COMPARING TANDEM-X INSAR FOREST STAND VOLUME PREDICTION MODELS TRAINED USING FIELD AND ALS DATA

Ritwika Mukhopadhyay[1], Mats Nilsson[1], Magnus Ekström[1,2], Eva Lindberg[1], Henrik J. Persson[1]

[1]Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden
[2]Department of Statistics, USBE, Umeå University, Sweden

ABSTRACT

Remote sensing (RS) techniques have been used for mapping forest variables, such as stem volume (important for forest management activities associated with timber production), over large areas which can be updated more frequently than with field inventory (FI) data. In this study, wall-to-wall TanDEM-X synthetic aperture radar images were used as auxiliary RS data for model-based prediction of stand-level volumes for two models, trained using volumes computed from FI (A) and airborne laser scanning estimations (B), respectively. The models were validated with harvester data available for independent stands. It was observed that the performance of model B was slightly better compared to model A based on adjusted R^2 and root mean squared error values. Therefore, it can be concluded that a completely RS based approach for prediction and mapping of stand volumes would be as promising as a method based on FI data along with being cost- and labour-efficient.

Index Terms— Airborne laser scanning, harvester data, stand volume, synthetic aperture radar, TanDEM-X.

1. INTRODUCTION

The use of auxiliary remote sensing (RS) data has been increasing over the past decades. RS data have been paired with field reference datasets for estimating forest variables such as, height, volume and aboveground biomass (AGB), in, e.g., [1], [2], [3]–[10], [11]–[17]. Stand- and tree-level volume estimates have been more extensively used by forest owners for measuring merchantable timber and forest products [18]. The collection of field inventory data is more cost and labour intensive when compared to acquiring aerial and spaceborne RS data, especially, in remote and inaccessible terrains.

Airborne laser scanning (ALS) data have been previously used for large scale mapping of AGB and volume [19], [20], but, using a wall-to-wall satellite image for large scale mapping of such forest variables would be even more time efficient [21], [22]. The German synthetic aperture radar (SAR) mission – TanDEM-X constitutes of a pair of satellites (launched in 2007 and 2010) carrying X-band SAR sensors, flying as an interferometer with global coverage. Several studies have been conducted using TanDEM-X for model-based inferences of forest AGB, volume and tree heights, in, e.g., [18], [23]–[29]. The use of TanDEM-X enables extraction of the vegetation height using the interferometric phase height (ph) and coherence [22]. The ph contains information about tree heights as well as forest density, which can be correlated to AGB and volume estimations [18], [22]. In some previous studies by [18], [30], these interferometric variables have been used to estimate forest volume at the stand level. Until now, manual field inventory data have been used as reference data for model-based estimations of volume, except for a handful of studies [2]. Still, models entirely based on RS data have not been implemented for large-scale mapping of forest variables.

Therefore, the main objective of this study was to use wall-to-wall TanDEM-X SAR data for modelling and mapping of stand-level forest volume across the entire test site and evaluating the performance of models A and B trained with stand-level volumes estimated from forest inventory data and volumes estimated based on ALS and national forest inventory (NFI) data, respectively. Stand volumes for independent forest stands within the same test site, estimated from harvesters during felling activities, were used for validating the models.

2. MATERIALS AND METHOD

2.1. Study area

The test site located in central Sweden, covers an area of 50,000 ha approximately (as shown in Fig 1a-b), majorly dominated by Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*), Lodgepole pine (*Pinus contorta*) and other deciduous tree species.
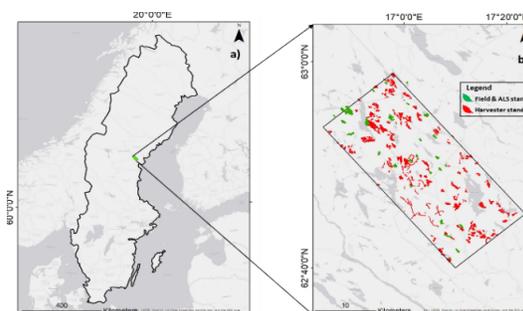
Figure 1. Represents the test site constituting the forest stands used for training models A and B (marked with 'green') and forest stands used for validation of the models available from the harvester data (marked with 'red').

Pine constituting around 50%, spruce around 44% and deciduous and other tree species constituting about 6% of the total species composition [2].

2.2. Field and Remote Sensing data
The field inventory data were acquired in 2019. Thirty stands were inventoried within the test site (marked with 'green' in Fig 1b). The inventory was done with an average of 8 circular plots (with 8m radius) distributed systematically across the stands. The distance between the plots and the number of plots varied for each stand depending on the stand area. The field inventory data were used as reference to select these 30 stands as training data since the field inventory data constituted observations over entire range of age-classes in the stands.

The ALS data were acquired in 2019 by a Leica ALS80 sensor from a 3000 m flight height and having an average point density of 1.5 points/m² scanning over entire Sweden. The mean volumes (in m³ha⁻¹) for the same individual 30 stands were extracted from the volumes estimated using the ALS metrics and the corresponding NFI plots over the scanned regions.

The TanDEM-X dataset was acquired on 14th November 2015 for HH (horizontally transmitted and horizontally received) polarization in strip-map mode. The SAR data specifications have been mentioned in Table 1.

Table 1. TanDEM-X SAR data specifications.

Polariza-tion	Pixel resolution		Multilook factor
	SLC*	Resampled	
HH	2.5(slant)×3.3 (azimuth) m²	10×10 m²	5×5

*SLC – single look complex.

The pre-processing was done as explained in [22] and similar image variables were derived, namely backscatter, ph and corrected coherence (c_coh). The complex interferogram was obtained as

$$\tilde{\gamma} = \frac{E[s_1 s_2^*]}{\sqrt{E[|s_1|^2]E[|s_2|^2]}}$$ (1)

where, $\tilde{\gamma}$ is the complex correlation co-efficient, $E[.]$ is the expectation value, * is the complex conjugate and s_1 and s_2 are the Hermitian product of the two complex SAR images [18], [31].

A minimum cost flow function was used for unwrapping the phase followed by a phase-to-height

sensitivity raster to obtain the height from the interferometric phase information [22].

The validation dataset consisted of 151 stands with volumes estimated from the harvester data (marked with 'red' in Fig 1b) acquired between 2019 and 2022. The average stand-level volume for the entire test site is 172.5 m³ha⁻¹ based on the harvester data accounting for only matured trees in stands. These 151 stands were checked and categorised into thinned (28) and clear-felled (123) stands. The thinned stands were discarded from the dataset to avoid representing over-estimated stand-volumes for such stands.

2.3. Volume estimation models
All the parameters derived from the TanDEM-X data were tested for statistical significance as model co-efficients of the explanatory variables for both model A and B. The final regression models A and B, represented in Eq. 2,

$$vol = \beta_0 + \beta_1 ph^{0.5} + \varepsilon$$ (2)

where, 'vol' represents the response variable (stand volume in this case), β_0 and β_1 are the model coefficients and ε is the random error.

The ph values ranged between [-5, 28]. Models A and B were compared based on adjusted co-efficient of determination (adj-R^2) and root mean squared error (RMSE) values with

$$RMSE = \sqrt{\frac{1}{n} \Sigma_{i=1}^{n}(y_i - \hat{y}_i)^2}$$ (3)

where y is the reference values, $\hat{y}$ is the predicted values, and n is the number of stands in the validation dataset.

3. RESULTS, DISCUSSION AND CONCLUSIONS

The results show that for both models A and B, expected value of stand volumes had dependency on the ph with a power of 0.5, as represented in Table 2. The power value of 0.5 for ph was a slight deviation compared to previous studies, e.g., [18], [22], [28]. The corresponding model coefficients of ph and c_coh were statistically significant in the models but c_coh did not contribute in improving the model prediction accuracy, as observed similar to [22], and was therefore, not included in the models. Also, both models A and B were formulated with intercepts non-significantly different from 0. The statistical summary of model validation has been presented in Table 3. The predictions of model B were slightly more accurate based on the adj-R^2 and RMSE values when compared to that of Model A. From Fig 2, there is no distinct difference between models A and B in the trend of the two plots representing the relation between the predicted stand volumes against the observed stand volumes. The predicted values in the lower range of volume were over-estimated in both cases. The overestimation of these stands with low

stand volumes might be due to that thinning activities were carried out which could not be filtered out during the categorising of the validation dataset or might be due to measurement errors in the harvester volumes. The relationship between the predicted and the observed volumes are almost linear for stand volumes ranging between 150 m^3ha^{-1} and 400 m^3ha^{-1} . Stand volume maps for the entire test site were predicted for both the models, as presented in Fig 3.

The approach based on laser scanning estimates (method B) appeared as accurate and promising as the field inventory based approach (method A) for mature stands. Therefore, this study indicates that dedicated field inventories intended to train models to estimate forest volume wall-to-wall based on TanDEM-X could be eliminated and replaced by estimates from low-resolution LiDAR and NFI data. This reduces costs and makes forest planning and decision making more efficient without compromising the accuracy of the stand volume estimates. The approach needs to be tested further in other sites and the temporal robustness of the model parameters needs to be further analysed.

Table 2. Summary of model parameters.

Model	Intercept	$ph^{0.5}$
A	-82.37˙	113.2***
B	-62.30˙	104.3***

˙= p>0.05 and ***=p≤0.001.

Table 3. Summary of statistics of volume prediction model validation.

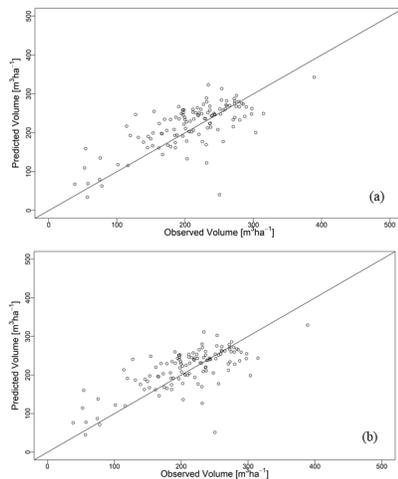| Model	adj-R^2	RMSE (m^3ha^{-1})	n
A	0.58	46.5 (22.3%)	123
B	0.60	44.6 (21.4%)	123

Figure 2. Scatterplots of observed volume vs predicted volume for: (a) model A and (b) model B.
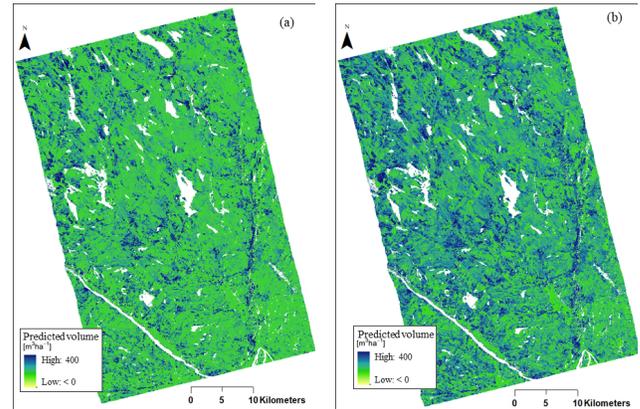

Figure 3. Wall-to-wall prediction maps of volume: (a) model A and (b) model B.

4. ACKNOWLEDGEMENTS

5. REFERENCES

[1]   J. Holmgren, "Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning," *Scand. J. For. Res.*, vol. 19, no. 6, pp. 543–553, 2006, doi: 10.1080/02827580410019472.

[2]   H. J. Persson, K. Olofsson, and J. Holmgren, "Two-phase forest inventory using very-high-resolution laser scanning," *Remote Sens. Environ.*, vol. 271, p. 112909, Mar. 2022, doi: 10.1016/J.RSE.2022.112909.

[3]   M. J. Soja, H. J. Persson, and L. M. H. Ulander, "Estimation of forest biomass from two-level model inversion of single-pass InSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 5083–5099, 2015, doi: 10.1109/TGRS.2015.2417205.

[4]   S. Saarela *et al.*, "Model-assisted estimation of growing stock volume using different combinations of LiDAR and Landsat data as auxiliary information," *Remote Sens. Environ.*, vol. 158, pp. 431–440, Mar. 2015, doi: 10.1016/J.RSE.2014.11.020.

[5]   R. O. Dubayah *et al.*, "Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica," *J. Geophys. Res. Biogeosciences*, vol. 115, no. G2, p. n/a-n/a, Jun. 2010, doi: 10.1029/2009JG000933.

[6]   M. Egberth *et al.*, "Combining airborne laser scanning and Landsat data for statistical modeling of soil carbon and tree biomass in Tanzanian Miombo woodlands," *Carbon Balance Manag.*, vol. 12, no. 1, pp. 1–11, Apr. 2017, doi: 10.1186/S13021-017-0076-Y.

[7]   R. Økseter, O. M. Bollandsås, T. Gobakken, and E.

Næsset, "Modeling and predicting aboveground biomass change in young forest using multi-temporal airborne laser scanner data," *Scand. J. For. Res.*, vol. 30, no. 5, pp. 458–469, Jul. 2015, doi: 10.1080/02827581.2015.1024733.

[8] T. Gobakken, O. M. Bollandsås, and E. Næsset, "Comparing biophysical forest characteristics estimated from photogrammetric matching of aerial images and airborne laser scanning data," *Scand. J. For. Res.*, vol. 30, no. 1, pp. 73–86, Jan. 2014, doi: 10.1080/02827581.2014.961954.

[9] E. Næsset, "Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 88–99, Apr. 2002, doi: 10.1016/S0034-4257(01)00290-5.

[10] J. Esteban, R. E. McRoberts, A. Fernández-Landa, J. L. Tomé, and E. Næsset, "Estimating Forest Volume and Biomass and Their Changes Using Random Forests and Remotely Sensed Data," *Remote Sens. 2019, Vol. 11, Page 1944*, vol. 11, no. 16, p. 1944, Aug. 2019, doi: 10.3390/RS11161944.

[11] F. Garestier, P. C. Dubois-Fernandez, D. Guyon, and T. Le Toan, "Forest biophysical parameter estimation using L- and P-band polarimetric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 10, pp. 3379–3388, Oct. 2009, doi: 10.1109/TGRS.2009.2022947.

[12] M. A. Stelmaszczuk-Górska, M. Urbazaev, C. Schmullius, and C. Thiel, "Estimation of above-ground biomass over boreal forests in Siberia using updated in Situ, ALOS-2 PALSAR-2, and RADARSAT-2 data," *Remote Sens.*, vol. 10, no. 10, Oct. 2018, doi: 10.3390/rs10101550.

[13] M. J. Soja, H. J. Persson, and L. M. H. Ulander, "Estimation of forest height and canopy density from a single InSAR correlation coefficient," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 646–650, 2015, doi: 10.1109/LGRS.2014.2354551.

[14] J. Carreiras, J. Melo, and M. Vasconcelos, "Estimating the Above-Ground Biomass in Miombo Savanna Woodlands (Mozambique, East Africa) Using L-Band Synthetic Aperture Radar Data," *Remote Sens.*, vol. 5, no. 4, pp. 1524–1548, Mar. 2013, doi: 10.3390/rs5041524.

[15] T. Mette, K. Papathanassiou, and I. Hajnsek, "Biomass estimation from polarimetric SAR interferometry over heterogeneous forest terrain," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2004, vol. 1, pp. 511–514, doi: 10.1109/igarss.2004.1369076.

[16] E. Næsset *et al.*, "Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: A case study from a boreal forest area," *Remote Sens. Environ.*, vol. 115, no. 12, pp. 3599–3614, Dec. 2011, doi: 10.1016/J.RSE.2011.08.021.

[17] C. Thiel and C. Schmullius, "The potential of ALOS PALSAR backscatter and InSAR coherence for forest growing stock volume estimation in Central Siberia," *Remote Sens. Environ.*, vol. 173, pp. 258–273, Feb. 2016, doi: 10.1016/J.RSE.2015.10.030.

[18] H. J. Persson and J. E. S. Fransson, "Comparison between TanDEM-X and ALS based estimation of above ground biomass and tree height in boreal forests," *Scand. J. For. Res.*, vol. 32, no. 4, pp. 306–319, 2017, doi: 10.1080/02827581.2016.1220618.

[19] M. Nilsson *et al.*, "A nationwide forest attribute map of Sweden predicted using airborne laser scanning data and field data from the National Forest Inventory," *Remote Sens. Environ.*, vol. 194, pp. 447–454, Jun. 2017, doi: 10.1016/j.rse.2016.10.022.

[20] M. Maltamo, O. M. Bollandsås, T. Gobakken, and E. Næsset, "Large-scale prediction of aboveground biomass in heterogeneous mountain forests by means of airborne laser scanning," *Can. J. For. Res.*, vol. 46, no. 9, pp. 1138–1144, 2016, doi: 10.1139/CJFR-2016-0086.

[21] H. J. Persson, M. J. Soja, J. E. S. Fransson, and L. M. H. Ulander, "National Forest Biomass Mapping Using the Two-Level Model," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 6391–6400, 2020, doi: 10.1109/JSTARS.2020.3030591.

[22] H. J. Persson, H. Olsson, M. J. Soja, L. M. H. Ulander, and J. E. S. Fransson, "Experiences from Large-Scale Forest Mapping of Sweden Using TanDEM-X Data," *Remote Sens. 2017, Vol. 9, Page 1253*, vol. 9, no. 12, p. 1253, Dec. 2017, doi: 10.3390/RS9121253.

[23] J. Praks, M. Hallikainen, O. Antropov, and D. Molina, "Boreal forest tree height estimation from interferometric TanDEM-X images," *Int. Geosci. Remote Sens. Symp.*, pp. 1262–1265, 2012, doi: 10.1109/IGARSS.2012.6351309.

[24] F. Kugler, D. Schulze, I. Hajnsek, H. Pretzsch, and K. P. Papathanassiou, "TanDEM-X Pol-InSAR performance for forest height estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6404–6422, 2014, doi: 10.1109/TGRS.2013.2296533.

[25] J. I. H. Askne, J. E. S. Fransson, M. Santoro, M. J. Soja, and L. M. H. Ulander, "Model-Based Biomass Estimation of a Hemi-Boreal Forest from Multitemporal TanDEM-X Acquisitions," *Remote Sens. 2013, Vol. 5, Pages 5574-5597*, vol. 5, no. 11, pp. 5574–5597, Oct. 2013, doi: 10.3390/RS5115574.

[26] A. Olesk, J. Praks, O. Antropov, K. Zalite, T. Arumäe, and K. Voormansik, "Interferometric SAR coherence models for Characterization of hemiboreal forests using TanDEM-X dssata," *Remote Sens.*, vol. 8, no. 9, Sep. 2016, doi: 10.3390/RS8090700.

[27] K. Karila, M. Vastaranta, M. Karjalainen, and S. Kaasalainen, "Tandem-X interferometry in the prediction of forest inventory attributes in managed boreal forests," *Remote Sens. Environ.*, vol. 159, pp. 259–268, Mar. 2015, doi: 10.1016/J.RSE.2014.12.012.

[28] S. Solberg, R. Astrup, J. Breidenbach, B. Nilsen, and D. Weydahl, "Monitoring spruce volume and biomass with InSAR data from TanDEM-X," *Remote Sens. Environ.*, vol. 139, pp. 60–67, Dec. 2013, doi: 10.1016/J.RSE.2013.07.036.

[29] S. Abdullahi, F. Kugler, and H. Pretzsch, "Prediction of stem volume in complex temperate forest stands using TanDEM-X SAR data," *Remote Sens. Environ.*, vol. 174, pp. 197–211, Mar. 2016, doi: 10.1016/J.RSE.2015.12.012.

[30] J. Rahlf, J. Breidenbach, S. Solberg, E. Næsset, and R. Astrup, "Comparison of four types of 3D data for timber volume estimation," *Remote Sens. Environ.*, vol. 155, pp. 325–333, Dec. 2014, doi: 10.1016/J.RSE.2014.08.036.

[31] J. Sen Lee and E. Pottier, *Polarimetric radar Imaging - from Basics to Applications*. New York, USA: Taylor & Francis Group, 2009.