ESTIMATION OF BOREAL FOREST BIOMASS FROM TWO-LEVEL MODEL INVERSION OF INTERFEROMETRIC TANDEM-X DATA

Maciej J. Soja[1), Henrik Persson[2), and Lars M. H. Ulander[1,3)]

[1)Chalmers University of Technology, Gothenburg, Sweden

[2)Swedish University of Agricultural Sciences, Umeå, Sweden

[3)Swedish Defence Research Agency, Linköping, Sweden

ABSTRACT

A new model for aboveground biomass estimation from forest height and canopy density estimates obtained from the inversion of a two-level model (TLM) is presented and studied using data from the hemi-boreal test site Remningstorp, situated in southern Sweden. Three bistatic-interferometric TanDEM-X acquisitions from the summers of 2011, 2012, and 2013 and with heights-of-ambiguity (HOAs) 49 m, 32 m, and 63 m, respectively, are used. An external, high-resolution digital terrain model (DTM) is used as ground reference during interferogram flattening. Model parameters are estimated for each acquisition separately, and the model is evaluated on all three acquisitions, to examine both its explanatory and predictive values. Residual root-mean-square errors (RMSEs) are 14%-19% and the model explains 67%-84% of the variance in the data. Prediction RMSE is 20% for the two images with the highest HOA, but much higher for the third image.

Index Terms— forest biomass, two-level model inversion, TanDEM-X interferometry

1. INTRODUCTION

There is an increasing need for accurate, global, high-resolution mapping of forest, mainly in terms of aboveground biomass. Biomass is one of the largest uncertainties in the current carbon cycle models, which affects climate change prediction.

One promising approach is based on high-resolution mapping of forest height with the TanDEM-X (TDM) system, which consists of two, almost identical X-band SAR satellites positioned in a tight tandem formation with a small spatial baseline [1]. The main task of the TDM mission is the creation of a first, global, high-resolution digital elevation model (DEM), that is a map of the mean scattering center. If a high-resolution digital terrain model (DTM) is available, the ground component can be removed and a map of the scattering center relative ground can be obtained. In forests, this height is related to forest height [2, 3], but the relation is non-trivial as the location of the scattering center is determined by three different factors: interferometric imaging geometry, forest and ground structure, and their dielectric properties. Factors, which are not directly related to forest height and biomass (imaging geometry, ground structure, and dielectric properties) will thus create a bias, if the interferometric height is to be used as a forest height or biomass predictor.

One way to compensate for the influence of the non-related factors is through an interferometric model. One of the most common interferometric models is the random volume over ground (RVOG) model [4]. In the common formulation, the RVOG has four unknown parameters: ground phase, extinction coefficient, ground-to-volume ratio, and volume height, and it models a complex correlation coefficient (one complex value). An inversion problem using one single acquisition will thus be underdetermined. Therefore, the inversion of RVOG requires additional acquisitions (more baselines and/or polarizations), or otherwise the model needs to be simplified [5].

In this work, the RVOG will be simplified and slightly modified. The ground phase will be taken from an external digital terrain model (DTM), volume penetration will be neglected by choosing an infinite extinction coefficient, and canopy gaps will be introduced. The new two-level model (TLM) will have two unknown parameters: a level distance and an area-weighted backscatter ratio; it will thus be possible to invert the model using one single acquisition only.

The inverted area-weighted backscatter ratio and level distance will be used to model biomass for three acquisitions made over the hemi-boreal test site Remningstorp during the summers of 2011, 2012, and 2013, at different baselines. The estimated parameters will be studied, and used to estimate biomass both for the same acquisition, and for other acquisitions. This way, both the explanatory and predictive values of the model will be studied. The approach presented here is an excerpt from an on-going study [6]. The inversion of the two-level model has been studied in [7, 8].

2. BIOMASS MODEL

A multiplicative model is here studied:

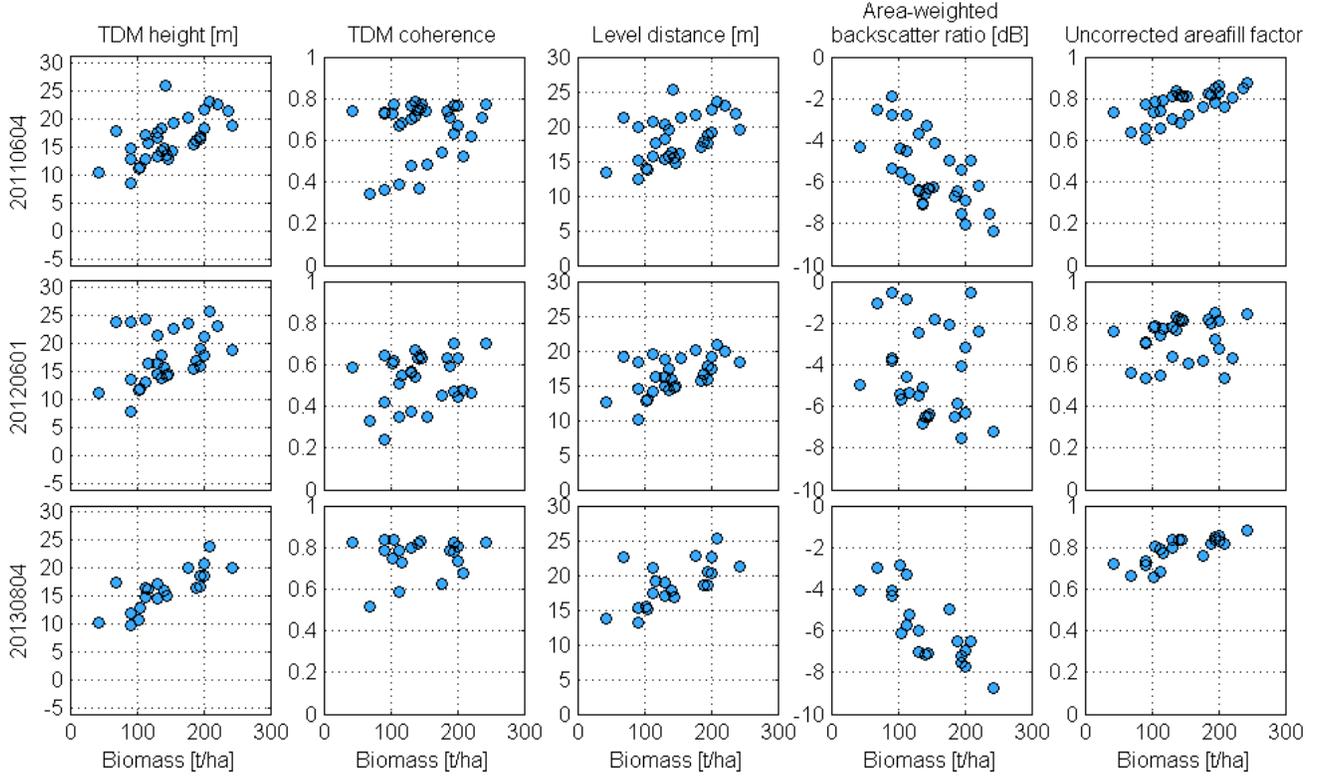$$(1) \qquad AGB = K \cdot \Delta h^{\alpha} \cdot \eta_0^{\beta}$$

Figure 1   **Plots of TDM height, coherence, the inverted TLM parameters level distance and area-weighted backscatter ratio, the derived uncorrected areafill factor against biomass for three acquisitions, from top to bottom: 20110604 (HOA=49 m), 20120601 (HOA=32 m), 20130804 (HOA=63 m).**

where AGB is above-ground biomass in tons per hectare, η_0 is the uncorrected areafill factor:

$$\eta_0 = \frac{1}{1+\mu},$$

Δh is the level distance in meters, and K, α, and β are unknown parameters. Area-weighted backscatter ratio μ and level distance Δh are obtained by solving the following equation:

$$(2) \quad \tilde{\gamma}_{gc} = \frac{\mu + e^{ik_z \Delta h}}{\mu + 1}$$

where $\tilde{\gamma}_{gc}$ is the value of the DTM-corrected correlation coefficient averaged over the entire plot, and k_z is the vertical wavenumber for each plot:

$$(3) \quad k_z = \frac{2\pi B_\perp}{\lambda R \sin \theta_0},$$

where $B_\perp$ is the perpendicular baseline, λ is the wavelength, R is the range, and θ_0 is the angle of incidence. The formula on the right-hand side of (2) is the expression for the two-layer model. The choice of the multiplicative form shown in (1) can be motivated by the fact that vegetation ratio is related to the amount of vegetation sensed within the observed field and height is related to the size of trees.

3. EXPERIMENTAL DATA

In this study, the hemi-boreal test site Remningstorp is used. The site is situated in southern Sweden, and it is a production forest consisting primarily of Norway spruce, Scots pine, and different birch species. Field inventories were conducted in autumn 2010 and spring 2011. 32, 0.5-hectare, circular forest plots were selected. For these plots, stem diameter of all trees has been measured, together with height for a subset of approximately 10% of the trees. Biomass was estimated using the allometric equations presented in [9], with an error lower than 5% at the time of acquisition [10].

Several forest management procedures have been conducted during the time of the study, resulting in a loss of biomass. Affected plots have been identified from optical imagery, field visits, and activity logs provided by the manager. The affected plots have been removed from the data set. Therefore, there are only 29 unaltered plots in 2012 and 21 in 2013.

Since this study covers a relatively short period of up to three growing seasons, growth will be neglected here.

The TanDEM-X data have been interferometrically processed using an in-house developed algorithm [3], and based on [11]. For ground information, a digital terrain model
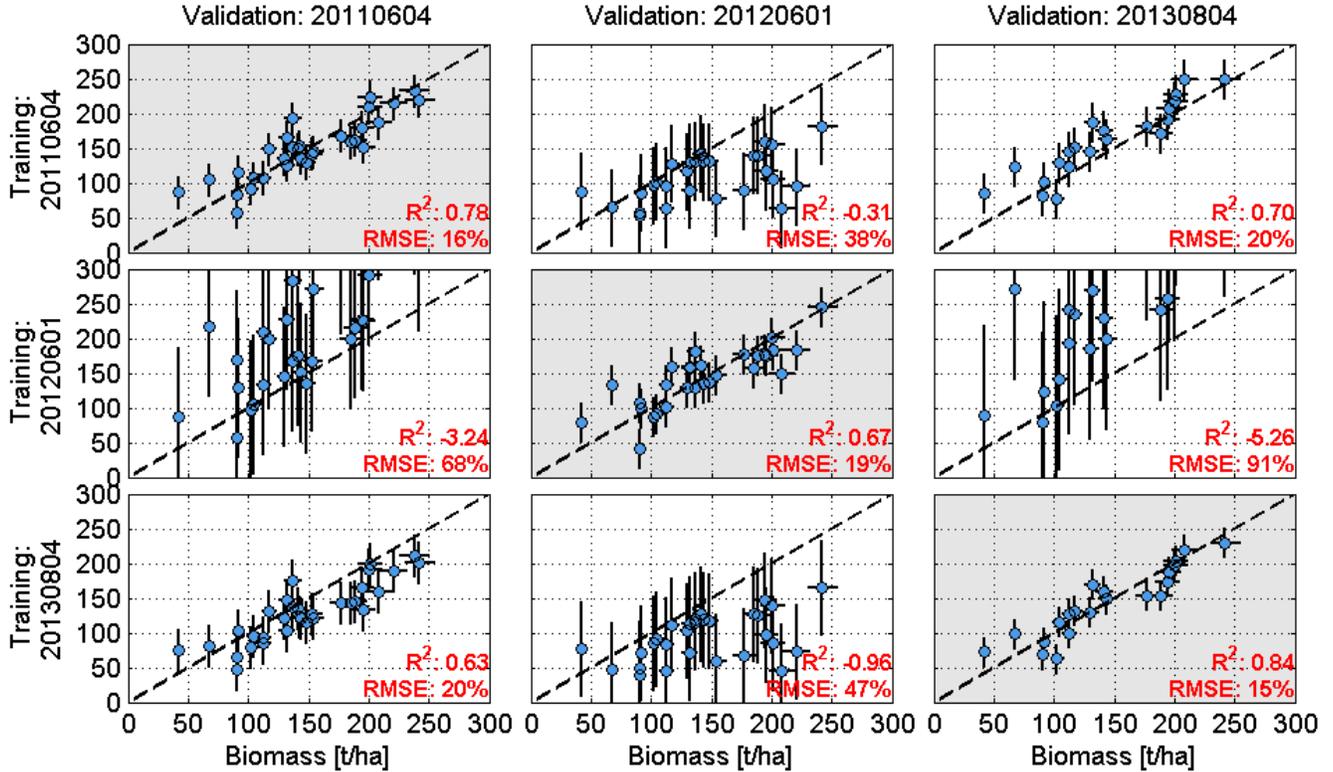
Figure 2 **Scatter plots for biomass estimation for different training/validation scenarios. Error bars are also shown: for in-situ data, a 5% error estimate was used (percentage of plot biomass), for modelled data, the corresponding RMSE value was used.**

(DTM) obtained using airborne lidar has been used. The DTM has a grid spacing of 2 m × 2 m and vertical accuracy better than 0.5 m, and it has been obtained from Swedish Land Survey, as part of the new national DTM. Ground phase has been subtracted from the TDM interferogram to create maps of the interferometric height. All processing has been carried out in radar geometry, and effects such as the quasi-bistatic acquisition geometry and satellite displacement during signal propagation have been accounted for. Geocoding accuracy has been verified using two 5-meter trihedral corner reflectors installed in Remningstorp. Geocoding accuracy has been estimated from ten images to better than 2 m. The standard deviation of height measurements was found to be 10 cm.

For each plot, a buffer zone of 5 m was added, and one single complex correlation coefficient was computed. The TLM was then inverted, and level distance and area-weighted backscatter ratios were obtained for all plots.

4. RESULTS

In Figure 1, the interferometric height, coherence, and the inverted TLM parameters are plotted against biomass. As it can be observed, the interferometric height is highly variable between the acquisitions at different HOAs. The inverted parameters are more stable, although at HOA=32 m, there is

a group of plots for which the uncorrected areafill decreases significantly. Note, that in 2012 and 2013, the amount of plots decreased from 32 to 29 and 21, due to harvesting and management procedures, which are not studied here.

Table 1 **The estimated parameter values are here shown, together with 2σ confidence interval.**

Training:	K	α	β
20110604	9.4 ± 4.5	1.2 ± 0.2	2.7 ± 0.4
20120601	0.2 ± 0.2	2.5 ± 0.4	1.8 ± 0.3
20130804	10.8 ± 7.8	$1.1 + 0.2$	3.2 ± 0.5

The model (1) was fitted to the data using non-linear least squares, which was chosen instead of linear regression of the log-transformed version of (1), to avoid logarithmic bias. In Table 1 Table 1, the estimated parameter values are shown, together with their 2σ confidence intervals ($\pm\sigma$, where σ is the standard deviation). A very high significance of both the whole model and each of the parameters was found.

In Figure 2, scatter plots for the residual study, in which the same data are used for training and validation, are shown on the diagonal. Fitting results are good for all three acquisitions.

Scatter plots for biomass prediction are shown off-diagonal in Figure 2. A different data set was used for training and validation. It can be observed that prediction results are very good for the combinations including the two acquisitions made in 2011 and 2013, while the prediction results for combinations including the acquisition made in 2012 are worse.

In Figure 2, root-mean-square errors (RMSEs) and coefficients of determination R^2 for all nine combinations are also shown. The model is able to explain between 67% and 84% of the variance in the data. Residual RMSE is between 15% and 19%, while prediction RMSE is good for acquisitions with the two largest HOAs. The results for combinations including the image with the lowest HOA are much worse.

It can be observed that the performance of the model may be dependent on HOA. A possible explanation for this could be that the exact modeling of the vertical scattering profile becomes more significant at low HOA, when phase variations in the vertical direction are faster. In that case, the two-level approach may be insufficient, and a more advanced vertical structure function should be used.

On the other hand, a study of meteorological data shows that the acquisition from 2012 was made after heavy rain. Therefore, this could also be a moisture effect. Note, that only for a subset of plots, the estimated uncorrected areafill factor drops significantly between the acquisitions from 2011 and 2012. Since μ is influenced by both the ground-to-vegetation backscatter ratio, and the canopy density [12], it is possible that different types of plots are affected differently by the rain, due to different soil types, for example.

However, since only one acquisition with low HOA has been studied, it is difficult to conclude which of the two effects is most responsible for the observed behavior.

A zero-intercept linear model that scales the interferometric height to biomass has been proposed in [2, 13]. Applied to the data used in this study, the residual RMSE is between 23% and 33%, and the prediction RMSE is between 23% and 36%. The linear model can explain only 40% and 59% of the variance in the images from 2011 and 2013, respectively, and 0% in the case of the image from 2012. The model proposed in this study shows significantly better performance.

REFERENCES

1. Krieger, G., et al., *TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry.* IEEE Transactions on Geoscience and Remote Sensing, 2007. **45**(11): p. 3317-3341.

2. Solberg, S., et al., *Estimating Spruce and Pine Biomass with Interferometric X-Band SAR.* Remote Sensing of Environment, 2010. **114**: p. 2353-2360.

3. Soja, M.J. and L.M.H. Ulander. *Digital Canopy Model Estimation from TanDEM-X Interferometry using High-Resolution Lidar DEM.* in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS).* 2013. Melbourne, Victoria, Australia.

4. Cloude, S.R. and K.P. Papathanassiou, *Polarimetric SAR Interferometry.* IEEE Transactions on Geoscience and Remote Sensing, 1998. **36**(5): p. 1551-1565.

5. Kugler, F., et al., *TanDEM-X Pol-InSAR Performance for Forest Height Estimation.* IEEE Transactions on Geoscience and Remote Sensing, 2014.

6. Soja, M.J., H. Persson, and L.M.H. Ulander. *Estimation of Boreal Forest Biomass from Two-Layer Model Inversion of Interferometric TanDEM-X Data.* in *accepted for Proceedings of IEEE Geoscience and Remote Sensing Symposium (IGARSS) 2014.* 2014.

7. Soja, M.J., H. Persson, and L.M.H. Ulander, *Estimation of Forest Height and Canopy Density from a Single InSAR Correlation Coefficient.* Geoscience and Remote Sensing Letters, 2014.

8. Soja, M.J. and L.M.H. Ulander. *Two-Level Forest Model Inversion of Interferometric TanDEM-X Data.* in *accepted for publication in proceedings for EUSAR 2014.* 2014.

9. Petersson, H., *Biomassafunktioner for tradfraktioner av tall, gran och bjork i Sverige*, 1999, SLU, Institutionen för skoglig resurshållning och geomatik, Arbetsrapport 59, Umeå.

10. Sandberg, G., et al., *L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest.* Remote Sensing of Environment, 2011. **115**(11): p. 2874-2886.

11. Duque, S., et al., *TanDEM-X Payload Ground Segment, CoSSC Generation and Interferometric Considerations*, 2012, Remote Sensing Technology Institute, German Aerospace Center (DLR).

12. Askne, J.I.H., et al., *Model-based biomass estimation of a hemi-boreal forest from multitemporal TanDEM-X acquisitions.* Remote Sensing, 2013. **5**: p. 5574-5597.

13. Solberg, S., et al., *Monitoring spruce volume and biomass with InSAR data from TanDEM-X.* Remote Sensing of Environment, 2013. **139**: p. 60-67.