

Article

## Model-Based Biomass Estimation of a Hemi-Boreal Forest from Multitemporal TanDEM-X Acquisitions

Jan I.H. Askne <sup>1,\*</sup>, Johan E.S. Fransson <sup>2</sup>, Maurizio Santoro <sup>3</sup>, Maciej J. Soja <sup>1</sup>  
and Lars M.H. Ulander <sup>1,4</sup>

<sup>1</sup> Department of Earth and Space Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden; E-Mails: maciej.soja@chalmers.se (M.J.S.); lars.ulander@foi.se (L.M.H.U.)

<sup>2</sup> Department of Forest Resource Management, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden; E-Mail: johan.fransson@slu.se

<sup>3</sup> Gamma Remote Sensing, Worbstrasse 225, CH-3073 Gümligen, Switzerland; E-mail: santoro@gamma-rs.ch

<sup>4</sup> Department of Radar Systems, Swedish Defence Research Agency (FOI), SE-581 11 Linköping, Sweden.

\* Author to whom correspondence should be addressed; E-Mail: jan.askne@chalmers.se; Tel.: +46-703-495-795; Fax: +46-31-772-1884

Received: 24 June 2013; in revised form: 22 October 2013 / Accepted: 23 October 2013 /

Published: 29 October 2013

---

**Abstract:** Above-ground forest biomass is a significant variable in the terrestrial carbon budget, but is still estimated with relatively large uncertainty. Remote sensing methods can improve the characterization of the spatial distribution and estimation accuracy of biomass; in this respect, it is important to examine the potential offered by new sensors. To assess the contribution of the TanDEM-X mission, eighteen interferometric Synthetic Aperture Radar (SAR) image pairs acquired over the hemi-boreal test site of Remningstorp in Sweden were investigated. Three models were used for interpretation of TanDEM-X signatures and above-ground biomass retrieval: Interferometric Water Cloud Model (IWCM), Random Volume over Ground (RVoG) model, and a simple model based on penetration depth (PD). All use an allometric expression to relate above-ground biomass to forest height measured by TanDEM-X. The retrieval was assessed on 201 forest stands with a minimum size of 1 ha, and ranging from 6 to 267 Mg/ha (mean biomass of 105 Mg/ha) equally divided into a model training dataset and a validation test dataset. Biomass retrieved using the IWCM resulted in a Root Mean Square Error (RMSE) between 17%

and 33%, depending on acquisition date and image acquisition geometry (angle of incidence, interferometric baseline, and orbit type). The RMSE in the case of the RVoG and the PD models were slightly higher. A multitemporal estimate of the above-ground biomass using all eighteen acquisitions resulted in an RMSE of 16% with  $R^2 = 0.93$ . These results prove the capability of TanDEM-X interferometric data to estimate forest aboveground biomass in the boreal zone.

**Keywords:** TanDEM-X; InSAR; forestry; boreal; biomass estimation; model-based; allometry

---

## 1. Introduction

Forest above-ground dry biomass (AGB, herewith simply referred to as biomass) is an important variable for the global carbon budget, not only due to the uptake of carbon dioxide in the process of photosynthesis, but also because forests store huge amounts of carbon, which are eventually released into the atmosphere following a disturbance [1]. Accurate and timely mapping of forest AGB is therefore crucial to support carbon cycle modeling. Traditional methods based on forest inventories and aerial photography, and more recently, LiDAR campaigns, give accurate estimates of AGB; however, such methods are expensive and become inefficient whenever frequent and large-scale mapping is needed. Therefore, there is a need for development of alternative methods for frequent and large-scale biomass mapping [2].

One of the more promising techniques for above-ground dry biomass mapping is Synthetic Aperture Radar (SAR), *cf.* [3]. Being an active sensor, radar is independent of weather and external illumination. Spaceborne SAR missions currently in operation are characterized by an image resolution on the order of meters. In addition, interferometric SAR, InSAR, offers the possibility to exploit two further observables besides the radar backscatter, namely the coherence and the interferometric phase. These are affected by the forest structure and, thus, are related to forest variables such as tree height, and stem volume or AGB. In a single-pass acquisition scenario, the association between InSAR observables and forest variables is expected to be maximized because temporal decorrelation can be assumed to be negligible. Experimental evidence on the suitability of single-pass InSAR to estimate forest variables at X-band (wavelength of approximately 3 cm) was provided by data acquired by airborne sensors [4–6], and during the Shuttle Radar Topography Mission (SRTM) [7].

In June 2010, the TanDEM-X (TerraSAR-X add-on for Digital Elevation Measurement) satellite was launched. Together with the almost identical twin-satellite TerraSAR-X (launched in June 2007), the first satellite-based single-pass SAR interferometer was formed. In the bistatic mode of the TanDEM-X mission (consisting of the TanDEM-X and TerraSAR-X satellites), only one satellite is used for transmission while both satellites are used for reception. For simplicity, we will refer to this mission as the TDM mission. In TDM data, temporal decorrelation is limited to a minimum because of the small along-track baseline between the sensors. The primary objective of TDM is to obtain a global Digital Elevation Model (DEM) with an absolute height accuracy better than 10 m and an equatorial spatial resolution of 12 m [8]. Because of the limited penetration of microwaves into the canopy, X-band interferograms over forests are characterized by an elevation offset which is dependent on

forest canopy height and density [9]. This offset suggests exploiting TDM imagery to estimate tree height although a reference for the ground elevation is needed. Since X-band microwaves do not significantly penetrate the closed canopy of a dense forest, a Digital Terrain Model (DTM) for the ground surface needs to be provided by some other, independent method, for example P-band SAR [10], or LiDAR [6,7]. Besides forest height estimation, retrieval of above-ground dry biomass was also investigated in some studies. In [10], a Root Mean Square Error (RMSE) of 46.1 Mg/ha (biomass range up to  $\approx 360$  Mg/ha) was obtained for biomass in a tropical forest using airborne SAR in X- and P-band, and in [7] RMSE = 19% was obtained using SRTM in X-band.

The objective of this study was to develop and assess estimation methods based on models linking X-band InSAR observations to forest biomass. For this, single polarized (VV), bistatic interferometric TanDEM-X data acquired between June 2011 and August 2012 over Remningstorp, a hemi-boreal test site situated in southern Sweden, were used. Three InSAR models were employed and evaluated: Interferometric Water Cloud Model (IWCM) [11–14], Random Volume over Ground model (RVoG) [15–17] and a simple model based on the penetration depth (PD) of X-band microwaves. As reference, biomass estimates derived from LiDAR scanning data acquired during the BioSAR 2010 campaign [18,19], performed within the BIOMASS phase-A study [20] were used. By means of 201 forest stands equally divided into a training and a validation dataset, properties of the model parameters were determined and biomass retrieval accuracy for the different models was quantified.

## 2. Test Site and Datasets

Remningstorp (58°30'N, 13°40'E) is an estate with 1,200 ha productive forest area in the hemi-boreal zone, which is the transition between the boreal and the temperate zone [21]. Forest species consist primarily of Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and birch (*Betula* spp.). The test site is fairly flat with elevations ranging from 120 m to 145 m above sea level.

### 2.1. Field Observations

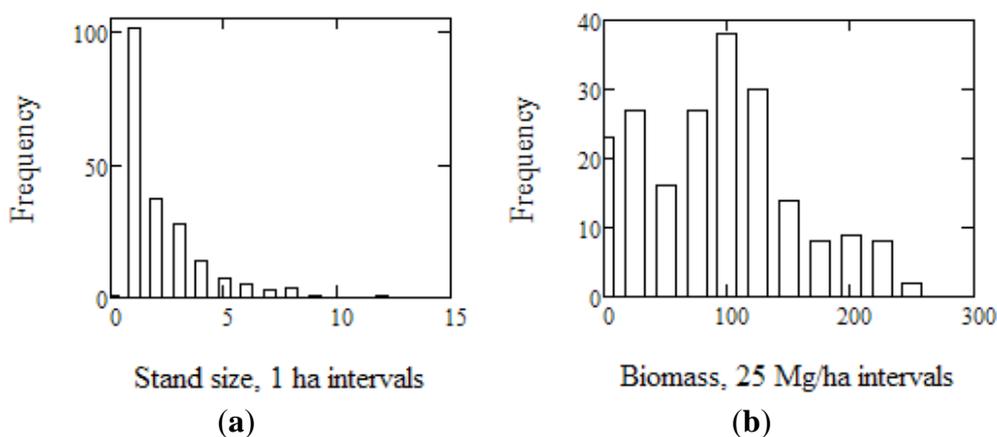
Field data used for this study were collected in 2010 [19] and consisted of 212 field plots with 10 m radius allocated over the estate. The survey assessed the stem volume, tree height, diameter at breast height (*i.e.*, 1.3 m above ground level), stem density, tree species composition based on proportion of total stem volume, and above-ground dry biomass, including stem, bark, branches and needles, but excluding stump and roots. Biomass was expressed in Mg of dry mass per hectare. The field survey and the biomass estimation for the field plots was carried out according to the Heureka forestry decision support system [22] with functions described in [23]. In addition, seven 80 m  $\times$  80 m field plots were inventoried *in situ* through single tree measurements (including all trees with diameter at breast height, DBH > 0.05 m). The measurements made on tree level included GPS position, DBH, species and height. Biomass and stem volume (including the stem above stump, and bark, and expressed in m<sup>3</sup>/ha) were estimated using functions developed in [24] and [25], respectively.

The biomass for the 212 field plots was then related to LiDAR metrics by least-squares regression in line with similar studies, *e.g.*, [26]. The LiDAR data were collected with a density of 69 returns/m<sup>2</sup> on average. The LiDAR metrics selected to establish the relationship with biomass were chosen based on studies of correlations and residual plots and included for example height percentiles of LiDAR

returns, vegetation densities from proportions of LiDAR returns, and tree species stratification information. The final model was tested for overfitting, and the coefficient of determination,  $R^2$ , between the biomass estimated for the 212 field plots and the biomass derived from the corresponding LiDAR metrics, was estimated to 0.81. Spatially explicit estimates of biomass were then derived from the function linking the LiDAR metrics to biomass. The LiDAR-based biomass was obtained for the entire forest estate and represented the reference biomass dataset for this study. The accuracy of the LiDAR-based estimates of biomass was determined by comparing it to the biomass for the seven  $80\text{ m} \times 80\text{ m}$  field plots which gave a RMSE of 12.7% [19].

The Remningstorp estate was divided into 665 delineated polygons, of which 562 consisted of forest stands, *i.e.*, areas of homogeneous tree cover, species composition and canopy structure. The remaining 103 polygons consisted of open fields, pastures, private lots, water, *etc.* The 562 forest stands, of which 201 stands were at least 1 ha large, were characterized by full LiDAR coverage and did not experience major forest cover changes between the LiDAR and the TDM acquisitions. The digital map with the forest stand boundaries was eroded with a 10-m buffer zone around the boundary of each stand to reduce border effects on the evaluation of the TDM data. The biomass of the 201 stands was between 6 and 267 Mg/ha with a mean of 105 Mg/ha. Figure 1 illustrates the distribution of stand sizes and biomass for the 201 forest stands. More than 90% of the forest stands were smaller than 5 ha (Figure 1a). The biomass presented an almost uniform distribution up to 150 Mg/ha; several stands were characterized by biomass above 200 Mg/ha (Figure 1b).

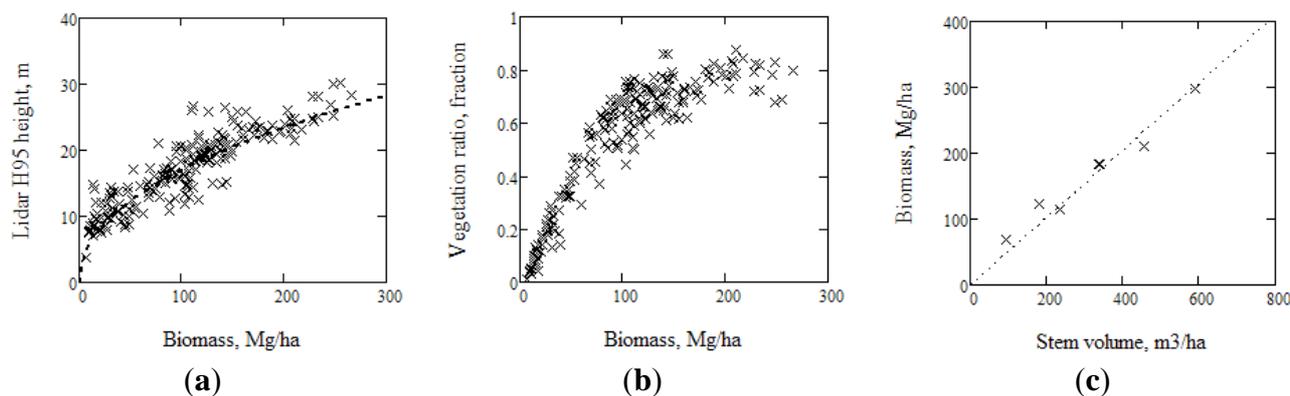
**Figure 1.** Distribution of forest stand size (a) and biomass (b) for the 201 forest stands  $\geq 1$  ha at the test site in Remningstorp.



Some of the LiDAR metrics and the LiDAR-based biomass for the 201 forest stands were used to support the interpretation of the TDM interferometric signatures and to support the modeling phase relating the interferometric data to biomass (see Section 3). In Figure 2a, stand-level averages of the 95th percentile of LiDAR return values above a height threshold of 1.0 m or 10% of the maximum height (H95) have been plotted against the LiDAR-based biomass to assess the validity of the allometric function to be then used in the modeling phase, see Equation (3). Vegetation ratio derived from the LiDAR data [19], *i.e.*, the ratio of LiDAR return values above a height threshold of 5.0 m and the total number of returns, provide information on canopy closure and can be considered as proxy for a similar parameter used in modeling, namely the area-fill factor (see Section 3). To understand the

relationship between canopy closure and biomass, Figure 2b shows the relationship between vegetation ratio and the LiDAR-based biomass for the 201 forest stands. Figure 2a,b shows strong correlations between the illustrated forest variables, indicating the suitability of empirical relationships to be implemented during the modeling phase of the interferometric signatures to express these solely as a function of forest biomass.

**Figure 2.** (a) Stand-wise values of LiDAR heights (H95) *versus* LiDAR-based biomass and an allometric equation relating basal-area-weighted mean height to biomass, see Equation (3); (b) LiDAR vegetation ratio *versus* LiDAR-based biomass, and (c) biomass *versus* stem volume for the seven 80 m × 80 m plots with single tree measurements.



The *in situ* information on biomass and stem volume from the seven 80 m × 80 m forest inventory plots was used to derive a linear equation linking the biomass,  $B$ , and the stem volume,  $V$ . In Equation (1), BEF represents a biomass expansion factor (expanding to include branches and needles as well as stem and bark) which was estimated to be  $0.512 \text{ Mg/m}^3$  with  $R^2 = 0.97$ , *cf.* Figure 2c,

$$B = BEF \cdot V \quad (1)$$

Although the BEF was estimated using a small dataset of samples and the relationship between biomass and stem volume is in theory dependent on tree species, age and local conditions, *cf.* [27] and [28], it was assumed that a single and approximate value was sufficient for the purpose of this investigation.

Furthermore, an allometric relationship between basal-area-weighted mean height,  $h$ , and stem volume,  $V$ , was considered

$$h(V) = (2.44 V)^{0.46} \quad (2)$$

Equation (2) was derived using measurements from test sites in Sweden and Finland and verified by means of 4,188 randomly chosen National Forest Inventory field plots located in different regions of Sweden [11,13]. Combining Equations (1) and (2), an allometric relationship between biomass and height was obtained:

$$B(h) = 0.21 h^{2.17} \quad (3)$$

illustrated in Figure 2a. Although stem volume and biomass also depend on other forest variables, for example tree species and tree number density, it was shown in [29] using a plant structure model that biomass can generally be modeled from height using the same functional form as Equation (3), where the exponent is determined by a scaling parameter related to species and thinning practices.

The allometric relationship, Equation (3), is not entirely optimal for the Remningstorp data, *cf.* Figure 2a, but was still used due to its general nature. The aim of the allometric relationship is to decrease the number of forest variables in the models relating these to TDM observables and have them expressed purely as functions of biomass.

## 2.2. TanDEM-X SAR Dataset

A large number of TanDEM-X InSAR acquisitions were available for the Remningstorp test site (Table 1). In Table 1, the stand-level mean values of the Height of Ambiguity (HOA) and the along-track baseline (ATB) are shown. HOA is the height interval corresponding to a phase difference of  $2\pi$ ; it is a measure for the sensitivity of the InSAR phase to elevation and is inversely proportional to the perpendicular component of the across-track baseline. ATB is the distance between the satellites along track. Since each pixel in the monostatic active and the bistatic passive image is focused in azimuth at its Zero Doppler Time [30] the two images can be looked as monostatic and ATB provides a measure for temporal decorrelation [31].

**Table 1.** TanDEM-X InSAR acquisitions and weather conditions at the time of image acquisition (HOA = Height of Ambiguity, ATB = Along-Track Baseline). The sign of HOA depends on the satellite positions, but has no relevance for the results in this paper.

Date	HOA (m)	ATB (m)	Incidence Angle	Temperature (°C)	Wind Speed (m/s)	Precipitation (mm)
2011-06-04	49	110	41°	24	2	0
2011-11-23	-185	4	34°	5	3	1.2
2011-12-26	-178	64	34°	6	5	0
2012-01-17	-172	-1	34°	-2	1	0
2012-01-28	-182	7	34°	-3	2	0
2012-02-01	80	267	41°	-4	1	0.2
2012-02-08	-179	29	34°	-3	1	0
2012-02-12	-79	-244	41°	-3	1	0.2
2012-02-19	-186	-8	34°	1	6	3.4
2012-02-23	79	232	41°	3	3	0
2012-03-01	-186	-11	34°	5	4	0
2012-03-12	-187	-11	34°	4	2	0
2012-03-23	-183	14	34°	-1	0	0
2012-05-28	349	262	34°	15	2	1.3
2012-07-22	339	262	34°	14	3	4.1
2012-08-02	315	233	34°	15	3	0.6
2012-08-13	358	229	34°	14	0	0
2012-08-24	301	208	34°	13	3	0.2

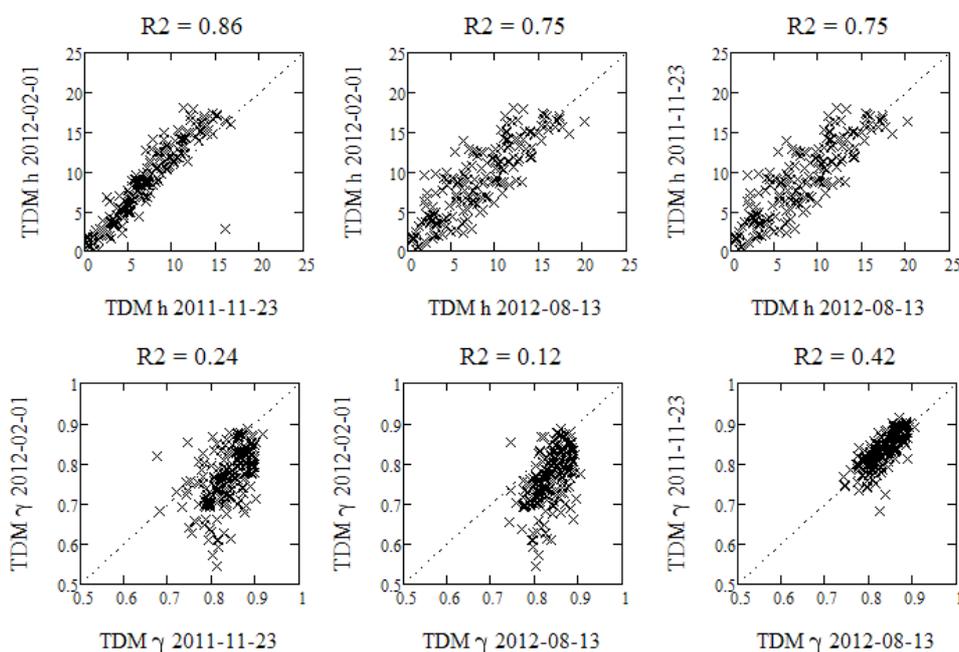
TDM images were provided by German Aerospace Center (DLR) in a co-registered single-look complex (SLC) format with common spectral filtering applied during pre-processing. Interferometric processing of TanDEM-X data were done with a Matlab-based algorithm [32] developed specially for interferometric processing of TanDEM-X data and based on [30]. The first step of interferometric processing consists of interferogram flattening for curved Earth and surface elevation. For this, InSAR

phase was simulated using an airborne LiDAR DTM with a spatial resolution of 2 m and mean height error lower than 0.5 m. The DTM was acquired by the Swedish National Land Survey (Lantmäteriet) within an ongoing nationwide LiDAR scanning campaign. The DTM was interpolated to range-azimuth coordinates (radar geometry) using satellite state vectors and look geometry information provided in the metadata of the images. For each stand, a complex average of all pixels within the stand was computed, and the corresponding phase was converted to elevation (in the following referred to as TDM height) by a multiplication with  $HOA/2\pi$ . Absolute height calibration was performed through the subtraction of the mean TDM height for non-forested stands (at least 0.5 ha in size and scattered over the entire estate) from the TDM heights. For a few stands, TDM heights were slightly below the mean TDM height for non-forest areas. In such cases, the absolute TDM height was set equal to zero. This applied also for one stand (2012-05-28) with an offset of  $-1.5$  m, probably due to the large HOA, which was 349 m.

Among the stands, one stand (ID = 189, 1.42 ha, 25.1 Mg/ha) presented TDM heights from  $-11.7$  m to 36.0 m. This stand also had a low coherence. High-resolution LiDAR data shows, that the stand is highly irregular and consists of several disjoint parts of high trees alternating with low vegetation or open ground. Such a stand is easily detected by comparing TDM observations in a multitemporal approach. In the following the TDM height of this stand has been put to zero, *i.e.*, the biomass will be assumed zero.

The temporal consistency of TDM heights was high (Figure 3, upper row), in particular between acquisitions with similar HOA ( $R^2 = 0.99$  for 2012-02-01 and 2012-02-23, for example). When acquisitions with different HOA were compared,  $R^2$  tended to decrease with increasing HOA for one or both of the acquisitions. Almost the opposite applied for coherence but then with much lower  $R^2$ . The dynamic range of coherence was found to decrease with increasing HOA. A large HOA corresponds to a short perpendicular baseline and thus to low volume decorrelation.

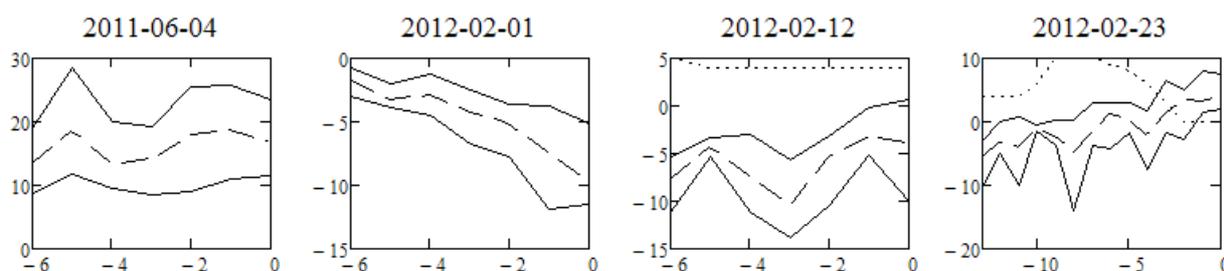
**Figure 3.** Temporal consistency visualized by means of scatterplots of stand-wise TDM height (TDM h, upper row) and TDM coherence (TDM  $\gamma$ , lower row). HOA = 80 m for 2012-02-01, HOA =  $-185$  m for 2011-11-23, and HOA = 358 m for 2012-08-13.



### 2.3. Meteorological Data

Table 1 lists temperature and wind speed measured within one hour of the satellite overpass at the closest official meteorological station, Hällum, situated 23 km from Remningstorp. Precipitation in Table 1 was recorded in Remningstorp for the date of acquisition. For the four image pairs with  $|\text{HOA}| \leq 80$  m showing the strongest agreement between TDM heights, the history of temperature and snow depth are further illustrated in Figure 4 in order to support the interpretation of the measurements. The history of the maximal, mean, and minimal temperatures has been plotted. For three of the acquisitions in Figure 4, temperature change for the six preceding days is shown. For the fourth acquisition (2012-02-23), data from two preceding weeks are shown to illustrate the more complex situation.

**Figure 4.** History of temperature for the Remningstorp test site prior to four TDM acquisitions with  $|\text{HOA}| \leq 80$  m. The units on the x- and y-axis are days and temperature ( $^{\circ}\text{C}$ ), respectively. The snow layer in cm is dotted for the acquisition dates 2012-02-12 and 2012-02-23.



## 3. Interferometric Forest Models

The models, used for interpretation of TDM observations and to explain their relation to forest height or biomass, were selected to be simple enough in order to make inversion possible. This means, that the forest properties can only be described by a few parameters. The Interferometric Water Cloud Model, the Random Volume over Ground model, as well as the simple model based on the penetration depth presented below fulfilled these requirements.

### 3.1. Interferometric Water Cloud Model

The Interferometric Water Cloud Model, IWCM, is a model for the complex coherence of a forest. The IWCM was introduced to explain the coherence of forest at C-band [11,12,14]. The model assumes that the medium, characterized by a certain forest height,  $h$ , and stem volume,  $V$ , can be described by a random vegetation layer like the Water Cloud Model [33], with uniform scatterer density but generalized to include gaps. The proportion of the area with vegetation relative the total area is denoted as the area-fill,  $\eta$ . Recently, and in agreement with IWCM, it was shown using observations of the spatial scattering spectrum of TanDEM-X data over tropical forests that forests cannot be modeled as a layered medium, but by a model with gaps or random scattering “clouds” [34].

The attenuation factor in the vegetation layer is described by  $\alpha$ , representing the mean attenuation for the idealized vegetation layer (either vegetation or no vegetation), and is a function of, for example,

the dielectric constant of the scatterers. According to [35], the forest backscatter,  $\sigma_{for}$ , can also be described by an empirical, exponential stem volume dependence characterized by  $\beta$ , as demonstrated by scatterometer measurements at X- and C-band. There are then two alternative expressions for the backscatter:

$$\sigma_{for}^0 = \eta[\sigma_{gr}^0 e^{-\alpha h} + \sigma_{veg}^0 (1 - e^{-\alpha h})] + (1 - \eta)\sigma_{gr}^0 = \sigma_{gr}^0 e^{-\beta V} + \sigma_{veg}^0 (1 - e^{-\beta V}) \quad (4)$$

In Equation (5),  $\sigma_{gr}^0$  represents the ground backscatter and  $\sigma_{veg}^0$  represents the vegetation layer backscatter.

The complex expression for the coherence of this random volume with gaps is then described by each of the independent scattering parts weighted by system and temporal decorrelation, the product of which will be represented by the coherence parameters  $\gamma_{gr}$  and  $\gamma_{veg}$ , for the ground and vegetation parts, respectively. For the vegetated part there is also volume decorrelation,  $\gamma_{vol}$ , related to the forest height. The complex coherence (assuming the phase of the ground surface has been compensated for) is then given by:

$$\begin{aligned} \tilde{\gamma} &= \frac{\{\eta[\gamma_{gr}\sigma_{gr}^0 e^{-\alpha h} + \gamma_{veg}\tilde{\gamma}_{vol}\sigma_{veg}^0 (1 - e^{-\alpha h})] + \gamma_{gr}(1 - \eta)\sigma_{gr}^0\}}{\sigma_{for}^0} \\ &= \frac{\gamma_{gr}\sigma_{gr}^0 e^{-\beta V} + \gamma_{veg}\tilde{\gamma}_{vol}\sigma_{veg}^0 (1 - e^{-\beta V})}{\sigma_{for}^0} \end{aligned} \quad (5)$$

For the two expressions in Equations (4) and (5) to agree, a requirement on the area-fill factor  $\eta$  is given by a relation between  $\alpha$  and  $\beta$  according to:

$$\eta(V) = \frac{1 - e^{-\beta V}}{1 - e^{-\alpha h(V)}} \quad (6)$$

with  $\eta \rightarrow 1$  when  $V \rightarrow \infty$ . According to Equation (6),  $\beta$  can be described by the area-fill  $\eta$ , the attenuation  $\alpha$ , and  $h(V)$ , *i.e.*, by the vegetation density and attenuation through the vegetation.  $h$  and  $V$  will, whenever needed, be related through the allometric relationship in Equation (2), *i.e.*,  $h(V)$  represents the basal-area-weighted mean forest height. Equation (4) shows that  $\beta$  characterizes the transition from dominant ground scattering to dominant vegetation scattering.

If the variation of scattering with height is only determined by the attenuation, the volume decorrelation is determined by, [11]:

$$\tilde{\gamma}_{vol} = \frac{\int_0^h e^{-\alpha(h-z')} \cdot e^{-jK(B_n)z'} dz'}{\int_0^h e^{-\alpha(h-z')} dz'} = \frac{\alpha}{\alpha - jK(B_n)} \frac{e^{-jK(B_n)h} - e^{-\alpha h}}{1 - e^{-\alpha h}} \quad (7)$$

For the TanDEM-X bistatic mode  $K(B_n) = 2\pi B_n / \lambda R \sin\theta = 2\pi / HOA$ , where  $B_n$  is the component of the baseline perpendicular to the line of sight,  $\lambda$  is the wavelength,  $R$  is the slant range distance, and  $\theta$  is the incidence angle.

Often an extinction coefficient,  $\kappa_{eff}$ , is used to define the attenuation through a homogeneous vegetation layer without gaps. From Equation (5),  $\kappa_{eff}$  is obtained by the definition  $\frac{2\kappa_{eff}h(V)}{\cos\theta_i} = \beta V$ . With  $h$  expressed by means of the allometric expression in Equation (2) and  $\beta$  expressed by Equation (6), it is obtained:

$$\kappa_{eff}(V) = -\frac{\cos\theta_i}{2h(V)} \ln [1 - \eta(V)(1 - e^{-\alpha h(V)})] \quad (8)$$

The expression for  $\kappa_{eff}$  in Equation (8) illustrates how the extinction coefficient is dependent on stem volume, forest height, area-fill, as well as temperature, humidity, *etc.*, through the attenuation coefficient  $\alpha$ . For dense vegetation layers,  $\eta(V) = 1$  and  $\kappa_{eff}(V)$  tends to  $\alpha \cos\theta/2$ .

Equation (4) can be rewritten as

$$\tilde{\gamma} = \frac{\gamma_{veg}\tilde{\gamma}_{vol} + \gamma_{gr}m}{1 + m} \quad (9)$$

where

$$m = \frac{\sigma_{gr}^0 e^{-\beta V}}{\sigma_{veg}^0 (1 - e^{-\beta V})} \quad (10)$$

is the ratio describing the relative amount of ground scattering compared to volume scattering (ground-to-volume ratio).

The TDM coherence,  $\gamma$ , and the estimated height,  $z_{est}$  are defined as

$$\gamma = |\tilde{\gamma}| \quad z_{est} = -\frac{\arg[\tilde{\gamma}]}{2\pi} HOA \quad (11)$$

where  $\arg$  stands for argument of the complex-valued term within brackets.

The IWCM has been used for C-band data in order to derive stem volume from coherence, and it was found to be suitable for the retrieval of stem volume at several test sites in Sweden and Finland. For the one-day repeat-pass interval of the European Remote Sensing ERS-1/2 mission, the coherence for stable winter conditions was found to be useful for stem volume estimation [13,14,36,37], while the interferometric phase height was found unstable [38]. As will be illustrated below, it is instead primarily the interferometric phase height which has the highest sensitivity to biomass for TanDEM-X.

### 3.2. Random Volume over Ground Model

The Random Volume over Ground, RVoG, was introduced for studies of polarimetric SAR interferometry, PolInSAR, [17]. The model can be obtained by excluding gaps in the analysis of IWCM, which means one less unknown parameter. This means that  $\eta \equiv 1$ ,  $\beta V = \alpha h$ , and  $\kappa_{eff} = \alpha \cos\theta/2$  is the extinction coefficient. The ground-to-volume ratio  $m$  then changes to

$$m = \frac{\sigma_{gr}^0 e^{-\alpha h}}{\sigma_{veg}^0 (1 - e^{-\alpha h})} \quad (11)$$

and the expression in Equation (9) for the complex coherence is used for cases when the temporal decorrelation can be neglected. RVoG has shown to be useful in PolInSAR height estimation without the need of training stands as long as the extinction coefficient can be assumed polarization-independent and a polarization combination with no ground contribution can be found [5,16]. RVoG was also used for the single polarized case, e.g., [5,6]. When  $m$  can be assumed negligible and a lidar DTM is available, the tree height and the extinction coefficient can be estimated without training stands. Biomass can then be estimated by means of an allometric relationship [39,40]. With training stands, the RVoG model parameters and model properties can be studied in detail.

### 3.3. Penetration Depth Model

If a medium is dense and attenuation is such that ground scattering is negligible, *i.e.*,  $\exp[-\alpha h] \ll 1$ , then  $m \approx 0$  and only volume decorrelation remains, if  $\gamma_{veg} = 1$ , see Equations (7), (9), and (12):

$$\gamma_{vol} = \frac{\alpha}{\alpha - jK(B_n)} \frac{e^{-jK(B_n)h} - e^{-\alpha h}}{1 - e^{-\alpha h}} \approx e^{-jK(B_n)[h - \frac{1}{\alpha}]} \tag{12}$$

If  $K(B_n)/\alpha \ll 1$  then the estimated interferometric phase height is approximated as  $h - 1/\alpha$ , while if  $K(B_n)/\alpha \approx 1$  a correction has to be introduced depending on  $\alpha$  and HOA, and  $h - 1/\alpha_{eff}$  represents the interferometric phase height. Setting  $h(V) - 1/\alpha_{eff}$  equal to the TDM height gives a simple model for the TDM height,  $H_{TDM}$ . This model is denoted as PD. Neglecting that the approximations made are not valid for small biomass, the biomass could be estimated from

$$B = 0.21(H_{TDM} + \alpha_{eff}^{-1})^{2.17} \tag{13}$$

### 3.4. Estimation of Model Parameters

The IWCM contains six unknown parameters ( $\sigma_{gr}^\circ$ ,  $\sigma_{veg}^\circ$ ,  $\gamma_{gr}$ ,  $\gamma_{veg}$ ,  $\alpha$ , and  $\beta$ ) that need to be determined (Table 2). The traditional estimation approach consists of a least-squares regression to reference data of the forest variable in the model and corresponding observations of backscatter and complex coherence. This is referred to as “training” the model. The models in Equations (4) and (5) assumed to be formulated as dependent on stem volume,  $V$ , have been transformed in this study to a dependence on biomass by means of Equation (1). The allometric expression in Equation (2),  $h(V)$ , and the BEF are assumed to be known *a priori*. In addition, there is a need for knowledge of the ground phase or, in an equivalent manner, of the TDM height for non-forest areas nearby the test site. All model parameters are estimated by fitting the IWCM to the sets of backscatter, coherence and InSAR height observations from the TanDEM-X dataset and the corresponding LiDAR-based values of biomass forming a training dataset. In addition, the estimates of vegetation ratio, which mimic the area-fill factor are used. It should be noted that the area-fill does not correspond exactly to the LiDAR measured vegetation ratio.  $1-\eta$  represents the fraction of gaps in the vegetation, and the gaps have to be larger for microwaves with longer wavelengths than for the LiDAR to propagate freely. The dielectric properties of the surrounding vegetation will also affect the wave propagation and transmission. Consequently, the area-fill can be expected to be higher than the LiDAR-based vegetation ratio and also vary with environmental conditions.

**Table 2.** Input and output of the model training phase for the three models.

Model	Input	Output
IWCM	Parameters of satellite look geometry, DTM, $h(B)$ , BEF, biomass, $\sigma_{TDM}^\circ$ , $\gamma_{TDM}$ and $h_{TDM}$ of training stands	$\sigma_{gr}^\circ$ , $\sigma_{veg}^\circ$ , $\gamma_{gr}$ , $\gamma_{veg}$ , $\alpha$ , $\beta$ ,
RVoG	Parameters of satellite look geometry, DTM, $h(B)$ , BEF, biomass, $\sigma_{TDM}^\circ$ , $\gamma_{TDM}$ and $h_{TDM}$ of training stands	$\sigma_{gr}^\circ$ , $\sigma_{veg}^\circ$ , $\gamma_{gr}$ , $\gamma_{veg}$ , $\alpha$ ,
PD	DTM, $h(B)$ , biomass and $h_{TDM}$ of training stands	$\alpha_{eff}$ ,

In the RVoG model, vegetation gaps are excluded, which reduces the number of unknowns to five since the  $\beta$ -parameter is eliminated (Table 2). The unknown model parameters were estimated in a similar manner as in the case of the IWCM.

In the PD model, only the  $\alpha_{eff}$ -parameter is unknown. It is noted that this model is a simple function of the TDM height. For estimating biomass, the dependence on the allometric relationship, which is indirect for the other models, is very clear in this case. The estimate of  $\alpha_{eff}$  is obtained by fitting the model to measurements of InSAR height for the training stands (Table 2).

### 3.5. Model Training and Inversion Procedure

The 201 forest stands with a size of at least 1 ha were divided in two datasets, one for training and one for validation. Then the datasets were interchanged and training and validation was repeated.

As a first-order approximation  $\sigma_{gr}^0$  and  $\sigma_{veg}^0$  in Equation (5) were estimated from the TDM backscatter measurements of the 20 stands with the smallest and largest biomass. Because of the large scatter of the measured backscatter with respect to biomass, it was assumed that a constant  $\beta = 0.007$  would return a realistic approximation of the two backscatter model parameters [41]. To correct for errors in the regression,  $\sigma_{gr}^0$  was manually adapted such that the model curve of Equation (4) would have been within the range of observations. The estimates of the coherence model parameters in Equation (5) were then obtained.  $\gamma_{gr}$  was determined as the mean of the TDM coherence for the ten stands with highest coherence and with biomass close to zero, while  $\gamma_{veg}$  was assumed equal to  $\gamma_{gr}$ .

In the case of IWCM, the values of the two remaining unknown parameters  $\alpha$  and  $\beta$  were estimated by means of least-squares fitting of the model to the observations in the training dataset using the Levenberg-Marquardt method together with the constraint that the area-fill (expressed by the LiDAR vegetation ratio) is  $<1$ . Since the TDM height was the most accurate SAR observation (Figure 3), the two unknown model parameters were estimated by minimizing the quadratic difference between the InSAR height predicted by the model ( $z_{est}$ ) and the corresponding observed TDM heights ( $H_{TDM,i}$ ):

$$\min \sum_i (z_{est}(B_i, \alpha, \beta) - H_{TDM,i})^2 \quad (15)$$

In Equation (15),  $B_i$  represents the LiDAR-based value of biomass of training stand  $i$ . A minimization similar to Equation (15) was also done for the coherence to determine a correction to  $\gamma_{veg}$ . The effect of the fine tuning of this parameters did not have any significant effect on the estimate of  $z_{est}$  and therefore on the estimation of the remaining model parameters. Once the model parameters  $\alpha$  and  $\beta$  were estimated, the model could be inverted to obtain estimates of biomass  $B_j$  for each forest stand  $j$  in the validation dataset by estimating the roots of the expression  $z_{est}(B_j, \alpha, \beta) - H_{TDM,j} = 0$ .

In the case of the RVoG model,  $m$  can sometimes be assumed negligible, but here  $m$  was included in the analysis since the contribution from ground is important for forest heights up to  $\approx 15$  m (depending on the attenuation observed in summer and in winter). The parameter  $\alpha$  was determined by fitting the model to TDM height by least-squares fitting. Biomass was then estimated in the same way as for the IWCM.

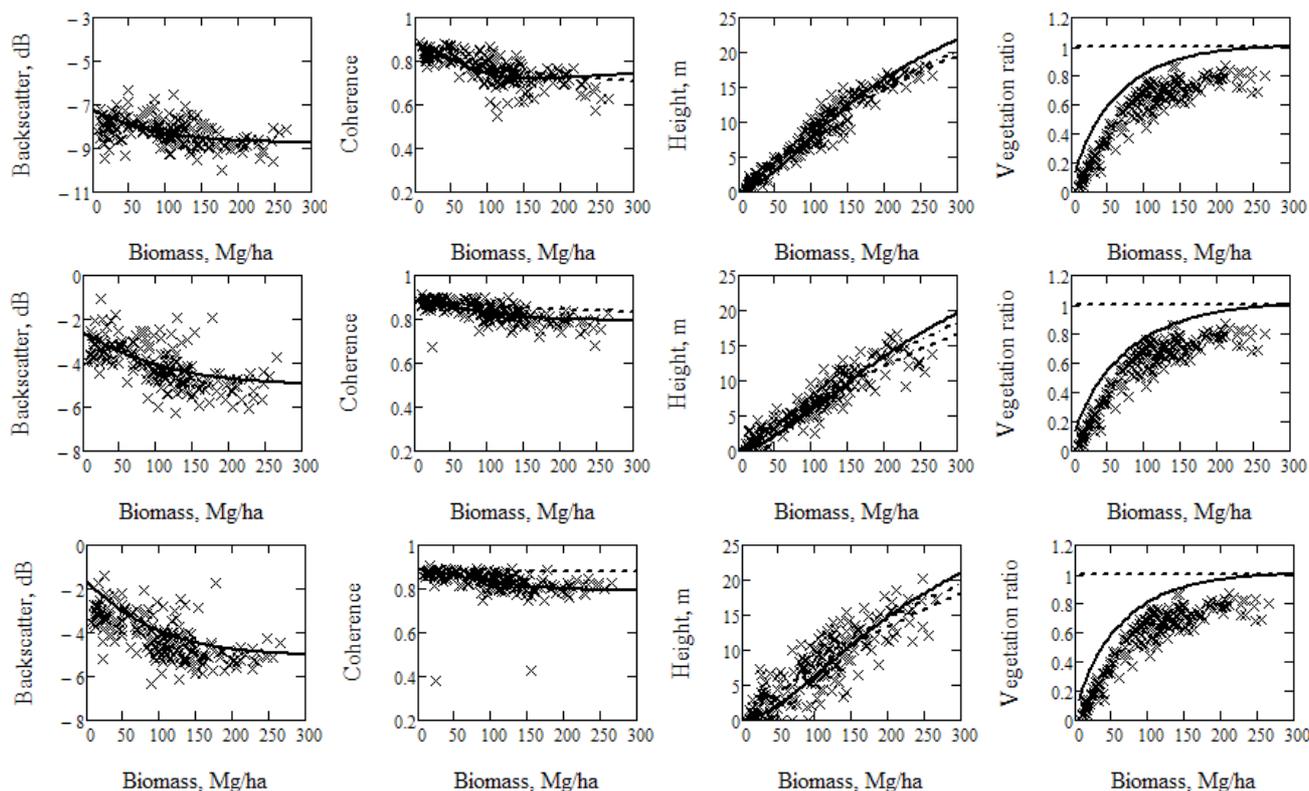
The simplest model, PD, is a model for the TDM height expressed by the allometric expression for the forest height and an attenuation coefficient for the penetration,  $h(B) - \alpha_{eff}^{-1}$ . From the training

stands an estimate for  $\alpha_{eff}$ , was obtained. The biomass for the validation stands was then estimated from Equation (14).

#### 4. Modeling and Biomass Retrieval Results

Results concerning the estimation of the model parameters are presented first (Section 4.1). Then, retrieval results are presented in the form of RMSE values of biomass estimation (Section 4.2). RMSE can be considered a measure of the usefulness of the different models as well as a measure of the biomass estimation accuracy.

**Figure 5.** Scatter plots and model results for three TDM acquisitions, in order from top to bottom 2012-02-01 (HOA = 80 m), 2011-11-23 (HOA = -185 m), and 2012-08-13 (HOA = 358 m). IWCM results for backscatter, coherence, TDM height and vegetation ratio (area-fill): solid line, RVoG results for coherence and TDM height: dashed line, and PD results for TDM height: dash-dotted line.



##### 4.1. Model Properties and Parameters Estimates

Figure 5 illustrates model-based backscatter, coherence, height and area-fill (vegetation ratio) as a function of biomass with respect to observations from the TanDEM-X dataset and corresponding LiDAR-based biomass. Three examples are shown representing a winter (2012-02-01), a fall (2011-11-23), and a summer (2012-08-13) acquisition as well as different HOA: 80 m, -185 m and 358 m, respectively. Backscatter and coherence presented weak sensitivity with respect to biomass. The backscatter was characterized by a dynamic range of 2–3 dB, with a clear decreasing trend for increasing biomass which can be connected with relatively rough and wet conditions of the soil.

Coherence was mostly above 0.7 and showed a decreasing dynamic range for increasing HOA (The two extreme coherence values in the bottom line are associated with stands crossed by a high voltage power line). Compared to backscatter and coherence, the InSAR heights showed the strongest sensitivity to biomass. The association strength between TDM heights and biomass however decreased for increasing HOA. The association between vegetation ratio and biomass was also strong. Saturation of the vegetation ratio slightly above 100 Mg/ha could be observed. The three models could reproduce the trend in the observations regardless of the observable and the acquisition date (Figure 5). Some discrepancies occurred in the range of the highest biomass values.

Estimates of the model parameters for each TDM acquisition are presented in Table 3 and illustrated in Figures 6, 7 and 8. Figure 6 illustrates the IWCM estimates of the two backscatter and the two coherence parameters for each acquisition. The ratio  $\sigma_{veg}^o/\sigma_{gr}^o$  indicates stronger sensitivity of the backscatter to biomass for shallow incidence angles (filled circles in Figure 6a). The backscatter ratio and  $\beta$  determine the biomass for which the backscatter from ground and vegetation are similar, and from this criterion as well as  $\exp[-\alpha h(V)] \approx 0.15$  it is found that the ground has an influence up to a biomass of 50–100 Mg/ha and forest heights of 12–17 m. The four cases with shallow incidence angle of  $41^\circ$  were also those with  $|HOA| \leq 80$  m. For these cases, the coherence was relatively noisy, in particular on 2011-06-04, whereas the coherence parameters  $\gamma_{veg}$  and  $\gamma_{gr}$  were almost equal, as expected for cases without temporal decorrelation (crosses and filled circles in Figure 6b). Images with  $|HOA| \geq 172$  m presented instead a slight difference with  $\gamma_{veg} < \gamma_{gr}$ . This difference could not be explained as an effect of ATB length or specific environmental conditions though.

**Table 3.** TDM acquisition date, model parameters  $\alpha$ ,  $\beta$ , and modeled TDM height in m at 150 Mg/ha (H150) for IWCM, and  $\alpha$  for RVoG and PD.

#	Date	$\alpha$ IWCM	$\beta$ IWCM	H150 IWCM	$\alpha$ RVoG	$\alpha_{eff}$ PD
1	2011-06-04	0.20	0.0093	15.0	0.14	0.17
2	2011-11-23	0.12	0.0056	9.4	0.09	0.10
3	2011-12-26	0.11	0.0053	8.9	0.08	0.10
4	2012-01-17	0.15	0.0070	11.2	0.11	0.12
5	2012-01-28	0.10	0.0049	8.5	0.07	0.09
6	2012-02-01	0.15	0.0070	12.0	0.11	0.12
7	2012-02-08	0.12	0.0056	9.2	0.09	0.10
8	2012-02-12	0.17	0.0080	13.1	0.12	0.14
9	2012-02-19	0.14	0.0068	10.9	0.10	0.11
10	2012-02-23	0.16	0.0078	12.8	0.12	0.13
11	2012-03-01	0.14	0.0066	10.4	0.10	0.11
12	2012-03-12	0.12	0.0058	9.6	0.09	0.10
13	2012-03-23	0.13	0.0061	9.8	0.09	0.10
14	2012-05-28	0.16	0.0076	11.8	0.12	0.12
15	2012-07-22	0.14	0.0068	10.8	0.10	0.11
16	2012-08-02	0.14	0.0068	10.9	0.10	0.11
17	2012-08-13	0.14	0.0066	10.3	0.10	0.11
18	2012-08-24	0.14	0.0065	10.6	0.10	0.11

**Figure 6.** (a) Estimates of  $\sigma_{veg}^o/\sigma_{gr}^o$  (cf. Equation (5)) ( $\times$  and  $\bullet$  for incidence angles of  $34^\circ$  and  $41^\circ$ , respectively) and (b) estimates of  $\gamma_{gr}$  ( $\times$ ) and  $\gamma_{veg}$  (cf. Equation (4)) ( $o$  and  $\bullet$  for  $|HOA| \leq 80$  m or above, respectively) versus TDM acquisition date (date order).

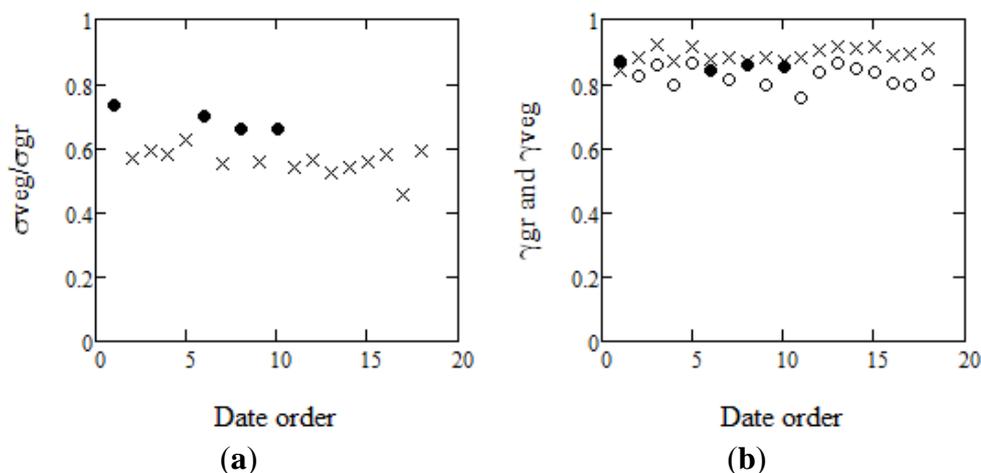


Figure 7 illustrates the range of modeled TDM heights with respect to biomass being delimited by the curves corresponding to the maximum (2011-06-04) and minimum (2012-01-28) model estimated TDM heights. For biomass of 150 Mg/ha Table 3 lists the corresponding modeled TDM height, which varied between 8.5 m and 15.0 m. Such variability must be compensated for by a model-based approach to retrieve biomass in order to correctly interpret the dependence of TDM height on biomass.

**Figure 7.** Range of model-based estimates of TDM heights versus biomass. The range is delimited by the curves (solid for IWCM, dashed for PD model) corresponding to the maximum and minimum of model estimated TDM heights versus biomass.

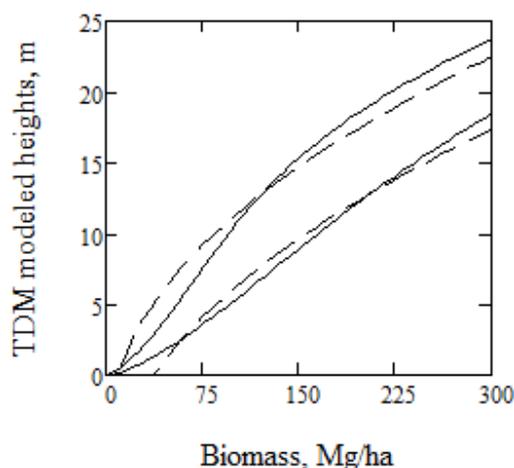
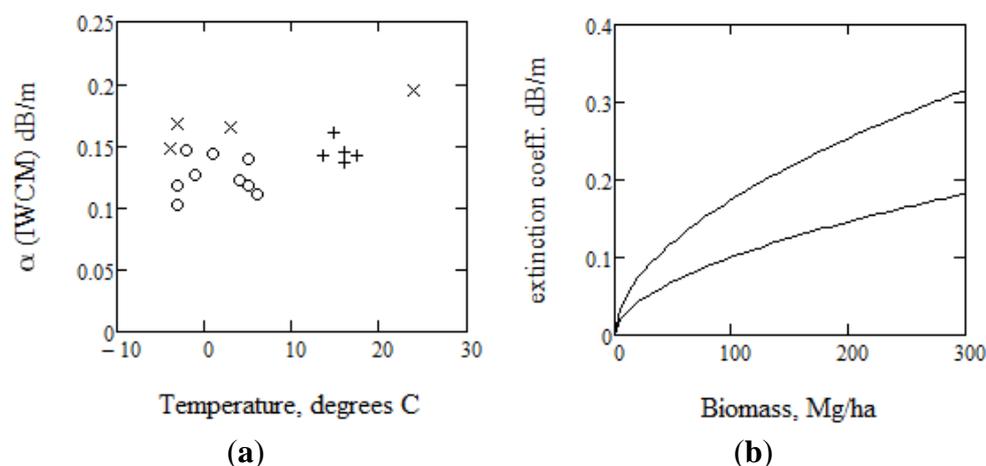


Figure 8 illustrates the estimates of the  $\alpha$  parameter in the case of the IWCM with respect to temperature and the corresponding range of the extinction coefficient  $\kappa_{eff}$  delimited by the two curves for the smallest and largest estimates of  $\kappa_{eff}$  using Equation (8). The attenuation in the vegetated fraction ( $\alpha$ ) did not present clear dependence on temperature (Figure 8), nor we could identify any dependence on HOA. This result is in contrast with the assumption that  $\alpha$  should be lower in case of sub-zero temperatures. However, since the few acquisitions characterized by frozen environmental conditions took place when temperature was close to  $0^\circ\text{C}$ , it is not possible to conclude that such an

assumption is incorrect. The increase of the extinction coefficient with biomass shown in Figure 8b also means an increase with area-fill (*i.e.*, LiDAR vegetation ratio). The range 0.1–0.3 dB/m confirm previous estimates derived in [5,6].

**Figure 8.** (a) Estimates of  $\alpha$  in the case of the IWCM with respect to temperature in °C (  $\times$  for  $|\text{HOA}| \leq 80$  m,  $\circ$  for  $|\text{HOA}| \approx 180$  m, and  $+$  for  $|\text{HOA}| \approx 330$  m) and (b) corresponding range of the extinction coefficient with respect to biomass as delimited by the curves corresponding to the maximum and minimum  $\alpha$  (acquisition dates 2011-06-04 and 2012-01-28, respectively).



#### 4.2. Biomass Estimation

The biomass of each stand in the validation dataset was estimated based on the trained models and compared with the LiDAR-based estimates of biomass. The biomass retrieval accuracy expressed in the form of the RMSE between the TDM biomass and the reference biomass is given in Table 4 for the different models. The estimates obtained with the IWCM presented slightly better accuracy compared to the retrieval based on the other models. Nonetheless, the difference between the models is relatively small, in particular between IWCM and PD.

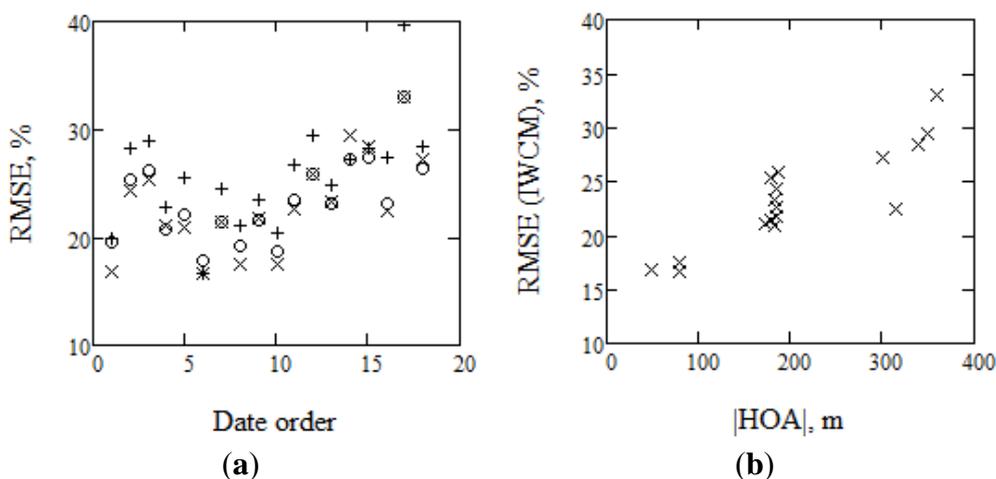
Table 4 indicates that winter-time data with long-lasting frozen conditions (February–March 2012, see also Figure 4) was more suitable for retrieval compared to other acquisitions. In Figure 9 the RMSEs are plotted with respect to date and to  $|\text{HOA}|$ . Figure 9b shows a clear difference depending on whether  $|\text{HOA}|$  was  $\leq 80$  m or above; in the former cases the retrieval RMSE was much lower. Since the TDM height is determined by the InSAR phase, and  $2\pi$  corresponds to a TDM height = HOA, a certain phase error will have increased effect on TDM height for increasing HOA.

The RMSE reported in Table 4 was obtained for the 201 forest stands larger than 1 ha. Taking into account smaller forest stands as well, resulted in larger retrieval errors as shown for the case of the IWCM-based retrieval in Table 5 for the two cases with lowest and highest RMSE, *i.e.*, for the TDM acquisitions on 2012-02-01 (HOA = 80 m) and 2012-08-13 (HOA = 358 m) respectively. The number of forest stands in the validation dataset increased to 315 when using a threshold on stand size of 0.5 ha (mean biomass of 110 Mg/ha). It further increased to 403 when the threshold was 0.25 ha (mean biomass of 112 Mg/ha).

**Table 4.** Single-image relative RMSE for the biomass estimated using the models IWCM, RVoG, and PD.

#	Date	RMSE %	RMSE %	RMSE %
		IWCM	RVoG	PD
1	2011-06-04	16.8	19.9	19.5
2	2011-11-23	24.3	28.2	25.4
3	2011-12-26	25.4	28.9	26.2
4	2012-01-17	21.1	22.8	20.7
5	2012-01-28	20.8	25.4	22.0
6	2012-02-01	16.7	16.7	17.9
7	2012-02-08	21.3	24.5	21.4
8	2012-02-12	17.5	21.1	19.1
9	2012-02-19	21.7	23.5	21.7
10	2012-02-23	17.5	20.5	18.7
11	2012-03-01	22.5	26.8	23.5
12	2012-03-12	25.9	29.3	25.8
13	2012-03-23	23.3	24.8	23.0
14	2012-05-28	29.5	27.2	27.2
15	2012-07-22	28.4	28.3	27.4
16	2012-08-02	22.5	27.4	23.1
17	2012-08-13	33.0	39.7	33.1
18	2012-08-24	27.2	28.3	26.4
	Mean RMSE	23.1	25.7	23.4

**Figure 9.** (a) Illustrating relative RMSE variation with TDM acquisition date (date order) × for IWCM, + for RVoG, o for PD. (b) Illustrating relative RMSE IWCM versus HOA.



**Table 5.** Relative RMSE in the case of the IWCM as a function of minimum forest stand size. The two acquisitions with the smallest and largest RMSE are shown for simplicity.

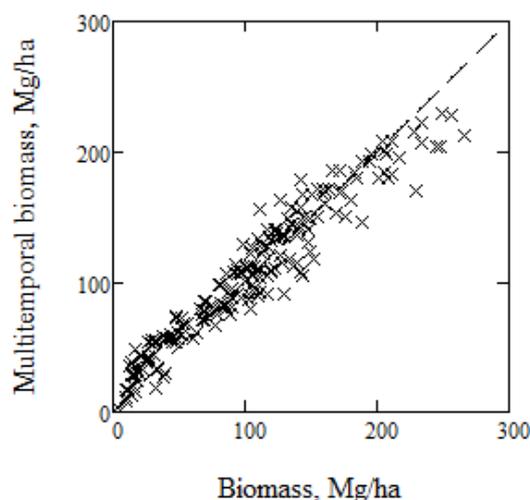
Date	RMSE %	RMSE %	RMSE %
	Stands ≥ 1 ha	Stands ≥ 0.5 ha	Stands ≥ 0.25 ha
2012-02-01	16.7	20.2	22.6
2012-08-13	33.0	35.2	38.3

It was previously observed (Figure 7) that the TDM height varied considerably between different acquisitions; however, for each acquisition the biomass of each stand,  $i$ , was rather constant thanks to the compensation embedded in the models. An important aspect of satellite observations is the possibility to further exploit repeated observations to reduce the uncertainties in the single-image estimates of the biomass [13,14,41]. A multitemporal combination of the different biomass estimates,  $l$ , is here proposed in which the weighting factor are based on the HOA since it has been shown above how the noise in the TDM height is increasing with HOA, *cf.* Figures 3, 5, and 9. The multitemporal estimate of the biomass of each stand,  $Bmt_i$ , is defined as

$$Bmt_i = \sum_{l=1}^{18} \frac{HOA_l^{-2}}{\sum_{l=1}^{18} HOA_l^{-2}} B_{l,i} \quad (14)$$

The multitemporal biomass determined by IWCM resulted in a RMSE of 16.5% or 17.3 Mg/ha and  $R^2 = 0.93$ , for stands  $> 1$  ha (Figure 10).

**Figure 10.** Scatterplot of biomass derived from a multi-temporal combination of 18 TanDEM-X InSAR pairs with respect to LiDAR-based biomass for 201 stands larger than 1 ha.



## 5. Discussion

The study on above-ground dry biomass retrieval with TanDEM-X interferometry follows a number of investigations on remote sensing data and retrieval techniques at the Remningstorp test site. Several of these studies have dealt with the use of multitemporal spaceborne and airborne SAR data. Interferometric SAR datasets acquired at C-band with one-day temporal separation were evaluated to retrieve forest stem volume in [13] achieving RMSE of 27%. In [27] L- and P-band SAR backscatter was used for biomass estimation with RMSE between 31% and 46% for L-band and 18% and 27% for P-band. In [42] a biomass model for P-band with training data from Krycklan, a test site in northern Sweden, was used, and validation data from Remningstorp resulted in RMSE 22–33%. For CARABAS VHF-band SAR single image estimates from different flight directions resulted in RMSE 11–25% [43]. However, it should be noted that RMSE is not the only way to compare different methods. In a first report on the use of TanDEM-X data from Remningstorp with the goal of biomass estimation [40], two monostatic acquisitions were studied with a delay of 3 s, which complicates the

analysis due to temporal decorrelation. For a 95% confidence interval four biomass classes up to 250 Mg/ha could be distinguished.

The RMSE obtained in the present study (16%) demonstrates the significant contribution of multitemporal TanDEM-X interferometric data to the quantification of forest aboveground biomass. Similar accuracy has earlier been found by means of other X-band satellite InSAR measurements; namely with SRTM [7], and TanDEM-X [44]. Since a dependence on tree species is expected [7,45], a knowledge about the species based on other data could improve RMSE. A shorter time difference between LiDAR measurements (August 2010) and TDM measurements (June 2011–August 2012) also could improve the RMSE, but the time difference is relatively short and the growth (approximately 5 Mg/ha/year) has not been compensated for. A local value of *BEF* *i.e.*, 0.512 was used, due to lack of a more general value. In the literature somewhat higher values can be seen, see *e.g.*, [46], and when testing *BEF* = 0.58 (determined from proportions of tree species of the seven 80 m × 80 m stands and *BEF*-factors for pine, spruce and deciduous according to [27]) the RMSE values were slightly changed (varying from 16.4% to 31.1%) but the mean value of RMSE for IWCM was unchanged at 23.1%. Since *BEF* is used for training as well as validation stands, the effect of *BEF* is small.

There is a close relationship between biomass and TDM height, which is determined by the forest height and the penetration depth. The latter is related to the vegetation density. The low extinction values, *i.e.*, <0.3 dB/m, which can be assumed to be related to gaps in the vegetation down to different levels, make the X-band microwaves to propagate up to 6–10 m into the vegetation according to the Penetration Depth model in the studied cases.

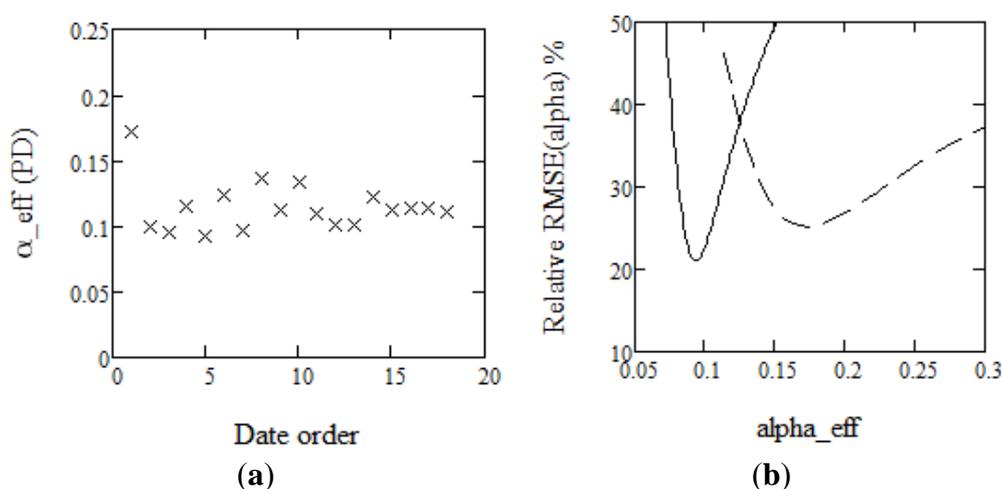
In the present study data the tendency to “saturate” at biomass > 200 Mg/ha in Figure 10 is probably not caused by the saturation effect observed in backscatter studies, since the measurement is based on the volume decorrelation and penetration depth of the upper vegetation layer, but may instead be caused by changes in the forest density, *i.e.*, number and dimension of gaps (Figure 2). This should be further investigated. The deviation from the reference line below 50 Mg/ha, could be related to the sensitivity of TDM data to specific conditions of the ground. In particular, the measured backscatter showed some deviations from the model in this region (Figure 5). However, it should also be noted that the reference biomass was obtained from LiDAR measurements which in turn were characterized by their own set of uncertainties and errors. The forest height varied more than 10 m in height in the biomass range < 50 Mg/ha (Figure 2) indicating certain complexity of the TDM and LiDAR metrics in this interval of biomass.

The number of training stands used in this study was half of the total number of stands, which in practical terms might be difficult to achieve when the aim is to map larger areas. With fewer training stands, it is assumed that the representativeness of the model parameters decreases in a manner related to forest homogeneity and measurement noise. To verify the impact of the number of stands used for model training on the retrieval, an extreme case was considered of a training dataset formed by only five stands chosen with approximate intervals of 50 Mg/ha. The RMSE for the best performing acquisition (2012-02-01) increased from 16.7% to 18.1% for stands larger than 1 ha, illustrating the possibility to limit the training dataset to a small number of training stands.

So far the accuracy of biomass estimation has been analyzed using models with parameters trained by stands having known properties. It would be very valuable if training stands could be avoided. The PD model is appealing since it contains only one unknown and the retrieval RMSE was close to the

result obtained with the IWCM (Table 4), in spite of the model approximation developed with emphasis on mature stands. Figure 11a shows that the estimate of the  $\alpha_{eff}$  parameter in the PD model obtained with training stands was relatively constant with one exception (see also Table 3). If  $\alpha_{eff}$  can be estimated from known conditions it may be possible to give a first-order estimate of forest height and biomass depending on how sensitive the estimates are for the correct  $\alpha_{eff}$ . A preliminary sensitivity analysis of the retrieval RMSE of biomass with respect to  $\alpha_{eff}$  is illustrated in Figure 11b and should be further investigated, when a wider range of environmental conditions have been studied.

**Figure 11.** (a) Estimates of  $\alpha_{eff}$  (PD model) versus acquisition date in date order and (b) sensitivity of the relative RMSE using PD to  $\alpha_{eff}$  for the two acquisitions with highest (dashed line, 2011-06-04) and lowest (solid line, 2012-01-28)  $\alpha_{eff}$ -values.



An analysis of two TanDEM-X acquisitions from a spruce dominated area in southeast Norway used a linear function [44] between biomass and TDM height without intercept,  $B \propto h_{TDM}$ . The relative RMSE at the stand level was 19% using a biomass increase of 14 Mg/ha per m increase of TDM height. Such a linear relation is in line with our results, *cf.* Figure 5. However, as shown by the variation of H150 (TDM height at 150 Mg/ha) there is a variation between the different acquisitions. Assuming a linear relation,  $B = \chi h_{TDM}$ , results in  $\chi$  varying between 10.3 and 16.4, with a mean of 13.3 Mg/ha per m TDM height.

The result in [44], and the analysis of PD show the importance of extending the TanDEM-X analysis to a wider range of environmental conditions and to investigate if a fixed value of a single parameter model results in a sufficiently high accuracy over a wide range of conditions. If so, the use of training stands can be avoided.

## 6. Conclusions

Eighteen interferometric TanDEM-X bistatic image pairs (VV-polarization) acquired between June 2011 and August 2012 over the test site of Remningstorp, situated in southern Sweden, have been studied in order to determine the potential of model-based above-ground dry biomass estimation. LiDAR-based estimates of biomass and vegetation ratio, acquired in August 2010 [19], were used as reference data. In order to interpret the TanDEM-X observations, the Interferometric Water Cloud

Model (IWCM), was primarily used. The Random Volume over Ground model (RVoG), was also studied, and a new model based on the penetration depth (PD) concept, was introduced. All the used models are based on physical principles, but in a significantly simplified form. Therefore, it should be stressed, that the parameters represent simplifications of complex phenomena. However, the physical relevance makes it possible to relate the variation of the parameters to other measurements and to environmental influence.

The relative RMSE of biomass associated with a retrieval based on the IWCM for forest stands  $\geq 1$  ha varied between 17% and 33% (relative to the mean value of 105 Mg/ha), with the best estimates obtained for small HOA. The relative RMSE for biomass retrieval based on the RVoG model varied between 17% and 40%. The relative RMSE for biomass retrieval based on the simple PD model was between 18% and 33%. Taking the mean of all 18 TDM estimates of stand biomass weighted inversely proportional to HOA<sup>2</sup> resulted in an RMSE of 16% in the IWCM case for forest stands larger than 1 ha. The environmental influence (temperature, humidity, rain *etc.*) on the TDM height resulted in a variation from 8.9 m to 15.2 m at a biomass of 150 Mg/ha, and this variation has to be taken care of by the model analysis.

The presented analysis demonstrates that TanDEM-X InSAR data together with an accurate high-resolution DTM, a fairly straightforward allometric expression, and forest stands for training model parameters, have a potential to estimate above-ground dry biomass with high accuracy in the case of forest conditions like those in Remningstorp. The results obtained by means of the bistatic TanDEM-X (VV-pol) are among the best remote sensing estimates of biomass obtained so far from Remningstorp. In a more general perspective, these results indicate the suitability of TanDEM-X data to retrieve boreal forest biomass with accuracy and spatial resolution as required by forest inventories, *cf.* [2].

## Acknowledgments

ESA and the BIOSAR2010 team are gratefully acknowledged for supporting the collection of forest inventory data. Leif Eriksson, Gustaf Sandberg, and Henrik Persson are acknowledged for support and discussions, and Heather Reese for proofreading the manuscript. TanDEM-X data were distributed through proposal XTI\_VEG0376 3D Forest Parameter Retrieval from TanDEM-X Interferometry. This work was partly supported by the Swedish National Space Board.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

1. Houghton, R.A. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* **2005**, *11*, 945–958.
2. Houghton, R.A.; Hall, F.; Goetz, S.J. Importance of biomass in the global carbon cycle. *J. Geophys. Res.: Biogeosci.* **2009**, *114*, doi:10.1029/2009JG000935.
3. Ouchi, K. Recent trend and advance of synthetic aperture radar with selected topics. *Remote Sens.* **2013**, *5*, 716–807.

4. Balzter, H.; Luckman, A.; Skinner, L.; Rowland, C.; Dawson, T. Observations of forest stand top height and mean height from interferometric SAR and LiDAR over a conifer plantation at Thetford Forest, UK. *Int. J. Remote Sens.* **2007**, *28*, 1173–1197.
5. Hajnsek, I.; Kugler, F.; Lee, S.; Papathanassiou, K.P. Tropical-forest-parameter estimation by means of Pol-InSAR: The INDREX-II campaign. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 481–493.
6. 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.
7. Solberg, S.; Astrup, R.; Gobakken, T.; Næsset, E.; Weydahl, D.J. Estimating spruce and pine biomass with interferometric X-band SAR. *Remote Sens. Environ.* **2010**, *114*, 2353–2360.
8. Krieger, G.; Moreira, A.; Fiedler, H.; Hajnsek, I.; Werner, M.; Younis, M.; Zink, M. TanDEM-X: A satellite formation for high-resolution SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3317–3341.
9. Weydahl, D.J.; Sagstuen, J.; Dick, Ø.B.; Rønning, H. SRTM DEM accuracy assessment over vegetated areas in Norway. *Int. J. Remote Sens.* **2007**, *28*, 3513–3527.
10. Neeff, T.; Dutra, L.V.; Santos, J.R.; Freitas, C.C.; Araújo, L.S. Tropical forest measurement by interferometric height modeling and P-band backscatter. *For. Sci.* **2005**, *51*, 585–594.
11. Askne, J.; Dammert, P.; Ulander, L.M.H.; Smith, G. C-band repeat-pass interferometric SAR observations of forest. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 25–35.
12. Askne, J.; Dammert, P.; Ulander, L.M.H.; Smith, G. Multi-temporal repeat pass SAR interferometry of boreal forests. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1540–1550.
13. Askne, J.; Santoro, M. Experiences in Boreal Forest Stem Volume Estimation from Multitemporal C-Band InSAR. In *Recent Interferometry Applications in Topography and Astronomy*; Padron, I., Ed.; InTech: Morn Hill, Winchester, UK, 2012; 169–194. Available online: [http://cdn.intechopen.com/pdfs/33100/InTech-Experiences\\_in\\_boreal\\_forest\\_stem\\_volume\\_estimation\\_from\\_multitemporal\\_c\\_band\\_insar.pdf](http://cdn.intechopen.com/pdfs/33100/InTech-Experiences_in_boreal_forest_stem_volume_estimation_from_multitemporal_c_band_insar.pdf) (accessed on 25 October 2013).
14. Santoro, M.; Askne, J.; Smith, G.; Fransson, J.E.S. Stem volume retrieval in boreal forests from ERS-1/2 interferometry. *Remote Sens. Environ.* **2002**, *81*, 19–35.
15. Cloude, S.R.; Papathanassiou, K.P. Polarimetric SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1551–1565.
16. Papathanassiou, K.P.; Cloude, S.R. Single-baseline polarimetric SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 2352–2363.
17. Treuhaft, R.N.; Madsen, S.N.; Moghaddam, M.; van Zyl, J.J. Vegetation characteristics and underlying topography from interferometric data. *Radio Sci.* **1996**, *31*, 1449–1495.
18. Ulander, L.M.H.; Gustavsson, A.; Dubois-Fernandez, P.; Dupuis, X.; Fransson, J.E.S.; Holmgren, J.; Wallerman, J.; Eriksson, L.; Sandberg, G.; Soja, M. BIOSAR 2010—A SAR Campaign in Support to the BIOMASS Mission. In *Proceedings of 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Vancouver, BC, Canada, 24–29 July 2011; pp. 1528–1531.

19. Ulander, L.M.H.; Gustavsson, A.; Flood, B.; Murdin, D.; Dubois-Fernandez, P.; Depuis, X.; Sandberg, G.; Soja, M.J.; Eriksson, L.E.B.; Fransson, J.E.S.; *et al.* *BIOSAR 2010 Technical Assistance for the Development of Airborne SAR and Geophysical Measurements during the BioSAR 2010 Experiment; Final Report*; ESA contract no. 4000102285/10/NL/JA/ef; ESA: Noordwijk, The Netherlands, 2011; p. 212. Available online: [http://earth.esa.int/campaigns/DOC/BioSAR\\_2010\\_final\\_report\\_v1.0.pdf](http://earth.esa.int/campaigns/DOC/BioSAR_2010_final_report_v1.0.pdf) (accessed on 25 October 2013).
20. ESA. *Report for Mission Selection: Biomass*; European Space Agency: Noordwijk, The Netherlands, 2012.
21. Ahti, T.; Hämet-Ahti, L.; Jalas, J. Vegetation zones and their sections in northwestern Europe. *Annales Botanici Fennici* **1968**, *5*, 169–211.
22. Wikström, P.; Edenius, L.; Elfving, B.; Eriksson, L.O.; Lämås, T.; Sonesson, J.; Öhman, K.; Wallerman, J.; Waller, C.; Klintebäck, F. The Heureka forestry decision support system: An overview. *Math. Comput. For. Nat.-Resour. Sci.* **2011**, *3*, 87–94.
23. Petersson, H. *Biomassfunktioner för Trädfraktioner av Tall, Gran och Björk i Sverige*; Department of Forest Resource Management, Swedish University of Agricultural Sciences: Umeå, Sweden, 1999.
24. Marklund, L.G. *Biomassfunktioner för Tall, Gran och Björk i Sverige*; Institutionen för Skogstaxering, Sveriges Lantbruksuniversitet: Umeå, Sweden, 1988.
25. Näslund, M. *Funktioner och Tabeller för Kubering av Stående Träd. Tall, Gran och Björk i Södra Sverige samt Hela Landet*; Statens Skogsforskningsinstitut: Stockholm, Sweden, 1947.
26. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **2002**, *80*, 88–99.
27. Sandberg, G.; Ulander, L.M.H.; Fransson, J.E.S.; Holmgren, J.; Toan, T.L. L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest. *Remote Sens. Environ.* **2011**, *115*, 2874–2886.
28. Jalkanen, A.; Mäkipää, R.; Ståhl, G.; Lehtonen, A.; Petersson, H. Estimation of the biomass stock of trees in Sweden: comparison of biomass equations and age-dependent biomass expansion factors. *Ann. For. Sci.* **2005**, *62*, 845–851.
29. Woodhouse, I.H. Predicting backscatter-biomass and height-biomass trends using a macroecology model. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 871–877.
30. Balss, U.; Breit, H.; Duque, S.; Fritz, T.; Rossi, C. TanDEM-X Payload Ground Segment: CoSSC Generation and Interferometric Considerations. Available online: [https://tandemx-science.dlr.de/pdfs/TD-PGS-TN-3129\\_CoSSCGenerationInterferometricConsiderations\\_1.0.pdf](https://tandemx-science.dlr.de/pdfs/TD-PGS-TN-3129_CoSSCGenerationInterferometricConsiderations_1.0.pdf) (accessed on 25 October 2013).
31. Huber, S.; Fiedler, H.; Krieger, G.; Zink, M. TanDEM-X Performance Optimization. In Proceedings of International Radar Symposium, Cologne, Germany, 5–7 September 2007.
32. Soja, M.; Ulander, L. Digital Canopy Model Estimation from TanDEM-X Interferometry Using High-resolution Lidar DEM. In Proceedings of 2013 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2013, Melbourne, VIC, Australia, 22–26 July 2013.
33. Attema, E.P.W.; Ulaby, F.T. Vegetation modelled as a water cloud. *Radio Sci.* **1978**, *13*, 357–364.
34. De Zan, F.; Krieger, G.; López-Dekker, P. On some spectral properties of TanDEM-X Interferograms over forested areas. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 71–75.

35. Pulliainen, J.T.; Heiska, K.; Hyyppä, J.; Hallikainen, M.T. Backscattering properties of boreal forests at the C-and X-band. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 1041–1050.
36. Fransson, J.E.S.; Smith, G.; Askne, J.; Olsson, H. Stem volume estimation in boreal forests using ERS-1/2 coherence and SPOT XS optical data. *Int. J. Remote Sens.* **2001** *22*, 2777–2791.
37. Santoro, M.; Shvidenko, A.; McCallum, I.; Askne, J.; Schmullius, C. Properties of ERS-1/2 coherence in the Siberian boreal forest and implications for stem volume retrieval. *Remote Sens. Environ.* **2007**, *106*, 154–172.
38. Santoro, M.; Askne, J.; Dammert, P. Tree height retrieval from ERS interferometric phase in boreal forest. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 207–217.
39. Mette, T.; Papathanassiou, K.P.; Hajnsek, I.; Zimmermann, R. Forest Biomass Estimation Using Polarimetric SAR Interferometry. In Proceedings of Workshop on POLinSAR—Applications of SAR Polarimetry and Polarimetric Interferometry, Frascati, Italy, 14–16 January 2003.
40. Caicoya, A.T.; Kugler, F.; Hajnsek, I.; Papathanassiou, K. Boreal Forest Biomass Classification with TanDEM-X. In Proceedings of 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012.
41. Santoro, M.; Beer, C.; Cartus, O.; Schmullius, C.; Shvidenko, A.; McCallum, I.; Wegmüller, U.; Wiesmann, A. Retrieval of growing stock volume in boreal forest using hyper-temporal series of Envisat ASAR ScanSAR backscatter measurements. *Remote Sens. Environ.* **2011**, *115*, 490–507.
42. Soja, M.J.; Sandberg, G.; Ulander, L.M.H. Regression-based retrieval of boreal forest biomass in sloping terrain using P-band SAR backscatter intensity data. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2646–2665.
43. Folkesson, K.; Smith-Jonforsen, G.; Ulander, L.M.H. Validating backscatter models for CARABAS SAR images of coniferous forests. *Can. J. Remote Sens.* **2008**, *34*, 480–495.
44. Solberg, S.; Astrup, R.; Breidenbach, J.; Nilsen, B.; Weydahl, D. Monitoring spruce volume and biomass with InSAR data from TanDEM-X. *Remote Sens. Environ.* **2013**, *139*, 60–67.
45. Ulander, L.M.H.; Askne, J.I.H.; Eriksson, L.E.B.; Fransson, J.E.S.; Persson, H.; Soja, M. Effects of Tree Species and Season on Boreal Forest Biomass Estimates from TanDEM-X. In Proceedings of 4th TanDEM-X Science Team Meeting, Oberpfaffenhofen, Germany, 12–14 June 2013; p. 15. Available online [https://tandemx-science.dlr.de/pdfs/ScienceMeeting2013\\_OralPresentations/TDX\\_8\\_Forest\\_Applications/3\\_Ulander.pdf](https://tandemx-science.dlr.de/pdfs/ScienceMeeting2013_OralPresentations/TDX_8_Forest_Applications/3_Ulander.pdf) (accessed on 25 October 2013).
46. Häme, T.; Salli, A.; Lahti, K. Estimation of Carbon Storage in Boreal Forests Using Remote Sensing Data, Pilot Study. In *The Finnish Programme on Climate Change, Progress Report*; Kanninen, M., Anttila, P., Eds.; Publications of the Academy of Finland; VAPK Publishing: Helsinki, Finland, 1992; Vol. 3/92; pp. 250–255.