.

In the two-level model (TLM) \[21\], \[22\], \[48\], forest is modeled as two scattering levels, ground and vegetation, at the respective elevations \(z_0\) and \(z_0 + \Delta h\), and with the respective backscattering coefficients \(\sigma^0_{gr}\) and \(\sigma^0_{veg}\). The vertical backscattering profile \(\sigma(z)\) is therefore:

\[
\sigma(z) = (1 - \eta)\sigma^0_{gr}\delta(z - z_0) + \eta \sigma^0_{veg}\delta(z - (z_0 + \Delta h)),
\]

where \(\eta\) is the area-fill factor (the fraction of the total area covered by the vegetation level) and \(\delta(\bullet)\) is the Dirac delta function.

Inserting \([8]\) in \([6]\) yields:

\[
\tilde{\gamma}_{vol} = e^{ik_z z_0} \frac{\mu + e^{ik_z \Delta h}}{\mu + 1},
\]

where

\[
\mu = \rho \cdot \frac{1 - \eta}{\eta}
\]

is the area-weighted backscatter ratio with the ground-to-vegetation backscatter ratio \(\rho\) defined as:
\[ \rho = \frac{\sigma_{gr}^0}{\sigma_{veg}^0}. \] (11)

The first exponential term in the expression for the TLM in (9) introduces a phase term related to ground topography \( z_0 \). This phase term can also be observed in the measured complex correlation coefficient \( \tilde{\gamma} \). If ground topography is known, for example from an external digital terrain model (DTM), then this exponential term can be used to compensate the complex correlation coefficient for ground topography, so that the ground-corrected complex correlation coefficient can be obtained:

\[
\tilde{\gamma}_{gc} = \frac{E[s_1 s_2^* e^{-ik_z z_0}]}{\sqrt{E[|s_1|^2] E[|s_2|^2]}}. \] (12)

From the phase of the ground-corrected complex correlation coefficient \( \tilde{\gamma}_{gc} \), the interferometric height (scattering center position above ground) can be obtained using:

\[
h_{gc} = \frac{\text{arg} (\tilde{\gamma}_{gc}) + 2\pi n}{k_z} = \text{HOA} \left( \frac{\text{arg} (\tilde{\gamma}_{gc})}{2\pi} + n \right), \] (13)

where \( \text{arg} (\bullet) \) is the argument operator and the integer \( n \) describes the ambiguity of the phase computation.

In the absence of decorrelation effects other than volume decorrelation, the measured ground-corrected complex correlation coefficient \( \tilde{\gamma}_{gc} \) can be modeled by the TLM expression in (9) with \( z_0 = 0 \):

\[
\tilde{\gamma}_{gc} = \frac{\mu + e^{ik_z \Delta h}}{\mu + 1}. \] (14)

Since this equation has two unknowns (\( \Delta h \) and \( \mu \)) and two observables (the real and imaginary parts of \( \tilde{\gamma}_{gc} \)), it can be solved without the need for multiple acquisitions. The TLM describes a circle in the complex plane with its center on the positive \( x \)-axis and passing through \( \tilde{\gamma}_{gc} \) and 1. The solutions for \( \mu \) and \( \Delta h \) are:

\[
\mu = \frac{1 - \gamma_{gc}^2}{1 - 2\text{Re} [\tilde{\gamma}_{gc}] + \gamma_{gc}^2}, \] (15)

\[
\Delta h = \frac{\tan^{-1} \left[ \frac{2\text{Im}[\tilde{\gamma}_{gc}][1-\text{Re}[\tilde{\gamma}_{gc}]]}{2\text{Re}[\tilde{\gamma}_{gc}][1-\text{Re}[\tilde{\gamma}_{gc}]] + \gamma_{gc}^2 - 1} \right] + \pi n}{k_z}, \] (16)

where \( \gamma_{gc} = |\tilde{\gamma}_{gc}| \) is the ground-corrected coherence, \( \text{Re} [\bullet] \) and \( \text{Im} [\bullet] \) are the real and imaginary part operators, respectively, and \( n \) is an integer describing the ambiguity of the inversion. The lowest positive \( \Delta h \) is chosen in cases when HOA is larger than forest height.
In [21], [22], it has been shown that $\Delta h$ estimated from VV-polarized TanDEM-X data is correlated with $H95$, which is a lidar metric for forest height (see Section III-E for the definition). It has also been shown in [21], [22] that the uncorrected area-fill factor defined as:

$$\eta_0 = \frac{1}{1 + \mu} = \frac{1}{2} \left( \frac{1 + \gamma^2}{1 - \text{Re}[\tilde{\gamma}]} \right).$$

(17)

is correlated with vegetation ratio, which is a lidar metric for canopy density (see Section III-E for a definition). The uncorrected area-fill factor can be obtained by solving (10) for $\eta$ under the assumption that $\rho = 1$. The validity of this assumption has been discussed in [21], and it has been concluded that at high frequencies, such as for the X-band data used in [21], [22], the ground- and vegetation-level scattering coefficients are similar as the wavelength is short compared to the size of the scatterers and the orientation of the scatterers can be considered random.

As mentioned earlier, the derivation of (15), (16), and (17) is based on the assumption that the total decorrelation is only caused by the volume effect. In the TanDEM-X system used in this study, the near-simultaneous, bistatic acquisition scenario minimizes the temporal decorrelation [49]. Common-band filtering of both master and slave images deals with most of the spatial decorrelation effect caused by different range and Doppler frequency bands [50]. Therefore, the two most significant decorrelation effects other than volume decorrelation are caused by the finite SNR and the system imperfections. In [49], the total coherence for soil and rock for VV-polarized TanDEM-X acquisitions in mid-swath and at a 41-degree incidence angle has been modeled to approximately 0.88 for an occurrence level of 50% and 0.82 for an occurrence level of 90%. In this study, however, all decorrelation effects other than volume decorrelation will be neglected for practical reasons, and the validity of this assumption will be discussed in Section V.

Due to the high resolution of TanDEM-X data and relatively large regions of interests, a large number of independent samples can be used during the estimation of the complex correlation coefficient, and the errors in coherence and phase estimation are negligible. An estimate of the number of looks used during the computation of the complex correlation coefficient will be given in Section III-G.

**C. Biomass Models**

The following biomass model, called the TLM biomass model or shortly the TBM, is introduced:

$$\hat{AGB} = K \cdot \Delta h^\alpha \cdot \eta_0^\beta,$$

(18)

where $K$, $\alpha$, and $\beta$ are unknown model parameters, and $\Delta h$ and $\eta_0$ are forest height and canopy density estimates obtained from the TLM inversion using (16) and (17), respectively.
The TBM will be compared to a linear, zero-intercept model, which scales the ground-corrected interferometric height to biomass:

$$\hat{\text{AGB}} = D \cdot h_{gc},$$  \hspace{1cm} (19)

where $D$ is a scaling factor, which needs to be estimated from the training data. This model has been proposed in [17]. In the following, it will be referred to as the scaling model, or simply the SM.

**D. Evaluation Strategy**

The models will be evaluated using multiple TanDEM-X acquisitions acquired over two, geographically separated test sites in Sweden, during three consecutive summers and at different HOAs. The models will be tested both for their explanatory values (that is how well they can be fitted to the data), and their predictive values (that is how well they can predict biomass from other data). The models will thus be tested for their robustness to the change of test site, acquisition year, and acquisition HOA. They will also be used to produce biomass maps, to see how well the spatial structures can be reproduced.

**III. DATA**

**A. Test Sites**

Remningstorp is a hemi-boreal test site situated in southern Sweden (58° 28’ N, 13° 38’ E), see Figure 2. It is fairly flat with ground slopes at stand level lower than 5° (computed from a 50 m × 50 m digital terrain model,
DTM). The test site covers approximately 1200 ha of productive forest land, and the forest consists primarily of Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and birch (*Betula* spp.). For a description of the test site, see [51].

Krycklan is a boreal test site located in northern Sweden (64° 14’ N, 19° 46’ E), see Figure 2. Krycklan is situated 720 km north-north-east of Remningstorp. Unlike Remningstorp, Krycklan has a strongly undulating topography with ground slopes on stand level up to 19° (again, computed from a 50 m × 50 m DTM). The forest is dominated by Norway spruce and Scots pine. For a description of the test site, see [52].

**B. In-Situ Data**

A set of 32 circular, 40-meter radius plots is available for Remningstorp. Field inventories were conducted during the autumn of 2010 and spring of 2011. For each plot, all trees with a diameter at breast height (dbh) higher than 5 cm were calipered and tree species were determined. Height was measured for a subset of roughly 10% of the trees. Out of the thirty-two plots, twenty-one are spruce-dominated (more than 2/3 of biomass), five are pine-dominated, and two are birch-dominated. Three plots consist of a mixed spruce and pine forest and one plot consists of a mixed forest with all three tree species.

In Krycklan, a set of 31 stands of irregular shape and sizes between 2.4 and 26.3 hectares were inventoried in the summer of 2008. Systematic grids of circular field plots (radius 10 m) were laid out in each stand. The spacing of each grid was selected to give between 8 and 13 field plots per stand. For each field plot, all trees with a dbh higher than 4 cm were calipered and the species were determined. Tree height and age were also measured for 1–2 randomly chosen sample trees in each field plot. Of the thirty-one stands, five are spruce-dominated, thirteen are pine-dominated, three are mixed coniferous, and the remaining ten are mixed forest stands.

**C. Biomass Estimates**

For both test sites, estimates of above-ground dry biomass have been made from the in-situ data using the Heureka system [53], which implements the allometric functions described in [23]. The allometric functions have been derived using multiple regression analysis of data from 1286 trees (Norway spruce, Scots pine, and birch) from 131 stands located across Sweden and described in [54], [55].

Stem volume growth has been modeled in Heureka using the radial growth functions described in [23]. Although SAR acquisitions have been made in the summer, which is in the middle of a growth season, biomass estimates for the end of the preceding growth season will be used throughout this study. The performance of the volume growth model used in Heureka has been evaluated in [56] using 1711 permanent plots from the National Forest Inventory (NFI) database. The prediction error (RMSE) for the stem volume has been found to be around 15%, and a small
underestimation (bias) of 2% has been observed for spruce. Due to the close relation between forest volume and above-ground dry biomass, similar errors are also expected for the AGB.

A realistic estimate of the uncertainty in the reference biomass data used in this study is 15%, primarily based on the results presented in [56] and the errors presented in [23]. Although the sampling procedures in Remningstorp and Krycklan include the measurement of the dbh for a large set of trees (all trees with dbh larger than 5 cm in Remningstorp), height has only been measured for a subset of trees, and thereafter extrapolated to the other trees using regression from the dbh. Since both the dbh and height are used for biomass estimation, the input variables to the allometric equations are correlated, which increases the uncertainty of the aggregated estimates. A possible bias will also occur when the models presented in [23] are used locally, on data which may deviate from the data used for the derivation of these models. Additional uncertainties, such as in-situ measurement errors and errors introduced during the determination of plot areas also contribute to the total error.

D. Forest Change Detection

After field measurements, several plots/stands have been altered through clearing, thinning, or clear-cutting. In Remningstorp, the altered plots have been identified using lists of management procedures provided by the managing company, SPOT-5 image analysis, and field visits. Three plots have been altered between the SAR acquisitions from 2011 and from 2012, and additional eight between the SAR acquisitions from 2012 and from 2013. In Krycklan, only SPOT-5 image analysis has been used. Two stands have been altered already before the first SAR acquisition in 2011, but no changes have been detected after that. Altered plots/stands have been disregarded in this study.

E. Lidar Data

Two lidar metrics have been extracted from 10 m × 10 m maps provided within the BioSAR 2008 and 2010 campaigns [51], [52]. The 95th-percentile forest height, called H95, has been computed as the 95th percentile of all lidar returns above a threshold of 1 m or 10% of the maximal return within a 10 m × 10 m cell. The lidar vegetation ratio, called VR, has been computed as the ratio between the number of returns from above that threshold to all returns. The VR is thus a measure of canopy density.

Biomass maps with a 10 m × 10 m resolution have been derived from multiple regression analysis of different lidar metrics and species stratification maps, see [51], [52]. In Remningstorp, 212 circular field plots with a radius of 10 m and distributed in a systematic grid over the entire test site, have been used for model training. In Krycklan, the previously mentioned field plots located within the 31 stands, together with additional 110 circular field plots surveyed with the same methodology and positioned within the central part of the test site have been used for model training. The uncertainty in the biomass maps is estimated to 20%, based on the uncertainties reported in [51], [52]. The maps will only be used in a qualitative, side-by-side comparison.
Note that the lidar data have not been corrected for growth. Since the lidar data are only used in qualitative comparisons, this does not affect the quantitative results presented in this study.

F. Digital Terrain Model

As ground level reference, the national, lidar-scanned digital terrain model (DTM) acquired by the Swedish Land Survey is used [14]. The DTM has a $2 \times 2$ m grid, with a mean height error lower than 0.5 meters. Lidar scanning has been performed from an airplane flying at an altitude between 1700 and 2300 meters. Point density on the ground is between 0.5 and 1 point per square meter. In the southern part of the country, lidar scanning has primarily been performed during non-vegetative periods to minimize the contribution of leaves, grass, crops, etc.

G. InSAR Data

TanDEM-X is a twin-satellite, X-band (9.65 GHz) SAR interferometer in which acquisitions are made almost simultaneously [49]. Bistatic-interferometric, VV-polarized, stripmap-mode TanDEM-X (TDM) acquisitions made at low HOA in the ascending mode are used in this study. The choice of the low-HOA data is motivated by the better sensitivity to forest height [57]. A summary of the data can be found in Table I where background color coding by HOA has been applied. Note that the data from 2012 feature lower HOAs than the data from 2011 and 2013. The nominal angle of incidence varies between 41.2 and 41.7 degrees for Remningstorp and between 40.4 and 41.0 degrees for Krycklan. For images 1 and 9 in Table I (the first acquisitions for each test site), the scene center resolutions provided by the DLR in the meta files are: 1.8 m in ground range and 6.6 m in azimuth. For the rest of the images, the ground range resolution is 2.7 m and the azimuth resolution is 3.3 m.

The InSAR data have been interferometrically processed using an in-house developed algorithm based on [50]. The raw interferograms have been ground-corrected in radar geometry using a linearly interpolated DTM and taking into consideration the quasi-bistatic acquisition geometry and satellite displacement between transmission and reception of the signals. A 5-meter buffer zone has been added prior to plot/stand-level averaging of the ground-corrected interferograms. The lowest number of looks has been estimated to 320, computed as the ratio between the area of the smallest plot/stand in the data set (excluding the buffer zone), and the ground range and azimuth resolutions for the image with the lowest resolution. Absolute phase calibration has been done using ground reference points derived from a non-forest mask. No unwrapping has been found necessary due to the limited height variations in the flattened interferogram. Geocoding error and height measurement errors have been estimated using two 5-meter trihedral corner reflectors situated within the Remningstorp site. The geocoding offset has been found lower than 2 m and the standard deviation of the measured elevation of the scattering center has been found lower than 10 cm. For the creation of the ground-corrected coherence and height images, a $5 \times 5$ averaging window has been used.
TABLE I

SUMMARY FOR THE EXPERIMENTAL DATA USED IN THIS STUDY. MEAN VALUES FOR ALL PLOTS ARE GIVEN. BACKGROUND SHADING HAS BEEN APPLIED ACCORDING TO HOA. N IS THE NUMBER OF AVAILABLE PLOTS/STANDS FOR EACH ACQUISITION.

<table>
<thead>
<tr>
<th>Nr</th>
<th>Site</th>
<th>Date</th>
<th>InSAR data</th>
<th>N-situ &amp; lidar data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$B_L$ [m]</td>
<td>HOA [m]</td>
</tr>
<tr>
<td>1</td>
<td>Remningstorp</td>
<td>20110604</td>
<td>282</td>
<td>49</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>20110809</td>
<td>266</td>
<td>52</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>20110820</td>
<td>258</td>
<td>54</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>20120601</td>
<td>432</td>
<td>32</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>20120828</td>
<td>370</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>20130702</td>
<td>270</td>
<td>51</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>20130724</td>
<td>226</td>
<td>61</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>20130804</td>
<td>220</td>
<td>63</td>
</tr>
<tr>
<td>9</td>
<td>Krycklan</td>
<td>20110617</td>
<td>258</td>
<td>52</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>20110720</td>
<td>250</td>
<td>54</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>20110811</td>
<td>242</td>
<td>55</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>20110822</td>
<td>240</td>
<td>56</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>20120717</td>
<td>374</td>
<td>36</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>20120808</td>
<td>360</td>
<td>37</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>20120819</td>
<td>350</td>
<td>39</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>20130601</td>
<td>270</td>
<td>50</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>20130623</td>
<td>260</td>
<td>52</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>20130726</td>
<td>216</td>
<td>62</td>
</tr>
</tbody>
</table>

Color coding by HOA: [30 m | 40 m | 50 m | 60 m]

IV. RESULTS

In this section, the two models will be fitted to the experimental data, and biomass will be estimated.

The significance of each model parameter will be studied using the Student’s $t$-test. This test evaluates the hypothesis that the expectation value of the normally distributed parameter estimate $\hat{\beta}$ is $\beta_0$. The $t$-statistic is computed as:

$$ t = \frac{\hat{\beta} - \beta_0}{\hat{\sigma}_\beta}, \quad (20) $$

where $\hat{\sigma}_\beta$ is the estimated standard deviation of $\hat{\beta}$. The Student’s $t$-test will here be used to test the hypothesis that $\beta_0 = 0$. For a known number of degrees of freedom, the probability $p$ of obtaining a certain $t$-statistic can be computed from the $t$-distribution. A low $p$-value means that $\beta$ is a significant parameter.

The goodness-of-fit of each model will be evaluated using the coefficient of determination $R^2$, which describes the fraction of the total variability observed in the data that can be explained by the model:
\[
R^2 = 1 - \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2},
\]

(21)

where \(Y_i\) are the observed values, \(\hat{Y}_i\) are the corresponding modeled values, and \(\overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i\) is the average observed value.

The model error will be evaluated using the root-mean-square error (RMSE), which is computed as:

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}.
\]

(22)

The fitting of the models has been done using non-linear least squares, as implemented in the \texttt{nls} function provided within the R-package \cite{58}. Although the use of linear least squares is possible for the TBM model after a logarithmic transform, non-linear regression has in many cases provided lower RMSE and larger \(R^2\), and it is therefore used throughout this study.

A. Interferometric Height and Coherence

In Figure 3(a), TanDEM-X interferometric height \(h_{gc}\) is plotted against lidar height \(H95\), separately for each year. A good correlation can be observed, but the interferometric height is approximately 5–10 m lower for almost all stands.

For some Remningstorp plots with \(H95\) just above 25 m, the interferometric height is approximately 5 m higher for the 2012 acquisitions (with HOAs equal to 32 m and 37 m) than for the 2011 and 2013 acquisitions (with HOAs around and above 50 m). This effect has previously been discussed in \cite{21}, where it has been concluded that it is caused by an interference effect occurring when ground- and vegetation-level scattering is of similar strength and when the distance between the respective scattering centers is around \(\text{HOA}/2\). The affected plots consist of former seed trees with new understorey vegetation. The trees are sparse, and allow for a significant penetration through the gaps, and the understorey vegetation layer boosts the ground-level scattering.

In Figure 3(b), ground-corrected TanDEM-X coherence \(\gamma_{gc}\) is plotted against lidar vegetation ratio, separately for each year. It can be observed that coherence is consistently lower for the acquisitions from 2012 than for the other two acquisitions, due to the larger baseline, and for some plots, the coherence falls below 0.3. It is noted that these plots are the same, sparse plots with former seed threes and rich understorey vegetation, and that this low coherence occurs due to the aforementioned interference effect.

In Figure 3(c), TanDEM-X interferometric height is plotted against reference biomass. For Krycklan, there is a good correlation between the interferometric height and biomass. For Remningstorp, however, there is a large height variance, especially in the case of the data from 2012, with low HOA. It is noted that the plots with relatively
low biomass but high interferometric height are the same sparse plots that have been discussed in the previous paragraphs. Note that one of these plots have been altered in 2012 and additional two in 2013, as described in [21], and they are not included in the scatter plots from 2012 and 2013.

B. TLM Inversion Products

In Figure 4(a) the level distance $\Delta h$ inverted using (16) is plotted against lidar height $H95$. It can be observed that the correlation between $\Delta h$ and $H95$ is better than between the interferometric height and $H95$, but the bias is different, as observed in [21], [22]. Note that the slope of the inverted level distance $\Delta h$ changes at low $H95$. The reason for this will be discussed in Section V.

In Figure 4(b) the area-weighted backscatter ratio $\mu$ inverted using (15) is plotted in decibels against lidar vegetation ratio. Although these two parameters measure different properties, a good correlation can be observed.

In Figure 4(c) the uncorrected area-fill factor $\eta_0$ inverted using (17) is plotted against lidar vegetation ratio. The correlation is good for most stands.

C. Parameter Estimation Results

In Table II estimates of $K$, $\alpha$, and $\beta$ are shown for the TLM biomass model (TBM). The TBM is able to explain between 65% and 89% of the variance observed in the data. It can be observed that $\alpha$ is similar for both test sites and for most acquisitions, with most values close to one. The other exponent, $\beta$, is similar for all acquisitions made over the same test site, but it changes between the test sites. For acquisitions made in Krycklan, it is close to one, whereas for those made in Remningstorp, it is closer to three. The third parameter, $K$, shows a large variance with values between 0.4 and 27.1.

Based on these observations, it is reasonable to let the exponents become constants. The exponent $\alpha$ can be fixed to the same value for both test sites, while the exponent $\beta$ must change between test sites. In Table III(a) regression results for the TBM with $\alpha$ fixed to 1.25, and $\beta$ fixed to 2.64 for Remningstorp and 1.16 for Krycklan are shown. The chosen values are all average values for the estimates presented in Table II. The estimated values of the slope constant $K'$, are more stable than the estimates of $K$, between 6.6 and 10.2, without any significant difference between the two test sites. Note that the lowest $R^2$ is obtained for the image from Remningstorp with the lowest HOA. Since the choice of the fixed parameters is based on the mean of all values, it is biased towards acquisitions with large HOA, as they are more frequent.

In Table III(b) regression results for the scaling model (SM) are shown. It can be observed that the slope is very stable for both Remningstorp and Krycklan, but it changes between the two sites. For Remningstorp, it is between 8.3 and 9.6, whereas for Krycklan, it is between 11.3 and 12.2. The two lowest values are obtained for the images
Fig. 3. TanDEM-X interferometric height and ground-corrected coherence are plotted against reference data. Acquisition year and HOA intervals are shown for each subplot. Note: several points may overlap. In Remningstorp, there are 88 points in 2011, 58 points in 2012, and 63 points in 2013. In Krycklan, there are 116 points in 2011 and 87 points in both 2012 and 2013.

Fig. 4. Inverted TLM parameters are plotted against reference data. Acquisition year and HOA intervals are shown for each subplot. Note: several points may overlap. In Remningstorp, there are 88 points in 2011, 58 points in 2012, and 63 points in 2013. In Krycklan, there are 116 points in 2011 and 87 points in both 2012 and 2013.
over Remningstorp acquired at the lowest HOA values. For these, the coefficient of determination is -0.07 and 0.13, and the SM is not able to explain the variance for these acquisitions.

Note that the fixed parameters chosen above as averages of the estimated parameters shown in Table II can also be estimated from regression of all relevant data but this approach has not been chosen here. The main purpose of this part is to show that fixed exponents provide a more stable model. The choice of exact parameter values is not of primary interest to this study.

D. Biomass Estimation Results

Residual scatter plots for the models are shown together with two-sigma error bars in Figure 5. It can be observed that only for the Remningstorp acquisitions from 2012, fixing of the exponents in the TBM significantly decreases model performance, as it also can be observed by comparing the $R^2$-values in Table II and Table III(a). The SM performs poorer in Remningstorp than in Krycklan.

Both residual and prediction root-mean-square error (RMSE) values are shown in Table IV for the TBM, in Table V for the TBM with fixed exponents, and in Table VI for the SM. As expected, the performance of the TBM is poorer in across-site evaluation and for large difference in HOA in Remningstorp, primarily due to the differences in the exponent $\beta$. It can also be observed that the TBM with fixed exponents gives a much lower and more stable prediction RMSE without increasing the residual RMSE significantly. In the case of the SM, both residual and prediction RMSEs are low in Krycklan, but higher in Remningstorp and across-sites.

In Figure 6 the dependence of the RMSE on the parameters $K$, $\alpha$, and $\beta$ is studied. The default values of the parameters are: $K = 7.42$, $\alpha = 1.25$, $\beta = 2.64$ for Remningstorp and $\beta = 1.16$ for Krycklan, and they are marked with vertical lines. Each parameter is varied around its default value and the RMSE is computed. While one parameter is varied, the other two are held constant at their default values. Remningstorp and Krycklan are shown separately, and color coding according to HOA has been applied. Note the significant difference between Remningstorp and Krycklan in model sensitivity to different parameter settings at different HOAs.

E. Biomass Mapping

In Figure 7 biomass maps obtained using the TBM and SM are shown for both Remningstorp and Krycklan, and compared to lidar-derived biomass maps. For both test sites, the first acquisition from 2011 has been used: nr 1 for Remningstorp and nr 9 for Krycklan), together with the respective parameters presented in Table II and Table III(b). Note that forest management procedures may have been conducted between the acquisition of lidar and TanDEM-X data. Note also that, in Remningstorp, regions not covered by lidar scanning have been masked out.
As observed earlier during the residual study, the TBM performs well both in Remningstorp and in Krycklan, if the parameters obtained from acquisitions made at similar HOA and within the same test site are used. Although the SM performs well in Krycklan, it can be observed that it cannot reproduce the variance of biomass in Remningstorp.

V. DISCUSSION

The TLM biomass model (TBM) is able to explain 65–89% of the variance observed in data, with a residual RMSE of 12-19% of the mean biomass, for 18 TanDEM-X images acquired over two test sites in Sweden. In cases when different data are used for training and validation, the model shows poorer results, with a prediction RMSE often exceeding 30% in across-site scenarios or when the difference in HOA is large. However, the TBM can be stabilized by fixing $\alpha$ and by letting $\beta$ be a site-dependent constant. In that case, the prediction RMSE is below 20% for most of the acquisitions.

The TBM is here compared to a linear, zero-intercept model, which scales the interferometric height to biomass. This scaling model (SM) has earlier been used in [17], where a stand-level residual RMSE of 19% has been obtained for a Norwegian test site using two images, one from the ascending orbit and one from the descending orbit, with the respective HOAs 23 m and 122 m. In [17], the scaling factor has been estimated to 14 t/ha/m, while in this study, the same factor is between 8.3 and 9.6 for Remningstorp and between 11.3 and 12.2 for Krycklan. A method for biomass change detection has also been proposed in [17], based on direct scaling of the change in the interferometric height to change in biomass, without the need for a high-resolution DTM. From the results of this study, it can be concluded that whereas Krycklan shows many similarities with the Norwegian test site, Remningstorp appears to be significantly different. The dependence of the interferometric height on HOA and the horizontal forest structure is especially large in Remningstorp, and the model presented in [17] does not function well in this test site.

In [19], an approach based on the interferometric water cloud model (IWCM) and multi-temporal averaging of stand-level biomass estimates from eighteen acquisitions made over Remningstorp at HOAs between 49 m and 358 m, in both summer and winter, gives an RMSE of 16%, whereas multi-temporal averaging of seven images acquired at temperatures below 3°C gives an RMSE of 14%. In the case of a single image, the RMSE is in the interval 17–33%. Additionally, a penetration depth (PD) model is introduced in [19], which uses a height-to-biomass allometric equation to compute biomass from the sum of the interferometric height and the penetration depth. The PD-based approach gives RMSE values in the interval 18–33%. The penetration depth is the only parameter that needs to be estimated (the allometric relation is assumed to be known). However, the model does not account for the horizontal structure of the forest, which can be problematic in the case of a commercial forest, where management activities such as thinnings and clearings affect the denseness of the forest and its biomass, but not necessarily its height.

July 11, 2014
In [20], an approach based on multiple regression of the interferometric height and the ground-corrected coherence, and their transformed versions gives an RMSE of 14% for one TanDEM-X image acquired over Krycklan (nr 9 in Table I) and 17% for one TanDEM-X image acquired over Remningstorp (nr 1 in Table I). The approach presented in [20] is based on multiple regression of interferometric observables, and the results may not be representative, especially considering the limited extent of the data used in the study.

A. TLM Inversion Requirements

As biomass predictors, the TBM uses two parameters obtained from the inversion of a two-level model (TLM). The inversion of the TLM requires a high-resolution digital terrain model (DTM), and it is based on the assumption that volume decorrelation is the dominant decorrelation effect.

In Sweden, there is a lidar DTM covering the whole country, and similar products are or will soon be available in many other countries. Therefore, the presented approach can already be used in many places and on a large scale, as the global TanDEM-X data used for DEM generation have been made available by the DLR for scientific use. Since ground surface is temporally stable in most forested areas, the availability of high-resolution DTMs will only increase with time. The exact requirements on the DTM have not been studied here, but it is possible that a coarser DTM than can be sufficient, as exemplified in [17]. Other techniques, such as low-frequency SAR, may also be used as ground reference.

In the presented approach, volume decorrelation has been assumed to be the dominant decorrelation effect. In the case of the single-pass interferometric, bistatic-mode TanDEM-X data used here, the most significant decorrelation sources other than volume effects are the thermal noise and system imperfections. These effects have been ignored throughout this study to keep the inversion process simple.

However, it has been observed in [21] that the TLM inversion provides unrealistically high $\Delta h$ for clear-cuts, and that the inverted level distance $\Delta h$ is affected by a HOA-dependent offset. In Figure 4(a), a change of slope has been observed for low H95. Additionally, a HOA dependence has been observed in the estimated model parameters $\alpha$, $\beta$, and $K$ presented in Table II and that this dependence is stronger in Remningstorp, where the forest is taller and the relative HOA is lower.

A probable cause for these effects is that the SNR and system decorrelation effects have not been considered in the TLM inversion process. Since the interferometric phase has been calibrated using non-forested areas, the complex correlation coefficient for low forest and open areas has a high, yet non-unitary real part and low imaginary part. The TLM cannot model a complex correlation coefficient with high coherence and low phase without making $\Delta h$ close to HOA/2. This introduces a HOA-dependent offset in the estimated $\Delta h$.

A solution for this issue can be obtained through the modeling of a real-valued system and SNR decorrelation.
term $\gamma_{sys}$, and replacing $\tilde{\gamma}_{gc}$ by $\tilde{\gamma}_{gc} \gamma_{sys}$ in (15), (16), and (17). If $\gamma_{sys}$ can be estimated, e.g., from the data, then an improvement of the TLM inversion performance can be expected.

A second probable cause for these effects is that at low HOA relative forest height, the modeling of the exact vertical distribution of scatterers becomes more important, as the phase change with height is larger. The assumption of two scattering levels may then become too simplistic.

Nevertheless, the issues related to HOA dependence can be avoided in practical use by a sensible choice of HOA. Moreover, $\eta_0$ can often compensate for the large $\Delta h$, as observed for Remningstorp in Figure 7, where biomass mapping in open fields is accurate.

B. Influence of Forest Structure

An interesting observation can be made about the estimated values for the exponent $\beta$, associated with the canopy density estimate $\eta_0$, which changes significantly between Remningstorp and Krycklan. This can be explained by different structure of the trees in Remningstorp and in Krycklan. It can be estimated using the Heureka system, that 68% of the total biomass in the 32 plots in Remningstorp is confined to the stem, while for the 31 stands in Krycklan, the same number is 76%. In [59], it is concluded that the trees in northern Sweden generally have smaller crowns than in southern Sweden. A larger $\beta$ in Remningstorp will reduce the contribution of the canopy density estimate $\eta_0$ to the total biomass, thus compensating for the fact that the forest in Remningstorp has in general denser canopies, at similar biomass and height.

The TLM can be used to study temporal change of canopy density from multi-temporal, single-pass InSAR acquisitions. By keeping $\Delta h$ constant for all acquisitions and letting $\eta_0$ vary, an over-determined equation system is obtained. A time series study of $\eta_0$ can provide information on the change of the canopy density, due to seasonal variations, management procedures such as clearing, thinning, and clear-cutting, or natural disasters, and, eventually, biomass change can be estimated as well.

C. Future Development

There is an ongoing debate about the mechanisms of microwave penetration into forest canopy, and whether the significant penetration of X-band SAR into the canopy is primarily due to the dielectric penetration through the scatterers or penetration through the canopy gaps [60]. In this study, it is shown that the inclusion of canopy gaps in an interferometric model can be beneficial for model inversion, but the dielectric penetration has been disregarded, and further discussion on the penetration mechanisms is left for future studies.

The presented approach has been evaluated on VV-polarized, X-band SAR data. An evaluation of this approach on other frequencies, for instance C-band, as well as other polarizations is of large interest. The presented approach is principally not restricted to the used frequency, although the TLM inversion process may need to be revisited at July 11, 2014 DRAFT
other frequencies, where a different choice of $\rho$ may be motivated. Also, as shown in [16], the difference between the interferometric height at HH- and VV-polarization can be several meters. Although the exact relation between the inverted parameters $\Delta h$ and $\eta_0$ and the lidar estimates of forest height and canopy density, and biomass will most likely be different at other polarizations and frequencies, the presented approach may still be useful.

This study has been restricted to data acquired at a 41-degree nominal angle of incidence. The influence of the incidence angle requires a separate study. An evaluation of the presented approach on tropical forest is also of interest. As the tropical forest is, in general, taller and denser, the penetration through canopy gaps is expected to be lower, which certainly will affect TLM inversion.

VI. Conclusions

A new biomass model is proposed, in which biomass is estimated from forest height and canopy density estimates obtained from the inversion of a two-level model (TLM) using single-pass interferometric SAR data. In this study, bistatic-mode, VV-polarized TanDEM-X data acquired at a 41-degree nominal incidence angle over two Swedish test sites separated by 720 km are used together with the national, digital terrain model (DTM) with a grid posting of $2 \times 2$ m and a vertical accuracy better than 0.5 m. Compared to other studies, the presented approach provides similar or better results in terms of biomass retrieval, a larger data set has been used for the evaluation, and across-site and across-acquisition biomass retrieval scenarios have been studied.

It is here concluded that the two test sites used in this study feature quite different forest, and regional training of the new model is required in operational use. However, only one of the three model parameters has been found significantly dependent on the test site, and the regional model training can be done using only a few data points, e.g., from the National Forest Inventory database. The HOA dependence most likely caused primarily by the lack of system and SNR decorrelation modeling can be suppressed either by choosing HOAs larger than approximately twice the forest height, which is the case for most of the global TanDEM-X acquisitions over boreal forests, or by the modeling of a real-valued system and SNR decorrelation term.

Since a high-resolution digital terrain model is required for TLM inversion, the presented approach is suitable for frequent mapping of large areas of forest in regions with known topography. However, the ground surface is most often temporally stable, and only one DTM acquisition is required. Thereafter, forest height, canopy density, and biomass mapping can be done using spaceborne SAR with large coverage, high resolution, and frequent acquisitions. Therefore, the presented approach is useful for the monitoring of national forest resources, and for improved forest management. With an access to the global TanDEM-X data, national maps of forest height, canopy density, and biomass can be created.
ACKNOWLEDGMENTS

The authors would like to thank the German Aerospace Center (DLR) for the TanDEM-X data and support provided within proposal XTI VEGE0376, the Swedish Land Survey for the DTM and SPOT-5 data, Skogssällskapet for the information on forest management activities, and the European Space Agency (ESA) for lidar data from the BioSAR 2008 and 2010 campaigns.

REFERENCES


[37] S. A. Soenen, D. R. Peddle, R. J. Hall, C. A. Coburn, and F. G. Hall, “Estimating aboveground forest biomass from canopy reflectance...


Maciej Jerzy Soja was born in 1985 in Warsaw, Poland. He received the B.Sc. and M.Sc. degree in engineering physics from Chalmers University of Technology, Gothenburg, Sweden, in 2008 and 2009, respectively. Since December 2009, he is pursuing his Ph.D. degree in remote sensing in the Radar Remote Sensing Group at the Department of Earth and Space Sciences at Chalmers University of Technology. His main research topic is synthetic aperture radar in forestry.
Henrik Persson was born in 1983. He received the M.Sc. degree in engineering physics from Luleå University of Technology in 2009. After two years of working at the ETH Zürich in Switzerland he started pursuing his Ph.D. degree in forestry remote sensing at the Swedish University of Agricultural Sciences in September 2011. His main research topic is modelling of forest variables from different satellite sensors.

Lars M. H. Ulander (S’86-M’90-SM’04) received the M.Sc. degree in engineering physics in 1985 and the Ph.D. degree in electrical and computer engineering in 1991, both from Chalmers University of Technology. Since 1995 he has been with the Swedish Defence Research Agency (FOI) in Linköping where he is Director of Research in radar signal processing. He is also Adjunct Professor in radar remote sensing at Chalmers University of Technology. His research areas are synthetic aperture radar (SAR), electromagnetic scattering models and remote sensing applications. He is the author or co-author of over 250 professional publications, of which more than 50 are in peer-reviewed scientific journals. He is the holder of five patents.
TABLE II
RESULTS FOR THE TLM BIOMASS MODEL (TBM). FOR EACH PARAMETER, THE ESTIMATED STANDARD DEVIATIONS \( \sigma \), t-STATISTICS, AND p-VALUES ARE SHOWN. FOR THE WHOLE MODEL, THE COEFFICIENTS OF DETERMINATION \( R^2 \) ARE SHOWN.

TBM: \( \hat{AGB} = K \Delta h^{\alpha} \eta^{\beta} \)

<table>
<thead>
<tr>
<th>Nr</th>
<th>Site</th>
<th>( K )</th>
<th>( \sigma )</th>
<th>( t )</th>
<th>( p )</th>
<th>( \alpha )</th>
<th>( \sigma )</th>
<th>( t )</th>
<th>( p )</th>
<th>( \beta )</th>
<th>( \sigma )</th>
<th>( t )</th>
<th>( p )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Remningstorp</td>
<td>9.4</td>
<td>4.87</td>
<td>1.9</td>
<td>6.4e-02</td>
<td>1.2</td>
<td>0.18</td>
<td>6.7</td>
<td>2.6e-07</td>
<td>2.7</td>
<td>0.37</td>
<td>7.3</td>
<td>4.5e-08</td>
<td>0.78</td>
</tr>
<tr>
<td>2</td>
<td>11.6</td>
<td>6.49</td>
<td>1.8</td>
<td>8.6e-02</td>
<td>1.1</td>
<td>0.19</td>
<td>6.0</td>
<td>3.3e-06</td>
<td>3.0</td>
<td>0.39</td>
<td>7.6</td>
<td>5.5e-08</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>27.1</td>
<td>15.39</td>
<td>1.8</td>
<td>9.1e-02</td>
<td>0.9</td>
<td>0.19</td>
<td>4.6</td>
<td>9.6e-05</td>
<td>3.0</td>
<td>0.38</td>
<td>7.9</td>
<td>3.0e-08</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.4</td>
<td>0.37</td>
<td>1.0</td>
<td>3.5e-01</td>
<td>2.4</td>
<td>0.39</td>
<td>6.0</td>
<td>2.2e-06</td>
<td>1.8</td>
<td>0.34</td>
<td>5.3</td>
<td>1.4e-05</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.7</td>
<td>1.13</td>
<td>1.5</td>
<td>1.5e-01</td>
<td>1.8</td>
<td>0.24</td>
<td>7.3</td>
<td>9.2e-08</td>
<td>2.0</td>
<td>0.30</td>
<td>6.5</td>
<td>7.4e-07</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7.4</td>
<td>5.63</td>
<td>1.3</td>
<td>2.0e-01</td>
<td>1.2</td>
<td>0.25</td>
<td>4.8</td>
<td>1.3e-04</td>
<td>2.6</td>
<td>0.44</td>
<td>5.8</td>
<td>1.8e-05</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>6.5</td>
<td>3.71</td>
<td>1.7</td>
<td>9.8e-02</td>
<td>1.3</td>
<td>0.19</td>
<td>6.8</td>
<td>2.4e-06</td>
<td>2.9</td>
<td>0.36</td>
<td>8.0</td>
<td>2.4e-07</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>20.1</td>
<td>2.54</td>
<td>1.3</td>
<td>2.1e-01</td>
<td>0.9</td>
<td>0.25</td>
<td>3.8</td>
<td>1.2e-03</td>
<td>3.2</td>
<td>0.55</td>
<td>5.8</td>
<td>1.7e-05</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Krycklan</td>
<td>15.0</td>
<td>11.34</td>
<td>1.3</td>
<td>2.0e-01</td>
<td>1.0</td>
<td>0.25</td>
<td>3.9</td>
<td>5.7e-04</td>
<td>1.4</td>
<td>0.26</td>
<td>5.5</td>
<td>9.4e-06</td>
<td>0.85</td>
</tr>
<tr>
<td>10</td>
<td>9.6</td>
<td>6.76</td>
<td>1.4</td>
<td>1.7e-01</td>
<td>1.1</td>
<td>0.23</td>
<td>4.8</td>
<td>5.8e-05</td>
<td>1.3</td>
<td>0.28</td>
<td>4.5</td>
<td>1.3e-04</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>13.7</td>
<td>9.25</td>
<td>1.5</td>
<td>1.5e-01</td>
<td>1.0</td>
<td>0.22</td>
<td>4.6</td>
<td>1.1e-04</td>
<td>1.4</td>
<td>0.27</td>
<td>5.2</td>
<td>2.2e-05</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8.5</td>
<td>5.09</td>
<td>1.7</td>
<td>1.1e-01</td>
<td>1.2</td>
<td>0.21</td>
<td>5.7</td>
<td>5.8e-06</td>
<td>1.2</td>
<td>0.24</td>
<td>5.0</td>
<td>3.3e-05</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>3.9</td>
<td>2.37</td>
<td>1.5</td>
<td>1.2e-01</td>
<td>1.5</td>
<td>0.21</td>
<td>6.9</td>
<td>2.4e-07</td>
<td>0.9</td>
<td>0.34</td>
<td>2.6</td>
<td>1.6e-02</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>4.6</td>
<td>2.96</td>
<td>1.6</td>
<td>1.3e-01</td>
<td>1.4</td>
<td>0.22</td>
<td>6.4</td>
<td>8.3e-07</td>
<td>0.9</td>
<td>0.35</td>
<td>2.6</td>
<td>1.5e-02</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>6.7</td>
<td>4.53</td>
<td>1.5</td>
<td>1.5e-01</td>
<td>1.3</td>
<td>0.22</td>
<td>5.7</td>
<td>4.9e-06</td>
<td>1.2</td>
<td>0.38</td>
<td>3.0</td>
<td>5.4e-03</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>8.9</td>
<td>6.47</td>
<td>1.4</td>
<td>1.8e-01</td>
<td>1.1</td>
<td>0.24</td>
<td>4.8</td>
<td>5.4e-05</td>
<td>1.1</td>
<td>0.35</td>
<td>3.1</td>
<td>4.8e-03</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>20.5</td>
<td>16.41</td>
<td>1.2</td>
<td>2.2e-01</td>
<td>0.9</td>
<td>0.26</td>
<td>3.4</td>
<td>2.1e-03</td>
<td>1.4</td>
<td>0.28</td>
<td>4.9</td>
<td>4.9e-05</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>8.8</td>
<td>6.21</td>
<td>1.4</td>
<td>1.7e-01</td>
<td>1.1</td>
<td>0.24</td>
<td>4.8</td>
<td>5.5e-05</td>
<td>1.0</td>
<td>0.26</td>
<td>4.0</td>
<td>5.0e-04</td>
<td>0.83</td>
<td></td>
</tr>
</tbody>
</table>

Color coding for acquisition number by HOA: 30 m 40 m 50 m 60 m
TABLE III
RESULTS FOR THE TLM BIOMASS MODEL (TBM) WITH FIXED EXPONENTS AND THE SCALING MODEL (SM). FOR EACH SLOPE PARAMETER, THE ESTIMATED STANDARD DEVIATIONS $\sigma$, $t$-STATISTICS, AND $p$-VALUES ARE SHOWN. THE COEFFICIENTS OF DETERMINATION $R^2$ ARE ALSO SHOWN.

(a) TBM with fixed exponents: 
\[
\hat{AGB_{Re}} = K' \Delta h^{1.25} \eta_2^{0.64},
\]
\[
\hat{AGB_{Kr}} = K' \Delta h^{1.25} \eta_1^{0.16}.
\]

<table>
<thead>
<tr>
<th>Nr</th>
<th>Site</th>
<th>$K'$</th>
<th>$\sigma$</th>
<th>$t$</th>
<th>$p$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Remningstorp</td>
<td>7.6</td>
<td>0.20</td>
<td>37.2</td>
<td>2.7e-27</td>
<td>0.78</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>7.5</td>
<td>0.21</td>
<td>35.5</td>
<td>3.4e-24</td>
<td>0.80</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>8.0</td>
<td>0.23</td>
<td>34.2</td>
<td>9.1e-24</td>
<td>0.79</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>10.2</td>
<td>0.55</td>
<td>18.5</td>
<td>3.0e-17</td>
<td>0.22</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>8.8</td>
<td>0.32</td>
<td>27.8</td>
<td>5.9e-22</td>
<td>0.64</td>
</tr>
</tbody>
</table>

(b) SM: 
\[
\hat{AGB} = D h_{gc}.
\]

<table>
<thead>
<tr>
<th>Nr</th>
<th>Site</th>
<th>$D$</th>
<th>$\sigma$</th>
<th>$t$</th>
<th>$p$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Remningstorp</td>
<td>9.0</td>
<td>0.40</td>
<td>22.2</td>
<td>1.4e-20</td>
<td>0.40</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>9.1</td>
<td>0.45</td>
<td>20.2</td>
<td>7.8e-18</td>
<td>0.41</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>9.4</td>
<td>0.44</td>
<td>21.2</td>
<td>2.3e-18</td>
<td>0.46</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>8.3</td>
<td>0.53</td>
<td>15.6</td>
<td>2.4e-15</td>
<td>-0.07</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>8.4</td>
<td>0.49</td>
<td>17.4</td>
<td>1.5e-16</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Color coding for acquisition number by HOA: 30 m | 40 m | 50 m | 60 m
Fig. 5. Residual scatter plots are shown. Acquisition year and HOA intervals are shown for each subplot. Horizontal error bars show the 15% uncertainty of the reference biomass estimates. Vertical error bars represent the residual RMSE values presented on the diagonal of Tables IV–VI. Note: several points may overlap. In Remningstorp, there are 88 points in 2011, 58 points in 2012, and 63 points in 2013. In Krycklan, there are 116 points in 2011 and 87 points in both 2012 and 2013.

(a) TBM: \( \hat{\text{AGB}} = K \Delta h^{0.8} \)  
(b) TBM with fixed exponents: \( \hat{\text{AGB}}_{Re} = K' \Delta h^{1.25} \eta^{2.64}, \) \( \hat{\text{AGB}}_{Kr} = K' \Delta h^{1.25} \eta^{1.16} \)  
(c) SM: \( \hat{\text{AGB}} = D h_{gc} \)
### TABLE IV

Residual and prediction RMSE values (in percent of the average biomass, which can be found in Table I) for the TBM. Residual RMSE values are marked in boldface characters and shown on the diagonal. Off-diagonal prediction RMSE values are shown.

\[
\hat{\text{AGB}} = K \Delta h^\alpha \eta^\beta
\]

\(K, \alpha, \text{and } \beta \text{ as in Table II)}

<table>
<thead>
<tr>
<th>Nr</th>
<th>Site</th>
<th>Training data</th>
<th>Validation data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Remningstorp</td>
<td>Krycklan</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>16 15 16 39 25 17 14 17</td>
<td>46 43 45 41 48 47 46 43 50 44</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>16 15 16 41 26 17 14 16</td>
<td>48 45 48 43 50 49 48 45 52 46</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>18 17 15 38 25 23 21 20</td>
<td>41 37 40 35 41 40 39 37 45 39</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>19 33 37 73 77 80 29 29 31 30</td>
<td>39 38 37 33 33 30</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>35 33 36 33 43 42 41 36 39 34</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>15 12 17</td>
<td>48 45 48 44 51 50 49 46 52 46</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>16 12 16</td>
<td>53 50 53 49 55 55 54 51 56 51</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>17 14 15</td>
<td>51 47 50 45 51 51 49 47 55 48</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>32 33 29 24 23 33 34 35 15</td>
<td>16 15 15</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>31 32 29 23 21 31 32 33</td>
<td>15 16 15 14 20 19 18 17 17 16</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>32 33 30 24 23 34 34 35</td>
<td>15 16 15 15 19 18 17 16 16 16</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>32 33 31 23 22 32 33 35</td>
<td>15 16 15 14 19 19 18 16 16 15</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>53 55 52 26 35 52 55 56</td>
<td>23 23 21 20 16 16 16</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>50 51 49 25 33 49 51 52</td>
<td>22 22 20 19 16 16 16</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>46 47 45 24 30 47 49 50</td>
<td>19 20 18 18 16 16 16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>35 36 34 23 24 35 36 37</td>
<td>16 17 16 15 18 17 17 16 15 16</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>36 37 33 26 27 38 38 39</td>
<td>17 19 16 18 18 18 17 17 14 17</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>31 32 30 23 22 31 32 33</td>
<td>16 16 15 14 19 18 18 16 16 15</td>
</tr>
</tbody>
</table>

Main color coding by RMSE value: 20% 40% 60% 80%

Color coding for the acquisition number by HOA: 30 m 40 m 50 m 60 m
RESIDUAL AND PREDICTION RMSE VALUES (IN PERCENT OF THE AVERAGE BIOMASS, WHICH CAN BE FOUND IN TABLE I) FOR THE TBM WITH FIXED EXPONENTS. RESIDUAL RMSE VALUES ARE MARKED IN BOLDFACE CHARACTERS AND SHOWN ON THE DIAGONAL. OFF-DIAGONAL, PREDICTION RMSE VALUES ARE SHOWN.

TBM with fixed exponents: \( \hat{AGB}_{Re} = K' \Delta h^{1.25} \eta_0^{3.64} \), \( \hat{AGB}_{Kr} = K' \Delta h^{1.25} \eta_0^{3.16} \) (\( K' \) as in Table III(a)).

<table>
<thead>
<tr>
<th>Nr</th>
<th>Site</th>
<th>Remningstorp</th>
<th>Krycklan</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>Remningstorp</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>13</td>
<td>Krycklan</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>19</td>
<td>19</td>
</tr>
</tbody>
</table>

Main color coding by RMSE value: 20% | 40% | 60% | 80%
Color coding for the acquisition number by HOA: 30 m | 40 m | 50 m | 60 m
TABLE VI
RESIDUAL AND PREDICTION RMSE VALUES (IN PERCENT OF THE AVERAGE BIOMASS, WHICH CAN BE FOUND IN TABLE I) FOR THE SM.
RESIDUAL RMSE VALUES ARE MARKED IN BOLDFACE CHARACTERS AND SHOWN ON THE DIAGONAL. OFF-DIAGONAL, PREDICTION
RMSE VALUES ARE SHOWN.

![Table VI](image)

SM: \( \hat{\text{AGB}} = Dh_{gc} \) (\( D' \) as in Table III(b))

Main color coding by RMSE value:
- 20%
- 40%
- 60%
- 80%

Color coding for the acquisition number by HOA:
- 30 m
- 40 m
- 50 m
- 60 m
Fig. 6. Sensitivity of the TBM to small changes of parameters $K$, $\alpha$, and $\beta$ around a default setup with $\alpha = 1.25$, $\beta = 2.64$ for Remningstorp, $\beta = 1.16$ for Krycklan, and $K = 7.42$ (being the mean of all $K'$-values in Table III(a)). The vertical lines show the default values and the green, horizontal lines mark the 20% error level.

(a) Sensitivity to $K$
(b) Sensitivity to $\alpha$
(c) Sensitivity to $\beta$

Color coding by HOA: 30 m 40 m 50 m 60 m
Fig. 7. Mapping results for (a)–(c) Remningstorp and (d)–(f) Krycklan. The lidar biomass maps are compared to biomass maps obtained using the TLM biomass model (TBM) and the scaling model (SM) with single TanDEM-X image pairs from the corresponding dates. Model parameters estimated using the corresponding plot/stand-level estimates data are used. Note, that growth has not been modeled in the lidar maps.