AUTOMATIC MODEL INVERSION OF MULTI-TEMPORAL C-BAND COHERENCE AND BACKSCATTER MEASUREMENTS FOR FOREST STEM VOLUME RETRIEVAL

Maurizio Santoro [(1)], Jan Askne [(2)], Christian Beer [(3)], Oliver Cartus [(4)],
Christiane Schmullius [(4)], Urs Wegmüller [(1)], Andrea Wiesmann [(1)]

[(1)] Gamma Remote Sensing, Gümligen, Switzerland;
santoro@gamma-rs.ch, wegmuller@gamma-rs.ch, wiesmann@gamma-rs.ch
[(2)] Department of Radio and Space Science, Chalmers University of Technology, Gothenburg, Sweden;
askne@rss.chalmers.se
[(3)] Max Planck Institute for Biogeochemistry, Jena, Germany; cbeer@bgc-jena.mpg.de
[(4)] Department of Earth Observation, Friedrich-Schiller University, Jena, Germany;
Oliver.Cartus@uni-jena.de, c.schmullius@uni-jena.de

ABSTRACT

Retrieval of forest stem volume from synthetic aperture (SAR) backscatter and interferometric SAR (InSAR) coherence is generally performed using a model-based approach, where *in situ* measurements are necessary to estimate the unknown model parameters. Problems arise when *in situ* data are either not available or of low quality or the observables present spatial variations. In this work we present three approaches for automatic modeling and inversion of forest backscatter and coherence to retrieve forest stem volume. The three approaches exploit statistical distributions of the observables to obtain estimates for the unknowns in the model. Results shows remarkable agreement with those obtained by means of traditional modeling approaches based on *in situ* data.

***Index Terms*—** ERS, ENVISAT, SAR backscatter, coherence, stem volume, MODIS VCF.

1. INTRODUCTION

Estimation of forest biophysical parameters from synthetic aperture radar, SAR, data is a major topic of investigation in remote sensing. This is due to the high demand of reliable biomass information from regional to global scale and the assessed capability through retrieval approaches to estimate forest biomass from SAR observables. Although low frequencies are indicated as most suitable for this purpose, the radar frequency for which the largest experience has been gathered on retrieval of forest parameters is C-band. This is primarily due to extensive database of images acquired by the European Remote Sensing Satellites ERS-1 and ERS-2 SAR, and ENVISAT ASAR sensors.

Investigations on ERS and ENVISAT data have focused on boreal forests and stem volume. The multi-temporal aspect of both SAR backscatter and InSAR coherence data have allowed detecting the best conditions for retrieval and the development of methods that exploit the variability of the environmental conditions to improve estimates of forest biomass based on one single image.

Typically the stem volume retrieval algorithms based on ERS and ENVISAT data exploit models of different sorts, all of which require a training phase before being usable for inversion [1-3]. Model training means to obtain estimates for a number of model parameters that are unknown *a priori*. Typically this requires the availability of training site(s), i.e. *in situ* measurements of stem volume, which however are not always available or can be obsolete, in particular in remote areas. In addition, spatial variations of the backscatter and/or coherence are not accounted for. These factors represent a major limit when it comes to exploiting SAR data for large-scale retrieval.

Recently we have developed approaches for automatic model training and inversion using ERS and ENVISAT observations. The main aspects of these approaches, their requirements and the results will be discussed in this paper. In Section 2 we present results obtained using a retrieval approach based on multi-temporal ENVISAT ASAR backscatter data. In Section 3 results for two retrieval approaches based on ERS tandem coherence data are reported. In Section 4 we present our main conclusions.

2. BACKSCATTER-BASED RETRIEVAL METHOD

C-band backscatter is known to present weak sensitivity to forest stem volume. Retrieval of stem volume by means of ERS backscatter data was found to provide acceptable estimates only when several images were combined in a multi-temporal approach [1]. Since 2002 ENVISAT ASAR has repeatedly acquired data in ScanSAR modes (Wide Swath and Global Monitoring), which reflected in the establishment of an extensive archive of multi-temporal observations. To exploit this large archive we developed a retrieval algorithm based on a Water-Cloud-like model [4] and a weighted combination of estimates from individual images [5]. The model contains two parameters unknown *a priori*, the backscatter of ground and vegetation (σ^0_{gr} and σ^0_{veg}). For each SAR backscatter image σ^0_{gr} coincides with the mean values of the backscattering coefficient for open areas whereas σ^0_{veg} is obtained from the backscatter of dense forests by compensating for residual ground contribution. Open areas and dense forests are identified by masking the SAR image with the MODIS Vegetation Continuous Fields (VCF) product. The weighted combination exploits the different sensitivity of the backscatter to stem volume depending on the environmental conditions: the greater the difference between the two model parameters for an image, the higher the weight given to the corresponding stem volume estimate.

In Fig. 1 *in situ* and retrieved stem volume for a 400,000 km^2 area in Central Siberia are compared. Fig. 2 shows validation results on an aggregated scale. These Figures highlight the great potential of the ASAR-based retrieval approach for mapping stem volume over large areas. Results improved when considering averages over large area (> 10x10 km^2). The good agreement with *in situ* observations at the km-scale therefore suggests the use of this method for deriving forest stem volume on regional to continental basis. This is of high importance for vegetation modeling approaches which are not capable (yet) to describe small scale variations of the forest biomass.

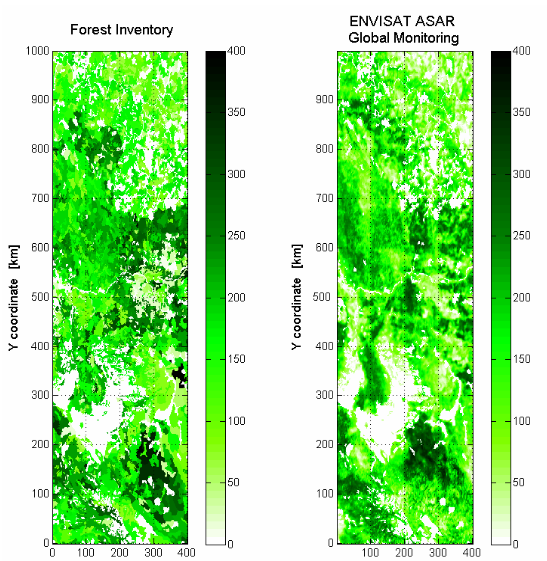

Fig. 1. In situ and retrieved stem volume for a 400,000 km^2 area in Central Siberia using a one-year dataset of ENVISAT ASAR Global Monitoring mode images. Values are in m^3/ha.

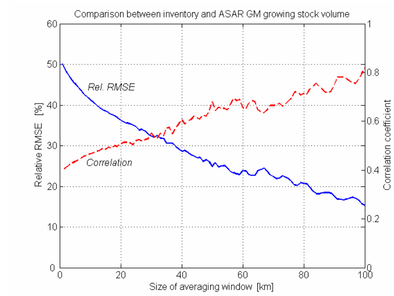

Fig. 2. Relative RMSE (solid line) and correlation coefficient (dashed line) as a function of level of aggregation of the maps shown in Fig. 1.

The main requirement for obtaining acceptable stem volume estimates is that the number of backscatter observations used in the multi-temporal retrieval must be fairly large (> 20). This critical mass can be easily obtained when a full dataset acquired during winter season, under dry conditions is available. This is the case for the entire boreal zone.

3. COHERENCE-BASED RETRIEVAL METHODS

The major limit in currently available retrieval methods based on ERS tandem coherence is that the relationship between the coherence and the forest stem volume can be quite variable in space and in time thus affecting in a negative sense the retrieval if one realization of the model used for the inversion is considered. To be able to capture these variations without the need of local *in situ* measurements for tuning the model, other methods have to be used. Here we will present two retrieval approaches, both of which make use of the Interferometric Water Cloud Model [4]. The model contains five parameters unknown *a priori*: temporal coherence and backscatter of ground (σ^0_{gr} and γ_{gr}) and vegetation (σ^0_{veg} and γ_{veg}), and the coefficient of the forest transmissivity (β). Their determination without the need of *in situ* data is the topic of the approaches described below.

3.1. Consistency plots

When a certain number of observations are available (at least two) it is possible to use the temporal consistency of the repeated coherence observations to estimate the IWCM parameters. A consistency plot (see Fig. 3) compares two sets of coherence observations from forests. For ideally same environmental conditions, the observations should lie along the 1:1 line. Curved trends are related to large-scale effects of different environmental conditions (e.g. summer v. winter conditions). If the spatial variability of the environmental conditions is strong the correlation between the observations is lost. In other words the environmental effects overshadow the effect of the forest properties.

The training method using consistency plots is described in [6]. From each consistency plot we determine what pixels fall in the upper part and the lower part (more noise influences the lower coherence values) and define these as potential pixels representing ground and dense vegetation respectively. We then define those pixels in common for all image pairs as "true" ground and dense vegetation pixels. From these we determine the average coherence and backscatter, thus directly determining σ^0_{gr} and γ_{gr}, and values for dense vegetation, σ^0_{dv} and γ_{dv}. These are compensated for by adding residual ground coherence and backscatter determined from the IWCM to obtain the vegetation model parameters σ^0_{veg} and γ_{veg}. After a first estimation, the model curves are adjusted in a refinement step to better follow the "ridges" of the *½n(n-1)* density functions of the *n* image pairs.

The method was applied at four test sites in Sweden. Fig. 3 shows the consistency plots for the Kättböle test site and the co-plotted IWCM curves obtained with the new method. Table 1 reports retrieval statistics using the traditional training method based on *in situ* data and the new approach based on consistency plots for the three best pairs in terms of RMSE and a multitemporal combination. The results seem to indicate that the new approach performs equally well compared to the traditional method. The accuracy of the results from Kättböle is favored by images with low temporal decorrelation and baselines up to 250 m.

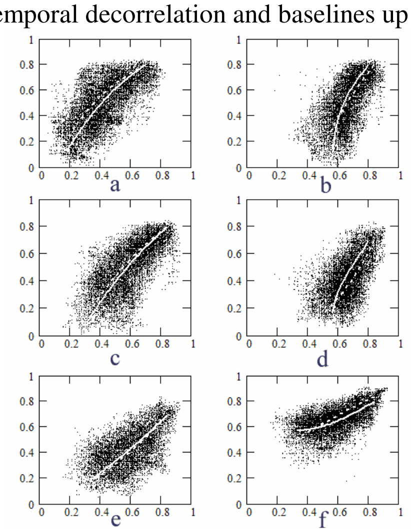

Fig. 3. Consistency plots for three coherence images covering the Kättböle test site. Each point represents a pixel. The dashed curves represent the co-plotted IWCM curves after the first estimation of the parameters. The solid curves represent the co-plotted IWCM curves as obtained after refinement of the model parameters.

Table 1. Relative RMSE for the Kättböle test site (42 stands, mean volume 135 m^3/ha, max volume 335 m^3/ha)

	Traditional method	New method
multitemporal	18.4 %	17.8 %
1996-03-12/13	19.8 %	22.0 %
1996-03-17/18	38.0 %	32.5 %
1996-04-21/22	31.6 %	31.5 %

The capability of the new approach to adapt to different environmental conditions was tested by comparing model realizations in Kättböle and for an area 25 km away. This area was known to us as being characterized by different coherence signatures as a consequence of different environmental conditions in some of the image pairs [2]. The model curves determined by means of the new approach are shifted from the model curve derived for the Kättböle area, in particular at the highest stem volumes. These changes could be interpreted as caused by a lower wind speed in the new area. We then conclude that we need investigations of such changes over the area of interest, and that we now have a method to derive the changes in the model curves and then the possibility to derive stem volume estimates independent of known training sites.

3.2. VCF-based retrieval

The previous approach requires at least two images, which furthermore should present a certain consistency of the observations. In case only one image is available or the (few) images present limited consistency this approach might encounter difficulties.

To get around the problem a retrieval approach similar to the backscatter-based method has been developed [7]. The σ^0_{gr} and γ_{gr} estimates correspond to the peak of the backscatter and coherence histograms for open areas as identified in the VCF product. The estimates of σ^0_{veg} and γ_{veg} are obtained by taking the peak of the backscatter and of the coherence histograms for dense forests and compensating these values for residual ground contribution. The approach has been developed on the basis of an ERS image frame, i.e. a 100x100 km^2 area. In case of strong spatial variations of the coherence, subsetting can be applied.

The method has been developed at test sites in Central Siberia. Fig. 4 shows examples of modeled coherence obtained with this automatic approach and with the traditional method based on *in situ* data at four test sites.

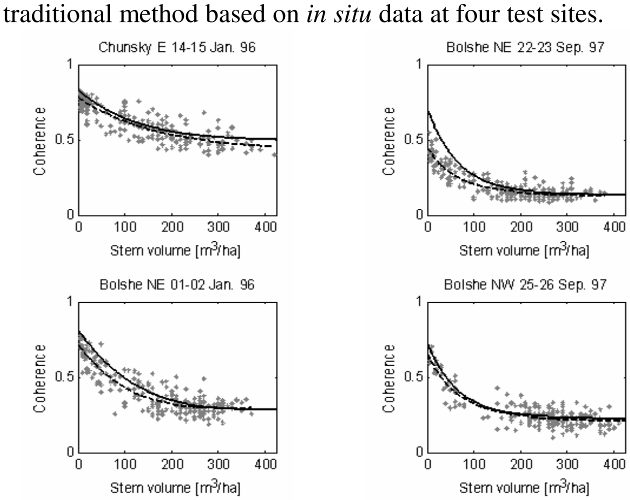

Fig. 4. IWCM Coherence as a function of stem volume obtained using in situ data (dashed) and the VCF-based approach (solid).

Mostly a good fit to the trend in the measurements was achieved and the modeled coherence obtained from the two methods was similar. Some differences were noticed for image pairs acquired after rain events (e.g. top right plot in Fig. 4). However, when we restricted the analysis to the area of the forest test sites, the modeled coherence values were similar.

With this approach stem volume has been classified in four classes (0-20, 20-50, 50-80 and > 80 m^3/ha). The accuracy assessment showed values similar to those obtained using a more traditional type of retrieval approach, partly based on *in situ* data [3, 8]. The highest accuracy was obtained for the class with highest and the lowest stem volume (~ 90% and 80% respectively). The accuracy for the intermediate classes was lower, in the range of 10-65%. This result is probably due to the low accuracy of the *in situ* data in this interval of stem volumes (cf. [8]).

4. CONCLUSIONS

We have introduced three approaches for retrieval of stem volume from SAR data that do not depend on *in situ* measurements. The SAR data consisted of multitemporal ENVISAT ASAR ScanSAR C-band backscatter data and ERS tandem coherence. The methods are however in theory independent from the specific frequency.

The approaches are rather similar in their rationale. A model has been validated separately and a number of unknown model parameters are determined by exploiting statistics of backscatter and coherence values for open areas and dense forests. Two ways have been devised to obtain these values: peaks of histograms based on image masking using the MODIS VCF product and through consistency plots of coherence. All approaches have the necessity of some knowledge of the typical distribution of stem volumes in the region of interest in order obtain the correct values for the vegetation model parameters σ^0_{veg} and γ_{veg}. The methods are in principle all spatially adaptive in the sense that they can run on a pixel-by-pixel basis.

An important point concerns the number of observations. We need an extensive dataset of C-band backscatter observations to obtain a satisfactory retrieval accuracy because of the weak sensitivity to stem volume. For the ERS coherence case few images, if not just one acquired under the “optimal” conditions (winter/frozen, with dry snow layer and moderate wind breeze, perpendicular baseline of about 100-250 m) are sufficient.

From these investigations we conclude that (i) synergy of radar and other remote sensing products can be exploited to develop retrieval algorithms independent from *in situ* data (as for all remote sensing these are valuable for validation); (ii) repeated winter-time observations of ERS tandem coherence-like images are probably the best choice for forest stem volume retrieval in boreal forest, (iii) with relation to available spaceborne SAR data ENVISAT ASAR ScanSAR is currently the only reliable source for obtaining large-scale estimates of biomass to be exploited in ecosystem modeling related to environmental change.

5. ACKNOWLEDGMENTS

Financial support has been provided through ESA (DRAGON Cooperation Programme, id 2583), Adlerbert Research Foundation and the Swedish National Space Board. ENVISAT ASAR data were acquired and distributed under AO-225 SIBERIA. ERS data were obtained through ESA Cat.1 projects ID 2583 and 3121. Collaboration with A. Shvidenko and I. McCallum, IIASA, is acknowledged. The VCF product is available at http://edcdaac.usgs.gov/modis/dataproducts.asp.

6. REFERENCES

[1] L. Kurvonen, J. Pulliainen, and M. Hallikainen, "Retrieval of biomass in boreal forests from multitemporal ERS-1 and JERS-1 SAR images," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 198-205, 1999.

[2] M. Santoro, J. Askne, G. Smith, and J. E. S. Fransson, "Stem volume retrieval in boreal forests with ERS-1/2 interferometry," *Remote Sens. Environ*., vol. 81, pp. 19-35, 2002.

[3] W. Wagner, A. Luckman, J. Vietmeier, K. Tansey, H. Balzter, C. Schmullius, et al., "Large-scale mapping of boreal forest in SIBERIA using ERS tandem coherence and JERS backscatter data," *Remote Sens. Environ*., vol. 85, pp. 125-144, 2003.

[4] J. Askne, P. B. G. Dammert, L. M. H. Ulander, and G. Smith, "C-band repeat-pass interferometric SAR observations of the forest," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 25-35, 1997.

[5] 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.

[6] J. Askne and M. Santoro, "Advances in boreal forest stem volume estimation from multitemporal C-band InSAR," *IEEE Trans. Geosci. Remote Sensing*, submitted.

[7] O. Cartus, M. Santoro, C. Schmullius, P. Yong, and Z. Li, "Creation of large area forest biomass maps for Northeast China using ERS-1/2 tandem coherence," *Proc. Envisat Symposium 2007*, Montreux, 23-27 April, 2007.

[8] H. Balzter, E. Talmon, W. Wagner, D. Gaveau, S. Plummer, J. J. Yu, et al., "Accuracy assessment of a large-scale forest cover map of central Siberia from synthetic aperture radar," *Can. J. Remote Sens.*, vol. 28, pp. 719-737, 2002.