POSSIBILITIES OF MAPPING FOREST GROWING STOCK VOLUME OF BOREAL FOREST USING MULTITEMPORAL ENVISAT ASAR WS AND GM OBSERVATIONS

Maurizio Santoro[(1)], Urs Wegmüller[(1)], Andreas Wiesmann[(1)]

[(1)] Gamma Remote Sensing, Worbstrasse 225, CH-3073 Gümligen, Switzerland,
Email: santoro@gamma-rs.ch, wegmuller@gamma-rs.ch, wiesmann@gamma-rs.ch

ABSTRACT

Estimation of forest growing stock volume, i.e. biomass, has been shown to be feasible in the boreal zone using coarse-scale multi-temporal ENVISAT ASAR Wide Swath (WS) mode data. The retrieval is based on a Water Cloud – like model, an automatic training procedure based on the MODIS Vegetation Continuous Field tree canopy cover fraction product and a multi-temporal combination of individual estimates. In this study we showed the feasibility of forest growing stock volume retrieval from a large stack of ENVISAT ASAR Wide Swath and Global Monitoring (GM) data at 1 km resolution for a 400.000 km^2 large study area in Central Siberia. The retrieval procedure is able to model forest backscatter as function of growing stock volume for both modes and under all environmental conditions. The volume estimates from the ASAR data agree remarkably well with the ground reference data also at the highest volumes (400 m^3/ha), this being far beyond expectations. Aggregation to form biomass maps at lower resolution results in improved estimates, the relative RMSE decreasing from approximately 50% at 1 km resolution to 25% at 50 km resolution and < 20% at 100 km resolution. WS- and GM-based estimates appear to be consistent and do not show any apparent sign of saturation up to 400 m^3/ha. The use of multi-temporal ASAR WS and GM data is therefore envisaged for coarse resolution forest biomass retrieval at regional scale in the boreal forest.

INTRODUCTION

Spaceborne SAR remote sensing has been demonstrated to allow the retrieval of forest biophysical parameters under the requirement that the SAR operates at a long wavelength. Unfortunately, current SAR systems in space do not fulfil this requirement entirely. The recently started ALOS PALSAR instrument has enormous potential thanks to the polarimetric capabilities and the dedicated acquisitions over the main forested regions of the world. Nonetheless it is too early to conclude on the effective accuracy of the biophysical parameters that can be retrieved from it. All other spaceborne SAR systems operate either at C-band (ERS-2, ENVISAT ASAR and RADARSAT-1) or at X-band (TerraSAR-X and COSMO-SkyMed). Of these, only ENVISAT ASAR acquires images at regional level in order to build up a consistent archive. Consistent acquisition occurs in particular in the Wide Swath (WS) and Global Monitoring (GM) modes. These modes are characterized by a large swath width (400 km) and long data stripes, covering up to 20 minutes of continuous data acquisition, which approximately corresponds to 8000 km in length. Compared to the GM mode, the WS mode has higher spatial resolution (75 m v. 1 km) and higher radiometric accuracy.

C-band backscatter is known to provide little information about forest parameters because of the weak sensitivity to the forest biophysical parameters and the high sensitivity to environmental conditions. For this reason it is commonly discarded when deciding which observable should be chosen for the retrieval of forest parameters. Nonetheless, if a large number of image acquisitions are available, multi-temporal combination might allow the estimation of forest biophysical parameters with a higher degree of accuracy. With respect to this ENVISAT ASAR WS and GM data become rather interesting because of the strong overlap of adjacent swaths, which implies that more backscatter measurement become available during one repeat-pass cycle (35 days). For example at the 60 degrees North latitude any point on the ground is imaged every third day (at different look angles) either on ascending or descending orbit. When both ascending and ascending passes are available one or two daily measurements are possible.

In this paper we illustrate some of the recent results obtained from ASAR WS and GM modes for mapping of growing stock volume in Central Siberia. We start with reporting on the retrieval method. This is a novel approach that runs automatically based on the information content of the MODIS Vegetation Continuous Fields product and does not make use of in situ data for training. We then discuss on the retrieved growing stock volume from ASAR WS and GM data. Particular attention is given to the effect of pixel size and the properties of the two acquisition modes on the retrieval accuracy. This paper represents an extension of first investigations on the use of ENVISAT ASAR WS for biomass retrieval in boreal forest [1, 2].

STUDY AREA

The area of interest was a 1000 km long and 400 km wide area (53-63° N, 91-99° E) in Central Siberia (EC FP5 SIBERIA-II project area). This region stretches from light coniferous taiga in the North to the dark coniferous taiga in the South, including the Western Sayani Mountains. Both the easternmost part of the Siberian Lowlands as well as a large part of the Yenisey Krijag, a hilly mountainous area, are included. The area is mostly forested, growing stock volumes decrease with increasing latitude (see also Fig. 5).

This subset was chosen based on a number of reasons: availability of a large multi-temporal and multi-angle WS and GM dataset, an acceptable level of radiometric compensation for local topography, reasonable quality of the ground reference data, possibility to test the sensitivity of the biomass upon the backscatter coefficient for a large range of forest conditions, and sufficient extent to be able to derive maps of growing stock volume at coarse resolution, up to 100x100 km^2 pixel size.

ENVISAT ASAR DATASET

For the study area ENVISAT ASAR was acquired extensively in WS mode between spring 2003 and summer 2004. Afterwards acquisitions were primarily in GM mode. As a trade-off between having a large number of images to perform multi-temporal retrieval and avoiding strong effects of forest dynamics (e.g. due to fire), we considered GM data acquired throughout the end of 2004 and the entire year 2005. Because of the large swath, neighbouring swaths strongly overlap which in turn means a high temporal sampling with significant multi-temporal datasets becoming available in relatively short time intervals. The availability of data acquired in ascending and descending orbits further increased the population of the SAR dataset.

Data processing and archiving has been described in [3]. Several hundred ASAR WS and GM image frames were radiometrically calibrated, geocoded and terrain corrected using SRTM-3 and GTOPO30 DEM data. Each frame was then split into 100 km x 100 km tiles. The WS and GM tiles had a pixel spacing of respectively 100m and 1000 m in easting and northing. All data were stored in stacks of co-localized data. To reduce the nonlinear dependence and the dependency of the backscatter from the incidence angle, the backscatter coefficient $\gamma=\sigma^0/\cos\theta$ was computed (σ^0 being the backscatter coefficient and θ the local incidence angle). For the data analysis γ was additionally corrected for pixel area size effects due to topography. This correction reduces the topographic information in the backscatter values and makes the different values more comparable. It should be noted that for the areas above 60° N, where only the GTOPO30 DEM was available, the effect of the local slope on the backscatter coefficient could only partially be compensated for. This caused the normalized backscatter coefficient to include a residual component due to the local topography not represented in the GTOPO30 DEM. Finally all WS data were multi-looked to obtain a pixel size of 1 km x 1 km, i.e. a size useful for the comparison with biomass from the inventory. In this way we also further filtered the images for speckle.

From an extensive GIS of this area at 1:1,000,000 scale we used growing stock volume information at polygon level, i.e. for areas of several hectares with similar properties in terms of species composition, biomass etc and homogeneous forest cover, as reference. Growing stock volume represents the total standing volume per unit area and is measured in m^3/ha. Conversion factors depending on species composition can be applied to derive above-ground and total forest biomass from the growing stock volume. Due to the vastness of the area and the remoteness of large parts of the region, the inventory data had different levels of accuracy and up-to-date status. From the GIS a map in raster format at 1 km resolution was obtained (see Fig. 5).

METHODOLOGY

Forest biomass retrieval using C-band backscatter is known to be of very limited potential due to the weak sensitivity of the backscatter to the forest structural properties. In addition, dielectric properties of the forest floor (soil moisture, wetness of snow cover) and soil roughness can strongly affect the relationship between the backscatter coefficient and the forest biomass. Correlation between backscatter and biomass goes from a slight positive correlation to no correlation up to a slight negative correlation for increasing wetness and/or roughness of the soil.

Several studies have indicated that by increasing the look angle the sensitivity of the backscatter to the biomass increases. Hence shallow look angles are suggested for forest applications. However, when we first looked at the data, we could notice that the temporal dynamics due to environmental conditions had a stronger impact on the relationship between the backscatter coefficient and the growing stock volume [2]. For example data acquired at steep look angle (e.g. 20 degrees) under winter/frozen conditions could show a wider dynamic range compared to data acquired at shallow look angle (e.g. 40 degree) during rainy periods. Since we could not detect any consistent behaviour for different viewing geometries, we decided to keep all images regardless of the look angle and develop a retrieval algorithm that would be independent from different look angles and tree attenuation.

The retrieval method has been introduced in [2] and therefore will be here only summarized. Modifications to the initial retrieval algorithm are instead described in more detail. Forest backscatter is modelled using the Water-Cloud like model presented in [4] and rewritten as a function of growing stock volume as in [5]:

$$\sigma^o_{for} = \sigma^o_{veg}\left(1 - e^{-\beta V}\right) + \sigma^o_{gr} e^{-\beta V} \tag{1}$$

In Equation (1) $\sigma_{gr}{}^0$ and $\sigma_{veg}{}^0$ represent the backscatter coefficient for a completely transparent and a completely opaque forest respectively and are a priori unknown. The term $e^{-\beta V}$ expresses the two-way forest transmissivity as a function of growing stock volume V. β is an empirically derived coefficient; it is here assumed to be constant (0.0055 ha/m^3).

To determine the two unknowns model training is needed. Traditionally, a set of in situ measurements and corresponding backscatter measurements are used to train the model. The main drawbacks of this method are the limited availability of reference data in particular in remote regions and the limited spatial extent for which the modelled backscatter can be considered representative of the forest backscatter. To overcome this difficulty and make model training more automatic we developed a training procedure that makes use of the MODIS Vegetation Continuous Fields (VCF) tree canopy cover product [6].

The forest backscatter is modelled on a tile-by-tile basis, i.e. for areas being 100x100 km^2 large. For each image in each cell, the coefficient is determined as the peak of the histogram of the SAR values obtained considering pixels with VCF < 25 %, i.e. with no or very little canopy cover. To obtain an estimate of the coefficient σ^0_{veg} at first we determine the backscatter coefficient corresponding to the peak of the histogram of the SAR values obtained for pixels with VCF between the maximum VCF value in the cell and 75% of it. Since this value theoretically represents the average backscatter for dense forests with small gap fraction, $\sigma^0_{av\_for}$, compensation for the residual ground backscatter is needed in order to obtain σ^0_{veg}. The compensation is described in (2) which is obtained by inverting (1) to σ^0_{veg}. In (2) the growing stock volume is set to a level typical of dense forests in the tile of interest, V'. This value can be obtained from coarse-scale records of inventory data.

$$\sigma^o_{veg} = \frac{\sigma^o_{av\_fpr} - \sigma^o_{gr} e^{-\beta V'}}{1 - e^{-\beta V'}} \tag{2}$$

A rather crucial aspect in the modelling is the correct estimation of σ^0_{gr}. In [2] we showed that when the percentage of "ground" pixels was high and the distribution presented a clear peak at a level the estimate of σ^0_{gr} was in line with the value we would have expected by looking at the measurements of the backscatter as a function of the growing stock volume.

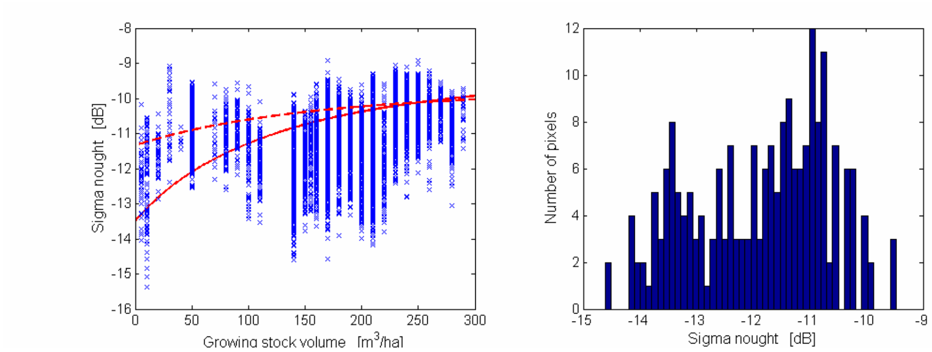

Fig. 1. The plot on the left shows modelled backscatter using automatic training procedure with (solid line) and without (dashed line) adjustment for bimodal peak in the distribution of the backscatter for "ground" pixels. This plot also shows the backscatter measurements as a function of growing stock volume. The plot on the right shows the histogram used to determine the σ^0_{gr}.

When developing the retrieval method some problems were encountered in two cases: (i) very few pixels identified as "ground" (ii) wide distribution of backscatter values. Typically such cases resulted in an overestimation of σ^0_{gr} (see dashed curve in the left plot of Fig. 1) and therefore in possible retrieval error. In the newer version of the algorithm, if the number of samples used to derive σ^0_{gr} is too low or the histogram spreads uniformly over several dB, the coefficient can be set manually by the operator by looking at the histogram. The improvement of the modelled backscatter can be seen in the left plot Fig. 1 (solid curve). It should be noted that σ^0_{gr} is determined by the operator without any knowledge concerning the distribution of backscatter as a function of growing stock volume.

Once the model parameters had been defined for each image in a tile, growing stock volume was retrieved for all points identified as forests in the reference inventory data. To improve the retrieval, multi-temporal combination of estimates from individual images was performed. As weights the span of the backscatter, i.e. the difference $\Delta\sigma=\sigma_{veg}^0-\sigma_{gr}^0$, was used. To avoid that cases with very little or no sensitivity of the backscatter to the growing stock volume would corrupt the multi-temporal estimate, images for which $\Delta\sigma<0.5$ dB were discarded. The reason for such a low threshold is the rather weak sensitivity of the backscatter to the growing stock volume, typically between 0 and 3 dB. A higher threshold would have resulted in rejecting many images

RESULTS AND DISCUSSION

Modelling

Although all images were used in the modelling part regardless of environmental conditions, it is interesting to observe that the sensitivity of the backscatter upon growing stock volume was characterized in most cases by a certain degree of seasonality. Fig. 2 shows typical temporal behaviour of weights used in multi-temporal combination. Since these weights are directly related to the difference $\Delta\sigma=\sigma_{veg}^0-\sigma_{gr}^0$, Fig. 2 can be interpreted as indicator of temporal variation of the forest backscatter dynamic range. For the area of interest, winter/frozen conditions last approximately from November to March. Thaw/freeze events occur in April/May and October. Summer/unfrozen conditions last from May to September. For GM data highest sensitivity was observed mostly for winter/frozen conditions, whereas during the summer the dynamic range was smaller. Winter conditions are characterized by frozen ground thus enhancing the forest/non-forest contrast. The variability of the weights in summer might be related to rather frequent rain events in the area of interest which causes continuous variations of the soil moisture level of the forest floor. Periods of freeze/thaw presented either very small or very large dynamic range because of continuous changes of the dielectric properties of the ground and the snow cover. For WS data highest sensitivity was found not only in winter but also for many summer images. Probably the very dry conditions occurred during summer 2003 caused the ground backscatter to be consistently low, which therefore caused the forest/non-forest contrast to be consistently higher.

Retrieval of growing stock volume

Fig. 3 shows the growing stock volume maps obtained using the retrieval algorithm described above for the ENVISAT ASAR WS and GM datasets. Overall the agreement between the two maps is rather high. A closer look at the maps revealed finer details in the WS-based map. This might be explained in terms of the higher radiometric accuracy and the more accurate radiometric compensation following the higher resolution of the WS data. The GM-based map shows fewer distortions at the tile borders (indicated by the grid lines) because a larger number of images could be used in the multi-temporal combination. Distortions appear in particular for tiles with a relatively small number of images used in the retrieval. Fig. 4 illustrates the number of images combined in the multi-temporal retrieval using the WS and the GM dataset respectively. For ENVISAT ASAR WS clear jumps appear in the top and bottom right corners. From Fig. 3 and 4 we may deduce that at least 20-30 images are necessary to avoid artefacts in the growing stock volume estimates. We might also expect that with a similar number of images as for the GM case the growing stock volume map produced from WS data would be of similar if not better quality compared to the GM-based map.

Since at this stage the GM-based results seem to be more correct, we will refer to these in the following discussions. Fig. 5 illustrates a comparison between the growing stock volume maps from the forest inventory and from the GM data. The agreement on a large scale is remarkable. Patterns of growing stock volume match and even areas with very high forest density are identified as such. Local disagreement between the maps can be explained in terms of a number of factors that need to be a thorough investigation (e.g. accuracy of VCF for automatic training, robustness of retrieval method, dependence of seasonal conditions on retrieval, different age of forest inventory data).

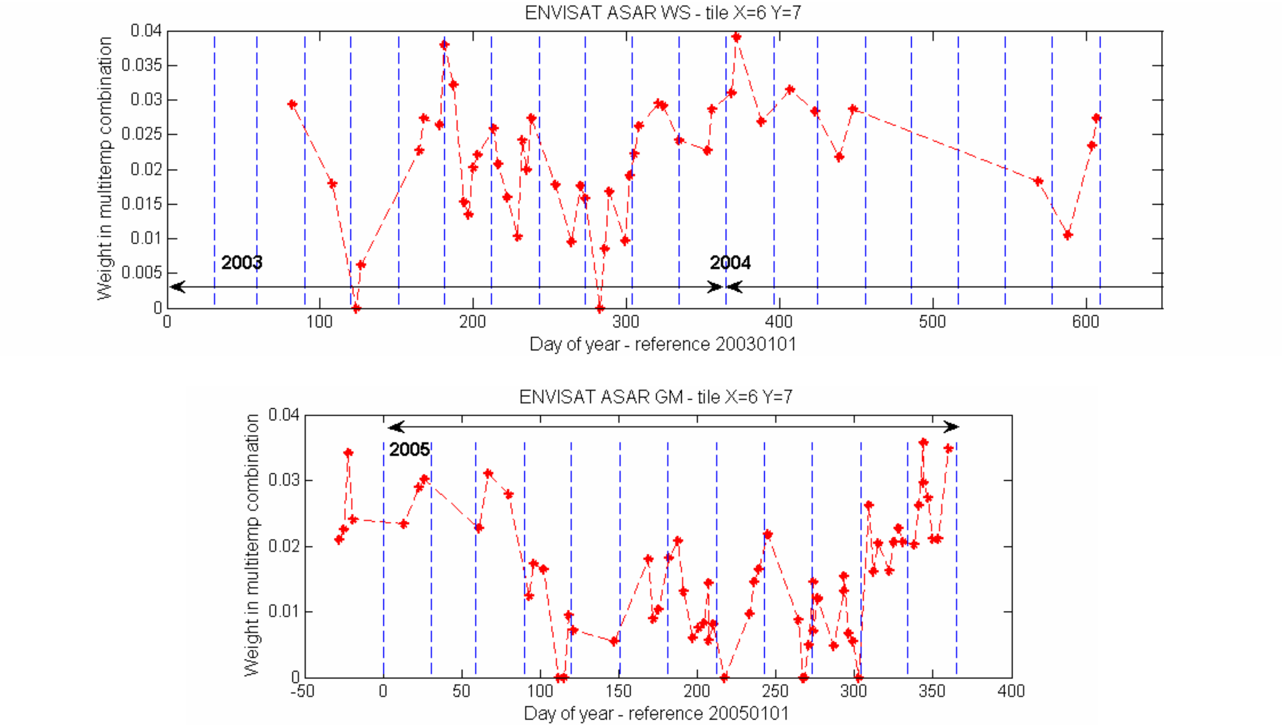

Fig. 2. Temporal behaviour of weights applied in the multi-temporal combination for one tile of co-localized data. The weights are directly proportional to the difference $\Delta\sigma=\sigma_{veg}^{0}-\sigma_{gr}^{0}$. The weights are plotted against day of year. Upper plot for WS data (March 2003 – August 2004), lower plot for GM data (November 2004 – December 2005).

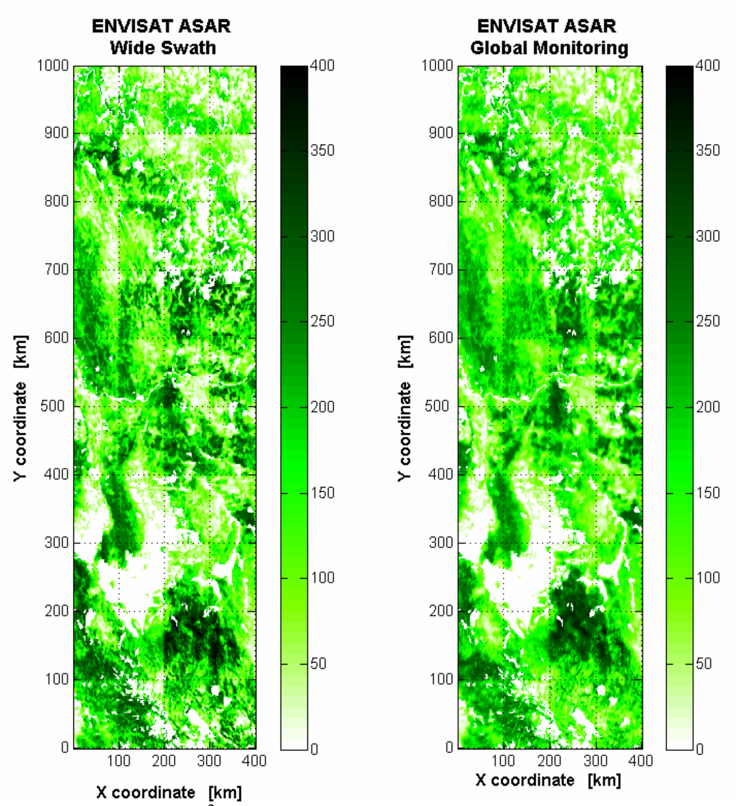

Fig. 3. Growing stock volume maps (in m^3/ha) obtained from multi-temporal ENVISAT ASAR WS and GM data.

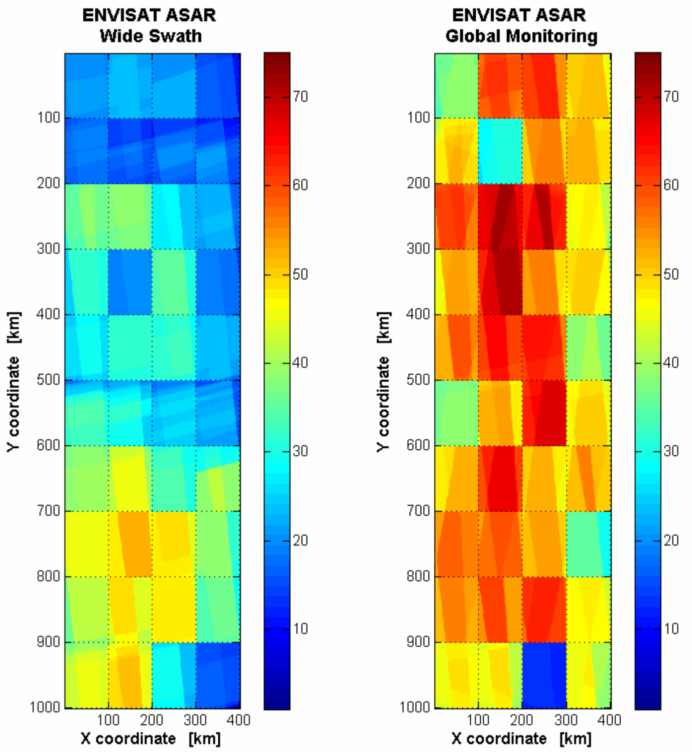

Fig. 4. Number of backscatter measurements used for multi-temporal retrieval of growing stock volume for the ENVISAT ASAR WS (left) and GM (right) datasets.

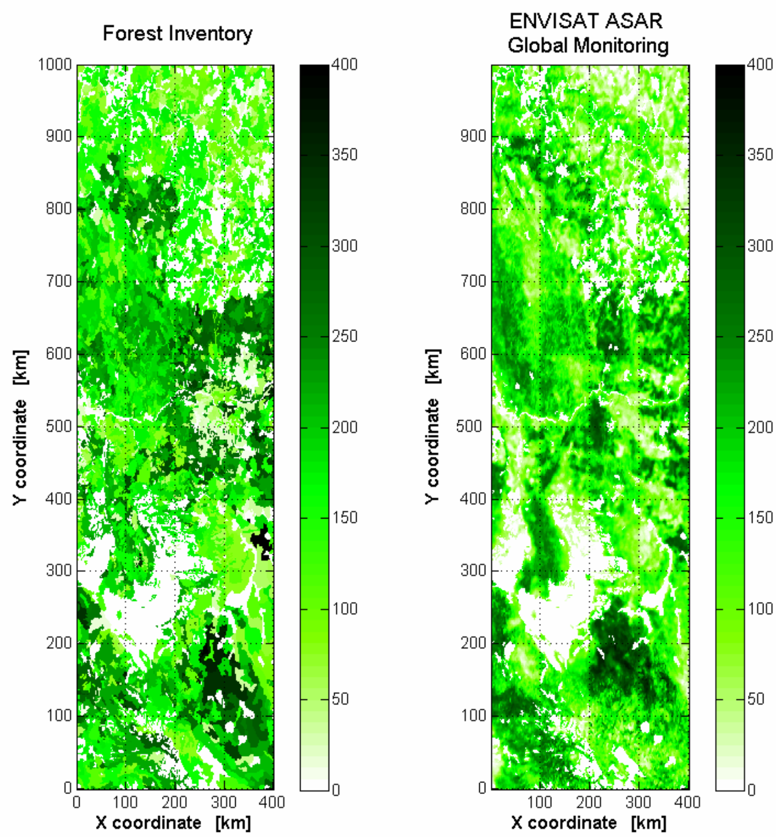

Fig. 5. Growing stock volume maps (in m^3/ha) obtained from forest inventory (left) and multi-temporal ENVISAT ASAR GM data (right).

The accuracy assessment consisted of computation of correlation coefficients, rms error, relative rms error and bias. It has however to be beard in mind that the figures presented below might be affected by errors for two reasons. The reference forest inventory data is at polygon level whereas the ASAR-based maps express growing stock volume at pixel level. Part of the inventory information might have become obsolete at the time of acquisition of the ASAR data. Table 1 shows the statistics for the WS- and the GM-based maps. Results from GM seem to be slightly better, which can be attributed to the better dataset available.

Table 1. Statistics of accuracy assessment for the 1-km resolution WS- and GM-based growing stock volume maps

	Bias (m^3/ha)	Correlation	RMSE (m^3/ha)	Relative RMSE (%)
WS	-20.0	0.4	90.9	54.6
GM	-15.0	0.39	83.4	50.1

Aggregated growing stock volume maps

When growing stock volume estimates were aggregated to form maps at lower resolution, the agreement between the inventory data and the ASAR-based estimates increased steadily. This is shown in Fig. 6 in which the correlation coefficient between inventory and GM-based estimates and the relative rms error have been plotted against the map spatial resolution.

For a resolution of 50 km we illustrate in Fig. 7 the comparison between the aggregated forest inventory data and the GM-based growing stock estimates together with a scatterplot that better shows the high level of agreement between the two maps. The scatterplot does not show any sign of saturation between 0 and 300 m3/ha, this being far beyond any expectation. There is a slight negative bias, i.e. the GM-based map underestimates the growing stock volume. This can be due to an imperfection in the retrieval algorithm. However, forest cover changes between the inventory date and the ASAR acquisition could be noticed at several spots, thus suggesting that probably the bias could also be realistic.

CONCLUSIONS

In two previous papers we have started analyzing the possibility of retrieving forest growing stock volume using ENVISAT ASAR Wide Swath (WS) data. The agreement between the radar-based and the in situ data for a test area in Central Siberia was remarkable even for dense forest areas. The accuracy was shown to improve when areas of increasing size were considered [1, 2]. In this paper we present results based on a refined version of the retrieval algorithm, which has been applied not only to multi-temporal ASAR WS data but also to multi-temporal ASAR Global Monitoring (GM) mode data.

The hypothesis that using a large stack of multi-temporal co-localized data and a rather simple retrieval algorithm can provide forest biomass estimates with good accuracy has been shown to be correct. At least 20-30 images are needed. Winter-time data seem to be more suitable although summer data should not be completely neglected. At 1 km resolution we obtained from both ASAR modes growing stock volume maps that clearly identify the patterns of the growing stock volume even for the highest forest densities in the area of interest (0-400 m^3/ha). It should also be remarked that the growing stock volume estimates from ASAR WS and GM do not present any apparent sign of saturation.

The accuracy assessment returned approximately 50% relative RMSE, which shows that probably such maps are not sufficient for updating a forest inventory at 1 km resolution. However when going to lower resolutions, the agreement between inventory and ASAR-based estimates improves significantly, reaching the 25% level at about 50 km grid spacing. Such maps become of tremendous importance for global vegetation and climate models (see for example [2]). The possibility to form a stack of multi-temporal data suitable for inversion in a short time period also implies that it becomes possible in theory to consider the production of such maps on a yearly basis or similar.

ACKNOWLEDGEMENT

ENVISAT ASAR data were acquired and distributed under AO-225 SIBERIA. Data processing and collection of inventory data were supported by the EC- FP5 Project SIBERIA-II (Contract No. EVG1-CT-2001-00048). Collaboration with Ch. Beer, MPI-BGC Jena, A. Shvidenko and I. McCallum, IIASA, and C. Schmullius, FSU Jena, is greatly acknowledged.

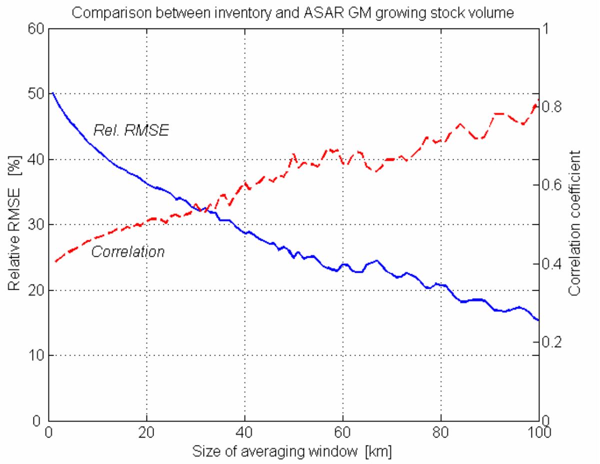

Fig. 6. Relative RMSE (blue solid line) and correlation coefficient (red dashed line) as a function of level of aggregation of the maps shown in Fig. 5. The level of aggregation is given in terms of spatial resolution.

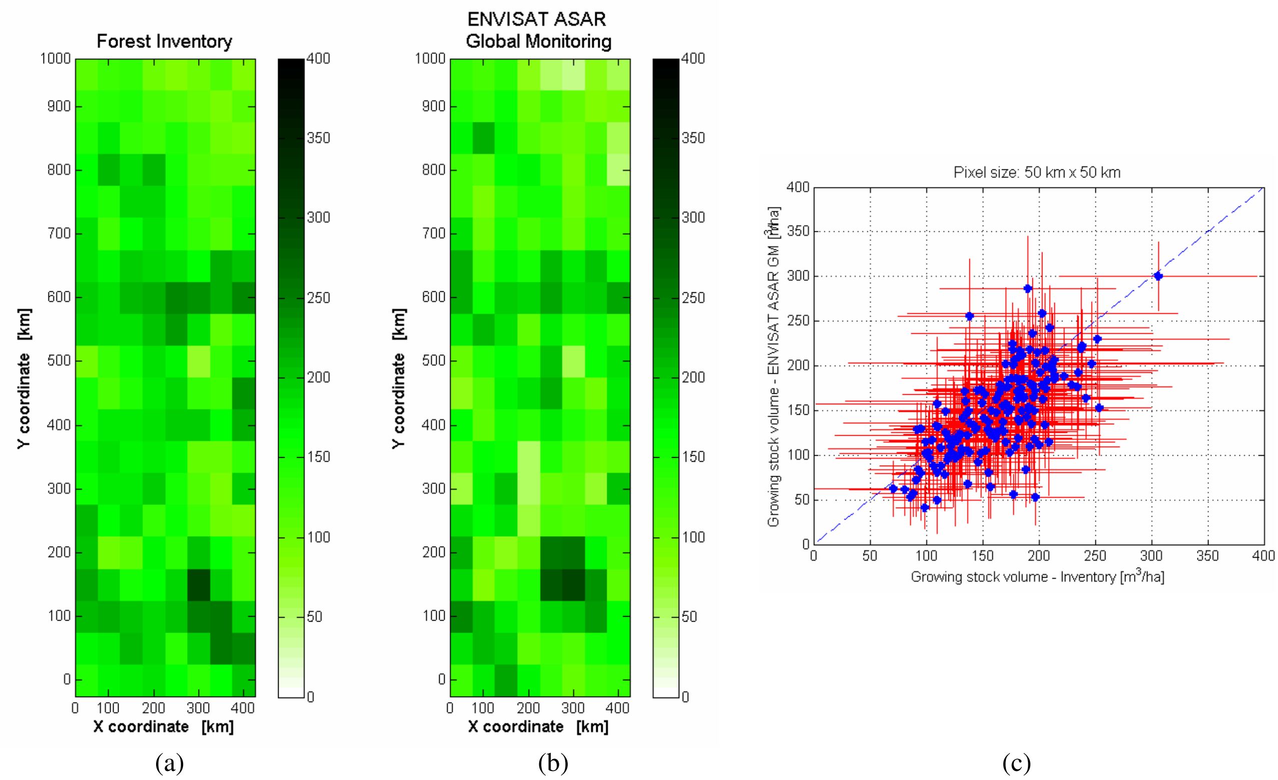

Fig. 7. Aggregated estimates of growing stock volume (in m^3/ha) from (a) inventory and (b) ASAR GM. The aggregated values are compared in the scatterplot in (c) together with the standard deviation.

REFERENCES

[1] U. Wegmüller, M. Santoro, and A. Wiesmann, "A novel methodology for parameter retrieval from multi-temporal data demonstrated for forest biomass retrieval from C-band SAR backscatter," *Proc. MULTITEMP-2007,* Leuven, 18-20 July, 2007.

[2] M. Santoro, C. Beer, A. Shvidenko, I. McCallum, U. Wegmüller, A. Wiesmann, and C. Schmullius, "Comparison of forest biomass estimates in Siberia using spaceborne SAR, inventory-based information and the LPJ Dynamic Global Vegetation Model," *Proc. Envisat Symposium 2007*, Montreux, 23-27 April, 2007.

[3] A. Wiesmann, U. Wegmüller, M. Santoro, T. Strozzi, and C. Werner, "Multi-temporal and multi-incidence angle ASAR Wide Swath data for land cover information," *Proc. 4th International Symposium on Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications*, Innsbruck, 16-19 November, 2004.
[4] J. Askne, P. Dammert, J. Fransson, H. Israelsson, and L. M. H. Ulander, "Retrieval of forest parameters using intensity and repeat-pass interferometric SAR information," *Proc. Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications*, Toulouse, 10-13 October, 1995.
[5] J. T. Pulliainen, K. Heiska, J. Hyyppä, and M. T. Hallikainen, "Backscattering properties of boreal forests at the C- and X-bands," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 1041-1050, 1994.
[6] M. C. Hansen, R. S. DeFries, J. R. G. Townshend, R. A. Sohlberg, C. Dimiceli, and M. Carroll, "Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data," *Remote Sens. Environ.*, vol. 83, pp. 303-319, 2002.