

BOREAL FOREST BIOMASS CLASSIFICATION WITH TANDEM-X

Astor Toraño Caicoya ^a, Florian Kugler ^a, Irena Hajnsek ^{a,b}, Kostas Papathanassiou ^a

^a German Aerospace Center (DLR), ^b ETH Zurich

ABSTRACT

High spatial resolution X-band interferometric SAR data from the TanDEM-X, in the operational DEM generation mode, are sensitive to forest structure and can therefore be used for thematic boreal forest classification of forest environments. The interferometric coherence in absence of temporal decorrelation depends strongly on forest height and structure. Due to the rather homogenous structure of boreal forest, forest biomass can be derived from forest height, on the basis of allometric equations with sufficient accuracy to be used for thematic classification applications. Two test sites in mid- and southern Sweden are investigated. A maximum of 4 biomass classes, up to 250 Mg/ha, are achieved. Larger spatial baselines result in better classification performances.

Index Terms— Biomass, boreal forest, TanDEM-X, RVoG.

1. INTRODUCTION

Boreal forests contain roughly 1/3 of the total Earth’s forest biomass, covering an area of 386 million of km². Thus, given the vastness and remoteness of the majority of its area, high-resolution satellite imagery is indispensable for mapping and monitoring these forests [1]. The main, existing land cover classifications are based on optical systems that are limited with respect to quantitative classifications, even in a low number of classes. Forest biomass classification can improve thematic mapping differentiating more classes within a qualitative forest type (e.g. biomass in coniferous forests) and, therefore, enhancing the understanding of forest carbon dynamics in the boreal region. In the SIBERIA project [1], a radar approach for forest classification was used to generate a land cover classification with three biomass classes (sensitive up to 81 Mg/ha), already showing the potential of synthetic aperture radar (SAR) for boreal forest biomass classifications.

The TanDEM-X (TDX) mission offers “single pass” interferometric X-band data from space for the first time, which can be achieved in monostatic and bistatic modes [3]. The primary objective of the mission is to generate a global Digital Elevation Model (DEM), with 2m of vertical accuracy. The operational DEM data acquisition is performed in a bistatic InSAR stripmap mode in HH polarization [3]. For the DEM mode two acquisitions with different baselines are foreseen: the first acquisition has a minimum height of ambiguity of 45m and the second of 30m. The incidence angle varies between 30° and 48°. The sensitivity of the InSAR coherence on forest structure, combined with a high spatial resolution, motivates the utilization of this global data set for updating, extending and/or improving existing thematic forest classifications.

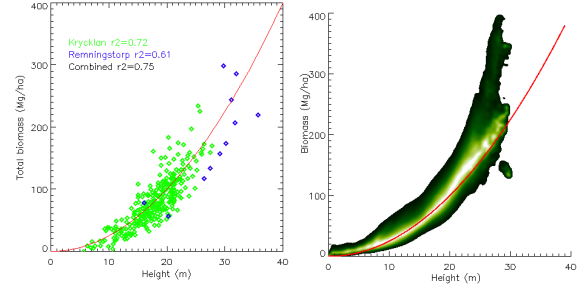


Figure 1: Height to biomass allometric relation. Left: ground truth plots in the test sites Krycklan (green) and Remmingstorp (blue). Right: Krycklan site for LiDAR height vs. LiDAR derived biomass (2-D histogram).

2. TEST SITES AND DATA

Two test-sites have been evaluated in this study: Krycklan located in middle Sweden (64°10’N and 20°01’E), a boreal forest site with a very hilly topography; and Remmingstorp in southern Sweden (58°25’N, 13°14’E), a hemi-boreal forest over a rather flat terrain. In Krycklan a bistatic acquisition acquired in the DEM standard mode and a monostatic one, acquired with 3 sec. of temporal baseline have been used. In Remmingstorp monostatic acquisitions (3 sec. temporal baseline) have been investigated. The used data specifications are summarized in Table 1.

Table 1: TanDEM-X data. \bar{k}_z is the mean vertical wave number, and H_{amb} the height of ambiguity.

Date	Test site	Mode	\bar{k}_z	H_{amb}
28-07-2101	Krycklan	Monostatic	0.186	33.8
17-12-2010	Krycklan	Bistatic	0.113	55.6
29-08-2010	Remmingstorp	Monostatic	0.151	41.6
20-09-2010	Remmingstorp	Monostatic	0.166	37.8

Validation for both tests sites is done against LiDAR derived forest height and biomass, and ground measurements. In Krycklan the airborne LiDAR data set was acquired in 2008 together with 289 ground-inventory plots. In Remmingstorp, the LiDAR data was acquired in 2007 with 10 inventory plots.

3. STRUCTURE IN THE BOREAL FOREST

Forest biomass can be derived from forest parameters such as forest height, using allometric equations. Height-to-biomass allometry works best for homogenous forest conditions i.e. even aged and single species forests [2]; however, this allometric

relation is limited in highly diverse forests, in terms of age and species, or forests with disturbances.

Boreal forests are characterized by a homogenous structure, which optimizes the usage of height-to-biomass allometric relations. In Figure 1 height vs. biomass from ground measurement plots is shown. The measurements for the two test sites, Krycklan in green and Remningstorp in blue, show a high correlation ($R^2 = 0.75$) and both follow the same allometric relation:

$$B = 0.25h_v^2 \quad (1)$$

where B is the biomass (Mg/ha) and h_v the forest height (m). In Figure 1 height-to-biomass, derived from the LiDAR data for the test site Krycklan, is shown. The relation follows the same allometric relation as in the ground measurements.

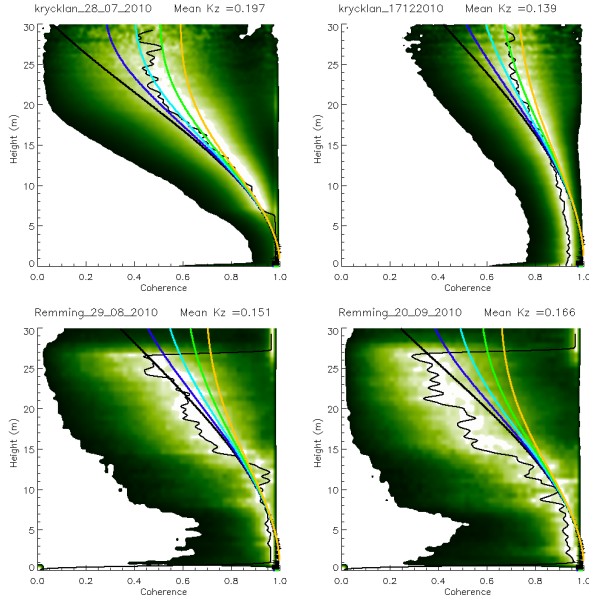


Figure 2: Height/Coherence histograms for Krycklan on the top and Remningstorp on the bottom. The curves show the height calculated from the RVoG model from a 0dB/m fixed extinction in black to a max of 0.4dB/m in yellow, in 0.1 dB/m steps, for the mean k_z of the scene. The black line follows the max occurrence.

4. BIOMASS INVERSION FROM X-BAND SINGLE CHANNEL

4.1 POLin-SAR Structure model

The interferometric coherence ($\tilde{\gamma}$) comprises several decorrelation contributions. The most important, in case of TDX, are volume ($\tilde{\gamma}_{vol}$), temporal ($\tilde{\gamma}_{temp}$) and noise ($\tilde{\gamma}_{SNR}$) decorrelations:

$$\tilde{\gamma} = \tilde{\gamma}_{vol} \cdot \tilde{\gamma}_{temp} \cdot \gamma_{SNR} \quad (2)$$

While in the operational bistatic mode the temporal baseline is in the order of 1/10 of second so that $\tilde{\gamma}_{temp} = 1$ in the early pursuit monostatic phase the temporal baseline was on the order of 3 seconds so that $\tilde{\gamma}_{temp} \leq 1$.

According to the random volume over ground model (RVoG) [4][5] the volume decorrelation contribution ($\tilde{\gamma}_{vol}(HH)$) of the interferometric coherence is given by:

$$\tilde{\gamma}_{vol}(HH) = \exp(ik_z z_0) \frac{\tilde{\gamma}_{v_0} + m}{1 + m} \quad (3)$$

where z_0 is the ground topography, k_z the vertical wave number and m the effective ground-to-volume ratio. $\tilde{\gamma}_{v_0}$ is the volume decorrelation caused in the absence of the ground layer and can be modeled [4][5]:

$$\tilde{\gamma}_{v_0} = \frac{\int_0^{h_v} \exp(ik_z z') \exp\left(\frac{2\sigma z'}{\cos\theta_0}\right) dz'}{\int_0^{h_v} \exp\left(\frac{2\sigma z'}{\cos\theta_0}\right) dz'} \quad (4)$$

where σ is the volume shape factor and θ_0 the local incidence angle.

A standard DEM TanDEM-X acquisition provides only one complex coherence (in HH polarization). A height estimate can be only obtained by assuming $m = 0$ and a fixed σ :

$$\min_{h_v} \left\| \left| \tilde{\gamma} \right| - \left| \tilde{\gamma}_{v_0}(h_v | k_z, \sigma = const) \right| \right\| \quad (5)$$

The measured coherence $\tilde{\gamma}$ needs to be corrected before inversion for non-volumetric decorrelation sources. The SNR correction is performed according to [6].

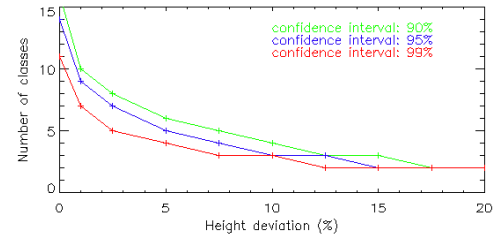


Figure 3: Biomass classification sensitivity for different confidence intervals (90%, 95% and 99%).

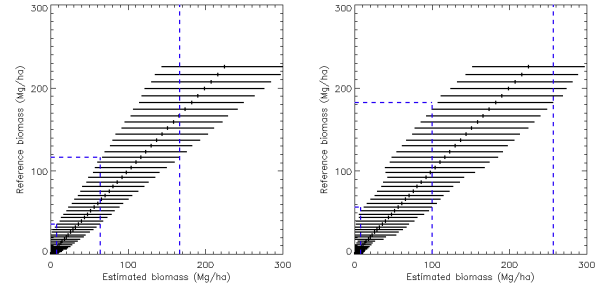


Figure 4: Classification performance for: Krycklan 28-07-2010, $\bar{k}_z=0.186$ (left), and 17-12-2010 (left), $\bar{k}_z=0.113$. Each horizontal line shows the total biomass deviation at each estimated biomass level (ΔB). The blue dashed lined represent the separation of biomass classes for a 95% confidence interval.

Now, the sensitivity of coherence to height is tested. For this, 2-D histograms between interferometric coherence and LiDAR height are plotted in Figure 2. In Krycklan the differences between the bistatic (winter) and the monostatic (summer) case are driven by the difference in baseline (k_z), as the temporal decorrelation in the monostatic is low. However, for Remningstorp the 3 sec. temporal baseline in the pursuit monostatic acquisition is enough to induce temporal decorrelation effects, i.e. the associated estimation biases on height and biomass levels. This effect is seen as a drop of coherence above ~ 15 m. Some strong deviations, observed in Remningstorp for heights above 30 m, correspond to stands that

have been harvested in the time between the LiDAR and the TDX acquisition

In each histogram the modeled height obtained from the RVoG (Eq. 4) for a mean k_z and for a range of fixed volume shape factor (0 to 0.4 dB/m, from black until yellow in 0.1 dB/m steps) is superimposed. A fixed mean extinction of 0.1 dB/m shows the best behavior for Krycklan while in Remningstorp 0 dB/m fits better, so these σ values will be used in Eq. 5 for the height estimation ($const = 0 ; 0.1$ dB/m). The low extinction values for X-band can be explained by the presences of an (unaccounted) ground contribution in the signal.

Finally, mean height and biomass have been calculated for a set of pre-defined homogenous stands in Krycklan for the summer acquisition (Figure 8) and are compared with the analogue biomass maps estimated from LiDAR (Figure 5). The height estimation accuracy lies within a range of +/-5 m while the estimated biomass in the majority of the stands lies within a range of +/-50 Mg/ha. Some stands show a high deviation in the high biomass range; however, these are stands affected by harvesting activities between the LiDAR and the TDX acquisitions.

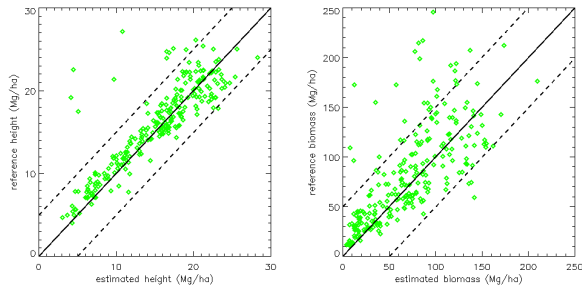


Figure 5: Stand level validation. Left: LiDAR height vs. estimated height; right: LiDAR reference biomass vs. estimated biomass.

4.2 Biomass inversion and classification performance

A biomass map is obtained in two steps: by first modeling height from coherence and then estimating biomass from height using Eq. 1. Thus, the accuracy of the measured biomass depends on two sources, one comes from the deviation between LiDAR and

estimated height (Δh_v); and the other between the biomass estimated from the allometric equation, and the LiDAR derived reference biomass (ΔB). Accordingly, the total biomass estimation sensitivity (ΔB_{tot}) can be expressed as following:

$$\Delta B_{tot} = \frac{\partial B}{\partial h_v} \Delta h_v + \Delta B \quad (5)$$

Where ΔH and ΔB are estimated from the data for one standard deviation interval and height dependent. Introducing Eq.1 in Eq.5 results in:

$$\Delta B_{tot} = 0.5H \Delta h_v + \Delta B \quad (6)$$

Since the allometric relation is fixed, the estimation sensitivity will depend on the height estimation accuracy. In Figure 3 the number of classes that can be estimated for confidence intervals of 90, 95 and 99% are shown against the height accuracy. For a 95% confidence interval, a maximum of 14 classes can be estimated assuming no height deviation degrading to 2 for $\Delta h_v = 15\%$. In both sites evaluated here, for a 95% confidence interval a maximum of 4 biomass classes are obtained ($\overline{\Delta h_v} \approx 5 - 10\%$). An example for Krycklan is shown in Figure 4. In Krycklan, the lowest deviation is seen in the summer acquisition (28-07-2010) with a \bar{k}_z of 0.186. The acquisition in winter (17-12-2010), with a shorter baseline ($\bar{k}_z=0.113$), has a higher height of ambiguity but a lower height sensitivity. Thus, biomass estimation with lower k_z is more accurate at higher biomass levels while larger k_z provide a better estimation at lower biomass levels.

4.3 Classification results

In our test sites, four biomass classes are defined (<10, 10-50, 50-150, >150 Mg/ha) to obtain the biomass classifications shown in Figure 6 (Krycklan) and Figure 7 (Remningstorp). Krycklan 28-07-2010 shows a better agreement with the LiDAR biomass, especially at the lower biomass classes (<50 Mg/ha in blue). The acquisition in 17-12-2010 shows an overestimation in the upper classes. This could be explained by the lower extinction due to frozen conditions and the associated higher ground contribution. In Remningstorp both acquisitions show similar performance compared to the LiDAR, with a poor sensitivity for the class between 10 and 50 Mg/ha. Harvested areas can be clearly distinguished between the LiDAR and the TDX images.

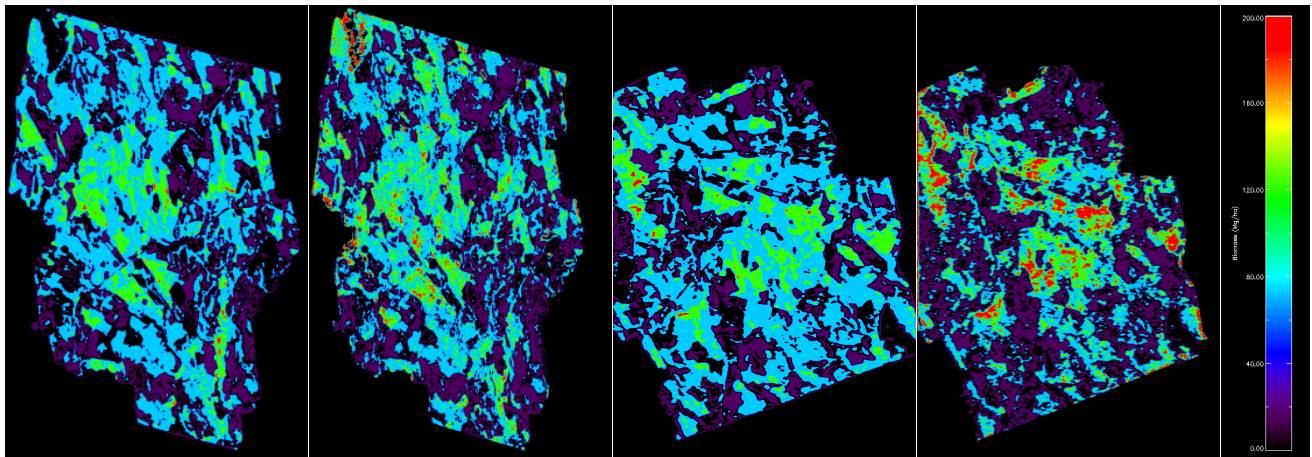


Figure 6: Biomass classification for Krycklan test site. LiDAR derived biomass is compared with the TanDEM-X biomass classification. From left to right: 28-07-2010 LiDAR and TanDEM-X classifications; and 17-12-2010 in same order .The color bar range moves from 0 to 200 Mg/ha.

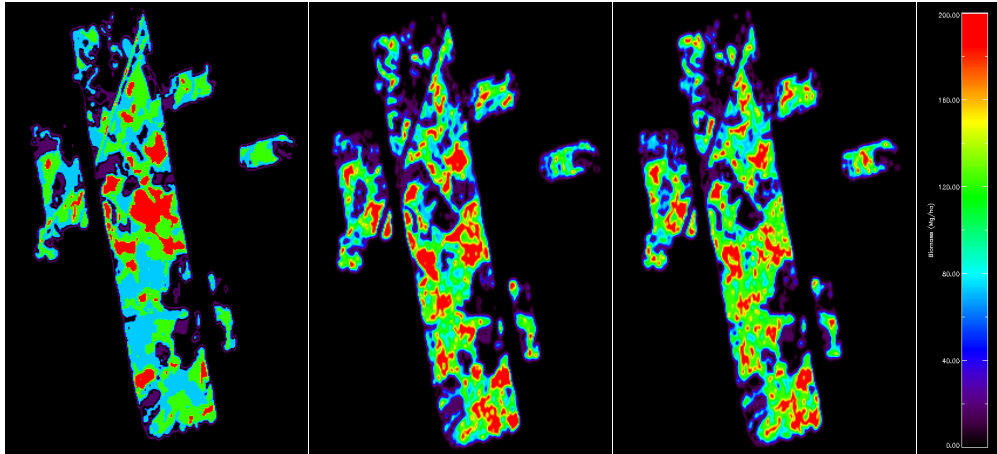


Figure 7: Biomass classification for Remningstorp test site. From left to right: LiDAR derived biomass, 29-08-2010 and 20-09-2010 TanDEM-X biomass classification. The color bar range moves from 0 to 200 Mg/ha.

For the test site Krycklan the resulting biomass classification has been compared with the European thematic classification CORINE (Figure 8-left). CORINE is a medium resolution classification (100*100m) while the TDX has a spatial resolution of 16m. Nevertheless, in CORINE non-forested areas, i.e. grasslands and agriculture, which are superimposed in white, always correspond to the lowest biomass class (<10 Mg/ha) in the TDX biomass map; and forested areas, classified in CORINE as coniferous forest (the rest of the site) is separated by TDX in the four biomass classes. Height-to-biomass allometric relations accuracy improves when increasing the number of used looks. Thus, as thematic classifications don't require the high spatial accuracy that TDX provides, a higher number of biomass classes could be estimated at the cost of spatial resolution.

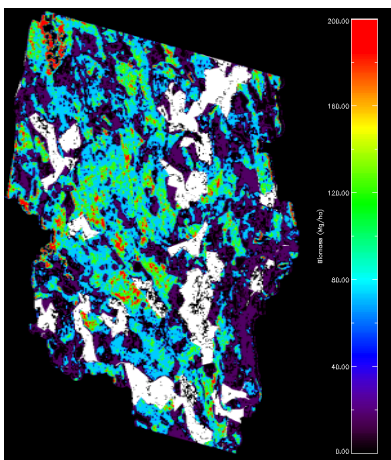


Figure 8: Biomass classification combined with CORINE land cover classification in Krycklan 28-07-2010. The white areas correspond to non-forested CORINE classes while the rest of the area is classified as coniferous forest.

5. CONCLUSIONS

The TDX DEM standard acquisition shows a great potential for forest biomass classification in the boreal regions as shown in our tests sites. Four biomass classes, up to biomass levels of 250 Mg/ha, could be distinguished. Longer baselines are more sensitive to low biomass levels while shorter baselines are more sensitive to higher biomass levels; thus, a combination of different baseline configurations could improve the biomass classification results (more biomass classes).

The obtained biomass classification maps can improve thematic mapping in forested areas as provided in classifications like CORINE, and is able to discriminate between different biomass levels within defined CORINE classes in the boreal forest (e.g. coniferous forest).

6. REFERENCES

- [1] W. Wagner et al., "Large-scale mapping of boreal forest in SIBERIA using ERS tandem coherence and JERS backscatter data", *Remote Sensing of Environment*, vol. 85, pp. 125-144, 2003.
- [2] C.A. Torano, F. Kugler, K. Papathanassiou, P. Biber, P. Pretzsch, "Biomass estimation as a function of vertical forest structure and forest height. Potential and limitations for Radar Remote Sensing". *Proceedings of EUSAR, Aachen*, 2010.
- [3] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, and M. Zink, "TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 11, pp. 3317- 3341, 2007.
- [4] K. P. Papathanassiou, and S. R. Cloude, "Single-baseline Polarimetric SAR Interferometry", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 11, pp. 2352-2363, 2001.
- [5] S.R. Cloude and K.P., Papathanassiou, "Three-stage inversion process for polarimetric SAR interferometry", *IEE Proceedings - Radar Sonar and Navigation*, vol. 150, no. 3, pp. 125-134, 2003
- [6] F. Kugler, I. Hajnsek, K. Papathanassiou, "Forest Parameter Characterization by means of TerraSAR-X and TanDEM-X (Polarimetric and) Interferometric data.", *Proceedings of PolInSAR 2011, 24th -28th of January Frascati, Italy*, 2011.