Title	Dual-band radar estimation of stem volume in Boreal forest
Author(s)	Rauste, Yrjö; Astola, Heikki; Ahola, Heikki; Häme, Tuomas; von Poncet, Felicitas
Citation	Proceedings of the ESA Living Planet Symposium, Bergen, Norway, 28 June - 2 July 2010, ISBN 978-92-9221-250-6
Date	2010
Rights	Copyright © 2010 ESA. This article may be downloaded for personal use only

VTT
http://www.vtt.fi
P.O. box 1000
FI-02044 VTT
Finland

DUAL-BAND RADAR ESTIMATION OF STEM VOLUME IN BOREAL FOREST

Yrjö Rauste[1], Heikki Astola[1], Heikki Ahola[1], Tuomas Häme[1], and Felicitas von Poncet[2]
[1] *VTT, P.O. Box 1000, 02044 VTT, Finland; yrjo.rauste@vtt.fi*
[2] *InfoTerra GmbH, 88039 Friedrichshafen, Germany*

ABSTRACT

Forest stem volume information is needed in planning of sustainable forestry, mapping of exploitable forest resources, carbon balance studies and many other environmental applications. Radar sensors offer an efficient and weather-independent means for stem volume mapping.

The radar dataset consisted of an ALOS/PALSAR dual-polarised scene from September 2008 and four TerraSAR-X spotlight scenes from February-March and July-August 2009. Ground data consisted of plot-wise data and stand-wise data. Regression models were developed with stand-wise training data where the stem volume varied between 0 and 390 m^3/ha. The best 3-predictor model - 2 ALOS-PALSAR amplitudes and the phase difference between HH and VV data in a TerraSAR-X scene - produced an RMSE of 46 m^3/ha (R^2 = 0.7) when evaluated against stands not used in the model training.

Stem volume estimation with plot-wise ground data produced lower estimation accuracies. The main reason was most likely mis-registration between the opposite-looking ALOS/PALSAR and TerraSAR-X scenes, which was caused by canopy height and which was not corrected by ortho-rectification with terrain elevation model. In future work, similar look directions should be used when no canopy surface model is available for ortho-rectification. A filtering approach was developed for using the stand-wise stem-volume model in areas with no forest stand information.

Key words: radar; forestry; biomass; TerraSAR-X; ALOS.

1. INTRODUCTION

As a weather-independent sensor, radar offers flexibility in forest inventory and monitoring applications. The sensitivity of radar to forest biomass has been documented in P-band (e.g. [1]), L-band (e.g. [2]), and X-band (e.g. [3]).

Project NewForest aims at comparing optical satellite imagery, radar satellite imagery, and air-borne lidar data in forest inventory applications. The objective of the radar study in project NewForest was to study the potential of TerraSAR-X and ALOS/PALSAR data in forest inventory in managed Boreal forest.

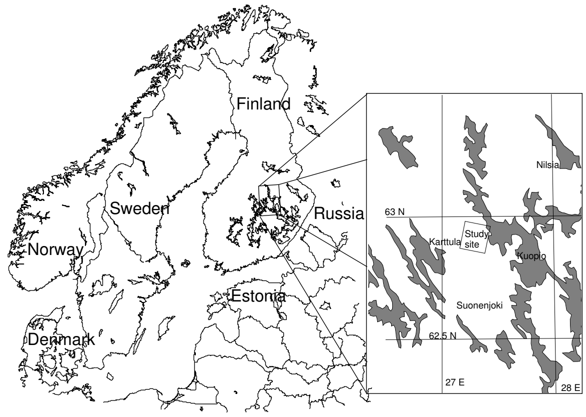


Figure 1. Location of study site in Karttula, Finland.

2. MATERIALS AND METHODS

2.1. Study Site and Ground data

The study site (centre 62°54'55".7 north, 27°17'37".6 east) was in Karttula (close to Kuopio) in Finland (Fig. 1). The area is mostly covered with forest (main tree species: spruce *Picea abies*, pine *Pinus sylvestris*, and birch *Betula pendula* and *Betula pubescens*. The dominating soil type is glacial drift, but marsh lands and lakes also occur. Agricultural areas cover only a small fraction of the study site.

Forest inventory ground data consisted of plot-wise data and stand-wise data. The plot-wise data consisted of two datasets: 1) a training set of 164 plots scattered across the area covered by the TerraSAR-X scenes, and 2) a reference set of 178 plots measured in 29 stands in the area (3 to 10 plots per stand). The average stem volume in the training set was 177 m^3/ha and in the reference set 203 m^3/ha. Both of these figures are above the average stem

Proc. 'ESA Living Planet Symposium', Bergen, Norway
28 June – 2 July 2010 (ESA SP-686, December 2010)

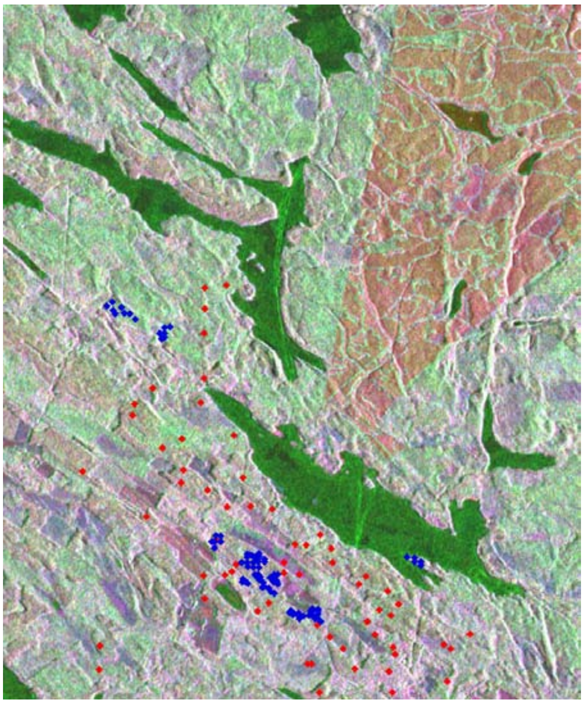


Figure 2. Plot-wise and stand-wise reference data overlaid with TerraSAR-X image. Blue dot = validation plot, red dot = training plot, red mask = stand data. Image data (image width = 5 km): R=t20090801VV, G=t20090228HH, B=t20090721HH. TerraSAR-X data InfoTerra GmbH 2009.

volume in managed forests in southern Finland. The forest in the area can be characterized as very variable. For instance, the stem volume varied a lot within the stands in the reference dataset. The standard deviation of stem volume varied from 9 percent of the mean to 43 percent of the mean stem volume.

The plot-wise ground data were screened for plots so close to a visible stand boundary (or some radar artifact area like radar shadow or strong double bounce from a forest edge) that it could affect the SAR data. After this screening, 94 plots remained in the training dataset and 118 plots in the reference dataset.

The stand-wise reference data covered 253 stands (of 2 ha or more), which were divided to two datasets: 127 stands for training (area-weighted mean stem volume 108 m^3/ha) and 126 for validation (mean stem volume 98 m^3/ha). The division was done by sorting the stands with stem volume, and selecting every second for training (so that the highest stem volume stand was included in the training set). The stand-wise data were provided by Tornator Oy. Fig. 2 shows a sample of ground data on top of a TerraSAR-X colour composite.

A digital elevation model (DEM) was obtained from Tapio Forest Development Centre. This DEM was derived from lidar data acquired for forest inventory purposes. The pixel spacing of the DEM was 2 metres, which was then interpolated (using cubic spline interpolation) to 6.25 metres before ortho-rectification of TerraSAR-X data. The pixel spacing of the DEM was further down-averaged to 12.5 metres for the ortho-rectification of ALOS/PALSAR data. Terrain elevation varied between 82 and 190 metres above sea level in the area of the TerraSAR-X images. The elevations in the DEM represented ground surface, not surface of the canopy. The DEM covered the coordinates N: 6971628.125 ... 6986709.375, E: 3507690.625 ... 3522171.875 in the Finnish coordinate system called “Yhtenaiskoordinaatisto” (KKJ uniform grid coordinate system).

2.2. SAR Data

Four dual-polarized (HH+VV) TerraSAR-X scenes were used as well as one ALOS/PALSAR dual-polarized (HH+HV) scene (Tab. 1). TerraSAR-X scenes were spotlight scenes (spot_049 with an incidence angle of 39°) with pixel spacing of 2.7 m in azimuth and 0.9 m in slant range (nominal single look resolution 3.2 m in azimuth and 1.9 m in ground range). The ALOS/PALSAR scene had an incidence angle of 39°and pixel spacings of 3.2 m in azimuth and 9.4 m in slant range in an SLC (Single Look Complex) product called Level 1.1 product (nominal single look resolution 4.5 m in azimuth and 15 m in ground range). The ALOS/PALSAR data were averaged over 5 lines in azimuth to obtain pixel spacings of 16 m in azimuth and 14.9 m in ground range before ortho-rectification. Re-sampling in connection with ortho-rectification was done by bi-linear interpolation, which introduced further averaging of the data.

Tab. 1 gives the acquisition times of all scenes in UTC. The orbit direction is also indicated: A = ascending, D = descending. Column Weather (from Kuopio airport appr. 20 km from the site) gives temperature in centigrade and snowing (s) or raining (r). In addition to TerraSAR-X scenes listed in Tab. 1, two more scenes were acquired in May 2010 for tree-height estimation by radargrammetric techniques. Since the acquisition of these scenes was delayed due to meteorological reasons (dry and snow-free canopy was required) the results of tree-height estimation could not be included in this paper.

Table 1. SAR data of Karttula/Kuopio study site.

Satellite	Acquisition Time (UTC)	Orbit	Weather
ALOS	2008-09-02 20:13	A	11
TsX	2009-02-28 04:39	D	3
TsX	2009-03-11 04:39	D	-6, s
TsX	2009-07-21 04:39	D	15, r
TsX	2009-08-01 04:39	D	18

2.3. Ortho-Rectification

TerraSAR-X scenes were ortho-rectified without ground control points or tie points between scenes. The ALOS/PALSAR scene was combined from two frames along the same orbit. A set of tie points were measured in the overlap area between the frames. A set of 8 ground control points were measured between the ALOS/PALSAR scene and Finnish topographic maps in a web service called “Kansalaisen karttapaikka” (map site of National Land Survey of Finland). The geo-coding by these GCPs was later revised (translation 21.2 m to north, 26.7 m to east) based on visual evaluation between the ortho-rectified PALSAR scene, ortho-rectified TerraSAR-X data, and the stand map of the stand-wise forest inventory data.

In connection with ortho-rectification, radiometric correction was made to correct for backscatter amplitude variation due to terrain topography [4].

2.4. Computation of SAR Features

In addition to the amplitude of all scenes, the following derived features were computed for each TerraSAR-X scene:

1. coherence magnitude between HH and VV bands, and
2. phase of the HH-VV coherence (= phase difference between the HH and VV bands).

All these derived features were computed in a 5-by-5 window, in the original resolution, after which these features were ortho-rectified the same way as the amplitude bands. Since the ALOS/PALSAR data did not include two co-polarized bands, no derived features were computed for PALSAR data.

ASCII text files (containing forest variable values and SAR feature values) were made (with plot-wise ground data) using a circular sampling window and a weighted average with Gaussian weights. The standard deviation of the weight was set to 18 metres for ALOS/PALSAR sampling and to 12 metres in TerraSAR-X sampling. Within each sampling window (and SAR feature), the highest and lowest value were discarded in the computation of the weighted average.

For stand-wise ground data, the SAR feature values were averaged within each stand. Only stands of 2 ha or more were included. The stand map was eroded by one pixel (6 m) before stand averaging.

2.5. Estimation and Classification Methods

Step-wise linear regression was used as the estimation method for forest stem volume. The statistical package R (function stepAIC from package MASS) was used to derive regression models.

Maximum likelihood classification was used for a land-cover map for class-wise filtering of SAR data outside the area of stand map data.

2.6. Evaluation of Estimation Performance

The estimation performance in stem volume was evaluated by two quantities: the residual mean square error (RMSE) and the coefficient of determination (R^2). RMSE was computed:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(e_i - g_i)^2}{n}} \qquad (1)$$

where e is the estimated stem volume, g is the ground-measured stem volume, and n is the number of observations. The coefficient of determination (R^2), which is usually used as a measure of goodness of fit for a regression model in a training dataset, was also applied to reference datasets:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(e_i - g_i)^2}{\sum_{i=1}^{n}(g_i - \bar{g})^2} \qquad (2)$$

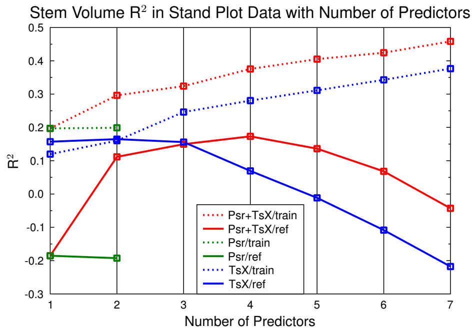
where $\bar{g}$ is the average ground-measured stem volume in the dataset.

3. RESULTS

3.1. Stem Volume Estimation with Plot Data

Plot data were processed in step-wise linear regression. Fig. 3 shows the development of R^2 and RMSE as a function of predictors, both for the training dataset and the reference dataset. Three series of predictors were used: 1) the whole set of TerraSAR-X variables and PALSAR variables, 2) TerraSAR-X variables only, and 3) PALSAR variables only. Negative R^2 values are impossible in the dataset that was used to derive the linear regression model. When the model is applied in a new dataset, the R^2 value can go negative. In all cases, the R^2 value of the reference dataset was below 0.2. The RMSE values in reference dataset were in most cases lower than those of the training set. This is most likely due to the higher variation in the training dataset, which was augmented with open area plots to have zero-stem volume plots in the dataset. The standard deviation of stem volume in the training set was 126 m^3/ha and in the reference set 104 m^3/ha. After three or four predictors, the difference in

R^2



RMSE

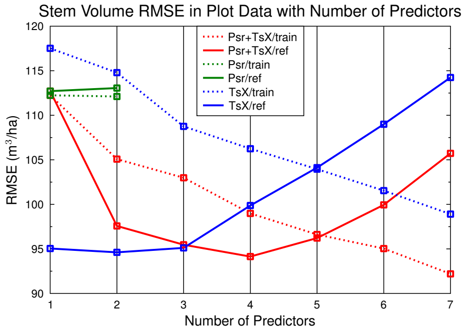


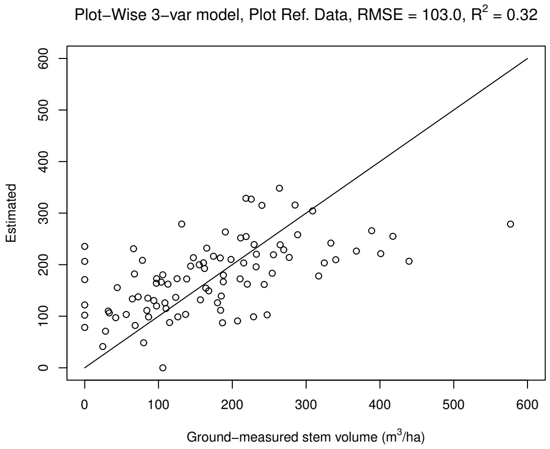
Figure 3. R^2 (upper) and RMSE (lower) in stem volume estimation in plot data as a function of number of SAR features used as predictors (Psr = Palsar, Tsx = TerraSAR-X.

training-set R^2 and the reference-set R^2 begins to grow, which can be seen as a sign of over-fitting.

Fig. 4 shows the scatter plot between estimated and ground-measured stem volume for the plot-wise training and reference datasets for a 3-predictor model.

One reason for the poor performance of SAR data in forest stem volume estimation in the plot-wise dataset is the uncertainty in image registration. The SAR dataset consisted of TerraSAR-X data with look direction from east to west and ALOS/PALSAR data with look direction from west to east. Fig. 5 shows the relief displacements of SAR images with opposite look directions. If terrain elevation is used in ortho-rectification, the relief displacement vectors (due to the uncorrected canopy height) point to opposite directions. The mis-registration can be twice the canopy height or even more with lower incidence angles.

The effects of opposite look directions could be corrected in ortho-rectification if a surface model (representing the top of the canopy) could be used instead of terrain elevation model. However, making an up-to-date forest surface model a pre-requisite for stem volume mapping is unrealistic in many cases. If the surface model is produced by costly air-borne lidar, the stem volume mapping can be made with the same lidar dataset. Surface models can be produced by single-pass interferometric SAR systems like Shuttle radar and the current TanDEM-X mission. Frequent repetition of such surveys is unlikely, however. The simplest way of avoiding co-registration problems is to use images with the same look direction in all images of a dataset. If the incidence angles are equal, the relief displacements are identical. If there is a difference in incidence angle, relief displacements differ but much less than in the case of opposite look directions.



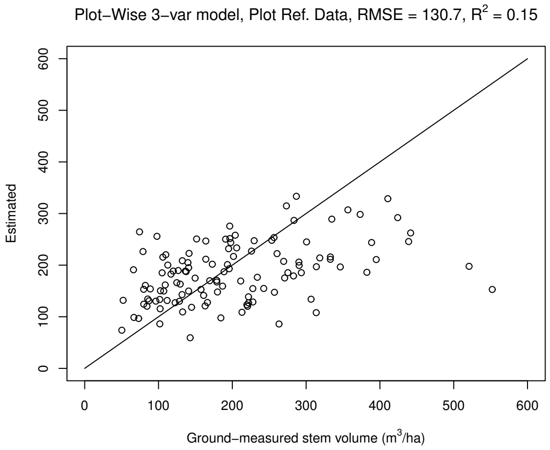


Figure 4. Estimated stem volume as a function of ground-measured stem volume using a linear regression model with 3 predictors (HV amplitude of Palsar and HH amplitudes of TerraSAR-X scenes 2009-03-11 and 2009-07-21) in the training dataset (upper) and in the reference dataset (lower).

3.2. Stem Volume Estimation with Stand Data

Stand-wise forest inventory data form a dataset where the stem volume (and other forest characteristics) are averaged over a much larger area than in plot-wise data. As shown in Fig. 6, this led to much better performing re-

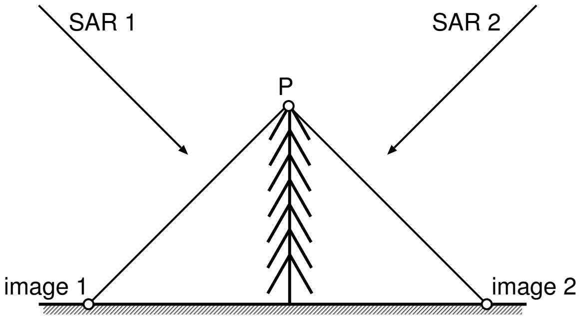
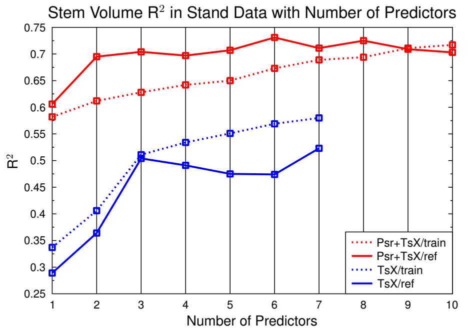


Figure 5. Relief displacements in case of opposite look directions in two SAR images. When using terrain elevation in ortho-rectification, the uncorrected canopy height causes relief displacement in opposite directions in the SAR images. If the incidence angle is 45°, the difference between the two image positions of point P is twice the canopy height. With lower incidence angles, the difference is even higher.

R^2



RMSE

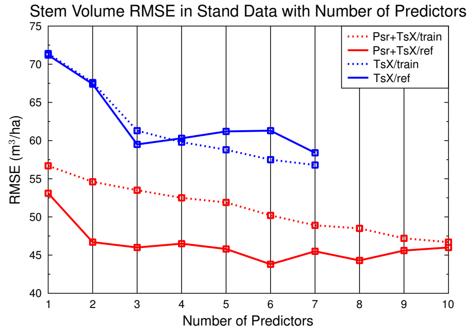


Figure 6. R^2 (upper) and RMSE (lower) in stem volume estimation in stand data as a function of number of SAR features used as predictors.

gression models in stem volume estimation. Fig. 6 does not show a separate line for the 2-predictor case of PALSAR data because the first two predictors in the combined dataset were the HV and HH amplitudes from PALSAR.

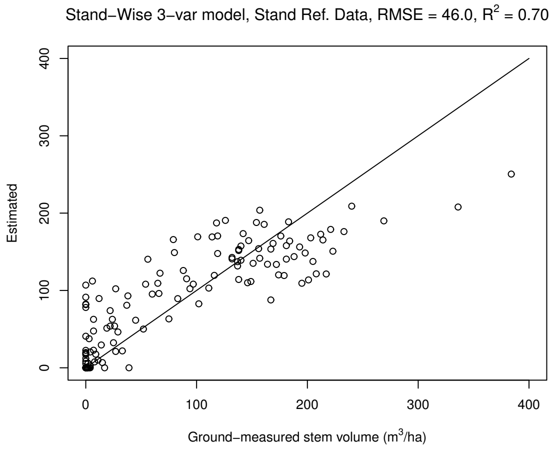


Figure 7. Forest stem volume results of a regression model including three predictors: Palsar/HV, Palsar/HH, and the TerraSAR-X phase difference between HH and VV on 2009-08-01. The reference data are the half of the stand-wise forest inventory data that was not used in derivation of the regression model.

Fig. 7 shows the scatter plot for the 3-variable model in the stand-wise reference dataset. The R^2 of the linear stem volume model was 0.7 and the RMSE was 46 m^3/ha in a dataset where the stem volume ranged from 0 to 390 m^3/ha. The predictor seems to saturate at a level close to 200 m^3/ha.

Testing with plot-wise reference data produced a negative R^2 value even when averaging over plots within the same forest stand. This can be expected in this type of small stands with irregular shapes. Averaging over a small stand does not remove the uncertainty in pixel-level image registration, which is caused by the uncorrected tree height in opposite-looking SAR images.

3.3. Stem Volume Estimation for Areas with no Stand Map

As shown in the previous sections, a reasonable forest stem volume estimation with the current SAR dataset of the study site was possible only when estimating stand mean stem volumes. In view of using radar in estimating stem volume for larger, unknown areas, this approach would require a stand mask covering the whole area. A stand mask could be generated by using some segmentation algorithm on optical satellite imagery. As radar techniques are mostly needed in areas where availability of optical satellite imagery is poor, a pure radar approach should also be available.

What the stand-wise estimation method does is averaging of backscatter observations (and stem volume values) over a large area in such a way that the averaging area is meaningful for stem volume mapping. The same averaging effect can be achieved also without a stand map if averaging over boundaries of spatial units with highly different stem volume can be avoided. The following algorithm was developed for stem volume estimation over the whole area of the TerraSAR-X scenes:

- a radar-based map was made including the land cover classes: forest, young forest, and non-forest,
- the SAR scenes were filtered in a moving window, averaging only within those pixels in the window that belonged to the class of the centre pixel, and
- the stand-developed stem-volume estimation model was applied to the filtered SAR images.

The algorithm above was applied by making a maximum likelihood classification using ALOS/PALSAR bands HH and HV and the first principal component over all amplitude bands in the dataset. Training areas were manually selected for two young forest classes, two forest classes, one agriculture class, and two water classes. The averaging was done in a 15-by-15 window with Gaussian weights. The standard deviation of the Gaussian weight was set to 7 pixels for those pixels belonging to the same class (from {forest, young forest}) and to 4 pixels for those pixels belonging to the opposite class (from {forest, young forest}).

Fig. 8 shows the classification map and the stem volume estimate map using the filtered SAR dataset. In qualitative visual evaluation, the stem volume map mostly seems to correspond to known stem volume data within the stand-wise estimation accuracy. The classification map includes some obvious classification errors (for example some shore line areas classified as agriculture), but this has very little effect on the stem volume estimate (these isolated error pixels and small patches are practically not filtered at all). The resulting stem volume estimate includes some radar artefacts - like high stem volume estimates along shorelines (due to strong backscatter from the water-tree-stem double bounce return).

4. DISCUSSION

Stand-wise ground data produced a level of stem volume estimation performance suitable for wide-area applications, where accuracy requirements are not very stringent.

The combination of ALOS/PALSAR data and TerraSAR-X data produced a better stem volume estimation performance than either of the sensors alone. The spatial resolution of the stem volume map was better than similar maps produced with ALOS/PALSAR data only.

Plot-wise stem volume estimation produced poor accuracies. The main reason was the uncertainty in image registration caused by canopy height and the opposite look directions of the ALOS/PALSAR and TerraSAR-X data. When planning radar acquisitions for multi-satellite methods, it is beneficial if the look direction of all satellites is close to each other. In SAR combinations that include ALOS/PALSAR, the acquisitions should be planned on ascending orbits because descending-orbit PALSAR data is very rare due to choices made in the systematic acquisition planning for ALOS data. Identical or close look directions between all SAR datasets minimizes the effects of canopy height on the co-registration between radar datasets.

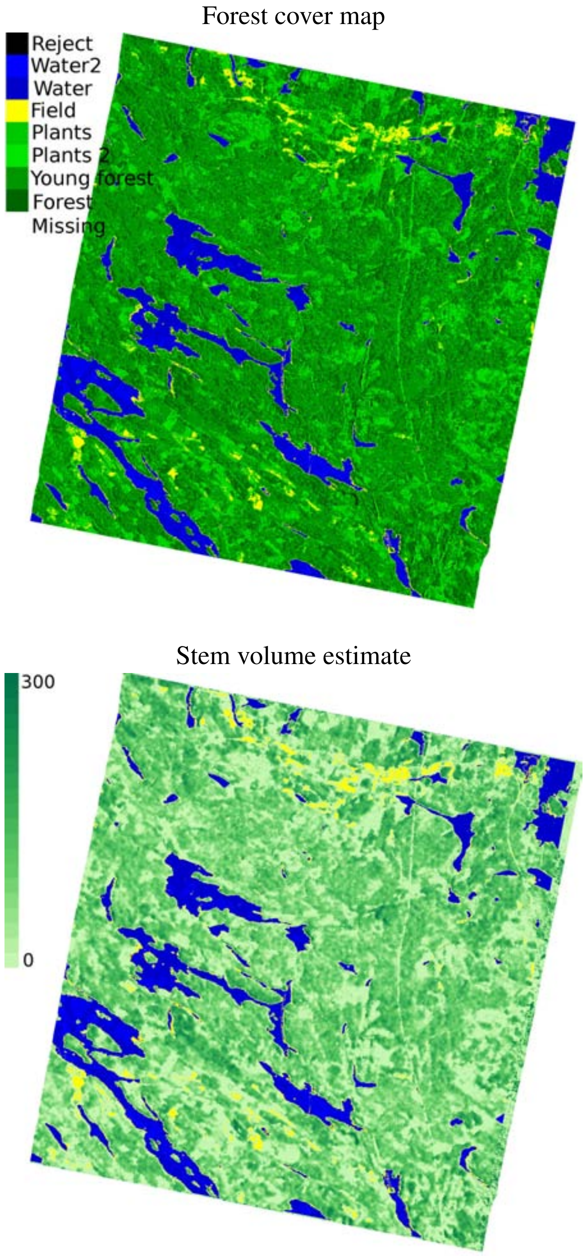


Figure 8. SAR-derived forest cover map (upper) and stem volume estimate using class-wise smoothed SAR data (lower).

The produced forest stem volume maps included effects of high double-bounce backscatter at forest/non-forest boundaries. If the stem volume map is only used visually, an expert user can make real-time interpretation and ignore high stem volume values e.g. along lake shores on the lake side opposite to the radar look direction. If the stem volume maps are used in computerized contexts the anomalies due to edge effects should be removed from the stem volume maps. Algorithms could be devised to adjust the observed backscatter in places where the back scatter has a systematic sharp increase when going from near pixels to far pixels in a non-rectified SAR scene. These algorithms should tolerate noise, which in many forest areas can be significant even after moderate averaging of SAR data.

The performance of radargrammetric techniques in tree height estimation will be evaluated in future.

ACKNOWLEDGEMENTS

This work was carried out in project NewForest, which was funded by the Finnish Funding Agency for Technology and Innovation (TEKES). TerraSAR-X data were provided by InfoTerra GmbH as a part of project "TerraSAR-X Data Evaluation in the context of boreal forest inventories". Alos/Palsar data, which were acquired by the Japanese space agency JAXA, were provided by the European Space Agency in the context of ALOS/Aden AO project AOALO.3713.

REFERENCES

[1] Thuy Le Toan, Andre Beaudoin, J. Riom, and D. Guyon. Relating forest biomass to SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2):403–411, 1992.

[2] Yrjö Rauste. Multi-temporal JERS SAR data in boreal forest biomass mapping. *Remote Sensing of Environment*, 97(2):263–275, 2005.

[3] Yrjö Rauste and Anne Lönnqvist Heikki Ahola. Mapping boreal forest biomass with imagery from polarimetric and semi-polarimetric sar sensors. *Ambiencia*, 4(Edicao Especial 2008):171–180, 2008.

[4] Yrjö Rauste, Anne Lönnqvist, Matthieu Molinier, Jean-Baptiste Henry, and Tuomas Häme. Orthorectification and terrain correction of polarimetric SAR data applied in the ALOS/Palsar context. In *Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS 2007*, pages 1618–1621, Barcelona, Spain, 23-27 July 2007. IEEE.