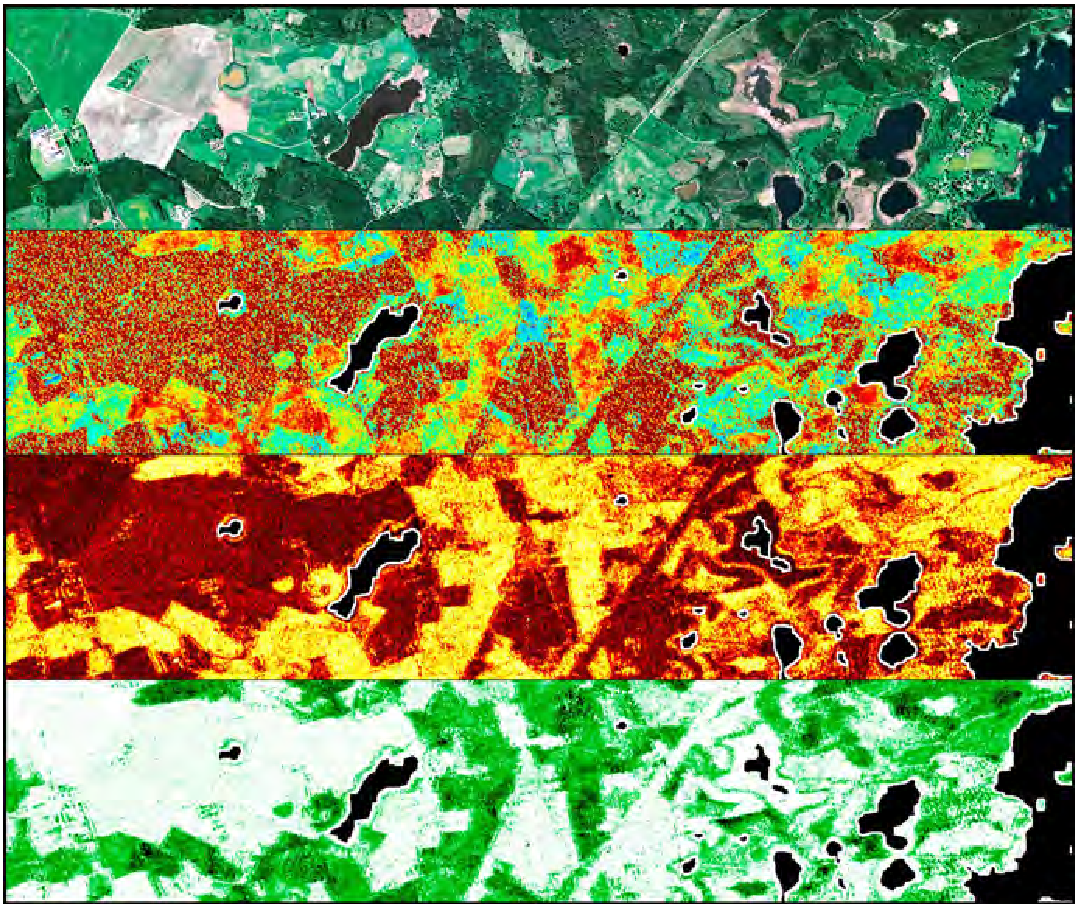
CHALMERS



Modelling and Retrieval of Forest Parameters from Synthetic Aperture Radar Data

MACIEJ JERZY SOJA

Department of Earth and Space Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2014

Thesis for the Degree of Doctor of Philosophy

Modelling and Retrieval of Forest Parameters from Synthetic Aperture Radar Data

Maciej Jerzy Soja

CHALMERS

Department of Earth and Space Sciences
Chalmers University of Technology
Gothenburg, Sweden 2014

Modelling and Retrieval of Forest Parameters from Synthetic Aperture Radar Data
Maciej Jerzy Soja
ISBN 978-91-7597-076-9

Doktorsavhandlingar vid Chalmers Tekniska Högskola
Ny serie nr 3757
ISSN 0346-718X

Radar Remote Sensing Group
Department of Earth and Space Sciences
Chalmers University of Technology
SE–412 96 Gothenburg, Sweden
Phone: +46 (0) 31 772 1000

Cover: Four images showing the same part of the Remningstorp test site, situated in southern Sweden, in the province of Västergötland, approximately 150 km north-east from Gothenburg. The topmost image is an aerial photograph, provided by the Swedish Land Survey (Lantmäteriet). The second image from the top shows forest height estimate Δh, obtained from two-level model inversion of a VV-polarised TanDEM-X acquisition from 2011-06-04 (incidence angle: 41°, height-of-ambiguity: 49 m), provided by the German Aerospace Center (DLR). The third image from the top shows the canopy density estimate η_0, also obtained from two-level model inversion. Finally, the bottom image shows a biomass estimate obtained from Δh and η_0. The root-mean-square errors are roughly 10% for forest height and canopy density estimates (compared to airborne lidar scanning data), and approximately 15% for the biomass (compared to *in situ* measurement-based biomass estimates). Note that lakes have been masked out. The legends are shown below.

Printed by Chalmers Reproservice
Chalmers University of Technology
Gothenburg, Sweden 2014

to my mom, my siblings, and Monika

Modelling and Retrieval of Forest Parameters from Synthetic Aperture Radar Data

Maciej Jerzy Soja
Department of Earth and Space Sciences
Chalmers University of Technology

Abstract: Frequent, high-resolution mapping of national and global forest resources is needed for improved climate modelling, degradation and deforestation detection, natural disaster management, as well as commercial forestry. Synthetic aperture radar (SAR) is an active radio- or microwave-frequency imaging sensor, which can be optimised to fit specific needs through the choice of the centre frequency. In particular, P-band SAR, with wavelengths around 70 cm, is a promising tool for biomass mapping due to the high sensitivity to tree trunks, whereas X-band SAR, with wavelengths around 3 cm and larger available bandwidths, is a promising tool for high-resolution mapping of forest canopies.

Papers A and B summarise the results obtained within the feasibility study for the European satellite BIOMASS, which is planned to become the first spaceborne P-band SAR system. In Paper A, a forward model relating relevant forest and system parameters to SAR observables is presented and evaluated. In Paper B, a new model for biomass estimation is proposed, in which the significant influence of topographic and moisture variations is treated using empirical corrections. The new model can be used with the same model parameters in two boreal test sites in Sweden, separated by 720 km, with a root-mean-square error (RMSE) of 22–33% of the mean biomass.

In Papers C, D, and E, X-band SAR data acquired with the twin-satellite, single-pass interferometric system TanDEM-X are studied. Using the principles of across-track interferometry, the position of the scattering centre is estimated from the phase difference between two SAR images. With a high-resolution digital terrain model, the interferometric data are ground-corrected, and the elevation of the scattering centre above ground is determined. In Paper C, boreal forest biomass is estimated for one test site in Sweden from ground-corrected TanDEM-X data using three models with tree canopies represented by a random volume, but with different assumptions of the ground component. The best results, with an averaged RMSE of 16%, are obtained with a model accounting for canopy gaps. Based on this observation, a two-level model (TLM) is introduced, in which forest is modelled as two discrete scattering levels: ground and vegetation, the latter with gaps. In Paper D, it is shown that TLM inversion of single-polarised, ground-corrected TanDEM-X data can provide forest height and canopy density estimates, with RMSE values below 10% for a boreal test site in Sweden. In Paper E, biomass is estimated from the inverted TLM parameters, with an RMSE in the interval 12–19% for eighteen acquisitions from two boreal test sites in Sweden.

Keywords: synthetic aperture radar (SAR), forestry, above-ground biomass, forest height, canopy density, P-band, X-band, BIOMASS, TanDEM-X

APPENDED PAPERS

The following papers are the main scientific outcome of this thesis and they have been appended in Part II:

Paper A **M. J. Soja** and L. M. H. Ulander, "Polarimetric-Interferometric Boreal Forest Scattering Model for BIOMASS End-to-End Simulator," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Quebec City, QC, Canada, 13–18 July 2014, pp. 1061–1064

Paper B **M. J. Soja**, G. Sandberg, and L. M. H. Ulander, "Regression-Based Retrieval of Boreal Forest Biomass in Sloping Terrain using P-band SAR Backscatter Intensity Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 2646–2665, May 2013

Paper C 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 Multi-Temporal TanDEM-X Acquisitions," *Remote Sensing*, vol. 5, no. 11, pp. 5574–5597, October 2013

Paper D **M. J. Soja**, H. Persson, and L. M. H. Ulander, "Estimation of Forest Height and Canopy Density from a Single Complex Correlation Coefficient," accepted for publication in *IEEE Geoscience and Remote Sensing Letters*, August 2014

Paper E **M. J. Soja**, H. Persson, and L. M. H. Ulander, "Estimation of Forest Biomass from Two-Level Model Inversion of Single-Pass InSAR Data," submitted to *IEEE Transactions on Geoscience and Remote Sensing*, July 2014

Related Publications

Apart from the appended papers, the author has also contributed to several other publications, including the following reports:

- L. M. H. Ulander, A. Gustavsson, B. Flood, D. Murdin, P. Dubois-Fernandez, X. Depuis, G. Sandberg, **M. J. Soja**, L. E. B. Eriksson, J. E. S. Fransson, J. Holmgren, and J. Wallerman (2011), “BioSAR 2010 Technical Assistance for the Development of Airborne SAR and Geophysical Measurements During the BioSAR 2010 Experiment: Final Report,” Technical report, European Space Agency, contract no.: 4000102285/10/NL/JA/ef, `https://earth.esa.int/c/document_library/get_file?folderId=87248&name=DLFE-1322.pdf`.

- ESA (2012), “Report for Mission Selection: Biomass,” ESA SP-1324/1 (3 volume series), European Space Agency, Noordwijk, The Netherlands, `http://esamultimedia.esa.int/docs/EarthObservation/SP1324-1_BIOMASSr.pdf`.

- K. Papathanassiou, H. Baltzer, L. E. B. Eriksson, A. Gustavsson, F. Kugler, S. K. Lee, S. Sauer, G. Sandberg, **M. J. Soja**, R. Scheiber, T. Le Toan, L. M. H. Ulander, and L. Villard (2012), “Development of Algorithms for Biomass Retrieval,” Technical report, European Space Agency, contract no.: 4200023081/NL/AF.

Preface & Acknowledgements

The story of Rapa Nui, or the Easter Island, is probably as interesting as it is tragic. Once an idyllic island in the Pacific, unknown to humanity and almost entirely covered by a palm forest inhabited by several endemic bird species, the arrival of the first Polynesian settlers around 1200 AD[1] drastically changed its landscape. Trees were felled, to be used as fuel, construction material for shelters and boats, and to give land for agriculture. The population of the island grew rapidly, perhaps even beyond ten thousand inhabitants. 887 monolithic statues, called moai and weighing up to 90 tonnes each, were raised. Soon, the lack of forests caused soil erosion and the leaching of agricultural fields, and almost all animals disappeared from the island. The lack of wood made it impossible to build boats for fishing or for fleeing. Hunger struck the islanders; the misery drew them to wars, and possibly also cannibalism[2]. By the arrival of James Cook in 1774, the population of the now barren island was below a thousand people, surrounded by almost equally many gigantic stone sculptures raised by their ancestors.

Although the understanding of our ecosystems has increased significantly since the arrival of the first Polynesians on Rapa Nui, deforestation and degradation of forests are still a clear and present danger to the entire global ecosystem. Moreover, due to the vastness and diversity of the global forests, many forests remain unexplored, making it difficult to study the terrestrial ecosystems.

This doctoral thesis is focussed on synthetic aperture radar (SAR) remote sensing of forests. SAR is a promising tool for frequent, global, high-resolution mapping of forest resources, and it can be used both to improve the knowledge of the global forest

[1] T. L. Hunt and C. P. Lipo, "Late colonization of Easter Island," *Science*, vol. 311, no. 5767, pp. 1603–1606, 2006

[2] P. Rainbird, "A message for our future? The Rapa Nui (Easter Island) ecodisaster and Pacific island environments," *World Archaeology*, vol. 33, no. 3, pp. 436–451, 2002

resources and to enforce international anti-deforestation agreements. It is my goal for this thesis to bring new knowledge to the field of SAR remote sensing of forests and, in the long run, help to better understand and protect our precious forests.

☆☆☆

All research is collective, and also this work would not have been possible without the help and support of many fantastic people. In the next few paragraphs, I would like to thank them for their efforts and contributions.

First, I would like to show my deepest gratitude to my main supervisor and mentor, Lars Ulander, for his trust in me and guidance throughout these five years. Under his watchful eye, I was able to write a couple of articles and a thesis of which I am very proud of. It is a rare gift to be both a good supervisor and a knowledgeable researcher, and Lars is both.

Also, I would like to thank Gustaf Sandberg, my colleague, fellow doctoral student, and co-author of several papers, for the countless discussions, support, inspiration, and cooperation. Gustaf introduced me to the field of remote sensing through the supervision of my bachelor thesis, and, later, he introduced me to radar remote sensing, through the supervision of my master thesis. Thanks to Gustaf, I have been able to find the fascinating field of SAR remote sensing.

I would like to thank my co-supervisors: Leif Eriksson, for his tireless work as the head of the Radar Remote Sensing Group, and Johan Fransson from the Swedish University of Agricultural Sciences (SLU) in Umeå, for his expertise on forest matters. Likewise, I owe a lot of thanks to Jan Askne and Henrik Persson, my co-authors and discussion partners.

I would also like to thank Anders Berg, for being a superb colleague and the perfect room-mate at work. Additionally, I would like to acknowledge my other colleagues from the group and from the Department of Earth and Space Sciences, for making it a pleasant working environment, especially: Gisela Carvajal, Jan Torgrimsson, Malin Johansson, Erik Blomberg, Marston Johnston, Mattias Andersson, Johan Löfgren, Donal Murtagh, the administrative staff, as well as Jo Urban, whose sudden passing struck us all with great sorrow.

Special thanks to Swedish National Space Board (SNSB), European Space Agency (ESA), and European Union (EU) for the funding of my work. I would also like to thank my colleagues from the BIOMASS feasibility study group, other colleagues from SLU, FOI, DLR, as well as anonymous reviewers and committees whose work contributed to this thesis.

I would like to express my gratitude to all the good teachers at Chalmers, for actually teaching me something, and all the other teachers, for helping me to choose what not to study. Many thanks to Conny Modig, my high school physics and maths teacher at Västerhöjdsgymnasiet in Skövde, for making physics fun. Additional thanks to my primary school maths and science teachers in Warsaw, Poland, for force-feeding me with knowledge in such abundance, that I was able to focus my first school years in Sweden on Swedish and English, rather than physics and maths.

Last, but absolutely not least, I would like to thank my family and my closest friends for your love, support, and patience. Special thanks to my mom Małgorzata, for being my role model, my cool siblings Kasia, Jan, and Paweł, for all the laughs we have together, and my dearest Monika, for coming into my life when I needed you the most, with all your happiness and joy. I am looking forward to our future adventures together!

Maciej Jerzy Soja
June 20th, 2014

Credits

The panorama of Rapa Nui on page ix is an adapted version of the photo "Rano Raraku quarry" (http://commons.wikimedia.org/wiki/File:Rano_Raraku_quarry.jpg) by Wikimedia contributor Rivi, available under the license Creative Commons Attribution-Share Alike 3.0 Unported (http://creativecommons.org/licenses/by-sa/3.0/). All other graphics in this thesis are the author's original work.

- TanDEM-X and TerraSAR-X data have been provided by the German Aerospace Center (DLR) within proposal XTI_VEGE0376.
- BioSAR 2007, 2008, and 2010 data have been provided by the European Space Agency (ESA).
- Aerial photographs and lidar digital terrain models (DTMs) have been provided by the Swedish Land Survey (Lantmäteriet).
- Reference forest data have been provided by the Swedish University of Agricultural Sciences (SLU).
- Meteorological data have been provided by the Swedish Meteorological and Hydrological Institute (SMHI).

Contents

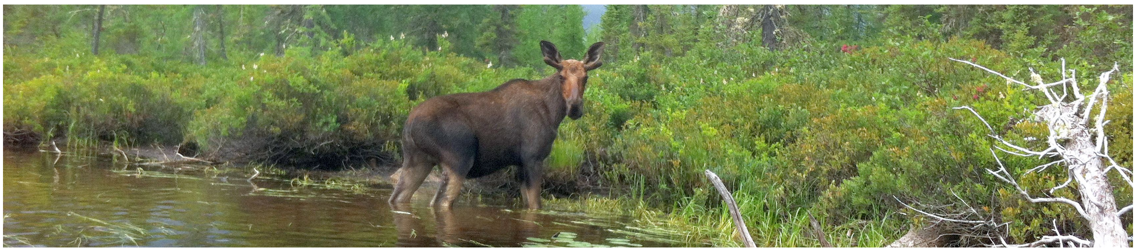
Part I

Introduction

List of Abbreviations

ALS	airborne lidar scanning
ASI	Italian Space Agency (*Italian:* Agenzia Spaziale Italiana)
ATI	along-track interferometry
BEES	BIOMASS End-to-End Simulator
CARABAS	Coherent All RAdio BAnd Sensing
CAST	Chinese Academy of Science and Technology
CDTI	Centre for the Development of Industrial Technology (*Spanish:* Centro para el Desarrollo Tecnologico Industrial)
CONAE	National Commision for Space Activities (*Spanish:* COmisión Nacional de Actividades Espaciales)
COSMO-SkyMed	COnstellation of Small satellites for the Mediterranean basin Observation
CSA	Canadian Space Agency
dbh	diameter at breast height
DCM	digital canopy model
DEM	digital elevation model
DLR	German Aerospace Center (*German:* Deutsches Zentrum für Luft- und Raumfahrt)
DSM	digital surface model
DTM	digital terrain model
EM	electromagnetic
ERS	European Remote Sensing satellite
ESA	European Space Agency
E-SAR	Experimental-SAR
FM	forward model
FOI	Swedish Defense Research Agency (*Swedish:* Totalförsvarets Forskningsinstitut)
GOM	geometrical optics model
HH	horizontal (receive), horizontal (transmit)
HOA	height-of-ambiguity
HV	horizontal (receive), vertical (transmit)
IEEE	Institute of Electrical and Electronics Engineers
InSAR	interferometric SAR

ISRO	Indian Space Research Organisation
ITU	International Telecommunication Union
IWCM	interferometric water cloud model
JAXA	Japan Aerospace Exploration Agency
lidar	light detection and ranging
NASA	National Aeronautics and Space Administration
PD	penetration depth
pdf	probability distribution function
PolInSAR	polarimetric-interferometric SAR
PolSAR	polarimetric SAR
radar	radio detection and ranging
RCS	radar cross section
RISAT	Radar Imaging Satellite
RMSD	root-mean-square difference
RMSE	root-mean-square error
RVoG	random volume over ground
SAR	synthetic aperture radar
SLU	Swedish University of Agricultural Sciences (*Swedish:* Sveriges lantbruksuniversitet)
SMHI	Swedish Meteorological and Hydrological Institute (*Swedish:* Sveriges Meteorologiska och Hydrologiska Institut)
SNR	signal-to-noise ratio
SPM	small perturbation model
SRTM	Shuttle Radar Tomography Mission
TanDEM-X	TerraSAR-X Add-oN for Digital Elevation Measurements
TBM	TLM biomass model
TDM	TanDEM-X (interferometer system)
TDX	TanDEM-X (satellite)
TLM	two-level model
TOPS	terrain observation by progressive scans
TSX	TerraSAR-X (satellite)
UHF	ultra high frequency
VH	vertical (receive), horizontal (transmit)
VHF	very high frequency
VV	vertical (receive), vertical (transmit)
XTI	across-track interferometry



Chapter 1

Forests

1.1 Multiple Roles of Forests

Forests play a vital role in the terrestrial ecosystems. Through the process of photosynthesis, trees and plants bind CO_2 from the atmosphere, part of which is transformed into carbon stock. Forests provide shelter to countless animal and vegetation species, housing around 80% of the terrestrial biodiversity [1]. They also take part in the water cycle, prevent soil from erosion, and clean water and air from pollutants. At the same time, forests are one of our greatest natural resources. Timber is used as a construction material, for paper production, and as a fuel. Animals and vegetation provide food. Also, forests have great recreational values.

For a long time, the global effects of human exploitation of forests were negligible due to the relatively small population and ineffective harvesting methods. However, during the last few centuries, the rapid growth of the human population caused an increased demand on forest products, which together with the excessive use of fossil fuels started to more significantly affect the global environment. In the second half of the 20th century, the signs of human influence on the global ecosystem could be observed, for example in the form of acid rains, ozone depletion, and probably also global warming. Although the public awareness of the environmental issues has increased during the last few decades and the first measures have been taken, there is still much to be learnt, and a lot of research is focussed on Earth system science and climate change.

One of the greatest concerns is the influence of deforestation on global CO_2 emissions. Some sources state, that as much as 20% of the global CO_2 emissions come from deforestation [2], but the exact numbers are unknown. One of the largest uncertainties in the current carbon cycle models is introduced by the inaccurate estimates of the terrestrial carbon stocks and fluxes, mainly associated with forests. The most relevant, measurable quantity directly related to the carbon distribution in the biosphere is biomass, which is the total dry mass of all organic tissue. Since roughly 50% of biomass is carbon, and forests account for around 80% of the terrestrial above-ground biomass [3], frequent and global mapping of forest biomass is needed for improved carbon cycle modelling and climate change prediction. High-resolution mapping of biomass and other forest parameters, e.g., forest height and canopy density, will also aid the detection of deforestation and forest degradation, improve natural disaster handling, and enable efficient and sustainable management in commercial forestry.

1.2 Synthetic Aperture Radar in Forestry

Forests cover more than 31% of the total land surface of the Earth [1], and spaceborne remote sensing is the only feasible method for frequent and global mapping of forests [3]. There are several different spaceborne remote sensing techniques which can be used for forest mapping. Optical methods have long been used for this task. However, these methods are sensitive to atmospheric conditions, which is especially problematic for the high-biomass tropical rainforests around the equator, where the cloud cover is the most persistent [3]. Airborne lidar scanning (ALS) is currently considered the most accurate remote sensing method for forest mapping [4]. However, spaceborne application of this technique is difficult, due to yet unresolved resolution, coverage, and technology limitations [3].

Synthetic aperture radar (SAR) does not suffer from the same disadvantages as the optical and lidar sensors. As an active radio- or microwave-frequency sensor, it provides its own illumination and it is generally less sensitive to clouds than the optical methods. Thanks to the synthetic aperture technique, the image resolution of a spaceborne SAR system can be of the order of metres. SAR is one of the most promising tools for forest remote sensing and many past and ongoing studies are dedicated to forest parameter estimation from SAR data [5]. A comprehensible introduction to SAR can be found in [6].

As the electromagnetic waves scatter more strongly from objects of sizes comparable to, or larger than the wavelength, SAR systems can be optimised to fit specific needs through the choice of the centre frequency [7], and they are commonly classified by the used frequency band. Two frequency bands often used for spaceborne SAR imaging of forests are the P- and X-bands. At P-band, the wavelength is around 70 centimetres and the strongest scatterers are large branches and tree trunks, which also contain most of the biomass. This causes the scattering centre to be located in the lower part of the canopy. At X-band, the wavelength is around 3 centimetres and scattering occurs even from the smaller branches, leaves, and needles in tree canopies. Therefore, the penetration capabilities are significantly lower than at P-band, and the scattering centre is located closer to the canopy top.

1.2.1 P-band

Many studies of VHF-band (30–300 MHz) data acquired with the Swedish SAR systems CARABAS-I and -II (20–90 MHz) have shown the excellent potential of VHF-band SAR for stem volume mapping [8–22], and, since stem volume is highly correlated with biomass, these conclusions also apply to biomass mapping. However, VHF-band SAR is not available for spaceborne use, primarily due to the large ionospheric distortions and lack of suitable frequency allocations [23].

A somewhat higher frequency band is the P-band (typically 420–450 MHz), for which numerous studies have shown good correlation between the backscattered signal strength and biomass over a wide biomass range [24–40]. It has also been shown that the temporal stability is high at P-band [41–43], making it possible to perform repeat-pass interferometry and tomography with a single satellite [41, 44–48]. This is an important result, as the interferometric and tomographic data may enable the estimation of forest height and the horizontal and vertical forest structure, which may both improve biomass estimation and provide other important forest parameters. Moreover, spaceborne signal transmission at P-band is since 2003 allowed for secondary use within a 6-MHz sub-band with a centre frequency of 435 MHz, and the ionospheric disturbances have been shown to be manageable [3, 49, 50].

In 2013, the European Space Agency (ESA) selected BIOMASS for the 7th Earth Explorer mission [51]. BIOMASS will be the first spaceborne P-band SAR system, designed to provide global biomass, biomass change, and forest height maps. The launch of BIOMASS is currently scheduled for 2020.

To be able to assess the error budget of the future BIOMASS mission, an end-to-end simulator has been implemented [52]. In the simulator, the processing chain of the satellite is modelled, including explicit error sources. Retrieval algorithms are applied to the modelled SAR data, and biomass estimation errors are studied against system parameters. An essential part of the entire simulator is the forward model, which generates the raw reflectivity data.

- In Paper A, a polarimetric-interferometric forward model used within the BIOMASS end-to-end simulator (BEES) is developed and evaluated for SAR image generation in boreal forest. The model predicts the extended covariance matrix from relevant forest and system parameters. It is a hybrid model, in which the polarimetric matrix is modelled with empirical relations, whereas the interferometric matrix is modelled using a theoretical model.

One of the difficulties at P-band is the significant influence of topographic and moisture variations on the backscattered signal. There is a need for a biomass retrieval model which accounts for the influence of these effects, and which can be used on large scales without requiring excessive training data. In the past studies, the proposed models have often been developed and evaluated using small amounts of experimental data, usually limited to acquisitions made in similar conditions over one single test site.

- In Paper B, a new biomass retrieval model is proposed. The model features empirical correction terms, which compensate for the influence of topographic and moisture variations. The model is compared to five other models, and evaluated using airborne SAR data acquired in different seasons and at different flight headings over two boreal test sites in Sweden separated by 720 km. It is concluded that the proposed model performs significantly better than the other models, and it can be successfully used in both test sites with the same model parameters.

1.2.2 X-band

Currently, the highest frequency allocations used for spaceborne SAR imaging of the Earth are located within the X-band (8–12 GHz). As mentioned earlier, the scattering centre is located in the upper part of the canopy, and the principles of SAR interferometry (InSAR) can be used to map the canopy height from small phase differences between two SAR images acquired at slightly different incidence angles [53, 54].

The TanDEM-X system, developed and operated by the German Aerospace Center (DLR), is an X-band SAR interferometer consisting of two almost identical satellites in a tight tandem formation [55]. Its primary goal is to provide the first global, high-resolution digital elevation model (DEM), which will replace the older DEM produced from the data acquired in February 2000 during the Shuttle Radar Topography Mission (SRTM) [56]. With the tight tandem formation, the temporal decorrelation is negligible for most land surfaces [55], and precise DEM estimation can be done even for the dynamic tree canopies.

One of the first models developed for forest parameter estimation from InSAR data is the interferometric water cloud model (IWCM), originally designed for stem volume estimation from multi-temporal, repeat-pass InSAR data acquired with the C-band systems ERS-1/2 [57, 58]. In the IWCM, forest canopy is modelled as a random volume with canopy gaps. To account for the temporal changes occurring between the two interferometric acquisitions, temporal decorrelation is modelled, separately for the ground and volume contributions. Model agreement with observations is achieved through the inclusion of empirical relations between stem volume and forest height, and between stem volume and backscatter intensity.

In the traditional use of the IWCM with single-polarised, single-baseline data, the number of model parameters is larger than the number of available observables. Therefore, model parameters are treated as constants for each acquisition, and estimated from training data. Using multi-temporal data, the influence of the acquisition geometry and environmental variables on model parameters can be studied, and biomass estimation algorithms can be developed.

- In Paper C, biomass is estimated from multi-temporal TanDEM-X acquisitions using the IWCM and two simpler models, one neglecting canopy gaps and one neglecting both canopy gaps and ground contribution. Model training and validation are conducted on two separate data sets. It is concluded that the most accurate biomass prediction can be achieved with the IWCM, i.e., when both canopy gaps and the ground contribution are modelled, and for InSAR data with large interferometric baselines.

The acquisition of training data is generally laborious and expensive, and, from the operational point of view, it is beneficial to decrease the requirements on field inventories. The random volume over ground (RVoG) model can be seen as a simplified version of the IWCM, with neglected canopy gaps and no temporal decorrelation modelling [59, 60]. Direct inversion of the RVoG model has been shown useful for the estimation of forest height and ground topography from interferometric and fully polarimetric, L- and P-band SAR data, without the need for model training [45, 60, 61]. For the standard, single-polarised TanDEM-X image pairs acquired within the global mapping campaign, direct inversion of the RVoG model requires further simplifications, to balance the number of parameters with the number of observables [62, 63].

By using an external, high-resolution digital terrain model (DTM), the number of model parameters can be reduced. With the general stability of the ground surface in forested areas, the increasing availability of DTMs acquired within national lidar scanning campaigns, as well as the future P- and L-band (1–2 GHz) InSAR missions (BIOMASS, and possibly also SAOCOM-CS [64] and TanDEM-L [65]), the availability of high-resolution DTMs will increase in time, making this approach useful from an operational point of view.

At X-band, canopy openings are often large in comparison to the wavelength, and the effective attenuation in the canopies is significant. It is therefore motivated to allow penetration to occur only through canopy gaps. This results in a two-level model, in which forest is represented by two discrete scattering levels, with their relative contributions to the total backscattered field determined by the canopy closure. With an external DTM, direct inversion of this two-level model is feasible even for the single-polarised case, and estimates of both forest height and canopy density may be obtained. This may also lead to improved biomass estimation performance.

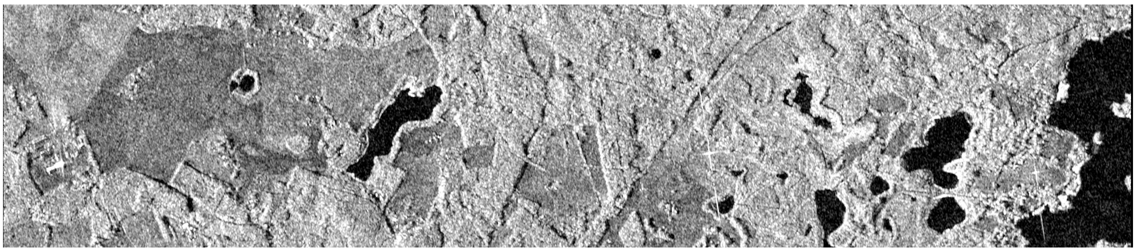
- In Paper D, it is observed that the ground-level contribution is significant in boreal forests at X-band. Based on that observation, an interferometric two-level model (TLM) is introduced, in which penetration through forest canopy can occur only through canopy gaps. The model has a simple form and can be inverted if a high-resolution DTM is available, without the need for multiple SAR acquisitions. It is shown that TLM inversion provides accurate estimates of both forest height and canopy density.

- In Paper E, estimates of forest height and canopy density obtained from the TLM inversion described in Paper D are used as biomass predictors. The new model is evaluated on experimental data from two boreal test sites in Sweden. It is concluded that the model performance is excellent within each test site (in fact, close to that achievable with ALS data). However, similarly to the case of ALS-based mapping, the same model parameters cannot be used in both test sites, due to the significant differences in the structure of forest canopies in the two test sites.

1.3 Thesis Scope and Structure

The main scope of this thesis is to develop methods for forest parameter estimation from SAR. At P-band, focus is put on the influence of topographic and moisture variations on biomass retrieval from SAR intensity data, in view of the future BIOMASS mission. At X-band, focus is put on the influence of the canopy gaps on scattering and the potential of X-band InSAR for large-scale mapping of forest height, canopy density, and biomass, for example with the existing global TanDEM-X data.

This thesis consists of two main parts: **Part I**, in which the motivation for this thesis has been given in this chapter, the basics of SAR are presented in the next chapter (Chapter 2), scattering from forests is studied in Chapter 3, the appended papers are summarised in Chapter 4, and some final conclusions and future prospects are presented in Chapter 5, and **Part II**, containing the appended papers, which are the main scientific outcome of this thesis.



Chapter 2

Synthetic Aperture Radar

In this chapter, synthetic aperture radar is introduced. First, the principles of the basic, one-dimensional radar measurements are presented. Thereafter, a second dimension is added, and high-resolution radar imaging is introduced. Lastly, more advanced techniques, expanding radar imaging capabilities beyond two dimensions, are presented.

2.1 1D: Radar Basics

Radar is an active remote sensing technique in which electromagnetic (EM) signals are transmitted, and the reflected echoes are detected, processed, and analysed [67–69]. The principles of radar are similar to the principles of echolocation, which is an ultrasonic navigation technique used by bats and toothed whales. Since radar is an active system, no external illumination is needed. Also, the terrestrial atmosphere is almost transparent to EM waves with frequencies up to approximately 10 GHz, meaning that radar systems can see through clouds [68, 70, 71]. Note, however, that the influence of the ionosphere increases with decreasing frequency, thus effectively reducing the potential usefulness of frequencies below a few hundred MHz for spaceborne radar [3, 71].

Although the development of radar-like systems started already in the beginning of the 20th century, it was first in 1940 that the term "radar" was introduced by

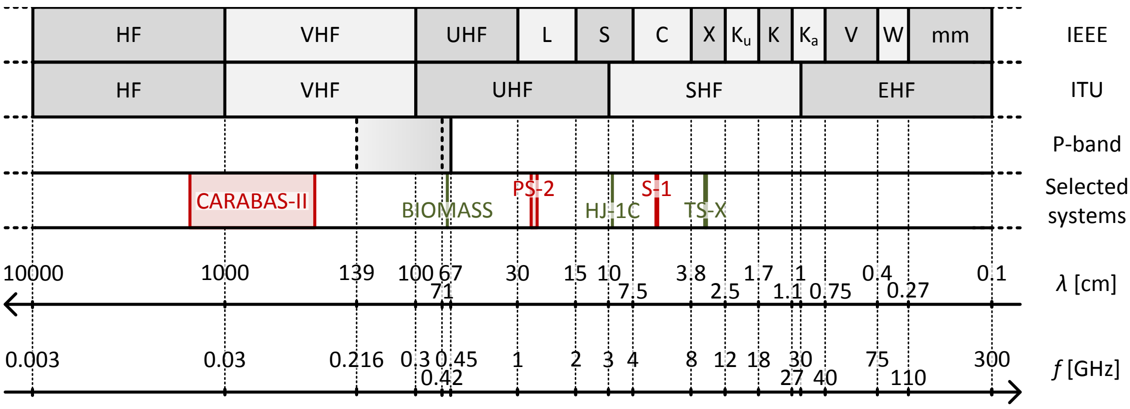


Figure 2.1: Radar frequency bands according to the IEEE and ITU standards. Logarithmic scales are used for frequency f and wavelength λ. Commonly, P-band is defined as the interval 420–450 MHz, but the broader interval 216–450 MHz can also be encountered in the literature. Frequency bands used by some selected SAR systems are also shown. CARABAS-II was an airborne system from the Swedish Defence Research Agency (FOI) [66]. The European BIOMASS satellite is planned to be the first P-band SAR in space, whereas the Japanese ALOS-2 PALSAR-2 (PS-2), Chinese HJ-1C, European Sentinel-1 (S-1), and German TerraSAR-X (TS-X) are examples of current spaceborne L-, S-, C-, and X-band SAR systems, respectively, see Section 2.4.

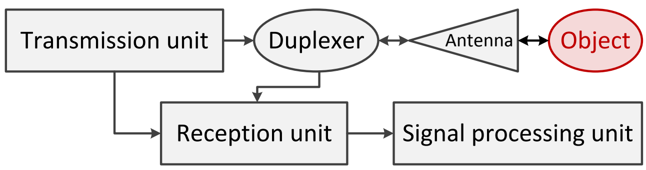


Figure 2.2: Typical setup for a monostatic radar system.

the US Navy, as an abbreviation for "*ra*dio *d*etection *a*nd *r*anging" [72, 73]. Modern radar systems are capable of not only detection and ranging, but also velocity measurements, shape and size determination of objects, angular measurements, and multi-dimensional mapping, and their applications include parking assistance in cars, traffic speed monitoring, airport surveillance, rain rate and wind mapping for weather forecasting, aircraft guidance and detection, planetary mapping from satellites, and space object monitoring from the ground [67, 68].

Since the scattering properties of objects are dependent on the frequency of the transmitted EM waves, the choice of the centre frequency determines the area of application of a radar system. To simplify the classification of radar systems, frequency bands have been introduced [67]. There are two main standards, one defined by the Institute of Electrical and Electronics Engineers (IEEE) and one by the International Telecommunication Union (ITU), and both nomenclatures are summarised in Figure 2.1. Although still used by the radar community, the P-band is an old band designation and it has been replaced by the UHF-band. P-band is commonly defined as the interval 420–450 MHz, although the broader interval 216–450 MHz can also be encountered in the literature [74, 75].

A radar system in which both transmission and reception are done from the same position is called monostatic. In case of multiple positions, the system is called multistatic (or bistatic in the case of two positions). The typical setup for a monostatic radar system is schematically depicted in Figure 2.2. In the transmission unit, an EM signal with the desired waveform and power is generated. The signal is then directed towards the antenna via a duplexer, which often is a switching device, alternating the functions of the radar unit between transmission and reception. The antenna transmits a wave into a medium (usually air), the wave is scattered from an object, and an echo is registered back at the antenna. The received signal passes again through the duplexer, which now guides it towards the reception unit. In the reception unit, the received signal is processed, sampled, and transmitted to a digital signal processing unit.

2.1.1 Range Measurements

The distance between the target and the antenna is called range and it can be estimated from the time delay between the signal transmission and the reception of the echo. If this time delay is denoted T, then the corresponding range is [76]:

$$R = \frac{cT}{2}, \tag{2.1}$$

where a factor of 2 accounts for the two-way propagation and c is the propagation velocity of the EM waves in the medium. In most radar applications, the latter is air and $c = c_0$ is often assumed, where c_0 is the speed of light in vacuum.

In the simplest radar systems, bursts of monochromatic EM waves are transmitted, and the best achievable range resolution is proportional to the pulse length [68]:

$$\delta_R = \frac{c\tau}{2}. \tag{2.2}$$

Thus, fine resolution can be achieved with short pulses, which decreases the signal-to-noise ratio (SNR) and aggravates signal detection. To avoid this problem, most modern radar systems use coded pulses with bandwidth B. For such systems, the best achievable range resolution is [68]:

$$\delta_R = \frac{c}{2B}. \tag{2.3}$$

The resolution is now inversely proportional to the signal bandwidth, and both fine resolution and high SNR can be achieved simultaneously.

2.1.2 Power Measurements

The signal-to-noise ratio (SNR) is a useful benchmark for system quality. The SNR is the ratio between the received signal energy and the noise power density, and for a monostatic system using coherent integration of multiple pulses, it is determined by [68]:

$$\mathrm{SNR} = \frac{P_{\mathrm{avg}}\lambda^2 G^2 t_{\mathrm{dwell}}\sigma}{(4\pi)^3 R^4 L C_B k_B T_s}, \tag{2.4}$$

where P_{avg} is the average power transmitted by the system during the dwell time (integration time) t_{dwell}, λ is the wavelength, G is the antenna gain, R is the range, L is a factor representing losses, C_B is the filter mismatch factor ($C_B = 1$ for a perfectly matched filter), k_B is Boltzmann's constant, T_s is the system noise temperature, and σ is the radar cross section (RCS). The RCS is the main observable in a radar system and it is the effective cross-section area of the target, measured in square metres. The RCS depends not only on the EM properties and the shape of the target, but also on system parameters such as polarisation, angle of incidence, and frequency.

2.1.3 Velocity Measurements

If a scatterer positioned within the antenna beam is moving radially relative the antenna, a frequency shift known as the Doppler shift will occur. The Doppler shift f_D can be computed using [68]:

$$f_D = -\frac{2\dot{R}(t)}{\lambda}, \tag{2.5}$$

where

$$\dot{R} = \frac{\mathrm{d}R(t)}{\mathrm{d}t} \tag{2.6}$$

is the radial velocity of the scatterer.

2.2 2D: Synthetic Aperture Radar (SAR)

Radar imaging can be achieved by sweeping the radar antenna, for example by mounting it in a side-looking configuration on an airborne- or spaceborne-platform. The target position is then determined by its range and by the along-track position of the antenna. Commonly, these two dimensions are referred to as slant range and azimuth, whereas the projected ground distance to the scatterer is called ground range. These concepts are illustrated in Figure 2.3.

For a radar antenna with along-track aperture size D_x such that:

$$D_x \gg \lambda, \tag{2.7}$$

the beamwidth is approximately [68]:

$$\Delta\phi \approx \frac{\lambda}{D_x}, \tag{2.8}$$

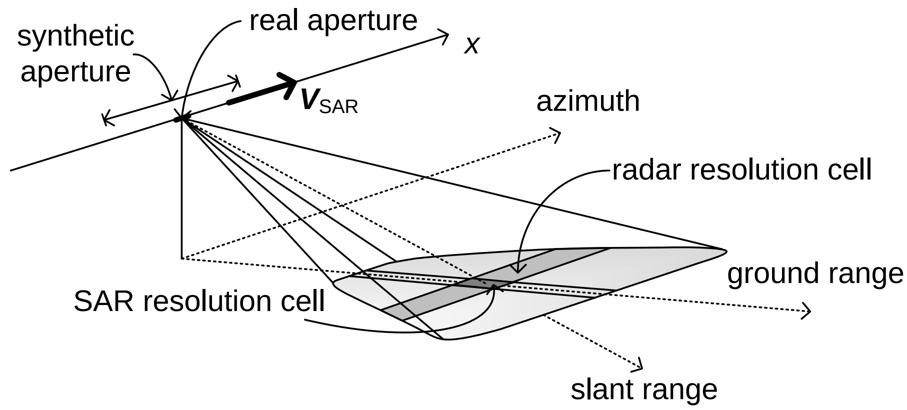


Figure 2.3: Basic imaging radar geometry under flat-earth approximation.

and the azimuth resolution at range R can be approximated by [68]:

$$\delta_x \approx \Delta\phi R \approx \frac{\lambda R}{D_x}. \tag{2.9}$$

As it can be observed in (2.9), the azimuth resolution is proportional to range and inversely proportional to the antenna aperture. Fine azimuth resolution therefore requires large antennas, which is both impractical and expensive, especially in the spaceborne case, for which a moderate resolution of 100 m at X-band would require an antenna with an aperture size of a few hundred metres.

Synthetic aperture radar (SAR) presents a solution to the aforementioned resolution problem. Through coherent signal processing, a larger, synthetic aperture can be created using the phase information from several consecutive pulse echoes.

This technique was first proposed in the early 50's, and further developed during the decades that followed [67]. In 1978, the first civilian, spaceborne SAR system called Seasat was launched by NASA [77], and in the early 90s, Europe (ESA), Russia, Japan, and Canada followed with their own systems. In the early 2000s, additional SAR satellites were launched by countries such as Germany, Italy, India, China, and South Korea, and both Spain and Argentina are planning to launch their first systems within a few years. Presently, more than 15 civilian SAR satellites are operating and around 10 are planned to be launched within the forthcoming five years [6, 78].

2.2.1 SAR Image Formation

The main concept of SAR is to synthesise a large antenna using multiple pulse echoes received by a smaller antenna moving along a known path [76,79]. From (2.9), a small

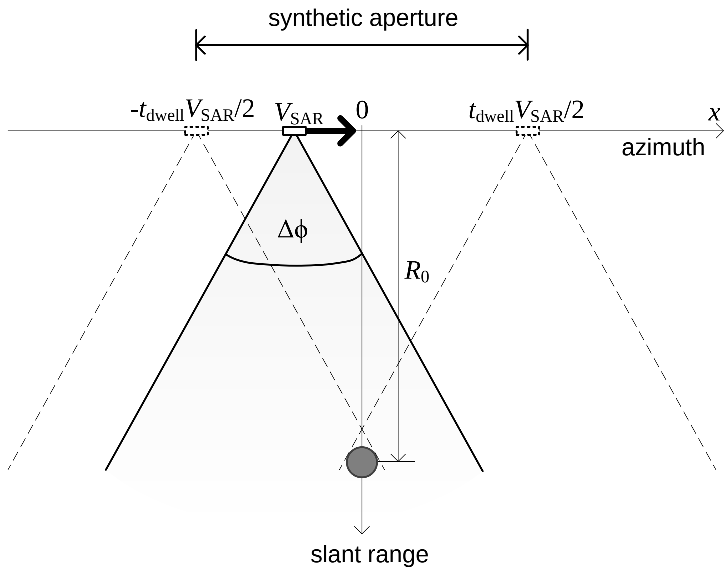


Figure 2.4: Basic slant range plane geometry of SAR.

radar antenna has a large footprint. An arbitrary scatterer on the ground is then covered by several consecutive pulse echoes received at different azimuth positions, and the relative radial velocity $\dot{R}$ of that scatterer is different at each azimuth position. Therefore, the Doppler shift induced by the relative motion of the scatterer changes with azimuth, creating a Doppler bandwidth. With coherent signal processing, azimuth resolution can be improved in the same way as range resolution is improved using coded pulses.

Assume that a radar antenna is travelling along the x-axis on a straight path with velocity V_{SAR}, according to the simplified geometry shown in Figure 2.4. Let R_0 be the range of closest approach to an arbitrary scatterer, and let t be the azimuth time such that $t = t_0 = 0$ is the time of closest approach. The instantaneous distance between the radar antenna and the scatterer can be computed using the Pythagorean theorem:

$$R(t) = \sqrt{R_0^2 + (V_{\text{SAR}}t)^2} \approx R_0 \left(1 + \frac{(V_{\text{SAR}}t)^2}{2R_0^2}\right), \tag{2.10}$$

where a Taylor expansion has been used under the assumption that $V_{\text{SAR}}t \ll R_0$.

The radial velocity of the scatterer relative the antenna can be computed from (2.10):

$$\dot{R}(t) = \frac{\mathrm{d}R(t)}{\mathrm{d}t} \approx \frac{V_{\text{SAR}}^2 t}{R_0} \tag{2.11}$$

and the corresponding Doppler frequency can be computed using (2.5) and (2.11):

$$f_D(t) = -\frac{2\dot{R}(t)}{\lambda} \approx -\frac{2V_{\mathrm{SAR}}^2 t}{\lambda R_0}. \tag{2.12}$$

For the simplified slant range geometry depicted in Figure 2.4, the integration time can be computed from the antenna footprint in (2.9) and the antenna velocity V_{SAR}:

$$t_{\mathrm{dwell}} \approx \frac{\lambda R_0}{V_{\mathrm{SAR}} D_x}. \tag{2.13}$$

The Doppler bandwidth is the difference between the maximal and minimal Doppler frequencies. It can be computed using (2.12) and (2.13) as:

$$B_D = f_D\left(t_{\mathrm{dwell}}/2\right) - f_D\left(-t_{\mathrm{dwell}}/2\right) \approx \frac{2V_{\mathrm{SAR}}}{D_x}. \tag{2.14}$$

Similar to the range resolution for a coded pulse with bandwidth B, see (2.3), SAR azimuth resolution can be computed from the ratio of the antenna velocity V_{SAR} and the Doppler bandwidth B_D [68]:

$$\delta_x = \frac{V_{\mathrm{SAR}}}{B_D} = \frac{D_x}{2}, \tag{2.15}$$

which means that the azimuth resolution of a SAR image can be as good as half the aperture length of the antenna.

The SAR mode presented above, with fixed antenna direction, is called stripmap SAR [68, 80]. Better resolutions, but lower coverage, can be achieved by focussing the antenna at the same point along the whole synthetic aperture, in a mode called spotlight SAR [68, 79]. A better coverage, but lower resolution, can be achieved by sweeping the antenna in different directions, in modes such as scan SAR and TOPS [68, 81].

Note that there is an essential difference in the way SAR images are resolved as compared to optical imagery. Optical imagery features constant resolution angle in both range and azimuth direction [68]. Far-range pixels are therefore resolved with lower resolution than the near-range pixels. In SAR, pixels are resolved at constant slant-range resolution. When projected to the ground, far-range pixels have better resolution than near-range pixels (assuming flat earth). For spotlight SAR, azimuth resolution is range-dependent, with better resolution achieved in near-range.

2.2.2 SAR Image Processing

The processing of the raw data acquired by a SAR platform can be summarised in three main steps: focussing, radiometric calibration, and geocoding. In the first step, high-resolution images are created. In the second step, pixel values are corrected so that they carry meaningful information. In the third step, the image is re-sampled to a cartographic projection so that it can be easily compared with other types of geographic information.

2.2.2.1 Focussing

Essentially, SAR image focussing consists of a 2D matched filtering, which removes the range and azimuth coding. Using the fast Fourier transform, matched filtering can be performed in the frequency domain at low computational costs. However, one of the main difficulties in frequency-domain processing is range cell migration (RCM), which is the movement of scatterers through resolution cells as the antenna moves along the synthetic aperture [79, 82]. Unless compensated for, the RCM will cause azimuth de-focussing.

In the beginning of digital SAR processing, computational costs were of great concern and many different frequency-domain algorithms have been developed [79,82]. The algorithms differ in the way they deal with the RCM, computational costs, and accuracy.

SAR focussing can also be achieved in the time domain, using back-projection algorithms [83–86]. Time-domain algorithms are generally easier to implement and the errors introduced by the uncertainties in the recorded flight path can often be treated using autofocus techniques. Traditionally, the large computational costs have been a major disadvantage of the time-domain algorithms, but the modern back-projection algorithms, such as the fast-factorised back-projection [86], are both accurate and computationally efficient.

2.2.2.2 Radiometric Calibration

To ensure that the focussed SAR image carries meaningful information, radiometric calibration needs to be performed [76, 87]. In this step, the effects of the range-dependent spreading loss, systematic variation in range due to residual effects of antenna pattern, as well as platform roll and yaw movements (around the velocity vector and the vertical axis, respectively) are treated. Moreover, the influence of system noise and SAR focussing also need to be considered during radiometric calibration.

There are two main approaches to radiometric calibration: internal and external [88, 89]. The internal calibration process is done using pre-flight and in-flight measurements of the effects of each element of the radar system. The external calibration process uses targets with known RCS positioned within the imaged scene, preferably at many different positions [88, 90]. The targets may be active (transmitters) or passive, discrete or distributed. Commonly used passive calibration targets are di- and trihedral corner reflectors, and dense forests.

For the imaging of distributed and dynamic targets, it is useful to average and normalise the measured reflectivity in order to reduce the stochastic variations and to remove the residual range dependence. Scattering coefficient sigma nought is often used for surface imaging, as it removes range-dependence caused by the fact that the resolution cell covers a larger ground area in near-range [91]:

$$\sigma^0 = \frac{\langle \sigma \rangle}{A_{\mathrm{GR}}}, \tag{2.16}$$

where A_{GR} is ground area covered by the resolution cell and $\langle \bullet \rangle$ denotes spatial average. The area A_{GR} can be computed from the slant range area A_{SR} using [92,93]:

$$A_{\mathrm{GR}} = \frac{A_{\mathrm{SR}}}{\cos \psi_i}, \tag{2.17}$$

where ψ_i is the angle between the image plane normal and the ground surface normal.

Scattering coefficient gamma nought is often used for the imaging of volumetric scatterers, as it compensates for the residual range-dependence in σ^0 caused by different penetration depths [94, 95]:

$$\gamma^0 = \frac{\sigma^0}{\cos \theta_i}, \tag{2.18}$$

where θ_i is the local angle of incidence.

2.2.2.3 Geocoding

In this step, the focussed SAR image is interpolated from radar geometry to a cartographic projection [70]. A geocoded reference height map is needed, e.g., in the form of a DEM, a geoid model, or an ellipsoid model. From the information about the antenna track, the range and azimuth positions of each pixel in the reference height model can be computed, yielding a geocoding look-up table, which can be used for the interpolation of the focussed image. The look-up table can be fine-tuned using image processing techniques, for example by cross-correlating the focussed intensity image with a synthetic intensity image simulated from a DEM with a simple scattering model.

2.2.3 SAR Image Properties

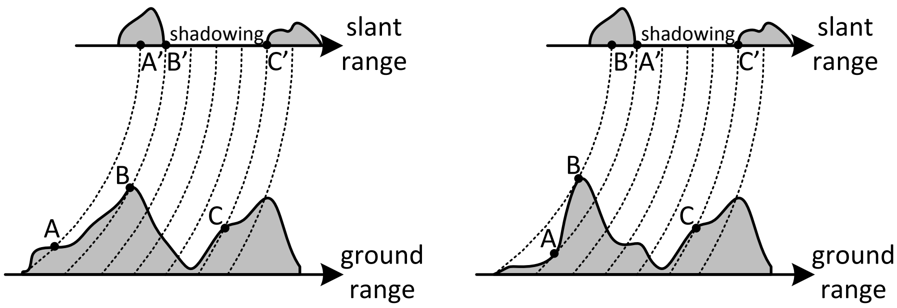
There are a few effects visible in SAR imagery which need to be considered during image analysis.

2.2.3.1 Geometric Distortions in SAR Images

SAR images are created in a side-looking geometry and they are projections of the three-dimensional world on the two-dimensional range-azimuth plane. Therefore, geometrical distortions and ambiguities are unavoidable. These distortions are especially visible in hilly and mountainous regions, as well as urban areas.

The effect of foreshortening occurs on sloping grounds, where two points, which may be significantly separated in the horizontal direction, appear closer on the radar due to the slope effect, see points A and B in Figure 2.5(a).

In the case of even steeper slopes, or steep incidence angle, the effect of layover may occur and cause the two points to appear in reversed order, as shown in Figure 2.5(b). This means that a mountain peak may sometimes appear closer than a mountain foot.



(a) Foreshortening (b) Layover

Figure 2.5: Three geometrical distortions visible in SAR images are here shown schematically. Foreshortening and layover are illustrated by points A and B in (a) and (b), respectively. Shadowing can be seen in (a) between points B' and C' and in (b) between points A' and C'.

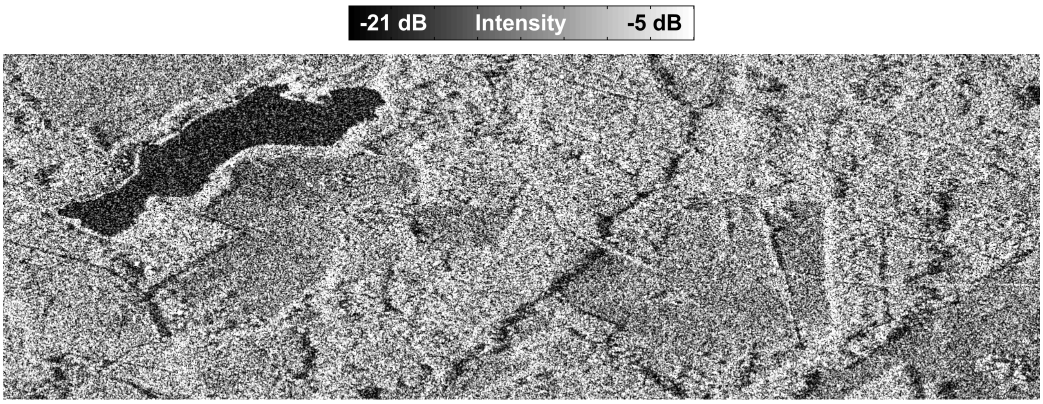


Figure 2.6: A single-look TerraSAR-X intensity image (in dB) over the Remningstorp test site. Speckle is the cause of the grainy texture in the image. Pixel size is 0.9 m in slant range (horizontal direction) and 6.6 m in azimuth (vertical direction), and the image has a size of 1500×500 pixels.

The effect of shadowing occurs when the radar signal is blocked by a large scatterer, e.g., a mountain, and a dark area is visible behind it, like between points B and C in Figure 2.5(a) and Figure 2.5(b).

2.2.3.2 Speckle and SAR Image Statistics

In SAR systems, coherent waves are used for imaging. If there are many scatterers within a resolution cell, interference between the scattered fields may cause an effect

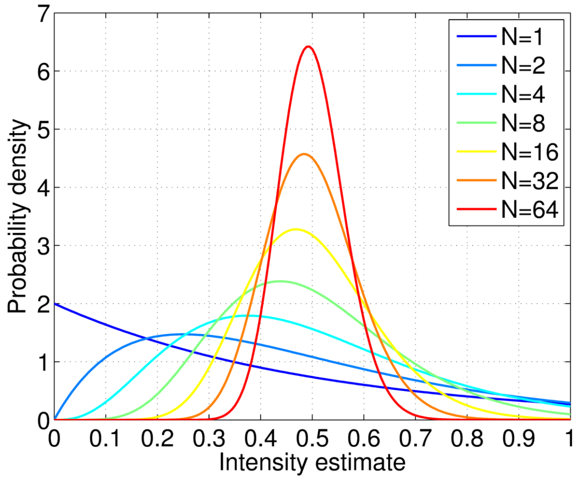


Figure 2.7: Pdf for the N-look RCS estimate $\hat{\sigma}_N$ ($\sigma = 0.5$).

called speckle. Speckle introduces a graininess in SAR images, which deteriorates image quality, see Figure 2.6

Speckle statistics can be studied in the case of fully developed speckle [76, 96], i.e., under the assumption that a resolution cell contains a large number of similar scatterers randomly distributed in range over an extent much larger than the wavelength. The central limit theorem implies that the real and imaginary parts I and Q of the total backscattered field, which is the sum of the individual signals backscattered by each scatterer, are independent, normally distributed random variables with zero mean and variance $\sigma/2$ (where σ is the RCS and the factor of 2 has been chosen to make the total variance equal to σ) [76, 96]. From this follows that the phase is uniformly distributed between $-\pi$ and π, the amplitude $\sqrt{I^2 + Q^2}$ has a Rayleigh probability distribution, and the intensity $\hat{\sigma}_1 = I^2 + Q^2$ has a negative exponential distribution with the following pdf:

$$p(\hat{\sigma}_1|\sigma) = \begin{cases} \frac{1}{\sigma}e^{-\frac{\hat{\sigma}_1}{\sigma}} & \hat{\sigma}_1 \geq 0 \\ 0 & \text{otherwise} \end{cases} . \tag{2.19}$$

As mentioned earlier, speckle appears in SAR imagery as a noise-like pattern, and it is most often an unwanted nuisance, aggravating the performance of segmentation algorithms and general image interpretation. One way to reduce speckle in intensity images is by non-coherent averaging, i.e., multilooking. There are several ways to achieve this: by splitting the range or Doppler spectrum into several parts, processing each part separately, and then averaging the final images; by spatial averaging of intensity (or amplitude, although intensity averaging has been shown superior [88]); or by averaging of several SAR images over the same scene (assuming stable scatterers). With non-coherent averaging, the variance in the image is decreased, but at the cost of resolution or additional acquisitions. Many studies have been devoted to speckle filtering and there are many different filtering algorithms available [96].

The N-look intensity $\hat{\sigma}_N = \frac{1}{N}\sum_{i=1}^{N}(I_i^2 + Q_i^2)$, where I_i and Q_i are assumed to be independent, normally distributed random variables, is a gamma distributed random variable with the following pdf [76, 96]:

$$p(\hat{\sigma}_N|\sigma, N) = \begin{cases} \frac{N^N \hat{\sigma}_N^{N-1}}{(N-1)!\sigma^N} e^{-\frac{N\hat{\sigma}_N}{\sigma}} & \hat{\sigma}_N \geq 0 \\ 0 & \text{otherwise} \end{cases}. \tag{2.20}$$

In Figure 2.7, the pdf for the N-look intensity estimate $\hat{\sigma}_N$ is plotted for $\sigma = 0.5$. It can be observed that with an increasing number of looks, the distribution becomes more symmetric around the mean, and the variance decreases.

In Figure 2.8, the effect of multilooking is visualised in a simulated SAR image. In Figure 2.8(a) and Figure 2.8(b), the real and imaginary parts are shown. The central part of the image has slightly larger variance, but it can barely be seen. In Figure 2.8(c), the complex phase is shown, and it is completely random. In Figure 2.8(d), the amplitude is shown, and a feature can be observed. In Figures 2.8(e)–2.8(i), the effect of number of looks on the variance is shown. The reference RCS is shown in Figure 2.8(j).

It can be shown that the expectation value and the variance of an N-look intensity estimate are [76, 96]:

$$\mathrm{E}\left[\hat{\sigma}_N\right] = \sigma, \tag{2.21}$$

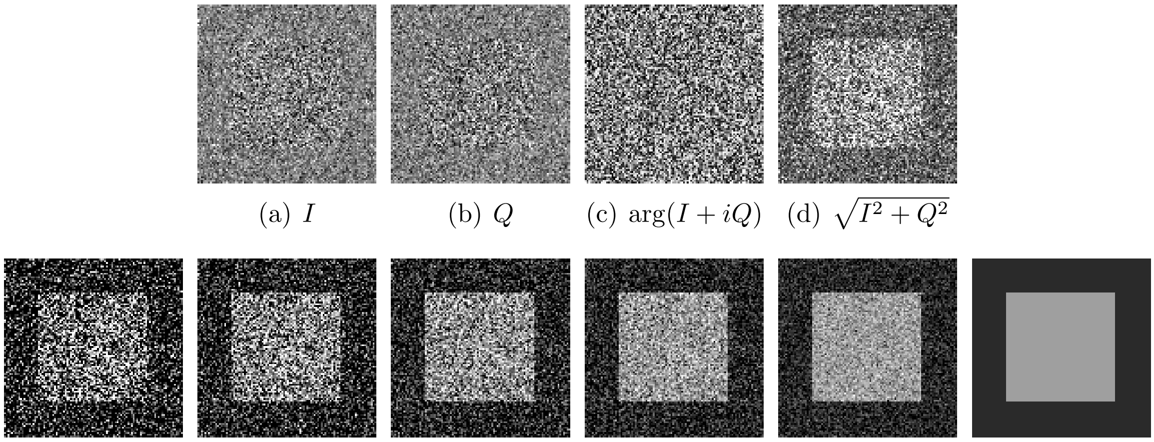
$$\mathrm{Var}\left[\hat{\sigma}_N\right] = \frac{\sigma^2}{N}. \tag{2.22}$$

The N-look intensity is therefore an unbiased estimate of the RCS, with variance decreasing as $\frac{1}{N}$.

2.3 3D: Advanced SAR Techniques

This far, the acquisition, processing, and properties of a single SAR image have been studied. However, if multiple acquisitions are available, additional information can be extracted from the data. Two advanced SAR techniques that will be discussed here are: SAR polarimetry (PolSAR) [96, 97], in which multiple polarisations are used to differentiate between different scattering mechanisms, and SAR interferometry (InSAR) [53], in which multiple SAR acquisitions made at slightly different incidence angles or at different occasions are used to study the position or the movement of the scattering phase centre.

PolSAR and InSAR form the basis for additional advanced SAR techniques: polarimetric SAR interferometry (PolInSAR) [59, 97], in which the principles of polarimetry and interferometry are combined; SAR tomography [98–100], in which a synthetic aperture is created in the vertical direction, and vertical scattering profiles



(a) I (b) Q (c) $\arg(I + iQ)$ (d) $\sqrt{I^2 + Q^2}$

(e) $\widehat{\sigma}_1$ (f) $\widehat{\sigma}_2$ (g) $\widehat{\sigma}_4$ (h) $\widehat{\sigma}_8$ (i) $\widehat{\sigma}_{16}$ (j) σ

Figure 2.8: Simulation showing the effect of number of looks on the intensity estimate. Figures (a)–(d) show the real and imaginary parts, phase, and amplitude of a one-look, complex Gaussian-distributed image. Figures (e)–(i) show the N-look intensity estimates in decibels, and Figure (j) shows the reference RCS, which is -3.0 dB for the central part and -6.7 dB for the outer part. Grayscale intervals are, from black to white: (a,b) $[-1, 1]$, (c) $[-\pi, \pi]$, (d) $[0, 1]$, and (e–j) $[-8\,\mathrm{dB}, 0\,\mathrm{dB}]$.

---

are estimated; polarimetric SAR tomography [47], in which the principles of polarimetry and tomography are combined; and SAR holography [101, 102], in which circular, tomographic acquisitions are used to create three-dimensional images of the scene. These techniques will not be discussed here, but a good overview can be found in [6].

2.3.1 SAR Polarimetry

One of the basic properties of an EM wave is its polarisation, that is the direction of the electric field oscillations. In the far field from the scatterer, the oscillations of the EM wave are perpendicular to the direction of propagation, and if two perpendicular polarisations are used in a radar system both at transmission and reception, then the full scattering properties of the target at that particular frequency and incidence angle can be measured [96].

One of the most common polarisation bases used in SAR imaging is the horizontal-vertical basis. If the transmission is done with a horizontally polarised antenna (H), and the reception is done with a vertically polarised antenna (V), the polarisation mode is then called VH. Similarly, HH means that horizontal polarisation is used both at transmission and reception. If a system is capable of measuring all four combinations (HH, HV, VV, and VH) at the same time, together with their phase information, it is called fully polarimetric.

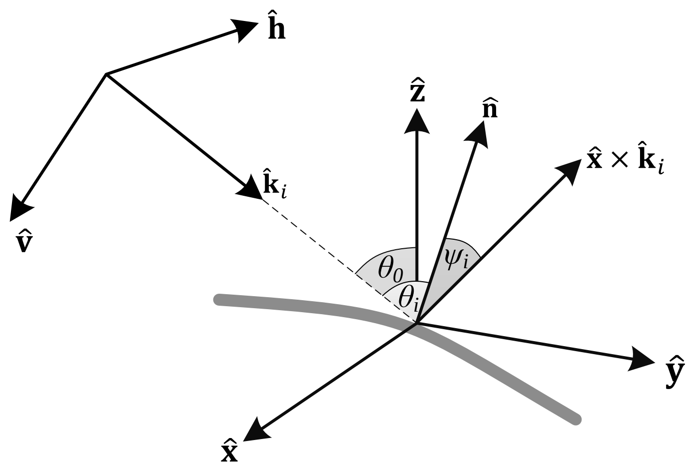


Figure 2.9: Basic scattering geometry. The incident wave is propagating in the direction of $\hat{\mathbf{k}}_i$, and oscillating in the plane defined by the horizontal and vertical vectors $\hat{\mathbf{h}}$ and $\hat{\mathbf{v}}$. The ground surface normal is $\hat{\mathbf{n}}$, θ_0 is the global incidence angle, θ_i is the local incidence angle, and ψ_i is the angle between the surface normal and image plane normal.

The electric field of an incident plane wave propagating in the direction of $\hat{\mathbf{k}}_i$, see Figure 2.9, can be expressed as a sum of two components, one in the horizontal direction and one in the vertical direction [96]:

$$\mathbf{E}^i = E^i_{\mathrm{H}}\hat{\mathbf{h}} + E^i_{\mathrm{V}}\hat{\mathbf{v}}, \tag{2.23}$$

where

$$\hat{\mathbf{h}} = \frac{\hat{\mathbf{z}} \times \hat{\mathbf{k}}_i}{|\hat{\mathbf{z}} \times \hat{\mathbf{k}}_i|} \tag{2.24}$$

is the horizontal unit vector, perpendicular both to the vertical axis and to the direction of propagation, and

$$\hat{\mathbf{v}} = \hat{\mathbf{h}} \times \hat{\mathbf{k}}_i \tag{2.25}$$

is the vertical unit vector, perpendicular both to the horizontal unit vector and the direction of propagation. Equivalently, the incident electric field can be written as a Jones vector [96]:

$$\mathbf{E}^i = \begin{bmatrix} E^i_{\mathrm{H}} \\ E^i_{\mathrm{V}} \end{bmatrix}. \tag{2.26}$$

Assuming plane waves, the Jones vector for the scattered field can be computed from the Jones vector for the incident field using [96]:

$$\mathbf{E}^s = \frac{e^{ikR}}{R}\,[S]\,\mathbf{E}^i, \tag{2.27}$$

where R is the distance between the target and the antenna, $k = 2\pi/\lambda$ is the wavenumber, and

$$[S] = \begin{bmatrix} S_{\mathrm{HH}} & S_{\mathrm{HV}} \\ S_{\mathrm{VH}} & S_{\mathrm{VV}} \end{bmatrix} \tag{2.28}$$

is the complex 2×2 scattering matrix. The scattering matrix fully describes scattering from the target at the current frequency and incidence angle.

The relation between the RCS and the scattering matrix elements can be obtained from the formal definition of the RCS [68]:

$$\sigma = \lim_{R\to\infty} 4\pi R^2 \frac{|\mathbf{E}^s|^2}{|\mathbf{E}^i|^2}, \tag{2.29}$$

which for polarisation PQ can be computed using (2.27):

$$\sigma_{\mathrm{PQ}} = 4\pi |S_{\mathrm{PQ}}|^2. \tag{2.30}$$

In remote sensing, most scatterers are not stable, fixed point targets, but they are distributed, dynamic targets stochastically changing in time and space. Such targets are best described using second order moments. The polarimetric covariance matrix contains all possible covariance combinations of the scattering matrix elements [96]. For a monostatic radar in a reciprocal medium, the cross-polarised terms of the scattering matrix are equal $S_{\mathrm{HV}} = S_{\mathrm{VH}}$ [103], and one of them is usually dropped or they are averaged to improve the SNR. To keep the total power invariant after dropping one element, the remaining cross-polarised term is usually scaled with $\sqrt{2}$ [96]. The polarimetric covariance matrix becomes:

$$[C] = \begin{bmatrix} \langle |S_{\mathrm{HH}}|^2 \rangle & \sqrt{2}\,\langle S_{\mathrm{HH}} S_{\mathrm{HV}}^* \rangle & \langle S_{\mathrm{HH}} S_{\mathrm{VV}}^* \rangle \\ \sqrt{2}\,\langle S_{\mathrm{HV}} S_{\mathrm{HH}}^* \rangle & 2\,\langle |S_{\mathrm{HV}}|^2 \rangle & \sqrt{2}\,\langle S_{\mathrm{HV}} S_{\mathrm{VV}}^* \rangle \\ \langle S_{\mathrm{VV}} S_{\mathrm{HH}}^* \rangle & \sqrt{2}\,\langle S_{\mathrm{VV}} S_{\mathrm{HV}}^* \rangle & \langle |S_{\mathrm{VV}}|^2 \rangle \end{bmatrix}, \tag{2.31}$$

where * is the complex conjugate operator.

Scattering coefficient sigma nought for polarisation mode PQ can be expressed in terms of the diagonal elements in (2.31) using (2.30) and (2.16):

$$\sigma^0_{\mathrm{PQ}} = \frac{\langle \sigma_{\mathrm{PQ}} \rangle}{A_{\mathrm{GR}}} = \frac{4\pi \langle |S_{\mathrm{PQ}}|^2 \rangle}{A_{\mathrm{GR}}} = \frac{4\pi \cos\psi_i \langle |S_{\mathrm{PQ}}|^2 \rangle}{A_{\mathrm{SR}}}, \tag{2.32}$$

where A_{SR} and A_{GR} are the areas of the resolution cells in slant range and ground range planes, respectively, and ψ_i is the angle between the image plane normal and ground surface normal, see Figure 2.9.

Scattering coefficient gamma nought can be computed from (2.32) using (2.18):

$$\gamma^0_{\mathrm{PQ}} = \frac{4\pi \cos\psi_i \langle |S_{\mathrm{PQ}}|^2 \rangle}{A_{\mathrm{SR}} \cos\theta_i}, \tag{2.33}$$

where θ_i is the local incidence angle, see Figure 2.9.

PolSAR data can be used to determine the dominant scattering mechanisms present within each pixel of the imaged scene using polarimetric decomposition theorems. There are several different types of decomposition theorems: coherent decompositions of the scattering matrix $[S]$ (e.g., the Pauli decomposition [96]), eigenvector and eigenvalue-based decompositions of the coherency matrix $[T]$ (e.g., the Cloude-Pottier or H/A/α decomposition [104]), and model-based decompositions of the covariance matrix $[C]$ or the coherency matrix $[T]$ (e.g., the Freeman-Durden three-component decomposition [105]). Polarimetric decomposition theorems are frequently used for land use classification, see [96].

2.3.2 SAR Interferometry

In SAR interferometry, the phase difference between two SAR images acquired over the same scene is used [53]. Two types of interferometry are common: across-track interferometry (XTI), in which the two images are acquired at slightly different incidence angles, and where the phase difference is used to estimate the scattering phase centre elevation, and along-track interferometry (ATI), in which the two images are acquired on different occasions, and where the phase difference is used to estimate the change in position of the scattering centre between the acquisitions. The most common application of the XTI technique is digital elevation model (DEM) creation, whereas the ATI technique can be used to estimate the radial velocity of, e.g., cars, ships, glaciers, ocean currents and waves, as well as ground surface deformations caused by earthquakes, volcanoes, landslides, etc.

The complex correlation coefficient (sometimes also called complex coherence) is the main interferometric observable. For two co-registered, complex SAR images S^1_{PQ} and S^2_{PQ}, it is defined as [53, 97]:

$$\widetilde{\gamma}_{=}\gamma e^{i\Delta\phi} = \frac{\mathrm{E}\left[S^1_{\mathrm{PQ}}S^{2*}_{\mathrm{PQ}}\right]}{\sqrt{\mathrm{E}\left[\left|S^1_{\mathrm{PQ}}\right|^2\right]\mathrm{E}\left[\left|S^2_{\mathrm{PQ}}\right|^2\right]}}, \tag{2.34}$$

where $\gamma = |\widetilde{\gamma}|$ is called coherence and $\Delta\phi$ is the interferometric phase, i.e., the phase difference between the two images. Coherence is a real valued quantity between 0 and 1 and it is a measure of the degree of similarity between the two images. The phase difference is related to the difference in range to the scattering centre between the two images.

The complex correlation coefficient can be described as a product of four separate decorrelation effects [87, 106, 107]:

$$\tilde{\gamma} = \gamma_{\mathrm{SNR}}\tilde{\gamma}_{\mathrm{sys}}\tilde{\gamma}_{\mathrm{sp}}\tilde{\gamma}_{\mathrm{temp}}, \tag{2.35}$$

where the terms marked with the tilde sign may attain complex values.

SNR decorrelation γ_{SNR} is caused by thermal noise in the images and it can be determined from the SNR using [6, 106]:

$$\gamma_{\mathrm{SNR}} = \frac{1}{1+SNR^{-1}}, \tag{2.36}$$

where it has been assumed that both images have the same SNR value.

System decorrelation $\tilde{\gamma}_{sys}$ is introduced by system imperfections, and it includes the effects of dynamic range, quantisation, misregistration, and ambiguities, as well as other errors introduced during SAR and InSAR processing [6, 55]. Phase offsets may also be introduced in radar hardware and during processing. System decorrelation can be minimised with optimised electronics, signal processing, and calibration.

Spatial decorrelation $\tilde{\gamma}_{sp}$ is due to geometric differences between the two images. It can be re-stated as a product of three decorrelation terms [55]:

$$\tilde{\gamma}_{sp} = \tilde{\gamma}_{\mathrm{rg}}\tilde{\gamma}_{\mathrm{az}}\tilde{\gamma}_{\mathrm{vol}}, \tag{2.37}$$

where the first and second terms are decorrelation effects caused by differences in the sampled range and Doppler frequency spectra, respectively, and the last term is volume decorrelation. The first two decorrelation effects can be minimised through common-band filtering of both images [53, 108, 109]. The last term is a very important factor in XTI, as it carries information about the vertical distribution of the scatterers.

Temporal decorrelation $\tilde{\gamma}_{\mathrm{temp}}$ is due to the temporal changes in the imaged scene between the two acquisitions [87, 106, 110]. In the case of single-pass XTI the temporal decorrelation is most often negligible. In the case of repeat-pass XTI, this term is usually a nuisance that causes a loss of quality in the estimated DEM. In repeat-pass ATI, this term carries the information about the phase difference between the two acquisitions.

In the absence of decorrelation effects other than the volume decorrelation $\tilde{\gamma}_{\mathrm{vol}}$, the phase difference $\Delta\phi$ is determined by the difference in range to the scattering phase centre, called ΔR in Figure 2.10, but it is also affected by a 2π-phase ambiguity. The phase difference can be re-stated in terms of three phase components: the phase introduced by the difference in range to a reference height model (ΔR_0), the phase introduced by the elevation of the scattering phase centre above the reference height model (Δh), and a 2π-phase ambiguity [68, 97]:

$$\Delta\phi = mk\Delta R + 2\pi n \approx k_z \left(\Delta R_0 \cos\theta_0 + \Delta h\right) + 2\pi n, \tag{2.38}$$

where m is equal to 2 for a monostatic system and 1 for a bistatic system, and where the vertical wavenumber is defined as [55, 111]:

$$k_z = \frac{mkB_\perp}{R\sin\theta_0}, \tag{2.39}$$

where $B_\perp$ is the perpendicular baseline, θ_0 is the average incidence angle, and R is the average range to the scattering phase centre, as defined in Figure 2.10. The vertical wavenumber is the number of 2π-cycles corresponding to a vertical height shift of one metre.

In order to estimate Δh from $\Delta\phi$, it is necessary to remove the first and last phase components on the right hand side of (2.38). The removal of the phase caused by the differences in range to the reference height model, i.e., the first term in the

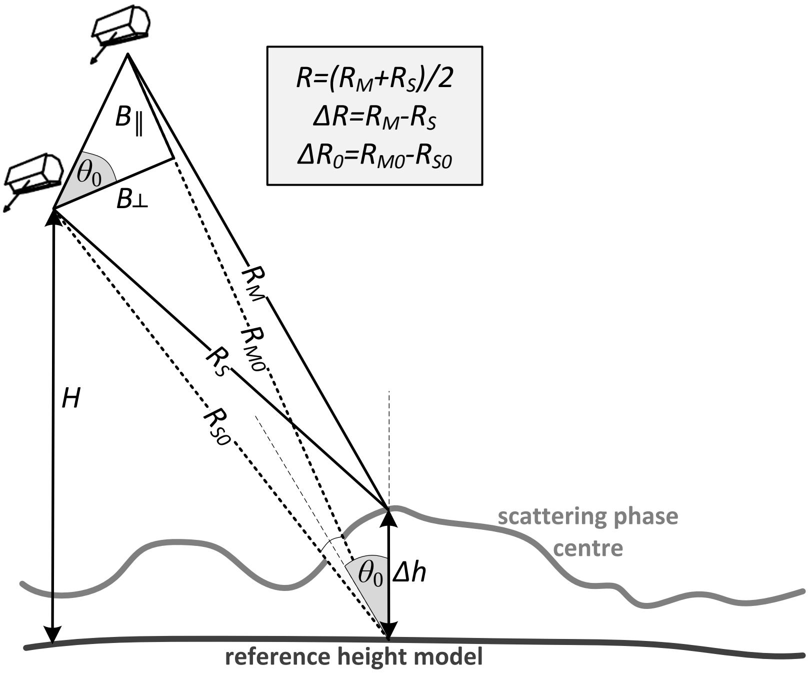


Figure 2.10: Simplified and exaggerated geometry for InSAR measurements. The reference height model may be flat earth, reference ellipsoid or geoid, or a DTM.

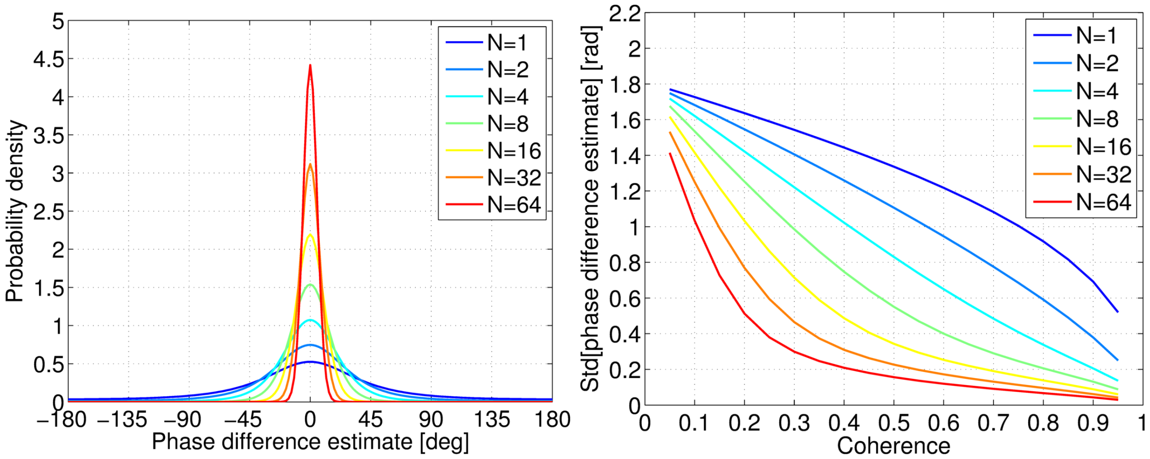
parenthesis in (2.38), is called flattening. Depending on application, different height models may be used: flat earth, geoid, ellipsoid, or a DTM. The removal of the 2π-phase ambiguity is called phase unwrapping. Phase unwrapping is often a non-trivial task and many different unwrapping algorithms have been developed [53].

Once both flattening and unwrapping have been performed, the elevation of the scattering centre above the reference height model can be estimated from the flattened and unwrapped phase difference $\Delta\phi'$ through a simple scaling:

$$\Delta h = \frac{1}{k_z}\Delta\phi'. \tag{2.40}$$

Instead of k_z, the more intuitive interferometric parameter height-of-ambiguity (HOA) is often used. It describes the vertical height shift equivalent to a 2π-phase shift and it is the maximal height difference that can be unambiguously resolved by an interferometric system. HOA is defined as:

$$\text{HOA} = \frac{2\pi}{k_z} = \frac{2\pi R \sin\theta_0}{mkB_\perp}. \tag{2.41}$$



(a) Pdf ($\gamma = 0.7$) (b) Standard deviation

Figure 2.11: Statistics for the N-look phase difference estimate $\widehat{\Delta\phi}_N$.

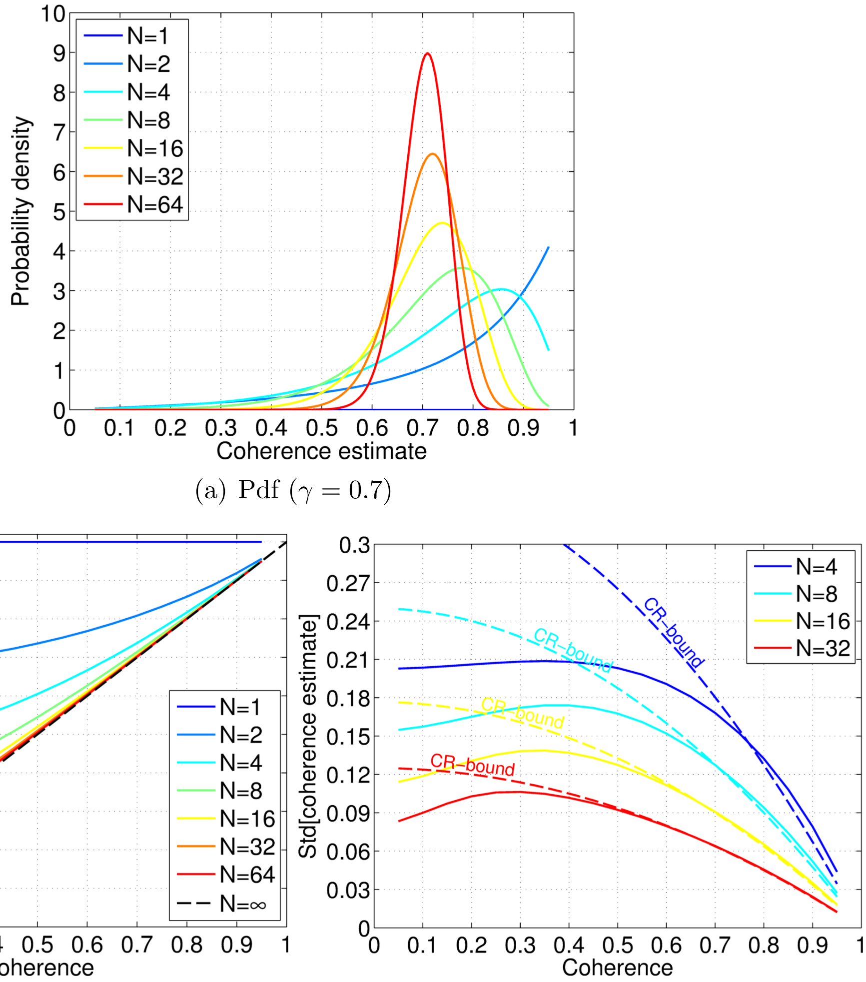
2.3.2.1 InSAR Image Processing

The processing of InSAR imagery is a multi-step process. The first step consists of image co-registration [112, 113], in which one of the images is usually re-sampled to the range-azimuth grid of the second image. This includes both range and azimuth interpolation, as well as spectral filtering of the two images so that they cover the same 2D-frequency spectrum. This filtering procedure minimises the range and azimuth decorrelation effects [53, 108, 109]. The next step consists of interferogram creation, in which the first image is multiplied with the complex conjugate of the second image. Thereafter, flattening is conducted, in which a reference phase is removed (corresponding to flat earth, ellipsoid, geoid, or a DTM). Next, the flattened interferogram is multilooked, and the phase is computed and unwrapped. Finally, the unwrapped phase is scaled to height using a k_z or HOA map computed for the current acquisition geometry. A height calibration may be conducted here, e.g., using ground reference points. After this step, geocoding can be performed, to obtain the final DEM in a cartographic projection.

2.3.2.2 InSAR Image Statistics

In applications, the complex correlation coefficient is estimated using spatial averaging of N samples [53]:

$$\widetilde{\widehat{\gamma}}_N = \widehat{\gamma}_N e^{i\widehat{\Delta\phi}_N} = \frac{\frac{1}{N}\sum_i^N \left(S^1_{i,\mathrm{PQ}} S^{2*}_{i,\mathrm{PQ}}\right)}{\sqrt{\left(\frac{1}{N}\sum_i^N \left|S^1_{i,\mathrm{PQ}}\right|^2\right)\left(\frac{1}{N}\sum_i^N \left|S^2_{i,\mathrm{PQ}}\right|^2\right)}}, \tag{2.42}$$



(a) Pdf ($\gamma = 0.7$)

(b) Expectation value

(c) Standard deviation

Figure 2.12: Statistics for the N-look coherence estimate $\widehat{\gamma}_N$. In (c), the dashed line shows the approximative standard deviation obtained with the Cramér-Rao (CR) bound.

---

where i is the sample index, and $\widehat{\gamma}_N$ and $\widehat{\Delta\phi}_N$ are N-look coherence and phase difference estimates.

The N-look phase difference estimate $\widehat{\Delta\phi}_N$ has the following pdf [53, 96]:

$$p(\widehat{\Delta\phi}_N|\Delta\phi, N, \gamma) = \begin{cases} \frac{\Gamma(N+\frac{1}{2})(1-\gamma^2)^N D}{2\sqrt{\pi}\Gamma(N)(1-D^2)^{N+\frac{1}{2}}} + \frac{(1-\gamma^2)^N}{2\pi} {}_2F_1(N, 1; \frac{1}{2}; D^2) & |\widehat{\Delta\phi}_N| < \pi \\ 0 & \text{otherwise} \end{cases}, \tag{2.43}$$

where $\Gamma(\bullet)$ is the gamma function, ${}_2F_1(\bullet)$ is the Gauss hypergeometric function, and

$$D = \gamma \cos(\widehat{\Delta\phi}_N - \Delta\phi). \tag{2.44}$$

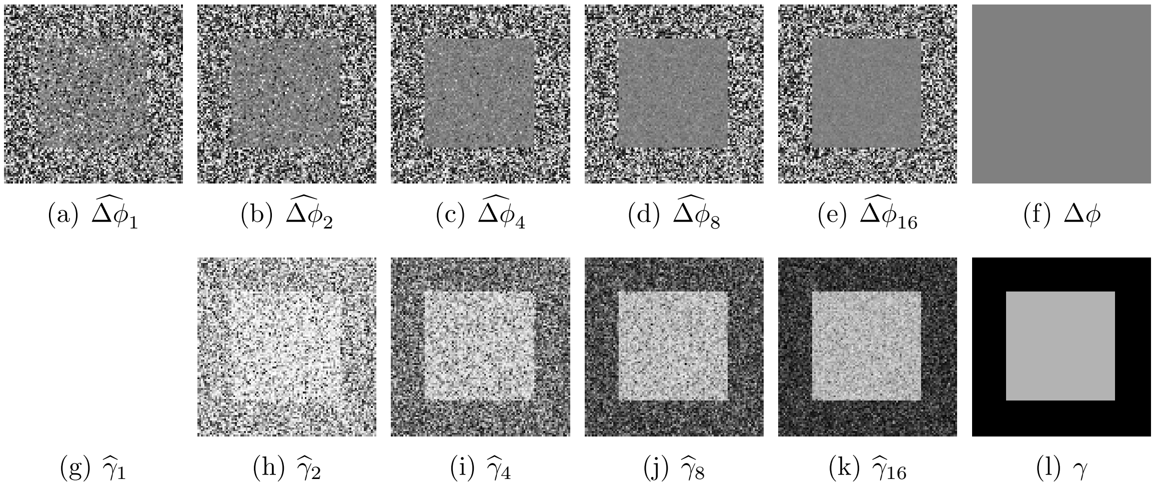


Figure 2.13: Simulation showing the effect of number of looks on the phase difference and coherence estimates. Figures (a)–(e) and (g)–(k) show the N-look phase difference and coherence estimates. Figure (f) shows the reference phase difference, which is zero for the entire image. Figure (l) shows the reference coherence, which is zero in the outer part and 0.7 in the central part. Grayscale intervals are, from black to white: (a–f) $[-\pi, \pi]$ and (g–l) $[0, 1]$. Note that the one-look coherence estimate is always equal to unity, which, for the chosen colour scale, results in a completely white image.

In Figure 2.11(a), the pdf for the N-look phase difference estimate $\widehat{\Delta\phi}_N$ is plotted for $\gamma = 0.7$ and $\Delta\phi = 0$. It can be observed that $\widehat{\Delta\phi}_N$ is an unbiased estimate of $\Delta\phi$, with standard deviation decreasing with increasing N. In Figure 2.11(b), the standard deviation is plotted against coherence for different numbers of looks N. The standard deviation decreases with increasing coherence and with increasing number of looks.

The N-look coherence estimate $\widehat{\gamma}_N$ has the following pdf [53]:

$$p(\widehat{\gamma}_N|\gamma, N) = \begin{cases} 2(N-1)(1-\gamma^2)^N \widehat{\gamma}_N (1-\widehat{\gamma}_N^2)^{N-2} {}_2F_1(N, N; 1; \widehat{\gamma}_N^2\gamma^2) & 0 \leq \widehat{\gamma}_N \leq 1 \\ 0 & \text{otherwise} \end{cases} . \tag{2.45}$$

In Figure 2.12(a), the pdf for the N-look coherence estimate $\widehat{\gamma}_N$ is plotted for $\gamma = 0.7$. It can be observed that for low N, coherence is overestimated and the standard deviation is high, but both the bias and standard deviation decrease with increasing N. In Figure 2.12(b), the expectation value for $\widehat{\gamma}_N$ is plotted against coherence for different numbers of looks N. The observed overestimation for low N is confirmed, and the bias is coherence-dependent. In Figure 2.12(c), the standard deviation of $\widehat{\gamma}_N$ is plotted against coherence for different N, together with the standard deviation given by the Cramér-Rao bound. As this is an approximation valid for unbiased estimates, it can only be used when both N and γ are high [53]:

$$\mathrm{Var}[\widehat{\gamma}] = \frac{(1-\gamma^2)^2}{2N}. \tag{2.46}$$

Simulation results for the N-look phase difference and coherence estimates are shown in Figure 2.13. In Figures 2.13(a)–2.13(e), the N-look phase difference estimates are shown. It can be observed that in the central region, the variance of the phase difference decreases with an increasing number of looks. However, for the outer part of the images, the phase difference is equally noisy for any number of looks. This is due to the fact that coherence in that region is zero. The reference phase difference is zero for the entire image, see Figure 2.13(f).

In Figure 2.13(g), the one-look coherence estimate is shown, and it is equal to unity for the entire image. For an increasing number of looks, the coherence overestimation decreases, see Figures 2.13(h)–2.13(k). The reference coherence is shown in Figure 2.13(l). In the central part, the coherence is 0.7, whereas in the outer part, it is zero.

2.4 Past, Present, and Future SAR Systems

As mentioned earlier, more than 15 civilian, spaceborne SAR systems are currently operational and around 10 are planned to be launched within the next five years [6,78]. Due to the major advantages of polarimetry and interferometry, most modern SAR systems are designed to be able to provide PolSAR and/or InSAR data.

Airborne systems are used to acquire SAR data in cases when satellite SAR data are unavailable or insufficient, e.g., in campaigns associated with preparatory studies for new satellite systems. Airborne SAR systems have the advantage of being easy to deploy and affordable on lower scales.

2.4.1 PolSAR Systems

Presently, two civilian, fully polarimetric SAR systems are operational: the C-band RADARSAT-2 system from the Canadian Space Agency (CSA) [114] and the L-band ALOS-2 PALSAR-2 system from the Japan Aerospace Exploration Agency (JAXA) [115]. Moreover, the X-band TerraSAR-X and TanDEM-X systems from the German Aerospace Center (DLR) are able to provide fully-polarimetric data in the experimental mode [55], whereas the C-band RISAT-1 system from the Indian Space Research Organisation (ISRO) is able to provide hybrid-polarimetric SAR data (circular polarisation on transmission, linear polarisations on reception) [116].

In May 2013, ESA selected BIOMASS for the 7th Earth Explorer mission [51]. BIOMASS will feature a fully-polarimetric SAR sensor operating at the centre frequency of 435 MHz with a bandwidth of 6 MHz, giving a nominal slant range resolution of 25 m [3]. It will be the first P-band SAR sensor in space, with the main goal to provide global biomass, forest height, and deforestation maps. Due to the low frequency, repeat-pass interferometry and tomography over forests will be feasible with BIOMASS. The launch of BIOMASS is currently scheduled for 2020.

Papers A and B appended to this thesis summarise the results obtained within the BIOMASS feasibility study. In these papers, P-band SAR data acquired with the airborne Experimental-SAR (E-SAR) system [117–119] from the DLR are used, acquired within the BioSAR 2007 and 2008 campaigns [41,120]. The centre frequency of this system was 360 MHz, which is lower than for BIOMASS, with a bandwidth of up to 100 MHz. The E-SAR system was decommissioned in 2008 and replaced by the F-SAR system [121].

2.4.2 InSAR Systems and Satellite Constellations

All modern SAR systems are coherent and repeat-pass interferometry can, in theory, be conducted using any satellite system, but the temporal decorrelation often limits the applications. Temporal decorrelation can be decreased either by using the lower frequency bands (as in the case of the P-band SAR system BIOMASS), or by using multiple sensors. Moreover, the use of multiple sensors in a constellation significantly increases both the coverage and acquisition frequency, which is beneficial for large-scale monitoring and reconnaissance purposes.

Two current SAR constellation missions capable of providing global imagery at daily rates are: the dual-purpose (civil and military) COSMO-SkyMed system of four X-band satellites from the Italian Space Agency (ASI) [122], and the military SAR-Lupe system of five X-band satellites from the DLR [123]. The civilian Sentinel-1a satellite, launched by ESA in 2014, is the first of two C-band SAR systems, which together will be able to provide SAR imagery of the entire Europe every third day [124]. Moreover, the S-band system HJ-1C, launched in 2012, is the first of four SAR satellites planned for a constellation mission from the Chinese Academy of Science and Technology (CAST) [125].

The X-band PAZ system, funded by the Spanish Center for Development of Industrial Technology (CDTI) and scheduled to be launched in 2014, is planned to operate in constellation with the almost identical satellites TerraSAR-X and TanDEM-X [126, 127]. The SAOCOM-1A/B mission from the Argentine National Commission for Space Activities (CONAE) will consist of two fully-polarimetric, L-band SAR satellites [128], planned to be launched in 2015 and 2016, respectively. The RADARSAT Constellation Mission of three C-band satellites from the CSA is planned to be launched in 2018, with the main task to provide daily imagery of the Canadian lands and oceans [129].

Single-pass interferometric SAR systems have the advantage of low temporal decorrelation, but they require multiple SAR sensors in a close formation. Currently, spaceborne single-pass InSAR data can be acquired only with the TanDEM-X system, consisting of the previously mentioned satellites TerraSAR-X and TanDEM-X in a tight helix formation [55]. The main goal of the TanDEM-X mission is to acquire the first, fully global DEM with a spatial resolution of $12\,\mathrm{m} \times 12\,\mathrm{m}$ and an absolute vertical accuracy better than $10\,\mathrm{m}$ [130], which will replace the older DEM acquired within the Shuttle Radar Topography Mission (SRTM) from the space shuttle Endeavour in February 2000. Papers C, D, and E appended to this thesis are focussed on forest parameter estimation from TanDEM-X data.

The proposed ESA SAOCOM-CS mission will feature a passive SAR sensor in a formation flight with the SAOCOM-1B satellite [64]. Also, a TanDEM-L mission has been proposed, consisting of two L-band satellites in a tight tandem formation and providing fully-polarimetric, high-resolution data in a wide-swath mode [65].

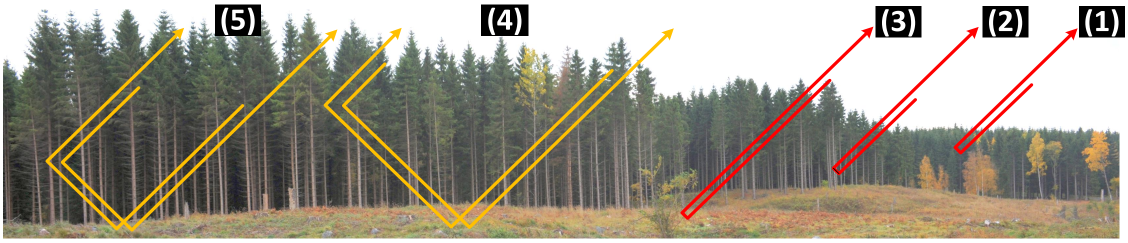


Figure 3.1: Some of the most significant scattering mechanisms in forest: direct backscatter from (1) tree canopies, (2) tree trunks, and (3) ground surface, and double-bounce interactions between (4) ground surface and tree canopies, as well as between (5) ground surface and tree trunks.

Chapter 3

Microwave Scattering from Forests

When studying forests with SAR imagery, it is important to be able to separate the influence of system parameters from the influence of geo- and biophysical forest parameters. Scattering models are important tools which can improve the understanding of the electromagnetic interactions.

Accurate modelling of electromagnetic scattering from forests can be achieved by means of computational electromagnetics, in which Maxwell's equations are solved numerically for a discretised forest model. However, the computational costs of this approach can be enormous, especially if multiple evaluations are required, e.g., in case when the influence of a system parameter on the scattering coefficient needs to be examined.

In many cases, the usefulness of a simplified model is of higher priority than extreme accuracy, e.g., when the model is to be used to explain the cause of a particular effect observed in the experimental data. In such cases, simplifications can be made by replacing complicated forest elements with simple objects, for which fast analytical solutions are available, and by reducing the number of modelled interactions.

The accuracy of such simplified models depends primarily on two factors: if all the most important scattering mechanisms are modelled, and how well each scattering mechanism is modelled.

3.1 Basic Scattering Mechanisms

In some of the most common models [132–136], forest is composed of three types of elements: ground surface, tree trunks, and tree canopies. Each of these elements contributes both on its own and in combination with the other elements to the total backscattered field. The most significant, low-order scattering mechanisms shown in Figure 3.1 are:

(1) direct backscatter from tree canopies,

(2) direct backscatter from tree trunks,

(3) direct backscatter from the ground surface,

(4) double-bounce interactions between the ground surface and tree canopies,

(5) double-bounce interactions between the ground surface and tree trunks.

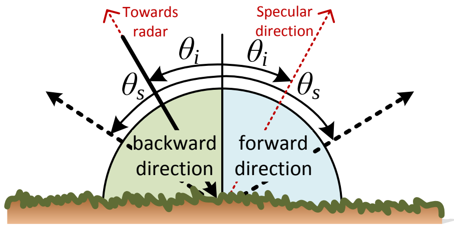
Higher-order mechanisms and near-field interactions can often be neglected. Due to reciprocity, the two double-bounce interactions between the same elements, but in reversed order (e.g., ground-trunk and trunk-ground), are equivalent and they add up in phase, acting as one single scattering mechanism [137]. In the following, oblique incidence angles and vertical or near-vertical tree trunks will be assumed.

3.1.1 Direct Backscatter

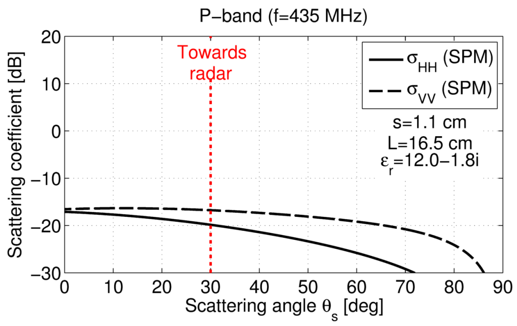
Direct backscatter is primarily caused by reflections from surfaces facing the radar antenna. Therefore, most of the direct backscatter occurs in the canopy, from rough surfaces, and from occasional slopes facing the radar antenna.

Canopies consist of more or less randomly oriented branches of different sizes, as well as needles or leaves. Canopy backscatter is generally stronger at high frequencies, when the scatterers are comparable to, or larger than the wavelength [95]. Consequently, penetration through the canopy is expected to decrease with increasing frequency. If the scatterers in the canopy have a preferred orientation, the backscattered field will show polarisation dependence.

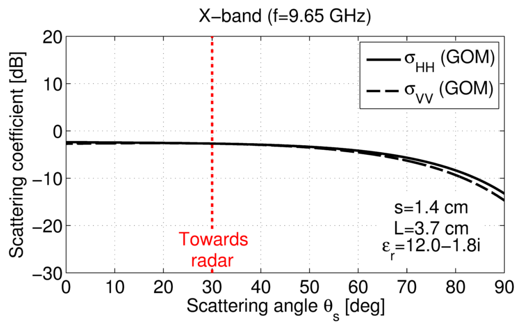
Surface roughness is characterised in relation to the wavelength, and a surface which is rough at high frequencies may be smooth at low frequencies [91, 138]. At oblique incidence, backscattering from a randomly rough surface is stronger than from a smooth surface. Many realistic surfaces, such as forest soil and bark, may be treated as slightly rough surfaces when the wavelength is much longer than the surface irregularities (e.g., at P-band) and very rough surfaces when the surface irregularities



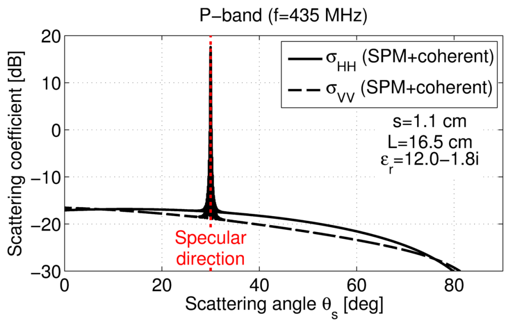
(a) Scattering geometry



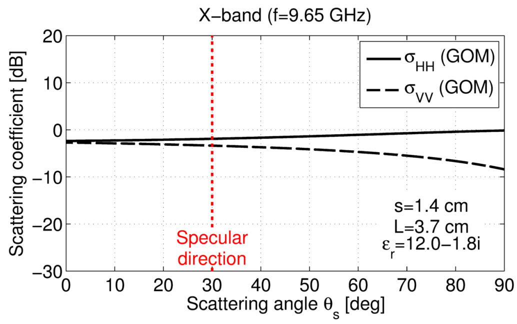
(b) Backward direction (slightly rough surface at P-band)



(c) Backward direction (very rough surface at X-band)

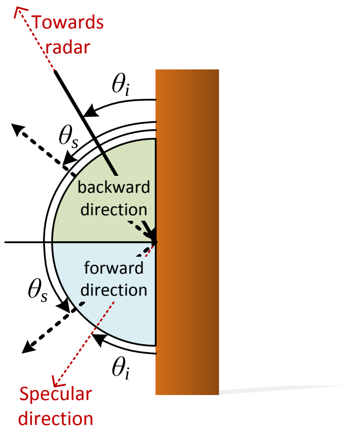


(d) Forward direction (slightly rough surface at P-band)

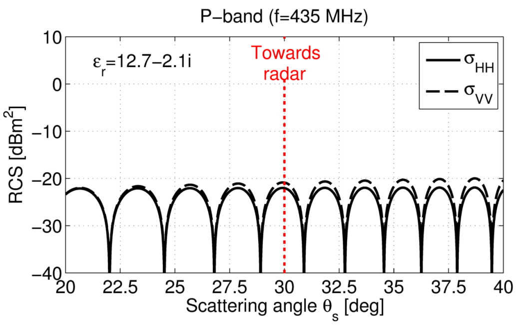


(e) Forward direction (very rough surface at X-band)

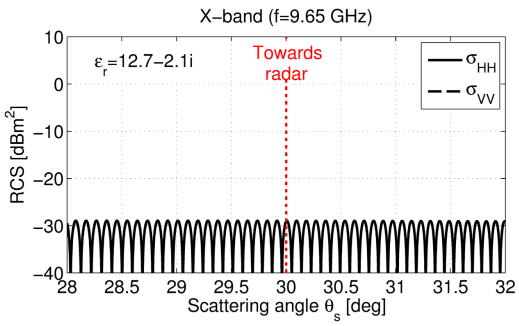
Figure 3.2: Scattering from a slightly rough surface (P-band) and a very rough surface (X-band), corresponding to low- and high-frequency regimes, respectively. In all cases, the incidence angle θ_i is 30° and only scattering in the incidence plane is studied. Non-coherent rough surface scattering has been modelled using the geometrical optics model (GOM) in the high-frequency regime and the small perturbation model (SPM) in the low-frequency regime, using the expressions found in [103]. Coherent scattering is only significant for the slightly rough surface in the forward direction, where it has been modelled using the expression presented in [91] and assuming a circular surface with a radius of 100 m. Parameters s and L are the standard deviation of the vertical variations and the correlation length of the rough surface, and together with the relative dielectric constant ε_r, their values have been chosen based on [131].



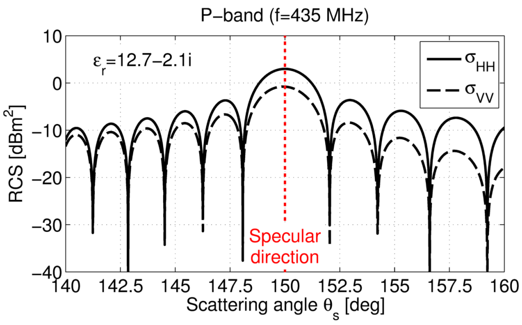
(a) Scattering geometry



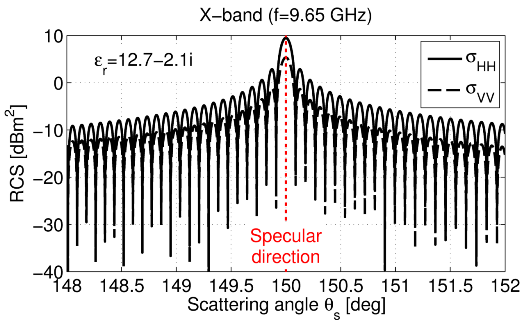
(b) Backward direction (P-band)



(c) Backward direction (X-band)



(d) Forward direction (P-band)



(e) Forward direction (X-band)

Figure 3.3: Scattering from a smooth vertical cylinder modelled using the infinite cylinder approximation [103]. In all cases, the incidence angle θ_i is 30° and only scattering in the incidence plane is studied. The cylinder has a length of 20 m, radius of 13 cm, and the relative dielectric constant ε_r chosen according to [17]. Note the significant difference in the scales of the x-axes.

are comparable to, or larger than the wavelength (e.g., at X-band). Therefore, direct backscatter from such surfaces is expected to be stronger at high frequencies.

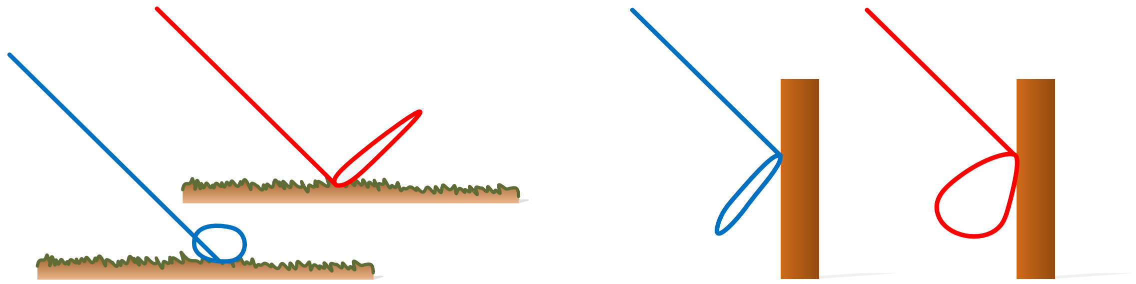
In Figure 3.2(b) and Figure 3.2(c), rough surface scattering in the backward direction is modelled at P- and X-bands using the small perturbation model (SPM) and the geometrical optics model (GOM), respectively. The SPM is a low-frequency approximation, valid when surface roughness is small in comparison to the wavelength, whereas the GOM is a high-frequency approximation, valid for very rough surfaces [103]. At P-band, with a centre frequency of 435 MHz and a wavelength of 69 cm, the chosen roughness parameters (standard deviation of vertical variations equal to 1.1 cm and the surface correlation length equal to 16.5 cm) correspond to a slightly rough surface. Backscattering is thus very weak, see Figure 3.2(b). At X-band, with a centre frequency of 9.65 GHz and a wavelength of 3.1 cm, the chosen roughness parameters (standard deviation equal to 1.4 cm and correlation length equal to 3.7 cm) correspond to a very rough surface. Backscattering is therefore much stronger, see Figure 3.2(c). Note that roughness parameters have been chosen according to [131], for optimal validity of the asymptotic models. The surfaces are therefore slightly different for the two studied cases (P- and X-band).

In Figure 3.3(b) and Figure 3.3(c), scattering in the backward direction from a smooth, vertical cylinder is modelled at both P- and X-bands using the infinite cylinder approximation [103]. As it can be observed, the backscattering coefficient is low in both cases. Note, however, that in more realistic scenarios, trunks are not smooth vertical cylinders, but they attain a rougher and possibly also curved form. Therefore, backscattering from a tree trunk will be stronger in reality.

3.1.2 Double-Bounce Interactions

Double-bounce interactions require a strong specular reflection from the ground as well as from the trunks and/or canopies. A smooth surface will generally have a stronger specular component than a rough surface [91, 138], and for many natural rough surfaces, double-bounce interactions will be more common for the lower frequencies. The strength of the double-bounce interactions between the ground surface and vertical trunks will be affected by the ground slope and by the relative length of the trunks. At lower frequencies, the trunks will be shorter relative the wavelength, resulting in a wider forward scattering lobe [139], and the double-bounce interaction will be less sensitive to ground slope. For a ground surface tilted in the azimuth direction, double-bounce interaction will be present also in the cross-polarised channel [17].

In Figure 3.2(d) and Figure 3.2(e), rough surface scattering in the forward direction is modelled at P- and X-bands using the SPM (with a coherent term) and the GOM, respectively, at an incidence angle of 30°. For P-band and the slightly rough surface, forward scattering is weak in all cases except when the scattering angle is around 30°, i.e., in the specular direction, when the coherent term becomes significant. The width of the coherent scattering lobe depends on the size of the surface; a larger



(a) Rough surface (b) Smooth cylinder

Figure 3.4: Simplified visualisations of rough surface scattering and reflection from a smooth cylinder in **high**- and **low**-frequency regimes (in blue and red, respectively).

surface results in a narrower lobe and a stronger specular reflection. For X-band and the very rough surface, forward scattering in non-specular directions is stronger than for a slightly rough surface, whereas the coherent component is negligible.

In Figure 3.3(d) and Figure 3.3(e), scattering in the forward direction from a smooth, vertical cylinder is modelled at both P- and X-bands using the infinite cylinder approximation. The forward scattering lobe is wider for P-band, although the scattering coefficient in the specular direction is lower.

Double-bounce interactions between the ground surface and tree canopies are often weak, due to the high penetration capabilities at low frequencies, when the specular reflection from the ground is the strongest.

3.1.3 Dominant Mechanisms at P- and X-bands

At P-band, direct backscatter from larger branches in the canopies and ground slopes facing the antenna, as well as the double-bounce interactions between ground and trunks are expected to be the strongest contributions, the latter due to the strong specular reflection from relatively smooth surfaces and wide forward scattering lobe from the trunks, see Figure 3.4. The strength of these contributions is related to the volume of the trunk, but it also depends on the dielectric properties of the elements, roughness of both ground and trunk surfaces, and the ground topography. Several models of dielectric cylinders over ground have been developed for lower frequencies [17, 140–143].

At X-band, direct backscatter is expected to be the main contribution, due to the random orientation of the canopy scatterers, and the relatively rough trunk and ground surfaces, see Figure 3.4. Direct backscatter from the rough trunk and ground surfaces is expected to be strong only in the case of sparse canopies, i.e., when penetration through canopy gaps is significant. In general, low sensitivity to biomass is expected for the backscattered signal at X-band due to the strong dependence on

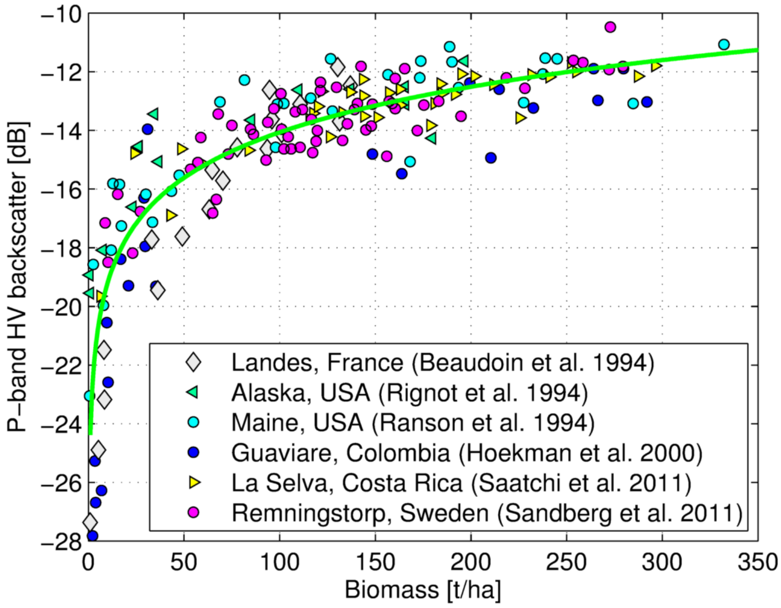


Figure 3.5: A power law function relating backscatter coefficient gamma nought to biomass is here plotted with the green, solid line, together with backscatter data from six airborne SAR campaigns. The referenced studies can be found in [26, 28, 32, 36, 37, 144]. The figure has been adapted from [38]. Note that the original data from La Selva and Remningstorp have been corrected with +5 dB and -3 dB, respectively, to compensate for the observed offset from the other four data sets.

surface roughness and the weak dependence on the trunk volume.

Polarisation and incidence angle dependence is expected to be stronger at P-band, due to the stronger contribution of the ground surface and tree trunks, which are often less randomly oriented than the canopy scatterers.

3.2 Models for Forest Scattering

Scattering models serve as a link between theory and empirical observations. Therefore, the development of models can be approached from both ends, and two general model types can be distinguished: empirical and theoretical.

3.2.1 Empirical Models

Empirical models are derived from observations in the experimental data. Using regression analysis, functions can be fitted to the data and used to explain the observed behaviour. As both the experimental data and model selection strategies may vary

(a) IWCM (b) RVoG

Figure 3.6: Two vertical backscatter profiles used by the IWCM and RVoG models.

significantly between different studies, the number of empirical models available in the literature is very large.

The power law is one of the more popular models, which can be used to explain many phenomena observed in nature [145,146]. In SAR remote sensing of forests, this function is often used to describe the relation between the backscattering coefficient and biomass, especially for HV-polarised, P-band data. In [38], the following power law function has been fitted to HV-polarised data from six airborne P-band SAR campaigns conducted in different biomes on three different continents:

$$\gamma^0_{\mathrm{HV}} = a\mathcal{B}^b, \tag{3.1}$$

where $\mathcal{B}$ is biomass in tons per hectare and a and b are model parameters. After correction of an observed offset in γ^0_{HV} in two of the six data sets, the same parameters a and b can be used for all six data sets, see Figure 3.5.

3.2.2 Theoretical Models

Theoretical models are created using simple objects for which approximative analytical solutions exist (cylinders, rough surfaces, discs, needles, dipoles, etc.). There are two main types of theoretical models: coherent and non-coherent.

Coherent models are based on wave propagation and Maxwell's equations, and the contributions from the different scattering mechanisms are added in phase. Common coherent models include models based on the cylinder-over-ground approximation, which are primarily used to model the trunk-ground interactions at lower frequencies [17, 140, 141, 143, 147–149], as well as more complex models with more elements and higher-order interactions [136, 150–152].

Non-coherent models are based on energy propagation and radiative transfer equations. Therefore, no correlation between the fields scattered by the different elements is assumed, and the contributions of the different scattering mechanisms are added in terms of power. The analytical treatment of these models is generally simpler. Many non-coherent models have been developed in the past [132, 133, 153–157].

3.2.2.1 Interferometric Models

In across-track interferometry, modelling of volume decorrelation can be done from the vertical backscattering profile $\sigma_\mathrm{v}(z)$ using [106, 110]:

$$\tilde{\gamma}_\mathrm{vol} = \frac{\int_{-\infty}^{\infty} \sigma_\mathrm{v}(z) e^{ik_z z} \mathrm{d}z}{\int_{-\infty}^{\infty} \sigma_\mathrm{v}(z) \mathrm{d}z}. \tag{3.2}$$

Two simple models have been frequently used since the late 90's for forest parameter estimation from InSAR data. In both models, vegetation canopy is modelled as a water cloud, as proposed in [94], and ground is modelled as an impenetrable surface. As the scattering centres of the double-bounce interactions are located at the ground level, the double-bounce interactions do not have to be modelled separately.

In the interferometric water cloud model (IWCM) [57, 58], vegetation is modelled as a homogeneous volume of randomly oriented scatterers located above a ground plane and covering a certain fraction of the total area, called the area-fill factor, see Figure 3.6(a). It has been shown in [94], that the effective attenuation of the random volume can be described by an exponential backscatter profile function. The vertical backscattering profile $\sigma_\mathrm{v}(z)$ can then be formulated as:

$$\begin{aligned} \sigma_\mathrm{v}(z) &= \eta \left[\sigma^0_\mathrm{veg} \alpha e^{-\alpha(z_0+h-z)} \Theta(z_0+h-z)\Theta(z-z_0) \right. \\ &\quad \left. + \sigma^0_\mathrm{gr} \delta(z-z_0) e^{-\alpha h}\right] + (1-\eta)\delta(z-z_0)\sigma^0_\mathrm{gr}, \end{aligned} \tag{3.3}$$

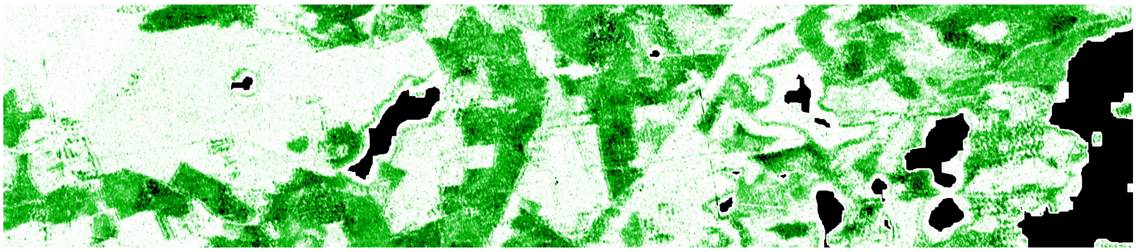
where η is the area-fill factor, σ^0_gr and σ^0_veg are the effective ground and vegetation backscattering coefficients, α is the extinction coefficient, z_0 is the ground elevation, h is the volume height, $\delta(\bullet)$ is the Dirac delta function, and $\Theta(\bullet)$ is the Heaviside step function. By inserting (3.3) into (3.2), the total volume decorrelation can be obtained.

The IWCM was originally developed for stem volume estimation from repeat-pass interferometric, C-band ERS-1/2 data [57, 58], where it was fitted to coherence and backscatter data, which were found stable in winter conditions in various forests [158–161] (the phase information was not used due to high temporal decorrelation [162]). Therefore, the full formulation of the IWCM also includes two temporal decorrelation terms (one for the ground and one for the volume parts) and an allometric relation between height and stem volume. Additionally, an empirical model relating stem volume to backscattering coefficient and first published in [163] is also included, to make backscatter modelling agree with the empirical observations. IWCM fitting is usually done collectively for all data points using training data, effectively making the model parameters into site-dependent constants.

In the random volume over ground (RVoG) model [59–61, 164], the profile used in the IWCM is simplified by neglecting the canopy gaps, see Figure 3.6(b):

$$\sigma_\mathrm{v}(z) = \sigma^0_\mathrm{veg} \alpha e^{-\alpha(z_0+h-z)} \Theta(z_0+h-z)\Theta(z-z_0) + \sigma^0_\mathrm{gr} \delta(z-z_0) e^{-\alpha h}. \tag{3.4}$$

The traditional use of the RVoG model includes forest height and ground elevation estimation from PolInSAR data [44,45,59–61,164,165]. Model fitting is generally done on a pixel-by-pixel basis, by balancing the number of observables with the number of parameters. In recent years, the RVoG model was further developed, and the estimation of vertical scattering profiles [99, 166] as well as temporal decorrelation [167] from PolInSAR data were introduced.



Chapter 4

Summary of the Appended Papers

This chapter begins with a short presentation of the reference data used in the appended papers. Thereafter, the appended papers are presented, first the two P-band papers and then the three X-band papers.

4.1 Experimental Data

Two test sites located in Sweden are used in the appended papers: Remningstorp and Krycklan, see Figure 4.1. The distance between these test sites is approximately 720 km.

Remningstorp (58° 28' N, 13° 38' E) is a hemi-boreal forest site situated in southern Sweden, approximately 150 km north-east of Gothenburg (Göteborg). The test site is fairly flat, with ground slopes at stand level lower than 5° (computed from a 50 m × 50 m DTM). Remningstorp covers approximately 1200 ha of productive forest land, and the forest consists primarily of Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and birch (*Betula* spp.). Remningstorp has been used in the two ESA-funded campaigns BioSAR 2007 and 2010 [41,42], conducted in support to the BIOMASS feasibility study [3].

Krycklan (64° 14' N, 19° 46' E) is a boreal forest site located in northern Sweden, approximately 50 km north-west of Umeå. Compared to Remningstorp, Krycklan has a more strongly undulating topography, with ground slopes on stand level reaching

Figure 4.1: The two test sites used in the appended papers, Remningstorp and Krycklan, separated by 720 km.

19° (computed from a 50 m × 50 m DTM). Krycklan covers approximately 6700 ha of forested land, which is dominated by Norway spruce and Scots pine. Krycklan has been used in the ESA-funded campaign BioSAR 2008 [120], conducted in support to the BIOMASS feasibility study [3].

For each test site, several sets with reference data have been provided by the Swedish University of Agricultural Sciences (SLU). The data can be divided in three categories: plot-level data, stand-level data, and maps.

Stands are relatively homogenous forest regions with similar species composition, biophysical characteristics (e.g., tree height and tree number density), and management procedures. They can vary in size and shape, and they are the main unit used for forest mapping and management [11]. Plots are usually smaller stand subsets of regular shape, which are used as within-stand samples. Often, they are distributed in a systematic grid covering the test site or a stand.

Although the exact methodologies used during field inventories may vary between the individual reference data sets, the general approach is similar. Commonly, the relevant and easily accessible parameters such as stem diameter at breast height (dbh) and tree species are sampled for all trees confined within plot borders and fulfilling a minimum dbh criterium (dbh larger than 4 cm in Krycklan; dbh larger than 5 cm in Remningstorp). For a subset of these trees, the more time-consuming measurements of, e.g., tree height and age, are made. Plot- and stand-level estimates are then computed from the sampled tree parameters, and biomass is estimated using allometric equations, for example the Marklund or Petersson formulas [168, 169]. In recent years, the forest management system Heureka [170] has been frequently used for forest parameter estimation from *in situ* data.

Maps of forest parameters are usually derived from airborne lidar scanning (ALS) data. In ALS, laser pulses transmitted downwards from an aircraft or a helicopter are used to sample canopy height at high vertical and horizontal resolutions. From the sampled pulses, different lidar-based estimates of forest parameters are obtained, and biomass maps are created using regression analysis and plot-level reference data.

The accuracy of plot-, stand-, and map-level biomass estimates depends on many factors: measurement error, the uncertainty in the allometric models, natural variation in the data, sampling density, etc. Moreover, correlation between error sources is not uncommon and need to be considered. When evaluating the performance of SAR-based biomass estimation algorithms, it is important to consider the uncertainties in the reference data.

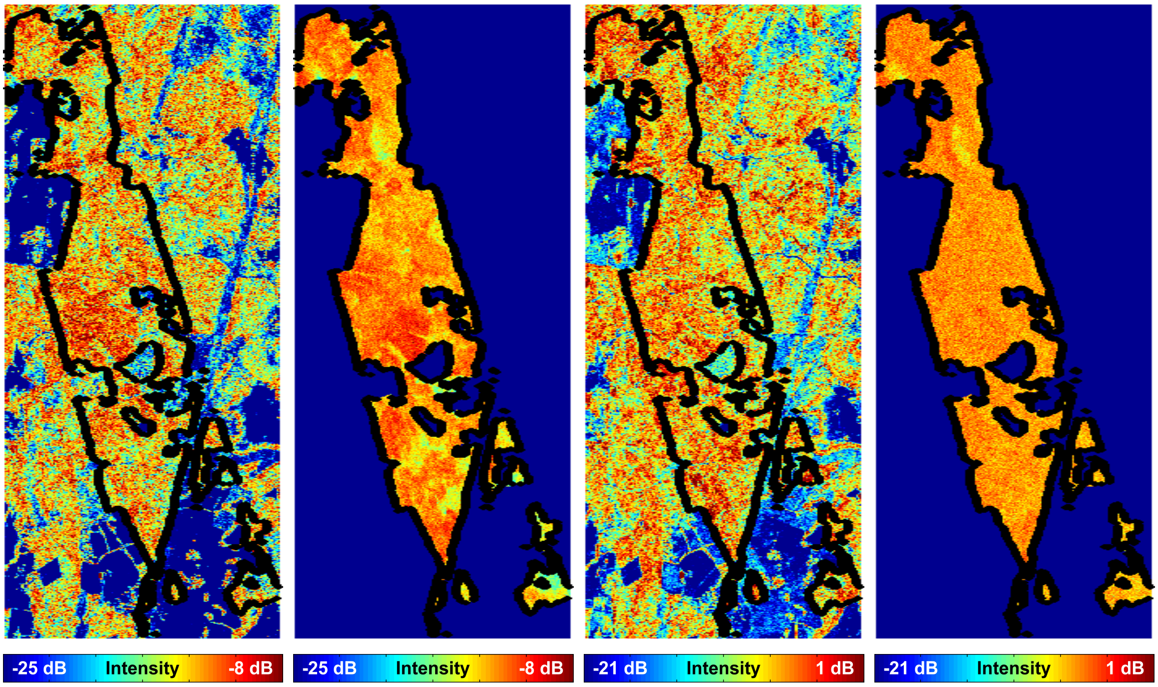
4.2 Paper A

In this paper, a P-band polarimetric-interferometric forward model (FM) is developed and used to model SAR imagery acquired with the airborne E-SAR system over Remningstorp. The FM is used within the BIOMASS end-to-end simulator (BEES), used by ESA to assess the error budget of the proposed (now selected) mission BIOMASS [3, 52].

In accordance with the requirements from ESA, the FM predicts the extended covariance matrix scaled to sigma nought on the diagonal from a small number of geo- and biophysical forest and system parameters. The influence of scene moisture and ground slopes is not modelled.

Four model scenarios have been developed for BEES, featuring two different backscatter profiles (exponential and truncated Gaussian) for two different biomes (tropical and boreal). However, only the boreal scenario with the exponential profile is studied in Paper A. For a full description of the model and the different scenarios, consult [171].

The extended covariance matrix is a matrix with all covariance combinations of the unique elements of two scattering matrices, one for each interferometric acquisition. Assuming a monostatic system, reciprocal and reflection symmetric medium, and identical polarimetric response at both ends of the baseline, the extended co-



(a) Reference (HV) (b) Modelled (HV) (c) Reference (VV) (d) Modelled (VV)

Figure 4.2: In Paper A, the extended covariance matrix is modelled for P-band from a few forest and system parameters. Here, backscatter coefficients sigma nought modelled from biomass and incidence angle maps are compared to the reference data acquired with the airborne E-SAR sensor. The test site is Remningstorp and the black contours mark the area covered by the biomass map. Outside, biomass has been set to zero.

variance matrix is a 6×6 matrix, and it can be re-stated in terms of three basic quantities: three real-valued backscatter coefficients sigma nought (one for each polarisation), three complex-valued interferometric correlation coefficients (one for each polarisation), and one complex-valued polarimetric correlation coefficient (between the two co-polarised channels).

Backscatter coefficients are for all polarisations modelled in decibels using a linear function of biomass, based on (3.1), and an additive, zero-mean Gaussian noise term. Model parameters and noise variance are estimated from the training data:

$$[\gamma^0_{\mathrm{HH}}]_{\mathrm{dB}} = -20.1 + 8.1 \log_{10} \mathcal{B} + \mathrm{N}(0, 1.3^2), \tag{4.1}$$

$$[\gamma^0_{\mathrm{HV}}]_{\mathrm{dB}} = -20.7 + 4.2 \log_{10} \mathcal{B} + \mathrm{N}(0, 0.7^2), \tag{4.2}$$

$$[\gamma^0_{\mathrm{VV}}]_{\mathrm{dB}} = -6.7 + 0.6 \log_{10} \mathcal{B} + \mathrm{N}(0, 1.2^2), \tag{4.3}$$

where $\mathcal{B}$ is the biomass in tons per hectare. As it can be observed, the highest sensitivity to biomass is obtained for the HH-channel, but with a larger variance. For the HV-channel, the sensitivity is still high, and the variance is lower. For the VV-channel, the sensitivity to biomass is very low.

Interferometric correlation coefficients are for all polarisations modelled from canopy height, ground elevation, incidence angle, and baseline information using the RVoG model presented in (3.4). All model parameters have been chosen to be normally distributed random variables, with statistics either estimated from training data or appropriately chosen based on the experience from earlier studies. Polarimetric coherence is modelled as a normally distributed random variable. The polarimetric phase difference has been found correlated with biomass and it is modelled using a linear function with an additive, zero-mean Gaussian noise term.

The model performance is evaluated in a side-by-side comparison of the modelled SAR images with the SAR images acquired by the E-SAR system. The same acquisition as used for model training is used in this evaluation. It is concluded that accurate modelling is achieved with the FM for the HH- and HV-polarised backscatter, the interferometric phase differences, and the polarimetric phase difference. However, modelling of the VV-polarised backscatter coefficient, the polarimetric coherence, and the interferometric coherences need to be studied further, as not all structures can be reproduced from the input data using the presented model.

In Figure 4.2, the results from backscatter modelling of the HV- and VV-channels from biomass are shown. Good agreement with the reference E-SAR image is observed for the HV-channel, where the correlation with biomass is high.

4.3 Paper B

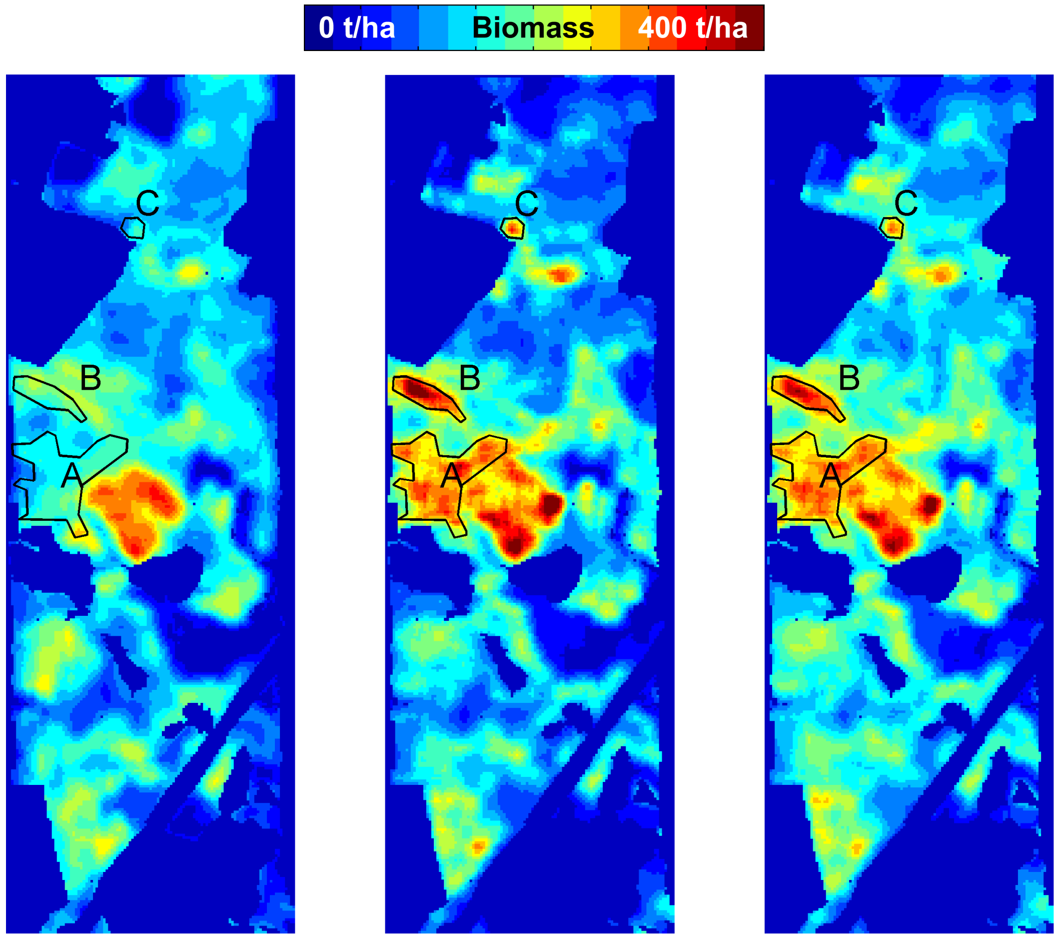
In this paper, a new biomass retrieval model is presented. The model includes terms which partially compensate for the influence of topographic and moisture variations. The model is evaluated on E-SAR data acquired in both Remningstorp and Krycklan. The model has been developed within the BIOMASS feasibility study and it has been included in the proposed biomass estimation algorithm for boreal forests [3].

The model is based on the power law function shown in (3.1), together with the backscatter ratio between the HH- and VV-channels, which has been found less susceptible to topographic and moisture variations, due to the often similar influence of these two effects on both co-polarised channels. A topographic correction is also included in the model:

$$\log_{10} \mathcal{B} = a_0 + a_1[\gamma^0_{\mathrm{HV}}]_{\mathrm{dB}} + (a_2 + a_3\theta_g)\left[\frac{\gamma^0_{\mathrm{HH}}}{\gamma^0_{\mathrm{VV}}}\right]_{\mathrm{dB}}, \tag{4.4}$$

where a_0, a_1, a_2, and a_3 are model parameters estimated from the training data and θ_g is the ground slope.

The model is evaluated together with five other models in a set of tests using the E-SAR data from the BioSAR 2007 and 2008 campaigns. The data have been acquired in different test sites, at different flight headings, and in different moisture



(a) Reference (lidar) (b) Estimated (E-SAR, Re) (c) Estimated (E-SAR, Kr)

Figure 4.3: In Paper B, biomass is estimated from P-band SAR backscatter using a new model with topographic and moisture correction. Three biomass maps for Remningstorp are here shown, estimated from: (a) lidar data, (b) SAR data, using the new model with parameters estimated in Remningstorp, and (c) SAR data, using the new model with parameters estimated in Krycklan, which is 720 km north-north-east of Remningstorp. Regions A, B, and C mark some disagreements, which are discussed in the text.

---

conditions. By using across-acquisition and across-site evaluation scenarios, it is possible to evaluate model susceptibility to topographic and moisture variations, as well as the potential of using the same parameter setup in different conditions.

It is concluded that the proposed model, with parameters estimated in Krycklan, can be used to estimate biomass in Remningstorp with a root-mean-square error of 40–59 tons/ha, or 22–32% of the mean biomass, which is significantly better compared to the other models. Since the two test sites are separated by 720 km and they feature quite different types of boreal forests, this is a very important conclusion for the future global mission.

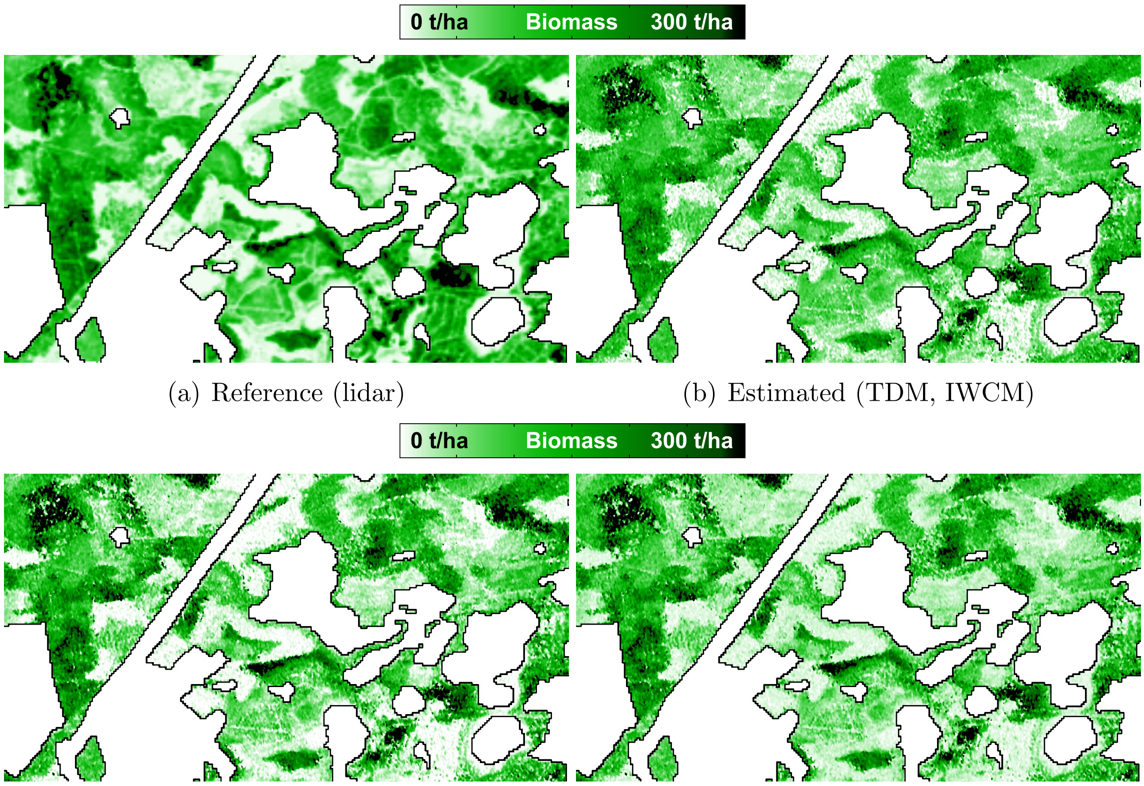
In Figure 4.3, biomass maps of Remningstorp, estimated using the proposed model with two parameter sets, one for Remningstorp and one for Krycklan, and compared to a reference biomass map estimated from lidar data. The performance is generally good, although some disagreements with respect to the lidar-based map can be observed. In region A, a significant understorey vegetation layer causes an increased HV backscatter, without contributing significantly to the total biomass. The disagreement in region B is caused both by an overestimation of biomass in the SAR map due to a strong double-bounce effect present in the HV-channel, and by an underestimation of biomass in the reference lidar data. The disagreement in region C is caused by an unusually strong double-bounce effect occurring for a group of tall trees surrounded by lower forest.

4.4 Paper C

In this paper, an approach based on the IWCM and developed for stem volume retrieval from repeat-pass interferometric, C-band ERS-1/2 data is used for biomass retrieval from single-pass interferometric, X-band TanDEM-X data, which have been ground-corrected using a high-resolution DTM. Multi-temporal data are studied, and the influence of both acquisition geometry and meteorological variables is assessed.

The development of an InSAR processing algorithm was an important part of the work conducted for this paper. Due to the quasi-bistatic acquisition geometry with one transmitting and two receiving satellites, and the helical orbit with dynamic baseline, a dedicated InSAR processing algorithm was developed for the TanDEM-X data. Using satellite state vectors, a geocoding look-up table was computed, and the high-resolution DTM was interpolated to radar geometry. The raw interferograms were then ground-corrected using the interpolated DTM and taking into consideration the quasi-bistatic acquisition geometry and satellite displacement between transmission and reception of the signals. A 5-metre buffer zone was added prior to plot- and stand-level averaging. Phase estimation errors were minimised by complex averaging of all relevant pixels within each plot/stand. Absolute phase calibration was done using ground reference points derived from a non-forest mask. Phase unwrapping was found unnecessary due to the limited height variations in the flattened interferogram. Conversion to height was done using a HOA map computed from the acquisition geometry. Geocoding and height estimation accuracies were evaluated using two 5-metre trihedral corner reflectors positioned within the Remningstorp test site. The standard deviation of height variations was found lower than 10 cm and the horizontal offset was found lower than 2 m.

In this study, the IWCM is compared to two other models, each being its simplified version: the RVoG, in which canopy gaps are neglected, and a new penetration depth (PD) model, in which both canopy gaps and ground contribution are neglected. The



(a) Reference (lidar) (b) Estimated (TDM, IWCM)

(c) Estimated (TDM, RVoG) (d) Estimated (TDM, PD)

Figure 4.4: In Paper C, biomass is estimated using three models: IWCM, RVoG, and a new penetration depth (PD) model. Above, the corresponding maps are compared to a lidar-derived reference map. The test site is Remningstorp and non-forested areas have been masked out. Note that there is one year time difference between the TanDEM-X (TDM) and lidar acquisitions, and growth and forest management have not been accounted for (e.g., in the cleared area slightly to the left from the bottom right corner of the images). The original TDM data have been acquired on 2011-06-04, with a HOA of 49 m, and at an incidence angle of 41°.

---

models are fitted to both intensity, coherence, and phase centre height data, and biomass is estimated.

The new PD model is a simplification valid for dense forests, and it requires one single parameter, the penetration depth, which is used to compensate the phase centre height for penetration. The compensated height is then converted to biomass using a height-to-biomass relation:

$$\mathcal{B} = 0.21 \left(h_{\mathrm{gc}} + \frac{1}{\alpha} \right)^{2.17}, \tag{4.5}$$

where the biomass $\mathcal{B}$ is measured in tons per hectare, h_{gc} is the phase centre elevation above ground (in metres) and α is the effective extinction coefficient (in m^{-1}).

The models are evaluated using eighteen VV-polarised TanDEM-X image pairs acquired over Remningstorp between 2011-06-04 and 2012-08-24, at HOAs between 49 m and 358 m, incidence angles between 34° and 41°, and in both ascending and descending modes. High-resolution DTM acquired within a national lidar scanning campaign is used as ground reference during InSAR processing. Meteorological data provided by the Swedish Meteorological and Hydrological Institute (SMHI) are also used in the study. The retrieval performance is assessed using 201 forest stands with a minimum size of 1 ha, and biomass in the interval 6–267 t/ha (mean: 105 t/ha), equally divided into two groups: one for training and one for validation. The root-mean-square error (RMSE) for the IWCM-based retrieval is between 17% and 33%, with the best results obtained for the low HOAs. For the RVoG and the PD models, the stand-level RMSE values are slightly higher. Biomass is also estimated using multi-temporal averaging from all eighteen acquisitions with a weighting factor inversely proportional to the square of HOA, with an RMSE of 16% and $R^2 = 0.93$.

In Figure 4.4, biomass maps obtained with each of the three models for an image from 2011-06-04 with a HOA of 49 m and an incidence angle of 41° are compared to a reference map estimated from lidar data. Good results are obtained for all models, except in the region that has been harvested between the lidar and SAR acquisitions.

4.5 Paper D

In this paper, the two-level model (TLM) is introduced and used for the estimation of forest height and canopy density from single-polarised TanDEM-X acquisitions in combination with a high-resolution DTM. With an access to the global TanDEM-X data, the presented approach can be used for frequent, large-scale, high-resolution mapping of forest height and canopy density in countries in which national lidar scanning campaigns have been conducted.

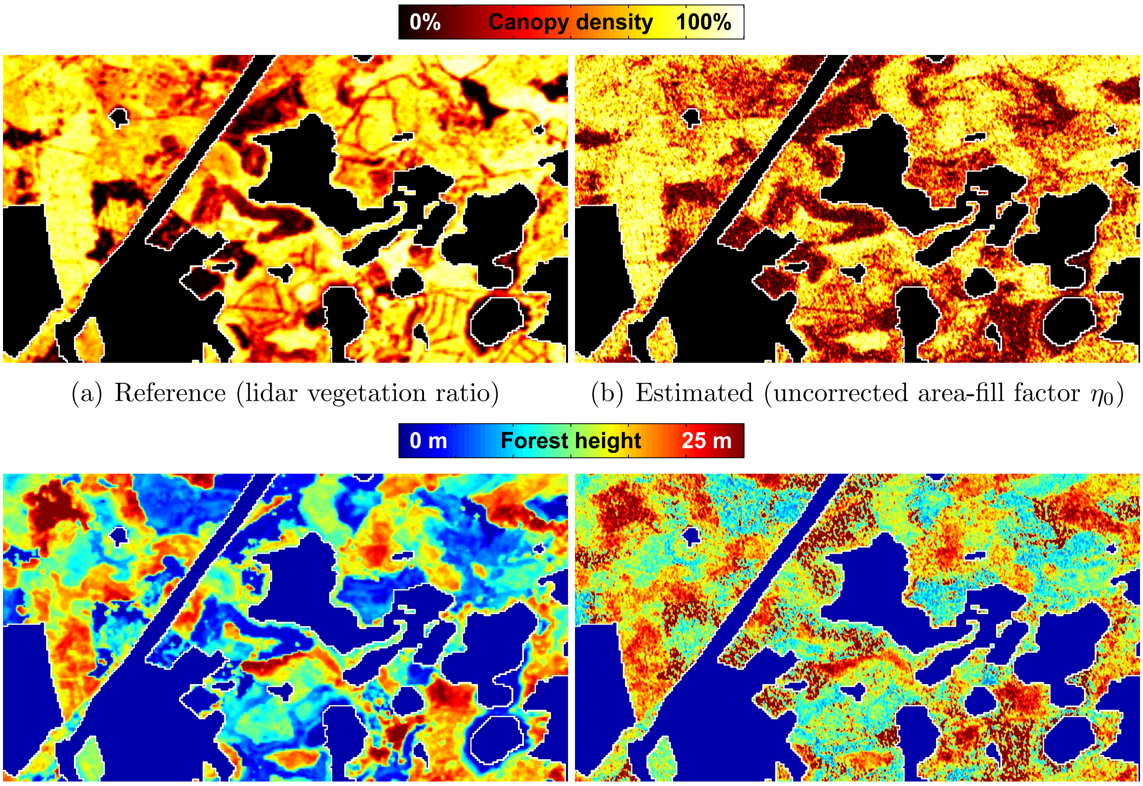
The TLM models forest as two discrete scattering levels: ground and vegetation, separated by a distance Δh and with canopy gaps described by the area-fill factor η, which is the fraction of the total area covered by the vegetation level. The two-level approach with canopy gaps is motivated by an interference effect observed in the data for sparse forest plots, for which the location of the scattering centre in the canopy is found sensitive to the interferometric baseline.

The ground-corrected complex correlation coefficient is modelled by the TLM as:

$$\tilde{\gamma}_{\mathrm{gc}} = \frac{\mu + e^{ik_z \Delta h}}{\mu + 1}, \tag{4.6}$$

where μ the area-weighted backscatter ratio:

$$\mu = \frac{\sigma^0_{\mathrm{gr}}}{\sigma^0_{\mathrm{veg}}} \frac{1 - \eta}{\eta}. \tag{4.7}$$



(a) Reference (lidar vegetation ratio) (b) Estimated (uncorrected area-fill factor η_0)

(c) Reference (lidar height H50) (d) Estimated (level distance Δh)

Figure 4.5: In Paper D, forest height and canopy density are estimated from the inversion of the TLM. Above, the corresponding maps are compared to lidar-derived reference maps. The test site is Remningstorp and non-forested areas have been masked out. Note that there is one year time difference between the TDM and lidar acquisitions, and growth and forest management have not been accounted for (e.g., in the cleared area slightly to the left from the bottom right corner of the images). The original TDM data have been acquired on 2011-06-04, with a HOA of 49 m, and at an incidence angle of 41°.

As the TLM requires only two parameters (Δh and μ), model inversion can be done individually for each ground-corrected complex correlation coefficient, without the need for additional SAR acquisitions. Analytical expressions for the computation of Δh and μ from a ground-corrected complex correlation coefficient are presented in the paper.

The model is evaluated using eight VV-polarised TanDEM-X acquisitions made at different baselines (HOAs between 32 and 63 metres) over Remningstorp in the summers of 2011, 2012, and 2013, and thirty-two, 0.5 hectare circular forest plots with different heights and canopy densities. The InSAR data have been processed using the same InSAR processing algorithm as described in Paper C.

It is concluded that level distance Δh can be used as an estimate of H50 (50th percentile of all lidar returns above 1 m or 10% of the maximal height) with a Pearson correlation coefficient of about 95% and a root-mean-square difference (RMSD) lower than 10% (or 1.8 m). It is also concluded that the uncorrected area-fill factor:

$$\eta_0 = \frac{1}{1+\mu} \tag{4.8}$$

can be used as an estimate of the vegetation ratio (the ratio between the number of lidar returns from above 1 m or 10% of the maximal height and all lidar returns) with a Pearson correlation coefficient better than 59% and RMSD around 10% (or 0.07). A HOA-dependent offset is observed for Δh, and it is most likely caused by residual SNR and system decorrelation effects, which have not been compensated for.

In Figure 4.5, maps of Δh and η_0 obtained from TLM inversion of a TanDEM-X acquisition from 2011-06-04 with a HOA of 49 metres and an incidence angle of 41° are compared to maps of H50 and VR, derived from lidar data acquired one year earlier. Forest changes such as growth and forest management procedures have not been accounted for.

4.6 Paper E

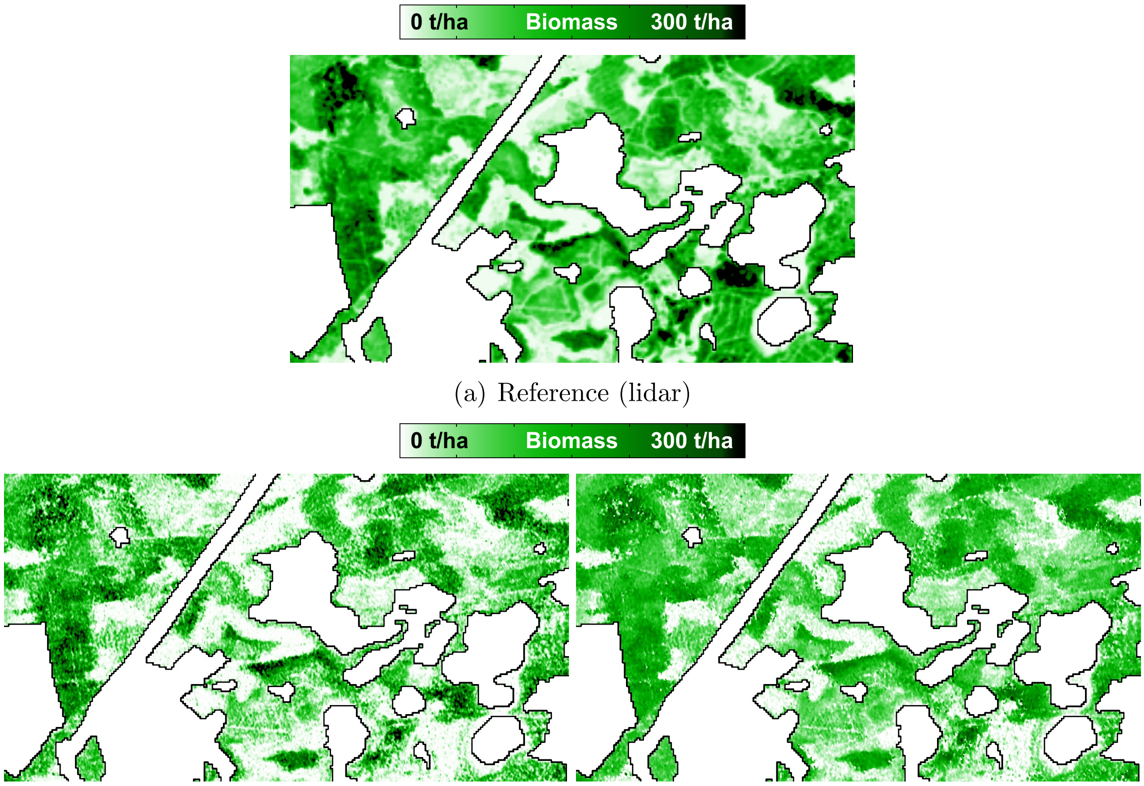
In this paper, biomass is estimated from forest height and canopy density estimates obtained from the inversion of the TLM presented in Paper D. With an access to the global TanDEM-X data, the presented approach can be used for frequent, large-scale, high-resolution mapping of biomass in countries in which national lidar scanning campaigns have been conducted.

The introduced TLM biomass model (TBM) is a power law function of the level distance Δh and the uncorrected area-fill factor η_0 obtained from TLM inversion:

$$\mathcal{B} = K\Delta h^{\alpha}\eta_0^{\beta}, \tag{4.9}$$

where $\mathcal{B}$ is the biomass in tons per hectare, and K, α, and β are model parameters. The power law form of the model is motivated by similar functions used in lidar-based biomass mapping.

The model is evaluated using eighteen VV-polarised TanDEM-X acquisitions made at different baselines (HOAs between 32 and 63 metres) over both Remningstorp and Krycklan in the summers of 2011, 2012, and 2013. Eight of these images have been used in the study summarised in Paper D. In Remningstorp, between 32 and 21 forest plots are used, whereas in Krycklan, 29 forest stands are used. The TBM is compared to a zero-intercept, linear scaling model (SM), in which biomass is estimated from a direct scaling of the ground-corrected interferometric height, as proposed in [172]. The models are evaluated in across-acquisition and across-site scenarios, to assess their operational values.



(a) Reference (lidar)

(b) Estimated (TDM, TBM)

(c) Estimated (TDM, SM)

Figure 4.6: In Paper E, biomass is estimated from forest height and canopy density estimates obtained from TLM inversion using the TLM biomass model (TBM), and compared to a scaling model (SM), which scales the ground-corrected interferometric height to biomass, as proposed in [172]. Above, the corresponding maps are compared to a lidar-derived reference map. The test site is Remningstorp and non-forested areas have been masked out. Note that there is one year time difference between the TDM and lidar acquisitions, and growth and forest management have not been accounted for (e.g., in the cleared area slightly to the left from the bottom right corner of the images). The original TDM data have been acquired on 2011-06-04, with a HOA of 49 m, and at an incidence angle of 41°.

The TBM can explain between 65% and 89% of the AGB variance observed in the data, with a residual root-mean-square error (RMSE) in the interval 12–19% (median: 15%). If model training and validation are carried out on different TanDEM-X acquisitions or different test sites, the prediction RMSE increases (12–80%, median: 30%). With α fixed and β a site-dependent constant, the prediction RMSE is lower (12–56%, median: 17%), while the residual RMSE is similar (12-29%, median: 16%). The SM shows similar performance when used on Krycklan data, whereas for Remningstorp data and across-site retrieval, the performance is poorer.

In general, the retrieval performance of the TBM with fixed α and β is good, with an RMSE below 20% for all acquisitions in Krycklan and for almost all acquisitions in Remningstorp with HOA above 40 m. In Remningstorp, where the forest is generally taller and with more complex horizontal structure (due to management), the performance of the model is decreased at HOAs below 40 m, where the errors caused by the insufficient modelling of the vertical structure have the strongest impact. It is also observed that the HOA-dependent offset in Δh, noted earlier in Paper D and most likely caused by the lack of SNR and system decorrelation modelling, causes reduced performance of the biomass model at low HOAs. A coherence calibration step is therefore proposed for the future.

In Figure 4.6, a biomass map for Remningstorp, estimated with the TBM using the forest height and canopy density estimates shown in Figure 4.5, is compared to a reference biomass map estimated from lidar data, and a biomass map created using the scaling model (SM), in which biomass is computed by scaling from the ground-corrected interferometric height (h_{gc}). The TBM-based estimate is better, as it is able to reproduce the full biomass variance. Note that the time difference between lidar and SAR acquisitions is one year, and forest changes such as growth and forest management procedures have not been accounted for.

Chapter 5

Conclusions

The main scope of the work conducted for this thesis has been to develop methods for forest parameter estimation from SAR imagery. By using P- and X-band, i.e., the lowest and highest frequency bands available and useful for spaceborne imaging of the Earth, this task has been studied in both the low- and high-frequency regimes.

Both empirical and theoretical models linking forest parameters to polarimetric and interferometric SAR observables have been developed. Within the models, some effects previously unaccounted for have been included. At P-band, an improved biomass retrieval model has been developed by including empirical corrections for the topographic and moisture variations. At X-band, the estimation of biomass has been improved and the estimation of canopy density has been made possible by introducing a model in which canopy gaps are used to explain volume decorrelation.

The models have been developed with their operational values in mind. At P-band, the proposed biomass model requires one single polarimetric acquisition, which is expected to be the standard acquisition mode for BIOMASS, and a ground slope map, which most likely can be derived from a standard TanDEM-X DEM. It has also been shown that the proposed model can be used in two geographically distant test sites with the same parameter setup, which reduces the requirements on training data.

At X-band, the proposed TLM inversion yields estimates of forest height and canopy density, and it can be performed using single-polarised TanDEM-X acquisitions, e.g., from the existing global data set, provided that a high-resolution digital terrain model (DTM) is available. With the temporal stability of the ground in most forested regions, the increasing popularity of national lidar scanning campaigns, and the upcoming P- and L-band InSAR missions (BIOMASS, and possibly also SAOCOM-CS and TanDEM-L), the availability of high-resolution DTMs will increase with time. The proposed biomass model requires local estimation of only one parameter related to canopy density, whereas the other parameters can be fixed for the two studied test sites, which reduces the requirements on training data.

5.1 Thesis Highlights

At P-band, the following findings can be considered the highlights of this thesis:

➥ The HH/VV backscatter ratio has been found useful for the compensation of topographic and moisture variations, which often have similar impact on the two co-polarised channels, and their influence is decreased when the ratio is formed.

➥ It has been shown that the same biomass model can be used with the same parameter values in both Remningstorp and Krycklan, which are two test sites separated by 720 km and featuring different types of boreal forest, due to the large contribution of tree trunks to the total backscatter.

At X-band, the most interesting findings can be summarised in the following highlights:

➥ For sparse plots with low canopy density, the location of the scattering phase centre has been found sensitive to the baseline. This is explained by an interference effect occurring when ground- and canopy-level contributions are of similar strength.

➥ The contribution of canopy gaps has been found significant at X-band, as shown in a comparative study of the three models: IWCM, RVoG, and a penetration depth model. Moreover, direct inversion of a two-level model (TLM), in which forest is modelled as two scattering levels and penetration can only occur through canopy gaps, can provide estimates of both forest height and canopy density.

➥ Biomass can be accurately estimated from forest height and canopy density estimates obtained from the inversion of the TLM using a power law model. However, the same exponent for the canopy density estimate cannot be used in both Remningstorp and Krycklan, due to the large difference in canopy structure.

5.2 Future Prospects

There are several topics that need to be studied in the future to further improve the results presented in this thesis. For all studies, one of the most important extensions is to apply the presented approaches on other data sets (other biomes, acquisition geometries, frequencies), in order to evaluate the generality of the findings and further improve the understanding of the governing processes.

At P-band, one of the largest difficulties encountered during forest parameter estimation is the influence of topographic and moisture variations on the backscattered signal, which is to a large degree still unknown. The forward model presented in Paper A does not account for these effects, and the biomass retrieval model presented in Paper B includes empirical correction terms, the HH/VV-ratio and the surface slope, which only partially compensate for these effects. Theoretical modelling is needed for better understanding of the scattering processes, for example using cylinder over ground models. Moreover, as noted in Paper A, the modelling of the interferometric coherence also needs to be improved, for example by using other vertical scattering profiles, or by improved modelling of the RVoG parameters. The inclusion of canopy gaps in P-band modelling should also be evaluated. Finally, the observed correlation between the HH-VV phase difference and biomass needs to be studied further using theoretical models. Also, its potential for biomass estimation needs to be evaluated.

At X-band, the TLM-based methods for forest height, canopy density, and biomass estimation are currently based on the assumptions of known topography, and negligible SNR and system decorrelation effects. Consequently, the first assumption makes the current approach unfeasible in regions with unknown topography, and the latter assumption introduces a HOA-dependent offset in regions with low volume decorrelation, as observed in Papers D and E. It would be interesting to extend the current model to the multi-polarised case and evaluate the possibility of reducing these requirements. Also, the inclusion of a coherence calibration step in the processing chain may improve TLM inversion performance in lower forests, and the HOA-dependent offset may be reduced. Additionally, the TLM inversion may be applied on multi-temporal and/or multi-baseline data, making it possible to study forest change and/or forest structure. The influence of meteorological conditions needs to be studied.

References

[1] World Wide Fund for Nature (WWF). Forests, jungles, woods & their trees. Retrieved from: http://wwf.panda.org/about_our_earth/about_forests/ (2014-08-08).

[2] C. Parker, A. Mitchell, M. Trivedi, and N. Mardas, *The Little REDD Book: A guide to governmental and non-governmental proposals for reducing emissions from deforestation and degradation*. Global Canopy Programme, John Krebs Field Station, Oxford OX2 8QJ, UK, 2008.

[3] ESA (2012), *Report for Mission Selection: BIOMASS*, ESA SP-1324/1 (3 volume series), European Space Agency, Noordwijk, The Netherlands.

[4] E. Næsset, T. Gobakken, S. Solberg, T. G. Gregoire, R. Nelson, G. Ståhl, and D. Weydahl, "Model-assisted regional forest biomass estimation using lidar and InSAR as auxiliary data: A case study from a boreal forest area," *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3599–3614, 2011.

[5] B. Koch, "Status and future of laser scanning, synthetic aperture radar and hyperspectral remote sensing data for forest biomass assessment," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 6, pp. 581–590, 2010.

[6] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 1, pp. 6–43, 2013.

[7] K. J. Ranson, G. Sun, R. H. Lang, N. S. Chauhan, R. J. Cacciola, and O. Kilic, "Mapping of boreal forest biomass from spaceborne synthetic aperture radar," *Journal of Geophysical Research*, vol. 102, pp. 29 599–29 610, 1997.

[8] H. Israelsson, L. M. H. Ulander, J. I. H. Askne, J. E. S. Fransson, P.-O. Frölind, A. Gustavsson, and H. Hellsten, "Retrieval of forest stem volume using VHF SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 1, pp. 36–40, 1997.

[9] H. Israelsson, "SAR remote sensing of forests from microwaves to VHF," Ph.D. dissertation, Chalmers University of Technology, 1998.

[10] F. Walter, "Extraction of forest tree volume from CARABAS SAR data," *Scandinavian Journal of Forest Research*, vol. 12, no. 4, pp. 370–374, 1997.

[11] J. E. S. Fransson, F. Walter, and L. M. H. Ulander, "Estimation of forest parameters using CARABAS-II VHF SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 720–727, 2000.

[12] G. Smith and L. M. H. Ulander, "A model relating VHF-band backscatter to stem volume of coniferous boreal forest," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 728–740, 2000.

[13] G. Smith, "Radar remote sensing of forests using CARABAS and ERS," Ph.D. dissertation, Chalmers University of Technology, 2000.

[14] P. Melon, J. M. Martinez, T. Le Toan, L. M. H. Ulander, and A. Beaudoin, "On the retrieving of forest stem volume from VHF SAR data: Observation and modelling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 11, pp. 2364–2372, 2001.

[15] J. E. S. Fransson, G. Smith, F. Walter, A. Gustavsson, and L. M. H. Ulander, "Estimation of forest stem volume in sloping terrain using CARABAS-II VHF SAR data," *Canadian Journal of Remote Sensing*, vol. 30, no. 4, pp. 651–660, 2004.

[16] B. Hallberg, G. Smith-Jonforsen, and L. M. H. Ulander, "Measurements on individual trees using multiple VHF SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 10, pp. 2261–2269, 2005.

[17] G. Smith-Jonforsen, L. M. H. Ulander, and X. Luo, "Low VHF-band backscatter from coniferous forests on sloping terrain," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 10, pp. 2246–2260, 2005.

[18] A. A. Kononov and M.-H. Ka, "Model-associated forest parameter retrieval using VHF SAR data at the individual tree level," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 1, pp. 69–84, 2008.

[19] B. Hallberg, "Synthetic-aperture radar imaging of forests in the VHF and UHF bands: Electromagnetic models and data analysis," Ph.D. dissertation, Chalmers University of Technology, 2007.

[20] K. Folkesson, G. Smith-Jonforsen, and L. M. H. Ulander, "Validating backscatter models for CARABAS SAR images of coniferous forests," *Canadian Journal of Remote Sensing*, vol. 34, no. 5, pp. 480–495, 2008.

[21] K. Folkesson, "Model-based stem volume retrieval and widthrow detection using CARABAS and P-band SAR," Ph.D. dissertation, Chalmers University of Technology, 2008.

[22] K. Folkesson, G. Smith-Jonforsen, and L. M. H. Ulander, "Model-based compensation of topographic effects for improved stem-volume retrieval from CARABAS-II VHF-band SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 4, pp. 1045–1055, 2009.

[23] M. Moghaddam, Y. Rahmat-Samii, E. Rodriguez, D. Entekhabi, J. Hoffman, D. Moller, L. E. Pierce, S. Saatchi, and M. Thomson, "Microwave Observatory of Subcanopy and Subsurface (MOSS): A mission concept for global deep soil moisture observations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 8, pp. 2630–2643, 2007.

[24] T. Le Toan, A. Beaudoin, J. Riom, and D. Guyon, "Relating forest biomass to SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 2, pp. 403–411, 1992.

[25] M. C. Dobson, F. T. Ulaby, T. Le Toan, A. Beaudoin, E. S. Kasischke, and N. Christensen, "Dependence of radar backscatter on coniferous forest biomass," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 2, pp. 412–415, 1992.

[26] A. Beaudoin, T. Le Toan, S. Goze, E. Nezry, A. Lopes, E. Mougin, C. C. Hsu, H. C. Han, J. A. Kong, and R. T. Shin, "Retrieval of forest biomass from SAR data," *International Journal of Remote Sensing*, vol. 15, no. 14, pp. 2777–2796, 1994.

[27] H. Israelsson, J. Askne, and R. Sylvander, "Potential of SAR for forest bole volume estimation," *International Journal of Remote Sensing*, vol. 15, no. 14, pp. 2809–2826, 1994.

[28] K. J. Ranson and G. Sun, "Mapping biomass of a northern forest using multifrequency SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 2, pp. 388–396, 1994.

[29] Y. Rauste, T. Häma, J. Pulliainen, K. Heiska, and M. Hallikainen, "Radar-based forest biomass estimation," *International Journal of Remote Sensing*, vol. 15, no. 14, pp. 1797–2808, 1994.

[30] M. L. Imhoff, "Radar backscatter and biomass saturation: Ramifications for global biomass inventory," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 2, pp. 511–518, 1995.

[31] E. J. Rignot, R. Zimmermann, and J. J. van Zyl, "Spaceborne applications of P band imaging radars for measuring forest biomass," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 5, pp. 1162–1169, 1995.

[32] D. H. Hoekman and M. J. Quiñones, "Land cover type and biomass classification using AirSAR data for evaluation of monitoring scenarios in the

Colombian Amazon," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2, pp. 685–696, 2000.

[33] J. R. Santos, C. C. Freitas, L. S. Araujo, L. V. Dutra, J. C. Mura, F. F. Gama, L. S. Soler, and S. J. S. Sant'Anna, "Airborne P-band SAR applied to the aboveground biomass studies in the Brazilian tropical rainforest," *Remote Sensing of Environment*, vol. 87, no. 4, pp. 482–493, 2003.

[34] G. Smith-Jonforsen, K. Folkesson, B. Hallberg, and L. M. H. Ulander, "Effects of forest biomass and stand consolidation on P-band backscatter," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 4, pp. 669–673, 2007.

[35] S. Saatchi, K. Halligan, D. Despain, and R. Crabtree, "Estimation of forest fuel load from radar remote sensing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 6, pp. 1726–1740, 2007.

[36] G. Sandberg, L. M. H. Ulander, J. E. S. Fransson, J. Holmgren, and T. Le Toan, "L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest," *Remote Sensing of Environment*, vol. 115, no. 11, pp. 2874–2886, 2011.

[37] S. Saatchi, M. Marlier, R. L. Chazdon, D. B. Clark, and A. E. Russell, "Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass," *Remote Sensing of Environment*, vol. 115, no. 11, pp. 2836–2849, 2011.

[38] T. Le Toan, S. Quegan, M. W. J. Davidson, H. Balzter, P. Paillou, K. Papathanassiou, S. Plummer, F. Rocca, S. Saatchi, H. Shugart, and L. M. H. Ulander, "The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle," *Remote Sensing of Environment*, vol. 115, no. 11, pp. 2850–2860, 2011.

[39] M. Neumann, S. S. Saatchi, L. M. H. Ulander, and J. E. S. Fransson, "Assessing performance of L- and P-band polarimetric interferometric SAR data in estimating boreal forest above-ground biomass," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 3, pp. 714–726, 2012.

[40] G. Sandberg, "Estimation of forest biomass and Faraday rotation using ultra high frequency synthetic aperture radar," Ph.D. dissertation, Chalmers University of Technology, 2013.

[41] I. Hajnsek, R. Scheiber, L. M. H. Ulander, A. Gustavsson, G. Sandberg, S. Tebaldini, A. M. Guarnieri, F. Rocca, F. Bombardini, and M. Pardini, "BioSAR 2007 technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2007 experiment: Final report without synthesis," ESA contract no. 20755/07/NL/CB, Tech. Rep., 2008.

[42] L. M. H. Ulander, A. Gustavsson, B. Flood, D. Murdin, P. Dubois-Fernandez, X. Depuis, G. Sandberg, M. J. Soja, L. E. B. Eriksson, J. E. S. Fransson, J. Holmgren, and J. Wallerman, "BioSAR 2010: Technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2010 experiment: Final report," ESA contract no. 4000102285/10/NL/JA/ef, Tech. Rep., 2011.

[43] P. C. Dubois-Fernandez, T. Le Toan, S. Daniel, H. Oriot, J. Chave, L. Blanc, L. Villard, M. Davidson, and M. Petit, "The TropiSAR airborne campaign in French Guiana: Objectives, description, and observed temporal behavior of the backscatter signal," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 8, pp. 3228–3241, 2012.

[44] F. Garestier, P. C. Dubois-Fernandez, and I. Champion, "Forest height inversion using high-resolution P-band Pol-InSAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3544–3559, 2008.

[45] I. Hajnsek, F. Kugler, S.-K. Lee, and K. P. Papathanassiou, "Tropical-forest-parameter estimation by means of Pol-InSAR: The INDREX-II campaign," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 2, pp. 481–493, 2009.

[46] 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 Transactions on Geoscience and Remote Sensing*, vol. 47, no. 10, pp. 3379–3388, 2009.

[47] S. Tebaldini and F. Rocca, "Multibaseline polarimetric SAR tomography of a boreal forest at P- and L-bands," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 232–246, 2012.

[48] D. H. T. Minh, T. L. Toan, F. Rocca, S. Tebaldini, M. d'Alessandro, and L. Villard, "Relating P-band synthetic aperture radar tomography to tropical forest biomass," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 2, pp. 967–979, 2014.

[49] N. C. Rogers, S. Quegan, J. S. Kim, and K. P. Papathanassiou, "Impacts of ionospheric scintillation on the BIOMASS P-band satellite SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 3, pp. 1856–1868, 2014.

[50] N. C. Rogers and S. Quegan, "The accuracy of Faraday rotation estimation in satellite synthetic aperture radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4799–4807, 2014.

[51] F. Heliere, F. Fois, M. Arcioni, P. Bensi, M. Fehringer, and K. Scipal, "Biomass P-band SAR interferometric mission selected as 7th Earth Explorer Mission,"

in *10th European Conference on Synthetic Aperture Radar (EUSAR)*, Berlin, Germany, 3–5 June 2014, pp. 1152–1155.

[52] P. Lopez-Dekker, F. De Zan, T. Borner, M. Younis, K. Papathanassiou, T. Guardabrazo, V. Bourlon, S. Ramongassie, N. Taveneau, L. M. H. Ulander, D. Murdin, N. Rogers, S. Quegan, and R. Franco, "Biomass end-to-end mission performance simulator," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Vancouver, BC, Canada, 24–29 July 2011, pp. 4249–4252.

[53] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse Problems*, vol. 14, pp. R1–R54, 1998.

[54] P. B. G. Dammert, "Spaceborne SAR interferometry: Theory and applications," Ph.D. dissertation, Chalmers University of Technology, 1999.

[55] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, and M. Zink, "TanDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 11, pp. 3317–3341, 2007.

[56] T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, and D. Alsdorf, "The Shuttle Radar Topography Mission," *Reviews of Geophysics*, vol. 45, no. 2, pp. 1–33, 2007.

[57] J. Askne, P. Dammert, L. M. H. Ulander, and G. Smith, "C-band repeat-pass interferometric SAR observations of the forest," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 1, pp. 25–35, 1997.

[58] J. Askne, M. Santoro, G. Smith, and J. E. S. Fransson, "Multitemporal repeat-pass SAR interferometry of boreal forests," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 7, pp. 1540–1550, 2003.

[59] S. R. Cloude and K. P. Papathanassiou, "Polarimetric SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 5, pp. 1551–1565, 1998.

[60] K. P. Papathanassiou and S. R. Cloude, "Single-baseline polarimetric SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 11, pp. 2352–2363, 2001.

[61] S. R. Cloude and K. P. Papathanassiou, "Three-stage inversion process for polarimetric SAR interferometry," *IEE Proceedings on Radar, Sonar & Navigation*, vol. 150, no. 3, pp. 125–134, 2003.

[62] J. Praks, O. Antropov, and M. T. Hallikainen, "LIDAR-aided SAR interferomety studies in boreal forest: Scattering phase center and extinction coefficient at X- and L-band," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3831–3843, 2012.

[63] F. Kugler, D. Schulze, I. Hajnsek, H. Pretzsch, and K. P. Papathanassiou, "TanDEM-X Pol-InSAR performance for forest height estimation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6404–6422, 2014.

[64] N. Gebert, B. Carnicero Dominguez, M. W. J. Davidson, M. Diaz Martin, and P. Silvestrin, "SAOCOM-CS – A passive companion to SAOCOM for single-pass L-band sar interferometry," in *10th European Conference on Synthetic Aperture Radar (EUSAR)*, Berlin, Germany, 3–5 June 2014, pp. 1251–1254.

[65] A. Moreira, G. Krieger, M. Younis, I. Hajnsek, K. Papathanassiou, M. Eineder, and F. De Zan, "Tandem-L: A mission proposal for monitoring dynamic earth processes," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Vancouver, BC, Canada, 24–29 July 2011, pp. 1385–1388.

[66] A. Gustavsson, B. Flood, P.-O. Frolind, H. Hellsten, T. Jonsson, B. Larsson, G. Stenstrom, and L. M. H. Ulander, "First airborne tests with the new VHF SAR CARABAS II," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Singapore, 3–8 August 1997, pp. 1214–1216.

[67] M. I. Skolnik, *Radar Handbook*, 2nd ed. McGraw-Hill, 1990.

[68] R. J. Sullivan, *Radar Foundations for Imaging and Advanced Concepts*. Scitech Publishing, 2004.

[69] M. A. Richards, J. A. Scheer, and W. A. Holm, Eds., *Principles of Modern Radar, Volume I: Basic Techniques*. Scitech Publishing, 2010.

[70] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing*. John Wiley & Sons, Inc, 1992.

[71] W. G. Rees, *Physical Principles of Remote Sensing*. Cambridge University Press, 2013.

[72] S. Kingsley and S. Quegan, *Understanding Radar Systems*. McGraw-Hill, 1992.

[73] D. K. Barton, *Radar System Analysis and Modeling*. Artech House, 2005.

[74] "IEEE standard letter designations for radar-frequency bands," *IEEE Std 521-2002 (Revision of IEEE Std 521-1984)*, 2003.

[75] J. Bruder, "IEEE radar standards and the radar systems panel," *IEEE Aerospace and Electronic Systems Magazine*, vol. 28, no. 7, pp. 19–22, 2013.

[76] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. SciTech Publishing, 2004.

[77] R. Jordan, "The Seasat-A synthetic aperture radar system," *IEEE Journal of Oceanic Engineering*, vol. 5, no. 2, pp. 154–164, 1980.

[78] World Meteorological Organisation (WMO). (2014) OSCAR: Observing systems capability analysis and review tool. Retrieved from: http://www.wmo-sat.info/oscar/instrumenttypes/view/23 (2014-09-10).

[79] W. G. Carrara, R. S. Goodman, and R. M. Majewski, *Spotlight Synthetic Aperture Radar: Signal Processing Algorithms*. Artech House, 1995.

[80] W. L. Melvin and J. A. Scheer, Eds., *Principles of Modern Radar, Volume II: Advanced Techniques*. Scitech Publishing, 2013.

[81] F. De Zan and A. Monti Guarnieri, "TOPSAR: Terrain observation by progressive scans," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 9, pp. 2352–2360, 2006.

[82] I. G. Cumming and F. H. Wong, *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*. Artech House, 2005.

[83] C. V. Jakowatz, Jr, D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*. Springer, 1996.

[84] M. Soumekh, *Synthetic Aperture Radar Signal Processing with MATLAB Algorithms*. Wiley-Interscience, 1999.

[85] A. F. Yegulalp, "Fast backprojection algorithm for synthetic aperture radar," in *IEEE Radar Conference*, Waltham, MA, USA, 20–22 April 1999, pp. 60–65.

[86] L. M. H. Ulander, H. Hellsten, and G. Stenström, "Synthetic-aperture radar processing using fast factorized back-projection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 3, pp. 760–776, 2003.

[87] L. M. H. Ulander and J. O. Hagberg, "Radiometric and interferometric calibration of ENVISAT-1 ASAR," Research Report No. 172, Department of Radio and Space Science, Chalmers University of Technology, Göteborg, (European Space Agency Contract Report: ASAR External Calibration and Validation Study, ESTEC Contract No. PO/133382), 1995.

[88] L. M. H. Ulander, "Radar remote sensing of sea ice: Measurements and theory," Ph.D. dissertation, Chalmers University of Technology, 1991.

[89] A. Freeman, "SAR calibration: an overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 6, pp. 1107–1121, 1992.

[90] M. Dobson, F. Ulaby, D. Brunfeldt, and D. Held, "External calibration of SIR-B imagery with area-extended and point targets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 24, no. 4, pp. 453–461, 1986.

[91] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing, Active and Passive, Volume II: Radar Remote Sensing and Surface Scattering and Emission Theory*. Addison-Wesley, 1982.

[92] L. M. H. Ulander, "Radiometric slope correction of synthetic aperture radar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 5, pp. 1115–1122, 1996.

[93] O. Frey, M. Santoro, C. L. Werner, and U. Wegmuller, "DEM-Based SAR pixel-area estimation for enhanced geocoding refinement and radiometric normalization," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 1, pp. 48–52, 2013.

[94] E. P. W. Attema and F. T. Ulaby, "Vegetation modeled as a water cloud," *Radio Science*, vol. 13, no. 2, pp. 357–364, 1978.

[95] F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing, Active and Passive, Volume III: From Theory to Applications*. Addison-Wesley, 1988.

[96] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. CRC Press, 2009.

[97] S. R. Cloude, *Polarisation Applications in Remote Sensing*. Oxford University Press, 2010.

[98] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 5, pp. 2142–2152, 2000.

[99] S. R. Cloude, "Polarization coherence tomography," *Radio Science*, vol. 41, no. 4, pp. 1–27, 2006.

[100] S. Tebaldini, "Single and multipolarimetric SAR tomography of forested areas: A parametric approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 5, pp. 2375–2387, 2010.

[101] O. Ponce, P. Prats-Iraola, R. Scheiber, A. Reigber, A. Moreira, and E. Aguilera, "Polarimetric 3-D reconstruction from multicircular SAR at P-band," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 4, pp. 803–807, 2014.

[102] O. Ponce, P. Prats-Iraola, M. Pinheiro, M. Rodriguez-Cassola, R. Scheiber, A. Reigber, and A. Moreira, "Fully polarimetric high-resolution 3-D imaging with circular SAR at L-band," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 52, no. 6, pp. 3074–3090, 2014.

[103] F. T. Ulaby and C. Elachi, Eds., *Radar Polarimetry for Geoscience Applications*. Artech House, 1990.

[104] S. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 1, pp. 68–78, 1997.

[105] A. Freeman and S. Durden, "A three-component scattering model for polarimetric SAR data," *IEEE Transaction on Geoscience and Remote Sensing*, vol. 36, no. 3, pp. 963–973, 1998.

[106] H. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 5, pp. 950–959, 1992.

[107] E. Rodriguez and J. Martin, "Theory and design of interferometric synthetic aperture radars," *IEE Proceedings F Radar and Signal Processing*, vol. 139, no. 2, pp. 147–159, 1992.

[108] F. Gatelli, A. M. Guarnieri, F. Parizzi, P. Pasquali, C. Prati, and F. Rocca, "The wavenumber shift in SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 4, pp. 855–865, 1994.

[109] S. Duque, U. Balss, C. Rossi, T. Fritz, and W. Balzer, "TanDEM-X payload ground segment, CoSSC generation and interferometric considerations," Remote Sensing Technology Institute, German Aerospace Center (DLR), Tech. Rep., 2012.

[110] J. O. Hagberg, L. M. H. Ulander, and J. I. H. Askne, "Repeat-pass SAR interferometry over forested terrain," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 2, pp. 331–340, 1995.

[111] J. O. Hagberg, "Repeat-pass satellite SAR interferometry," Chalmers University of Technology, Tech. Rep., 1995.

[112] Q. Lin, J. Vesecky, and H. Zebker, "New approaches in interferometric SAR data processing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 3, pp. 560–567, 1992.

[113] R. Scheiber and A. Moreira, "Coregistration of interferometric SAR images using spectral diversity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 5, pp. 2179–2191, 2000.

[114] L. C. Morena, K. V. James, and J. Beck, "An introduction to the RADARSAT-2 mission," *Canadian Journal of Remote Sensing*, vol. 30, no. 3, pp. 221–234, 2004.

[115] Y. Kankaku, Y. Osawa, S. Suzuki, and T. Watanabe, “The overview of the L-band SAR onboard ALOS-2,” in *Progress in Electromagnetics Research Symposium*, Moscow, Russia, 18–21 August 2009, pp. 735–738.

[116] R. K. Arora, C. V. N. Rao, B. V. Bakori, R. Neelakantan, and J. G. Vachchani, “Synthetic aperture radar payload on-board RISAT-1: Configuration, technology and performance,” *Current Science*, vol. 104, no. 4, pp. 447–461, 2013.

[117] R. Horn, “E-SAR – The experimental airborne L/C-band SAR system of DLR-HR,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Edinburgh, UK, September 12–16 1988, pp. 1025–1026.

[118] R. Horn, M. Werner, and B. Mayr, “Extension of the DLR airborne synthetic aperture radar, E-SAR, to X-band,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, College Park, MA, USA, May 20–24 1990, pp. 2047–2047.

[119] R. Horn, “The DLR airborne SAR project E-SAR,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Lincoln, NE, USA, May 27–31 1996, pp. 1624–1628.

[120] I. Hajnsek, R. Scheiber, M. Keller, R. Horn, S. Lee, L. M. H. Ulander, A. Gustavsson, G. Sandberg, T. Le Toan, S. Tebaldini, A. M. Guarnieri, and F. Rocca, “BioSAR 2008 technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2008 experiment: Final report – BioSAR campaign,” ESA contract no. 22052/08/NL/CT, Tech. Rep., 2009.

[121] R. Horn, A. Nottensteiner, A. Reigber, J. Fischer, and R. Scheiber, “F-SAR: DLR’s new multifrequency polarimetric airborne SAR,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Cape Town, South Africa, July 12–17 2009, pp. 902–905.

[122] F. Covello, F. Battazza, A. Coletta, E. Lopinto, C. Fiorentino, L. Pietranera, G. Valentini, and S. Zoffoli, “COSMO-SkyMed: An existing opportunity for observing the Earth,” *Journal of Geodynamics*, vol. 49, no. 3, pp. 171–180, 2010.

[123] G. Triltzsch, “Challenges of automated processing of spaceborne high resolution SAR data,” in *3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, 2011, pp. 1–2.

[124] R. Torres, P. Snoeij, D. Geudtner, D. Bibby, M. Davidson, E. Attema, P. Potin, B. Rommen, N. Floury, and M. Brown, “GMES Sentinel-1 mission,” *Remote Sensing of Environment*, vol. 120, pp. 9–24, 2012.

[125] X. Wang, G. Wang, Y. Guan, Q. Chen, and L. Gao, "Small satellite constellation for disaster monitoring in China," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Seoul, South Korea, July 25–29 2005, pp. 467–469.

[126] A. S. Gonzalez, M. Labriola, J. C. Soteras, and J. S. Palma, "PAZ instrument design and performance," in *3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, Seoul, South Korea, September 26–30 2011, pp. 1–4.

[127] J. Janoth, S. Gantert, T. Schrage, and A. Kaptein, "From TerraSAR-X towards TerraSAR next generation," in *European Conference on Synthetic Aperture Radar (EUSAR)*, Berlin, Germany, June 3–5 2014, pp. 1337–1340.

[128] D. D'Aria, D. Giudici, A. M. Guarnieri, P. Rizzoli, and J. Medina, "A wide swath, full polarimetric, L band spaceborne SAR," in *IEEE Radar Conference*, Rome, Idaly, May 26–30 2008, pp. 1–4.

[129] J. Colinas, G. Séguin, and P. Plourde, "Radarsat Constellation, moving toward implementation," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Honolulu, HI, USA, 25–30 July 2010, pp. 3232–3235.

[130] M. Zink, M. Bachmann, B. Brautigam, T. Fritz, I. Hajnsek, A. Moreira, B. Wessel, and G. Krieger, "TanDEM-X: The new global DEM takes shape," *IEEE Geoscience and Remote Sensing Magazine*, vol. 2, no. 2, pp. 8–23, 2014.

[131] T.-D. Wu, K.-S. Chen, J. Shi, H.-W. Lee, and A. K. Fung, "A study of an AIEM model for bistatic scattering from randomly rough surfaces," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 9, pp. 2584–2598, 2008.

[132] F. T. Ulaby, K. Sarabandi, K. McDonald, M. Whitt, and M. C. Dobson, "Michigan microwave canopy scattering model (MIMICS)," University of Michigan, Tech. Rep., 1988.

[133] F. T. Ulaby, K. Sarabandi, K. McDonald, M. Whitt, and M. C. Dobson, "Michigan microwave canopy scattering model," *International Journal of Remote Sensing*, vol. 11, no. 7, pp. 1223–1253, 1990.

[134] G. Sun and K. J. Ranson, "A three-dimensional backscatter model of forest canopies," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 2, pp. 372–382, 1995.

[135] M. Moghaddam and S. Saatchi, "Analysis of scattering mechanisms in SAR imagery over boreal forest: Results from BOREAS '93," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 5, pp. 1290–1296, 1995.

[136] M. L. Williams, *PolSARproSim: A Coherent, Polarimetric SAR Simulation of Forests for PolSARPro; Design Document and Algorithm Specification (u1.0)*, 2006.

[137] S. V. Nghiem, S. H. Yueh, R. Kwok, and F. K. Li, "Symmetry properties in polarimetric remote sensing," *Radio Science*, vol. 27, no. 5, pp. 693–711, 1992.

[138] A. K. Fung and K. S. Chen, *Microwave Scattering and Emission Models for Users*. Artech House, 2010.

[139] M. A. Karam, A. K. Fung, and Y. M. M. Antar, "Electromagnetic wave scattering from some vegetation samples," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 6, pp. 799–808, 1988.

[140] Y.-C. Lin and K. Sarabandi, "Electromagnetic scattering model for a tree trunk above a tilted ground plane," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 4, pp. 1063–1070, 1995.

[141] J. Lopez-Sanchez, H. Esteban-Gonzalez, M. Baquero-Escudero, and J. Fortuny-Guasch, "An electromagnetic scattering model for multiple tree trunks above a tilted rough ground plane," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, pp. 659–667, 1999.

[142] M. Brolly and I. H. Woodhouse, "A matchstick model of microwave backscatter from a forest," *Ecological Modelling*, vol. 237–238, pp. 74–87, 2012.

[143] E. Blomberg, M. J. Soja, and L. M. H. Ulander, "P-band polarimetric model of vertical tree stems on sloping ground," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Quebec City, QC, Canada, July 13–18 2014, pp. 80–83.

[144] E. Rignot, J. Way, C. Williams, and L. Viereck, "Radar estimates of aboveground biomass in boreal forests of interior Alaska," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 5, pp. 1117–1124, 1994.

[145] R. Freedman, "Size and shape in biology," *Scientific American*, vol. 224, 1971.

[146] G. B. West, J. H. Brown, and B. J. Enquist, "A general model for the origin of allometric scaling laws in biology," *Science*, vol. 276, no. 5309, pp. 122–126, 1997.

[147] Y. Dong and J. Richards, "Studies of the cylinder-ground double bounce scattering mechanism in forest backscatter models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 1, pp. 229–231, 1995.

[148] H. Israelsson, L. M. H. Ulander, T. Martin, and J. I. H. Askne, "A coherent scattering model to determine forest backscattering in the VHF-band," *IEEE*

Transactions on Geoscience and Remote Sensing, vol. 38, no. 1, pp. 238–248, 2000.

[149] B. Hallberg, G. Smith-Jonforsen, L. M. H. Ulander, and G. Sandberg, "A physical-optics model for double-bounce scattering from tree stems standing on an undulating ground surface," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 9, pp. 2607–2621, 2008.

[150] Y.-C. Lin and K. Sarabandi, "A Monte Carlo coherent scattering model for forest canopies using fractal-generated trees," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 1, pp. 440–451, 1999.

[151] Y.-C. Lin and K. Sarabandi, "Retrieval of forest parameters using a fractal-based coherent scattering model and a genetic algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1415–1424, 1999.

[152] L. Thirion, E. Colin, and C. Dahon, "Capabilities of a forest coherent scattering model applied to radiometry, interferometry, and polarimetry at P- and L-band," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 4, pp. 849–862, 2006.

[153] M. A. Karam, A. K. Fung, R. H. Lang, and N. S. Chauhan, "A microwave scattering model for layered vegetation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 4, pp. 767–784, 1992.

[154] C. C. Hsu, H. C. Han, R. T. Shin, J. A. Kong, A. Beaudoin, and T. Le Toan, "Radiative transfer theory for polarimetric remote sensing of pine forest at P band," *International Journal of Remote Sensing*, vol. 15, no. 14, pp. 2943–2954, 1994.

[155] M. L. Imhoff, "A theoretical analysis of the effect of forest structure on synthetic aperture radar backscatter and the remote sensing of biomass," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 2, pp. 341–352, 1995.

[156] G. Picard, T. Le Toan, S. Quegan, Y. Caraglio, and T. Castel, "Radiative transfer modeling of cross-polarized backscatter from a pine forest using the discrete ordinate and eigenvalue method," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1720–1730, 2004.

[157] P. Liang, M. Moghaddam, L. Pierce, and R. Lucas, "Radar backscattering model for multilayer mixed-species forests," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 11, pp. 2612–2626, 2005.

[158] J. E. S. Fransson, G. Smith, J. I. H. Askne, and H. Olsson, "Stem volume estimation in boreal forests using ERS-1/2 coherence and SPOT XS optical data," *International Journal of Remote Sensing*, vol. 22, no. 14, pp. 2777–2791, 2001.

[159] M. Santoro, J. Askne, G. Smith, and J. Fransson, "Stem volume retrieval in boreal forests from ERS-1/2 interferometry," *Remote Sensing of Environment*, vol. 81, no. 1, pp. 19–35, 2002.

[160] M. Santoro, A. Shvidenko, I. McCallum, J. Askne, and C. Schmullius, "Properties of ERS-1/2 coherence in the Siberian boreal forest and implications for stem volume retrieval," *Remote Sensing of Environment*, vol. 106, no. 2, pp. 154–172, 2007.

[161] J. Askne and M. Santoro, *Recent Interferometry Applications in Topography and Astronomy*. Morn Hill, 2012, ch. Experiences in Boreal Forest Stem Volume Estimation from Multitemporal C-Band InSAR.

[162] M. Santoro, J. I. H. Askne, and P. Dammert, "Tree height influence on ERS interferometric phase in boreal forest," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 207–217, 2005.

[163] J. T. Pullianen, K. Heiska, J. Hyyppä, and M. T. Hallikainen, "Backscattering properties of boreal forests at the C- and X-band," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 8, pp. 1041–1050, 1994.

[164] R. N. Treuhaft, S. N. Madsen, M. Moghaddam, and J. J. van Zyl, "Vegetation characteristics and surface topography from interferometric radar," *Radio Science*, vol. 31, pp. 1449–1485, 1996.

[165] J. Praks, F. Kugler, K. P. Papathanassiou, I. Hajnsek, and M. T. Hallikainen, "Height estimation of boreal forest: Interferometric model-based inversion at L- and X-band versus HUTSCAT profiling scatterometer," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 3, pp. 466–470, 2007.

[166] S. R. Cloude, "Dual-baseline coherence tomography," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 1, pp. 127–131, 2007.

[167] M. Lavalle, M. Simard, and S. Hensley, "A temporal decorrelation model for polarimetric radar interferometers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 7, pp. 2880–2888, 2012.

[168] L. G. Marklund, "Biomassafunktioner för tall, gran och björk i Sverige," Institutionen för skogstaxering, Sveriges lantbruksuniversitet, Umeå, Sweden, Rapport 45, 1988.

[169] H. Petersson, "Biomassafunktioner for trädfraktioner av tall, gran och björk i sverige," SLU, Institutionen för skoglig resurshållning och geomatik, Arbetsrapport 59, Umeå, Tech. Rep., 1999, (in Swedish).

[170] P. Wikström, L. Edenius, B. Elfving, L. O. Eriksson, T. Lämäs, J. Sonesson, K. Öhman, J. Wallerman, C. Waller, and F. Klintebäck, "The Heureka

forestry decision support system: An overview," *Mathematical and Computational Forestry & Natural-Resource Sciences*, vol. 3, no. 2, pp. 87–94, 2011.

[171] M. J. Soja, "Modelling and retrieval of forest parameters from synthetic aperture radar data," Chalmers University of Technology, Tech. Rep., 2012.

[172] S. Solberg, R. Astrup, J. Breidenbach, B. Nilsen, and D. Weydahl, "Monitoring spruce volume and biomass with InSAR data from TanDEM-X," *Remote Sensing of Environment*, vol. 139, pp. 60–67, 2013.

Part II

Appended Papers

Paper A

Polarimetric-Interferometric Boreal Forest Scattering Model for BIOMASS End-to-End Simulator

Authors:

M. J. Soja and L. M. H. Ulander

Status:

Reformatted and corrected version of the paper published in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Quebec City, QC, Canada, 13–18 July 2014, pp. 1061–1064

POLARIMETRIC-INTERFEROMETRIC BOREAL FOREST SCATTERING MODEL FOR BIOMASS END-TO-END SIMULATOR

Maciej J. Soja[1] *and Lars M. H. Ulander*[1,2]

1) Chalmers University of Technology, Gothenburg, Sweden
2) Swedish Defence Research Agency, Linköping, Sweden

ABSTRACT

A polarimetric-interferometric forward model (FM) for extended covariance matrix modeling is presented. The FM has been designed to be used within the end-to-end simulator for BIOMASS, a new ESA satellite mission aiming at the global mapping of above-ground forest biomass with P-band synthetic aperture radar (SAR). The FM uses linear regression models for prediction of backscatter intensity and HH-VV correlation coefficient, and the random volume over ground (RVoG) model for the prediction of the interferometric correlation coefficients. For boreal forest, parameter values for these sub-models have been derived using polarimetric-interferometric SAR data acquired within the BioSAR 2007 campaign over the Swedish test site Remningstorp. The FM is evaluated qualitatively in a boreal forest scenario through a side-by-side comparison with BioSAR 2007 data. The general agreement is good, although there are regions with structures which cannot be reproduced by the model, probably due to insufficient forest description by the input parameters.

***Index Terms*—** BIOMASS, forward model, extended covariance matrix

1. INTRODUCTION

In May 2013, European Space Agency (ESA) selected the BIOMASS satellite for the 7th Earth Explorer mission. The main goal of the mission is accurate, high-resolution mapping of global forest resources in terms of above-ground biomass (total mass of living forest tissue), biomass change, and forest height. This will aid global carbon cycle modelling, and eventually lead to improved climate change predictions [1].

BIOMASS will feature the first P-band synthetic aperture radar (SAR) in space, and also the lowest frequency SAR in space. The main advantage of P-band radar are its penetration capabilities. In forestry, this means that a P-band radar has the capability to see through the canopy and it is sensitive to scattering from trunks and large branches, which is where most biomass is stored. These structures are also significantly more stable in time (compared to the canopy), which means that temporal decorrelation at P-band is relatively low, and repeat-pass, multi-baseline interferometry and tomography will routinely be carried out. Also, with the fully polarimetric capabilities of BIOMASS, estimation of forest height will be done from polarimetric-interferometric SAR (PolInSAR) data.

In order to be able to evaluate the performance of the future BIOMASS satellite, a BIOMASS end-to-end simulator (BEES) has been implemented for both boreal and tropical forests [2]. Using the simulator, system effects can be modeled, and error budgets can be estimated. An important part of BEES is the forward model, which predicts the extended covariance matrix for different forest biomes from a small number of input parameters. A preliminary version of the model has been presented in [3]. In this paper, the boreal forest version of the forward model will be presented in its final version, and its performance in 2D modelling will be assessed qualitatively on data from BioSAR 2007.

2. DATA

SAR data were acquired with a flight heading of 200° over Remningstorp, a hemi-boreal test site located in southern Sweden, by the airborne ESAR system in May 2007 during BioSAR 2007 [4]. Small-footprint lidar-based estimates of biomass and forest height for 58 forest stands have been used for the estimation of model parameters. The errors of the FM have been estimated using ten $80\,\mathrm{m}\times 80\,\mathrm{m}$ forest plots, for which stem diameter has been measured for all trees, and height for a subset of trees [5]. For quantitative performance analysis, biomass and forest height maps derived from lidar data and species stratification information are used as input to the FM.

3. FORWARD MODEL

The model is designed to compute the extended covariance matrix for a polarimetric-interferometric pair. First, it is assumed that the backscatter signature is equal for both the master and slave images, which gives the following extended covariance matrix:

$$\hat{C}_6 = \begin{bmatrix} \hat{V} & \hat{K}_{12} \\ \hat{K}_{12}^H & \hat{V} \end{bmatrix}, \tag{1}$$

where ^{H} is the Hermitian (conjugate transpose) operator. $\hat{V}$ is the polarimetric covariance matrix, formulated as:

$$\hat{V} = \begin{bmatrix} \sigma^0_{\mathrm{HH}} & 0 & \tilde{\rho}\sqrt{\sigma^0_{\mathrm{HH}}\sigma^0_{\mathrm{VV}}} \\ 0 & 2\sigma^0_{\mathrm{HV}} & 0 \\ \tilde{\rho}^*\sqrt{\sigma^0_{\mathrm{HH}}\sigma^0_{\mathrm{VV}}} & 0 & \sigma^0_{\mathrm{VV}} \end{bmatrix}, \tag{2}$$

where the correlation between co- and cross-polarized channels has been shown to be zero for monostatic acquisitions [6]. The polarimetric-interferometric covariance matrix $\hat{K}_{12}$ can be formulated as:

$$\hat{K}_{12} = \begin{bmatrix} \tilde{\gamma}_{\mathrm{HH}}\sigma^0_{\mathrm{HH}} & 0 & \tilde{\rho}D \\ 0 & 2\tilde{\gamma}_{\mathrm{HV}}\sigma^0_{\mathrm{HV}} & 0 \\ (\tilde{\rho}D)^* & 0 & \tilde{\gamma}_{\mathrm{VV}}\sigma^0_{\mathrm{VV}} \end{bmatrix}, \quad (3)$$

where

$$D = \frac{\tilde{\gamma}_{\mathrm{HH}} + \tilde{\gamma}_{\mathrm{VV}}}{2}\sqrt{\sigma^0_{\mathrm{HH}}\sigma^0_{\mathrm{VV}}}. \quad (4)$$

Backscattering coefficient (σ^0) for polarization PQ is modelled in dB using a linear model with an additive error:

$$[\sigma^0_{\mathrm{PQ}}]_{\mathrm{dB}} = a_{\mathrm{PQ}} + b_{\mathrm{PQ}}\log_{10} B + 10\log_{10}(\cos\theta_i) + N(0, s^2_{\mathrm{PQ}}) \quad (5)$$

where θ_i is the local angle of incidence, B is the biomass in tons (Mg) per hectare (100 m x 100 m), and the last term is a normally distributed, zero-mean error, with standard deviation estimated using the 10 field plots. The parameter values estimated for the boreal data can be found in Table 1.

Table 1. Parameter values for backscatter model.

Polarization	a_{PQ}	b_{PQ}	s_{PQ}
HH	-20.1	8.1	1.3
HV	-20.7	4.2	0.7
VV	-6.7	0.6	1.2

The complex correlation coefficient between the HH and VV channels $\tilde{\rho}$ is for the boreal scenario modelled as:

$$\tilde{\rho} = (0.39 + N(0, 0.07^2)) \cdot e^{i(-41.5^\circ - 0.27B + N(0,(11.6^\circ)^2))} \quad (6)$$

where the last term in both magnitude and phase are normally distributed, zero-mean errors. The standard deviations have been estimated from the same field plots.

For the interferometric part, correlation coefficients ($\tilde{\gamma}_{\mathrm{PQ}}$) are modeled using the random volume over ground (RVoG) model with two different profile functions. Here, the exponential profile will be used, yielding:

$$\tilde{\gamma}_{vol} = \frac{\int_0^{h_{top}} f(z)e^{ik_z z}\mathrm{d}z}{\int_0^{h_{top}} f(z)\mathrm{d}z} = \frac{1}{1+\frac{ik_z\cos\theta_i}{2\sigma}} \cdot \frac{e^{\left(\frac{2\sigma}{\cos\theta_i}+ik_z\right)h_{top}}-1}{e^{\frac{2\sigma h_{top}}{\cos\theta_i}}-1} \quad (7)$$

where σ is the extinction coefficient, h_{top} is top forest height, and k_z is the vertical wave number. This is inserted in the general RVoG expression giving:

$$\tilde{\gamma}_{\mathrm{PQ}} = e^{ih_0k_z} \cdot \frac{\tilde{\gamma}_{vol} \cdot \gamma_{temp} + \mu_{\mathrm{PQ}}}{1 + \mu_{\mathrm{PQ}}}, \quad (8)$$

where $\gamma_{temp} = e^{-B_T/\tau_D}$ is a temporal decorrelation term, B_T is the temporal baseline, τ_D is decorrelation time, h_0 is ground height, and μ_{PQ} are ground-to-volume ratios.

In the boreal forest model, $\sigma = N(0.1, 0.1^2)$ dB/m has been chosen, based on results from PolInSAR height inversion, and $\mu_{\mathrm{HH}} = N(6.4, 1.3^2)$ dB, $\mu_{\mathrm{HV}} = N(-2.1, 0.7^2)$ dB, and $\mu_{\mathrm{VV}} = N(2.2, 0.7^2)$ dB were estimated from the data using polarimetric decomposition. h_0, h_{top}, k_z, B_T, and θ_i are known input parameters. τ_D is set through the choice of temporal decorrelation scenario.

4. RESULTS

The forward model is evaluated qualitatively for 2D mapping. Predictions of σ^0_{PQ}, $\tilde{\rho}$, and $\tilde{\gamma}_{\mathrm{PQ}}$ are made from biomass map, forest height map, and DTM, and compared to E-SAR data. Temporal decorrelation is neglected. The results are shown in Figures 1-4. The results are in general good, but in the case of VV-backscatter, HH-VV coherence, and interferometric coherences, the model does not predict some spatial changes, probably due to insufficient description of the scene with the input data. Information on, e.g., forest density or forest type would probably improve modeling.

5. REFERENCES

[1] ESA (2012), *Report for Mission Selection: BIOMASS*, Noordwijk, The Netherlands, 2012, (3 volume series).

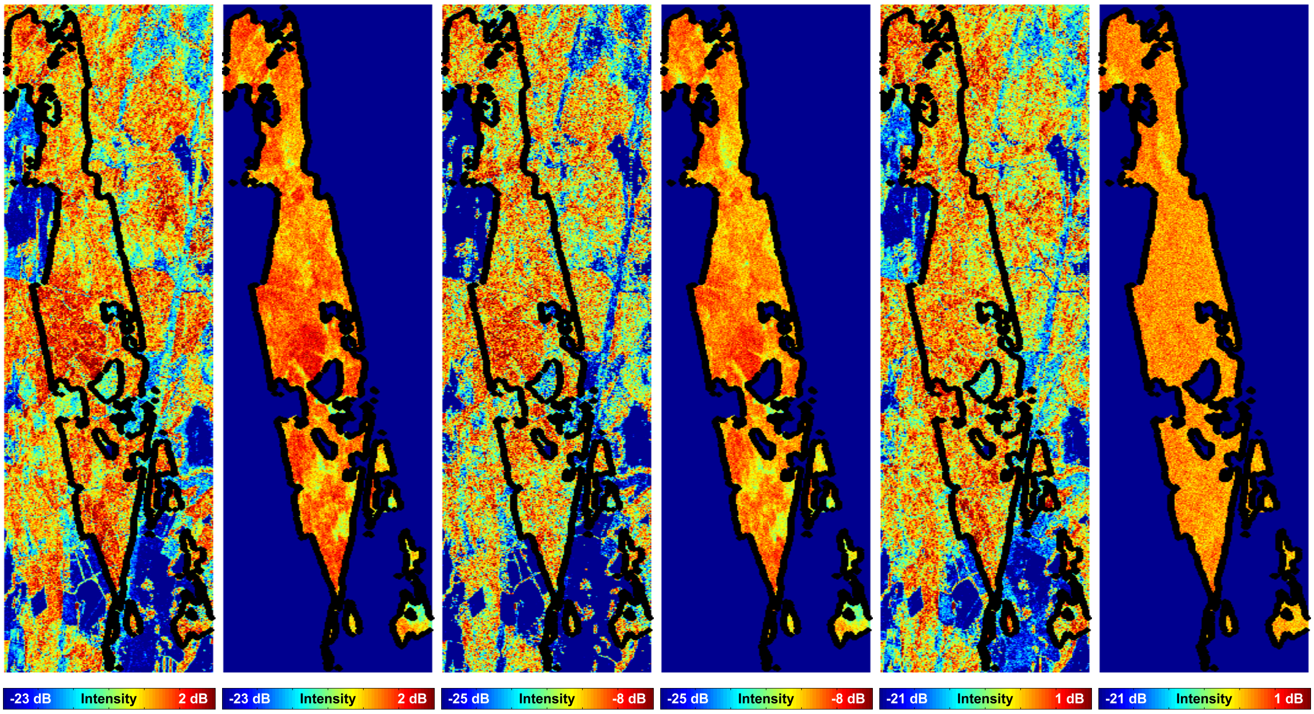
[2] P. Lopez-Dekker, F. De Zan, T. Borner, M. Younis, K. Papathanassiou, T. Guardabrazo, V. Bourlon, S. Ramongassie, N. Taveneau, L. Ulander, D. Murdin, N. Rogers, S. Quegan, and R. Franco, "Biomass end-to-end mission performance simulator," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vancouver, BC, Canada*, July 2011, pp. 4249–4252.

[3] Maciej J. Soja and Lars M. H. Ulander, "A hybrid model for interferometric and polarimetric p-band sar modelling of forests," in *Proceedings of PolInSAR 2011*, 2011, (ESA SP-695, March 2011).

[4] Irena Hajnsek, Rolf Scheiber, Lars Ulander, Anders Gustavsson, Gustaf Sandberg, Stefano Tebaldini, Andrea Monte Guarnieri, Fabio Rocca, Fabrizio Bombardini, and Matteo Pardini, "BioSAR 2007 technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2007 experiment: Final report without synthesis," Tech. Rep., ESA contract no. 20755/07/NL/CB, 2008.

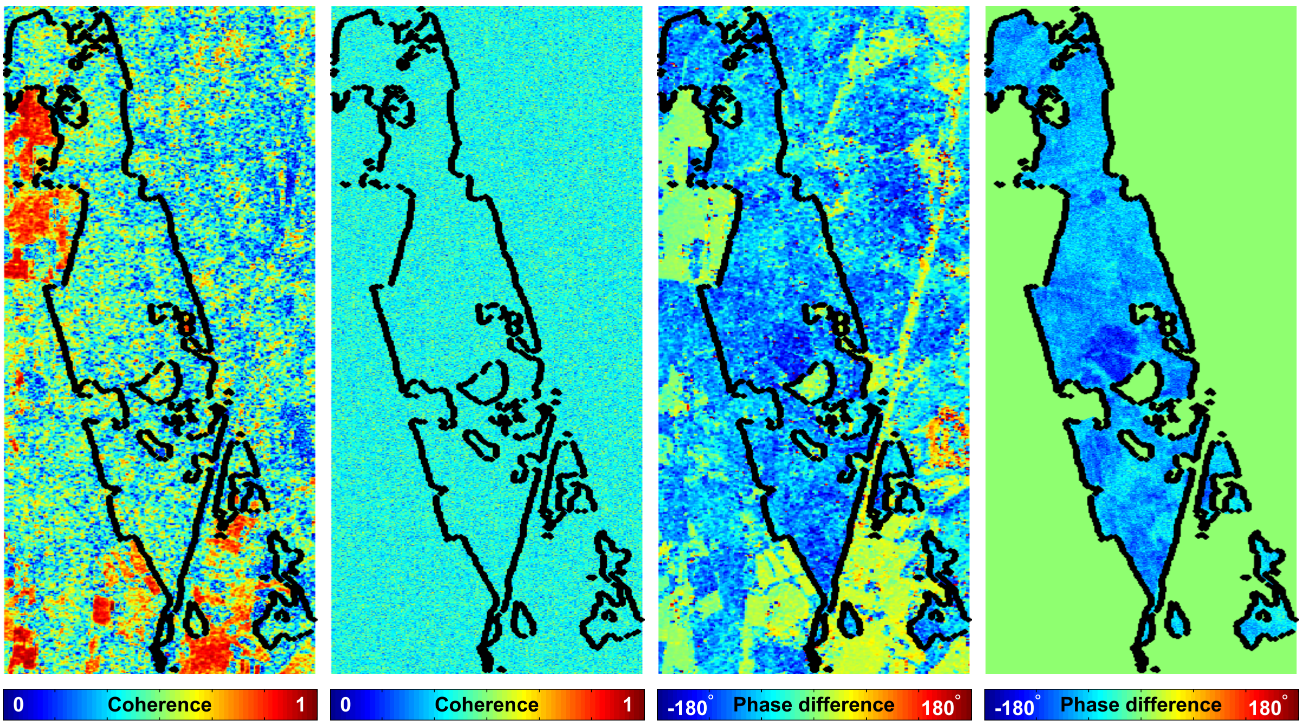
[5] G. Sandberg, L. M. H. Ulander, J. E. S. Fransson, J. Holmgren, and T. Le Toan, "L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest," *Remote Sensing of Environment*, vol. 115, no. 11, pp. 2874–2886, 2011.

[6] S. V. Nghiem, S. H. Yueh, R. Kwok, and F. K. Li, "Symmetry properties in polarimetric remote sensing," *Radio Science*, vol. 27, pp. 693–711, 1992.



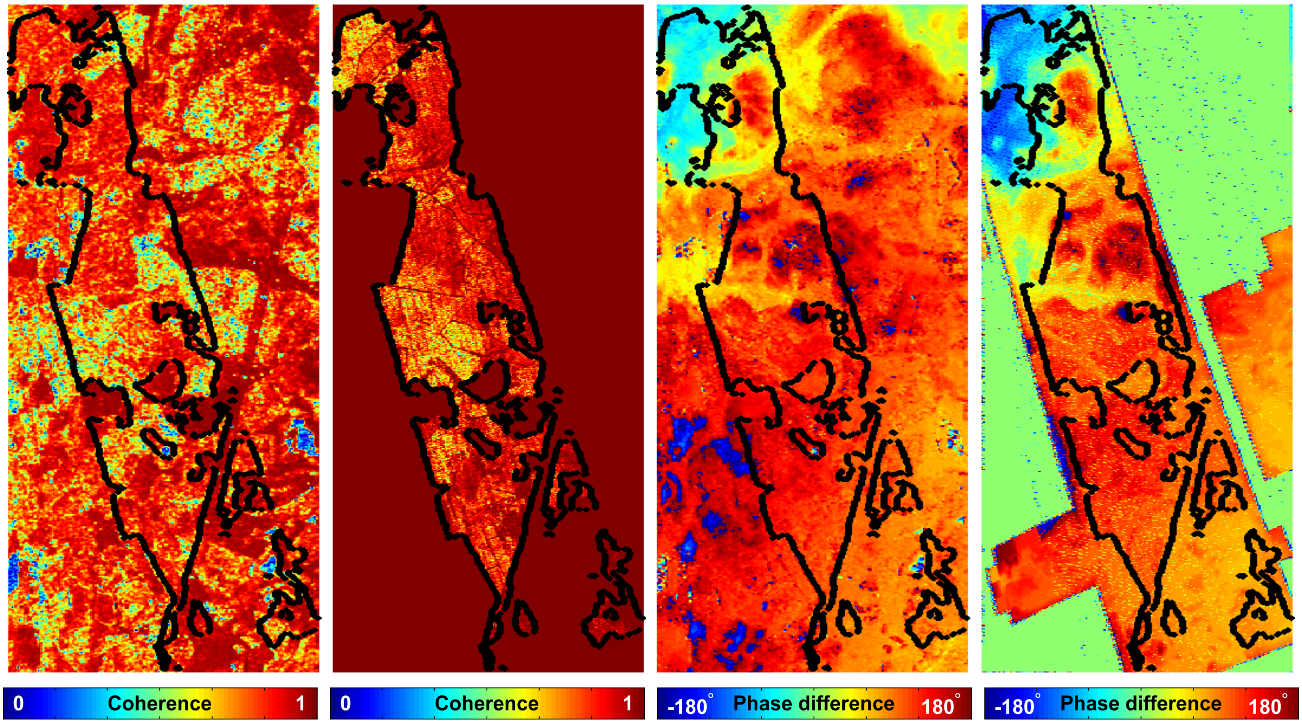
(a) E-SAR, HH (b) Model, HH (c) E-SAR, HV (d) Model, HV (e) E-SAR, VV (f) Model, VV

Fig. 1. Modeling results for backscattering coefficient compared to E-SAR data. All three polarizations are shown. The black outline marks the largest region covered by all required input data.



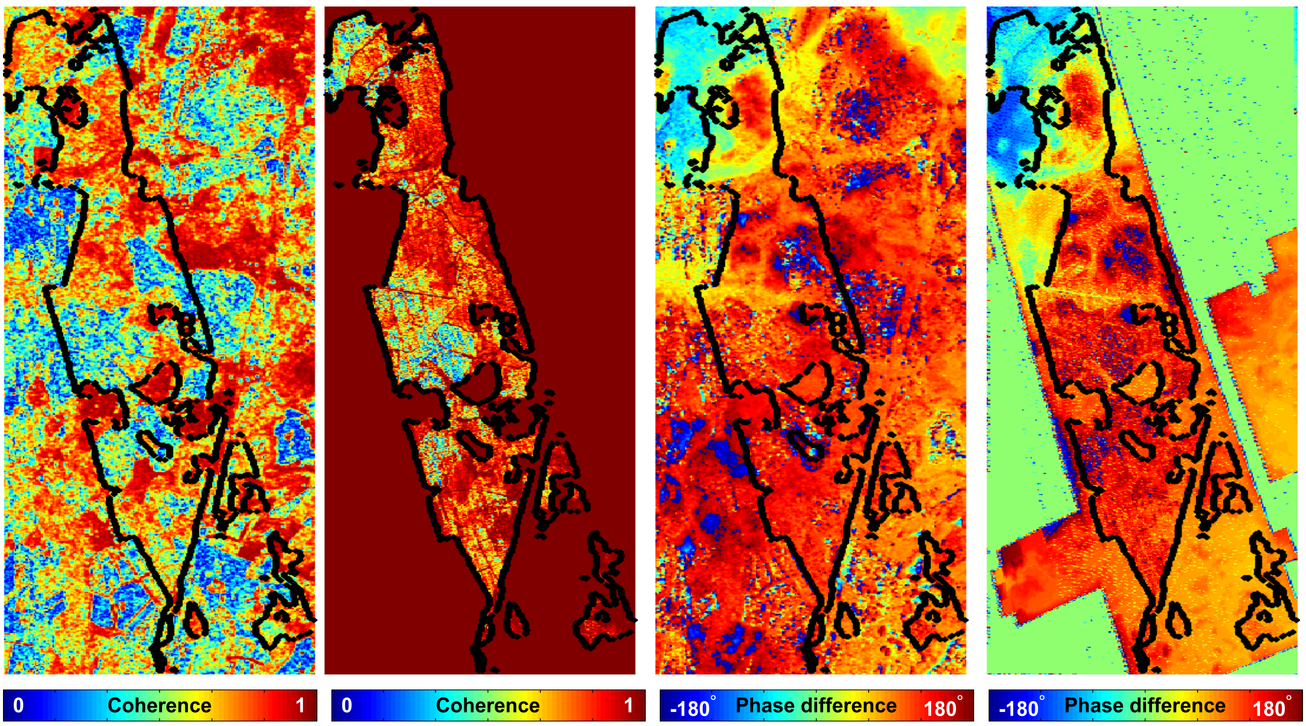
(a) E-SAR (b) Model (c) E-SAR (d) Model

Fig. 2. Modeling results for polarimetric coherence and phase compared to ESAR data. The black outline marks the largest region covered by all required input data.



(a) E-SAR, VV (b) Model, VV (c) E-SAR, VV (d) Model, VV

Fig. 3. Modeling results for interferometric coherence and phase compared to ESAR data. VV-polarization is shown here. The black outline marks the largest region covered by all required input data.



(a) E-SAR, HV (b) Model, HV (c) E-SAR, HV (d) Model, HV

Fig. 4. Modeling results for interferometric coherence and phase compared to ESAR data. HV-polarization is shown here. The black outline marks the largest region covered by all required input data.

Paper B

Regression-Based Retrieval of Boreal Forest Biomass in Sloping Terrain using P-band SAR Backscatter Intensity Data

Authors:

M. J. Soja, G. Sandberg, and L. M. H. Ulander

Status:

Published in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 2646–2665, May 2013

Regression-Based Retrieval of Boreal Forest Biomass in Sloping Terrain Using P-Band SAR Backscatter Intensity Data

Maciej Jerzy Soja, Gustaf Sandberg, and Lars M. H. Ulander, *Senior Member, IEEE*

***Abstract*—A new biomass retrieval model for boreal forest using polarimetric P-band synthetic aperture radar (SAR) backscatter is presented. The model is based on two main SAR quantities: the HV backscatter and the HH/VV backscatter ratio. It also includes a topographic correction based on the ground slope. The model is developed from analysis of stand-wise data from two airborne P-band SAR campaigns: BioSAR 2007 (test site: Remningstorp, southern Sweden, biomass range: 10–287 tons/ha, slope range: 0–4°) and BioSAR 2008 (test site: Krycklan, northern Sweden, biomass range: 8–257 tons/ha, slope range: 0–19°). The new model is compared to five other models in a set of tests to evaluate its performance in different conditions. All models are first tested on data sets from Remningstorp with different moisture conditions, acquired during three periods in the spring of 2007. Thereafter, the models are tested in topographic terrain using SAR data acquired for different flight headings in Krycklan. The models are also evaluated across sites, i.e., training on one site followed by validation on the other site. Using the new model with parameters estimated on Krycklan data, biomass in Remningstorp is retrieved with RMSE of 40–59 tons/ha, or 22–33% of the mean biomass, which is lower compared to the other models. In the inverse scenario, the examined site is not well represented in the training data set, and the results are therefore not conclusive.**

***Index Terms*—Biomass retrieval, boreal forest, P-band, synthetic aperture radar (SAR), topographic correction.**

I. Introduction

FACING the threat of global warming, one of the most important topics in climate research is understanding the terrestrial carbon cycle and predicting future climate changes. One of the major uncertainties in the current carbon cycle models lies in terrestrial ecosystems, in particular forests [1]. Moreover, up to 20% of the global emissions of carbon dioxide are estimated to come from deforestation [2]. Accurate, global-scale forest mapping is therefore one of the most important elements of climate modeling. Current global forest maps are simply too inaccurate for this task, creating a demand for the development of new tools.

The most relevant quantity directly related to the forestal carbon stock is aboveground dry biomass (further on simply called "biomass"). Biomass is the dry weight of aboveground forest, including stem, bark, branches, and needles/leaves, but excluding stump and roots. Biomass is usually measured in metric tons per hectare (1 ton/ha $= 0.1$ kg/m^2).

Currently, the most accurate technique for remote biomass mapping is small-footprint lidar scanning (see [3] and references therein). However, accurate lidar-based biomass estimation requires high-quality plot-level measurements for training. Biomass tends also to be underestimated as small trees may be covered by large trees blocking the laser beam. As with all optical methods, measurement accuracy is dependent on weather conditions. In reality, small-footprint lidar scanning is inefficient for global biomass mapping. Spaceborne lidar has been considered a possible alternative, but complications arise chiefly due to large footprint and low coverage, and there are currently no ongoing spaceborne lidar missions.

Synthetic aperture radar (SAR) is a high-resolution, microwave imaging sensor which is weather independent and provides its own illumination. Moreover, SAR systems can be customized to fit a particular task through the choice of system parameters (frequency, polarization, incidence angle, and imaging mode).

SAR imaging at low frequencies (here: below L-band) has proven itself particularly useful for biomass mapping due to its superior penetration capabilities and sensitivity to a wide range of biomass levels. Due to transmission restrictions, there neither are, nor have been, any satellites in Earth's orbit with a SAR sensor operating below L-band. Therefore, all low-frequency studies have been performed using data acquired with airborne platforms. The low VHF-band (20–90 MHz) SAR system CARABAS-II, run by the Swedish Defence Research Agency (FOI), has previously proven itself useful for accurate stem volume estimation (see [4] and references therein).

Also, several P-band (approximately 0.20–0.45 GHz) studies have been performed using data acquired with airborne SAR systems [5]–[17]. In all these studies, regression models relating biomass to SAR observables are derived (see Table I for a summary of these models). They all conclude that biomass and radar backscatter are correlated, but the presented functions and their regions of validity differ (due to different biomes and moisture conditions, different acquisition platforms, and

Manuscript received February 10, 2012; revised June 5, 2012; accepted August 30, 2012. Date of publication November 20, 2012; date of current version April 18, 2013. This work was supported by the Swedish National Space Board and the European Space Agency.

M. J. Soja and G. Sandberg are with the Department of Earth and Space Sciences, Chalmers University of Technology, 412 96 Gothenburg, Sweden (e-mail: maciej.soja@chalmers.se; gustaf.sandberg@chalmers.se).

L. M. H. Ulander is with the Department of Earth and Space Sciences, Chalmers University of Technology, 412 96 Gothenburg, Sweden. He is also with the Radar Systems Unit, Swedish Defence Research Agency, 581 11 Linköping, Sweden (e-mail: lars.ulander@foi.se).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2012.2219538

TABLE I
SUMMARY OF SOME PREVIOUS STUDIES IN WHICH BIOMASS OR STEM VOLUME WERE RELATED TO P-BAND BACKSCATTER THROUGH REGRESSION. MODELS INCLUDING BACKSCATTER AT OTHER FREQUENCIES THAN P-BAND WERE DISREGARDED. IN SOME CASES, MODELS WERE JUST HINTED (FOR EXAMPLE THROUGH A STUDY OF CORRELATION), BUT NOT EXPLICITLY DEFINED OR USED IN THE TEXT [5], [7]. NOTE, THAT IN THE CASES OF [10], [12], AND [13], FORWARD MODELS ARE DEFINED

Ref.:	Sensor (campaign):	Test site:	Forest type:	Model structure:	Modelled quantity:
[5]	AirSAR (MAESTRO 1)	Landes (southern France)	maritime pine	hinted: linear function of HV in linear units	trunk biomass
[6]	AirSAR (MAESTRO 1)	Landes and Duke University Forest, (North Carolina, USA)	maritime pine and lolloby pine, respectively	linear function of HV in linear units	aboveground biomass
[7]	AirSAR (MAESTRO 1)	Flevoland (the Netherlands)	poplar and ash studied here, also other deciduous and coniferous in test site	hinted: linear functions of single polarisation (HH, VV, HV, all in dB)	logarithm of stem volume
[8]	AirSAR (MAESTRO 1)	Freiburg (south-west Germany)	Norway spruce, Scots pine, silver fir, some deciduous	linear function of HV in dB	stem volume
[9]	AirSAR	Howland (Maine, USA)	boreal coniferous and northern hardwood	linear function of HV in dB	logarithm of aboveground biomass
[10]	AirSAR	Hawaii Volcanoes National Park (Hawaii, USA), Landes, and Duke University Forest	broadleaf evergreen, maritime pine, and lolloby pine, respectively	third order polynomials of aboveground biomass (evergreen and combined) or logarithm of aboveground biomass (pine)	backscatter in dB (HH, VV, HV)
[11]	AirSAR	Landes, Duke University Forest, Bonanza Creek experimental forest (Alaska, USA), and Manu National Park (Peru)	maritime pine, lolloby pine, boreal forest, and primary tropical rain forest, respectively	linear function of multiple polarisations (HH, VV, HV, all in dB), and their ratios and squares	logarithm of aboveground biomass
[12]	AirSAR	Guaviere (Colombian Amazon)	primary and secondary forest, recently cut forest, pastures	power function of aboveground biomass	backscatter in dB (HH, VV, HV)
[13]	SAR AeS	Tapajós River region (Pará State, Brazil)	primary rainforest and secondary succession	linear functions of logarithm of aboveground biomass and third order polynomials of aboveground biomass	backscatter in dB (HH, VV, HV)
[14]	AirSAR	Yellowstone National Park (USA)	coniferous (mainly lodgepole pine)	linear function of multiple polarisations (HH, VV, HV, all in dB), and their squared terms, all including topographic corrections	logarithm of aboveground biomass (separately for trunk and crown)
[15]	ESAR (BioSAR 2007)	Remningstorp (southern Sweden)	boreal (Norway spruce, Scots pine, birch)	linear functions of multiple polarisations normalised to gamma nought (HH, HV, both in dB)	square root of aboveground biomass
[16]	AirSAR	La Selva (Costa Rica)	lowland old growth and secondary tropical wet forest	linear functions of multiple polarisations (HH, VV, HV, all in dB) and InSAR height	square root of aboveground biomass
[17]	ESAR (BioSAR 2008)	Krycklan (northern Sweden)	boreal coniferous (mostly Norway spruce and Scots pine)	linear functions of multiple polarisations (HH, VV, HV, all in dB), InSAR height, and several PolInSAR indicators	aboveground biomass

changes in forest structure and surface topography). This means that the models derived in these papers usually have little or no application outside the studied test site. This is an obvious disadvantage when global biomass mapping is concerned.

At low frequencies, radio waves are generally scattered from larger objects, which in the case of trees means trunks and large branches. The increased temporal stability (as compared to for example X-band) makes it possible to perform repeat-pass polarimetric SAR interferometry (PolInSAR), which produces forest height estimates [18]–[20]. However, both PolInSAR-based height estimation and allometric height-to-biomass conversion are sensitive to parameters such as vertical structure, species composition, and management procedures [21]. Since it is not likely that these parameters can be estimated accurately with radar, accurate biomass estimation from PolInSAR is aggravated. Possible improvements include multi-baseline PolInSAR [22], [23] and different tomographic techniques [24]–[26]. However, these techniques require the acquisition of high-quality multi-baseline data, which is a very costly and time-consuming process.

Although the temporal stability and biomass sensitivity are both improved at low frequencies, a different problem occurs instead: ground topography. The double-bounce effect (scattering between ground and trunk, or vice versa) is very prominent at low frequencies, and ground tilt has an obvious influence. This issue has been addressed in [27], where a physical-optics model was successfully used to describe the influence of topography on radar backscatter from forests (at both VHF- and P-band). In [4] and [28], a simplified approach based on electromagnetic models like those described in [29]–[34] was used at VHF-band to reduce topographic influence, giving stem volume retrieval results comparable to those for flat ground. In this text, an even simpler approach will be used. The influence of topography will be examined as the change in model parameters for some reference models, and the most prominent factors will be included.

Due to the recent opening of the P-band at frequencies 432–438 MHz for spaceborne use (World Radiocommunications Conference 2003 [35]), a fully polarimetric P-band SAR satellite system called BIOMASS has been proposed to

European Space Agency (ESA) for the 7th Earth Explorer mission [35]–[39]. The system is planned to employ both intensity-based biomass retrieval and PolInSAR-based height retrieval. The two methods show different performance in different environments and are complementary, thus extending the capability of the proposed satellite.

In this paper, a new model for biomass retrieval from polarimetric SAR backscatter is presented. The model is tested for its sensitivity to site topography and for temporal change. Also, the model is compared to some previously published models and evaluated using two sets of test data. The data were acquired within two BioSAR campaigns performed in 2007 and 2008 in the two test sites Remningstorp and Krycklan, respectively, both situated in Sweden. The test sites are located 720 km apart and represent two different cases of boreal forest. In previous papers dealing with biomass retrieval from BioSAR data, the two test sites were treated separately [15], [17], [40]–[42]. In this paper, models fitted to data from one test site are evaluated on the other. In this manner, the model is validated independently of the training data set. An excerpt of the results presented here has been published in [43].

This paper begins with a brief description of the experimental data (Section II). Next, in Section III, the previously published models are presented, and the new model is introduced. Thereafter, the models are evaluated with respect to temporal change, topographic change, and across-site retrieval (Section IV). The results are summarized, and conclusions are drawn in Section V.

II. Experimental Data

The experimental data used in this paper were acquired within two BioSAR campaigns conducted by the airborne experimental SAR (ESAR) platform from the German Aerospace Center (DLR). Ground-truth data were collected and processed by Swedish University of Agricultural Sciences (SLU).

A. Test Sites

BioSAR 2007 was conducted in Remningstorp ($58°28'$ N, $13°38'$ E) situated in southern Sweden, see Fig. 1. Remningstorp is fairly flat with ground slopes at stand level less than 5° (computed from a 50 m × 50 m digital elevation model, DEM). The test site covers approximately 1200 ha of productive forest land, and the forest consists primarily of Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and birch (*Betula* spp.). For a thorough description of the campaign, see [15], [44].

BioSAR 2008 was conducted in Krycklan ($64°14'$ N, $19°46'$ E) located in northern Sweden, see Fig. 1. Krycklan is situated 720 km north-north-east of Remningstorp. Unlike Remningstorp, Krycklan has a strongly undulating topography with ground slopes on stand level up to 19° (again, computed from a 50 m × 50 m DEM). The forest is dominated by Norway spruce and Scots pine. For a thorough description of the campaign, see [45].

It is worth mentioning that a third BioSAR campaign has been conducted in Remningstorp in October 2010, aiming at the detection of long-term temporal changes in Remningstorp,

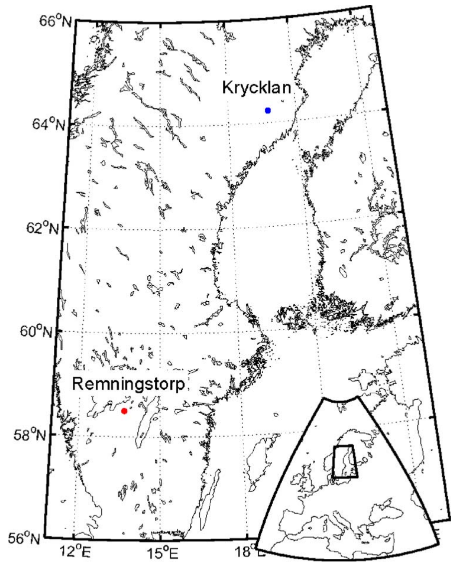


Fig. 1. Two test sites used in BioSAR 2007 and BioSAR 2008 campaigns are shown here. The test area in Remningstorp was covered by SAR imagery in the spring of 2007, whereas Krycklan was covered in October 2008. The distance between the two sites is 720 km.

see [46], [47]. However, data processing and analysis were not finished at the time of writing of this text, and this campaign is thus not included.

In the following text, the two test sites will sometimes be referred to as Re (Remningstorp) and Kr (Krycklan).

B. In-Situ and Laser Scanning Data

In conjunction with both BioSAR campaigns, plot-level *in-situ* data and airborne lidar scanning data were collected for the estimation of biomass. Species stratification information extracted from aerial photography was also used to aid biomass estimation. Biomass maps with 10 m × 10 m pixels were produced for both Remningstorp and Krycklan. Slightly different data collection strategies and estimation procedures were used for the two campaigns, and campaign reports should be consulted for a thorough description [44], [45].

Table II summarizes four reference data sets used in this work, together with their approximate error levels and their type. In forestry, a distinction between "plots" and "stands" is made. Stands are relatively homogenous forest regions with similar species composition, biophysical characteristics (e.g., height and tree number density), and management procedures. They can vary in size and shape, and they are the main unit used for forest mapping and management [48]. Plots are usually smaller stand subsets of regular shape, which are used as within-stand samples. They are usually distributed in a regular pattern. For each test site, two data sets are available. Here follows a short description of these data sets.

The first data set in Remningstorp consists of 10 80 m × 80 m plots [44]. Only trees with stem diameter at breast height (dbh, measured 1.3 m above ground level) larger than 5 cm were included in the measurements. Position, dbh, and species were measured for all relevant trees for all ten plots. Tree height was

TABLE II

Summary of Available Biomass Reference Data. Only Stands Completely Covered by P-Band SAR Data are Included. **SID** Stands for "Site ID" (Re for Remningstorp, Kr for Krycklan). **GID** Stands for "Group ID" and Refers to Type of Stand-Wise Data Set (Based on **Main data source**). **N** is the Size of Each Data Set. **Type** Refers to the Correct Denomination of the Data Points, as it Would be Referred to in Forestry. **Mean $\mathcal{B}$** and **$\mathcal{B}$ range** Refer to the Mean Biomass and Biomass Range for Each Data Set. **Area** Refers to the Stand Area (or Area Range) in Hectares. **Error** Refers to the Estimated Standard Biomass Error (if in %, Then Relative **Mean $\mathcal{B}$**, if a Percentage Interval, Then Different Percentage for Each Stand Relative Its Mean Biomass)

SID:	GID:	N:	Type:	Main data source:	Mean $\mathcal{B}$	$\mathcal{B}$ range:	Area range:	Error:
Re	INS	10	plots	plot-level measurements	181	52–267	0.66–0.69	a few percent
	LID	58	stands	stem volume map, species stratification info	129	10–287	0.50–9.4	25 tons/ha
Kr	INS	29	stands	plot-level measurements	95	23–183	1.5–22	4–21 %
	LID	97	plots	biomass map	76	8–257	0.79	16 %

TABLE III

Summary of Available Stand-Wise SAR Data. **SID** Stands for "Site ID" (Re for Remningstorp, Kr for Krycklan). **DID** Stands for "Date ID" and Refers to the Acquisition Date. **Image ID** Refers to the Image Identification Numbers as Defined in [44], [45]. **GID** Stands for "Group ID" and Refers to Type of Stand-Wise Data Set (LID for Lidar Measurement-Based Stands, and INS for *In Situ* Measurement-Based Stands). **Covered stands** Refers to the Number of Stands Covered for Each Scene, Respectively. **Biomass range** Refers to the Biomass Range of the Covered Stands (in tons/ha) for Each Scene, Respectively. Consult Also Table II for a Description of the Different Reference Data Sets

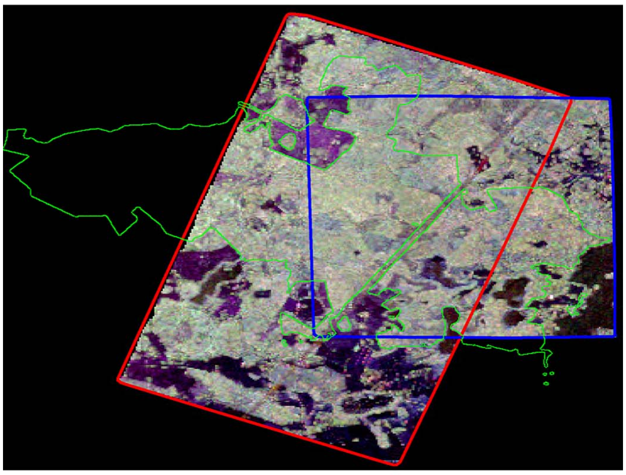
SID:	DID:	Heading:	Image ID:	GID:	Covered stands:	Biomass range:
Re	Mar	179°	0110	INS	9	52–267
				LID	46	10–287
		200°	0105, 0109	INS	10, 10	52–267, 52–267
				LID	58, 58	10–287, 10–287
	Apr	179°	0206	INS	9	52–267
				LID	46	10–287
		200°	0301, 0306	INS	10, 10	52–267, 52–267
				LID	58, 58	10–287, 10–287
	May	179°	0412	INS	9	52–267
				LID	46	10–287
		200°	0406, 0411	INS	10, 10	52–267, 52–267
				LID	58, 58	10–287, 10–287
Kr	Oct	43°	0304	INS	10	27–167
				LID	97	8–257
		134°	0104, 0303, 0305	INS	28, 9, 10	23–183, 27–167, 27–167
				LID	97, 97, 97	8–257, 8–257, 8–257
		314°	0103, 0302	INS	27, 10	27–183, 27–167
				LID	97, 97	8–257, 8–257
		358°	0301	INS	9	27–167
				LID	97	8–257

measured for all trees in four plots, and for a subset of trees in the other six plots. Biomass was then estimated for each single tree using Marklund's species specific allometric formulas, see [49]. The biomass estimation error (standard deviation of the residuals) computed using error estimates found in [49] is estimated to a few percent [15].

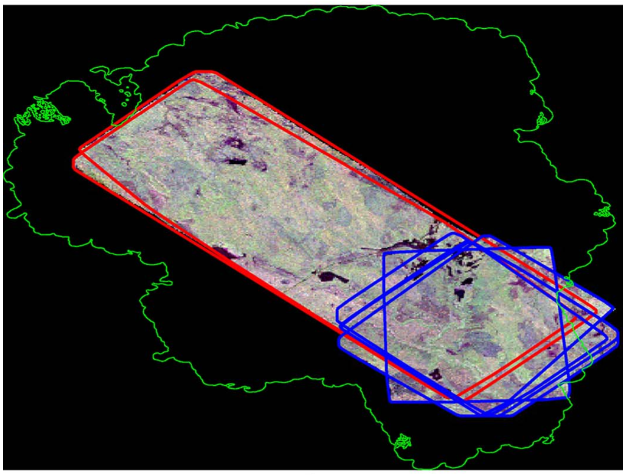
The second data set in Remningstorp consists of 58 stands of irregular shape and sizes between 0.5 and 9.4 ha [44]. A systematic grid of 849 circular field plots (radius 10 m) with a spacing of approximately 40 m was used. Within each field plot, all trees with dbh larger than 5 cm were calipered, and tree height was measured for approximately 10% of these trees. These data were then used together with lidar scanning data and species stratification information to obtain estimates of biomass for all 58 stands. The estimated standard biomass error for these 58 stands is 25 tons/ha, computed using validation against the ten plots described in the previous paragraph, see [15].

The first data set in Krycklan consists of 29 stands of irregular shape and sizes between 1.5 and 22 ha [45]. Systematic grids of circular field plots (radius 10 m) were laid out in each stand. The spacing of each grid was selected to give approximately ten field plots per stand. For each field plot, all trees with dbh larger than 4 cm were calipered, and the species was determined. Tree height and age were also measured for approximately 1.5 randomly chosen sample trees in each field plot. Biomass was then determined using Petersson's biomass functions [50]. The estimated standard biomass error was computed based on the number of field plots within each stand and the variation between these plots within each stand [45], [51]. This error estimate varies between 4 and 21%, depending on stand.

The second data set in Krycklan consists of 97 plots. This set has been introduced in [41] and it is based on data acquired from airborne lidar scanning. Functions estimating biomass from lidar observables were derived using multiple regression and studies of residuals based on field plot data (both from the previously mentioned field plots situated within stands and additional 110 field plots randomly positioned in the central part of the Krycklan test site). A biomass map was then created using lidar data with additional species information acquired from aerial stereo photography interpretation. Ninety-seven circular plots (radius 50 m) were selected within the region fully covered by the biomass map and SAR images for all four flight headings [see Section II-C and Fig. 2(b)], and mean biomass estimates were extracted from the biomass map. The plots were selected to have as constant ground slope as possible. The standard biomass error was here estimated to be



(a)



(b)

Fig. 2. Different acquisition headings are visualized here for (a) BioSAR 2007: Remningstorp, and (b) BioSAR 2008: Krycklan. In red, image frames for the main headings are shown (headings used for PolInSAR height retrieval and SAR tomography; 200° in Remningstorp, 134° and 314° in Krycklan). In blue, image frames for additional headings are shown (179° in Remningstorp, 43°, 134° (flown twice), 314°, and 358° in Krycklan). In green, the borders of the test sites are shown. As background, polarimetric SAR images are used (HH in the red channel, HV in the green channel, and VV in the blue channel; all channels are scaled for optimal viewing). ESAR is a left-looking system.

16%, which is equal to the error of the corresponding biomass map, for which it was computed by cross-validation against the previously mentioned 29 stands [45].

As it can be observed, biomass estimates for the data sets based on plot-level measurements are generally more accurate than for those based on maps and lidar data. In this text, the available reference data will therefore be divided in two groups. The stands and plots with biomass estimated only from plot-level *in-situ* measurements will be referred to as INS-stands, while the other data sets will be referred to as LID-stands, see Table II. Note, that although the stands can vary drastically in size (0.5–22 ha), the number of looks is at least 390 (for the 0.5-hectare stands, see [15]), which allows to disregard the variation in stand area in the further analysis.

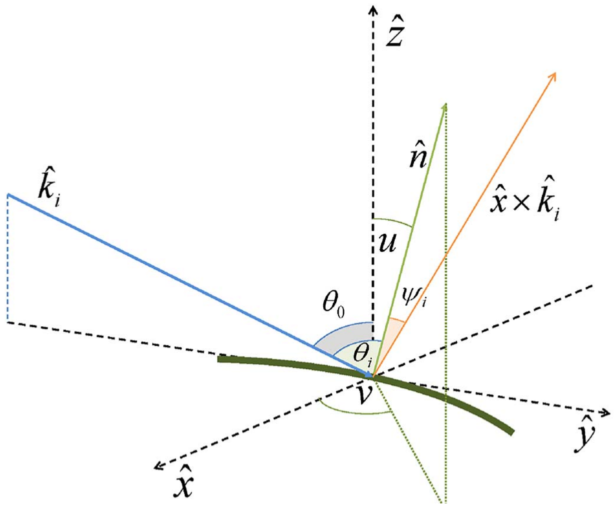


Fig. 3. Basic acquisition geometry. The ground normal is $\widehat{n}$, and the ground slope is defined by the two angles u and v. The incident unit wave vector $\widehat{k_i}$ is assumed to lie in the y–z-plane.

C. SAR Data

In Remningstorp, P-band SAR data were collected during three different periods of spring 2007: 3rd of March, 31st of March to 2nd of April, and 2nd of May. At each occasion, two flight headings were used for P-band: 179° and 200° relative north, marked in blue and red, respectively, in Fig. 2(a). The first track features steeper incidence angles for all stands, close to those expected for a spaceborne scenario (all stands lie in near range with nominal incidence angles between 26° and 35°). The second track features a wider range of incidence angles (between 30° and 50°). It was flown several times at each occasion at different baselines in order to provide PolInSAR and tomographic data. No precipitation was observed within 24 h prior to the acquisitions in the vicinity of the observation point ($58°27'$ N, $13°40'$ E, one automatic weather station maintained by the Swedish Meteorological and Hydrological Institute). Field notes and photography from March show, that the forest soil was often saturated by water and standing water on the surface was commonly observed, most often due to the recently melted snow present in these areas. In April and May, corresponding observations show, that the ground had dried out and the soil moisture was considerably lower. These observations are consistent with the fact, that May and June are generally the driest period of the year in the region.

In Krycklan, P-band SAR data were acquired during two days only: 14th and 15th of October 2008. The first day, the main flight track (134°) was flown several times at different baselines for PolInSAR use. The same area was also covered from the opposite direction (314°). The second day, SAR data of a smaller area were collected from four directions (headings: 43°, 134°, 314°, and 358° relative north). These additional flight tracks were selected in such a way that the regions with strongest topographic variability were covered by data from all flight tracks. In Fig. 2(b), an overview for the different acquisitions is shown. Image frames for the two main

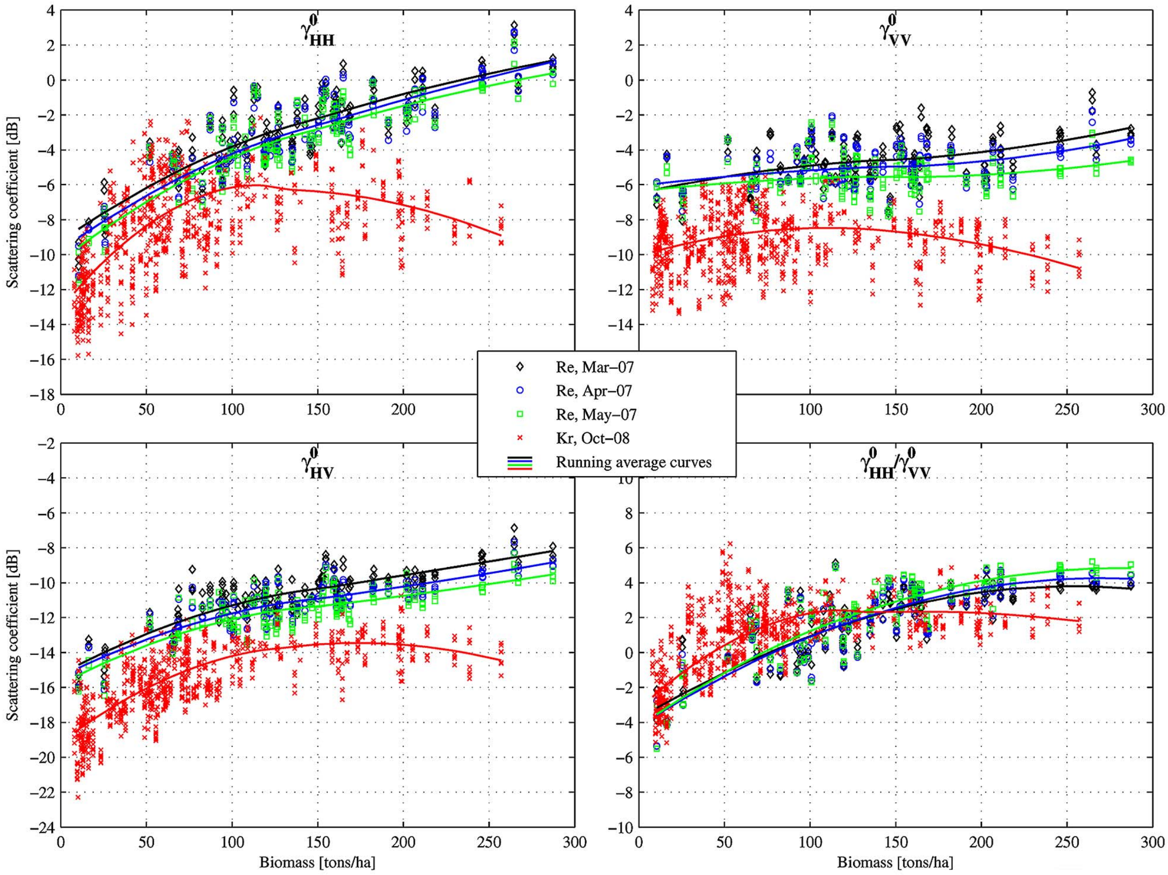


Fig. 4. Backscatter gamma nought for HH, HV, and VV, as well as HH/VV backscatter ratio are here plotted in dB for both Krycklan and Remningstorp. Data points are plotted in different colors and markers if they represent different acquisition time and site. Four running average curves are also plotted to simplify trend investigation. Their colors correspond to the colors of the data points. The grid spacing in y-direction is 2 dB in all four plots.

acquisitions are shown in red, whereas image frames for the additional acquisitions are shown in blue. No precipitation was observed at the test site before and during the acquisitions. Weather conditions were recorded using an operational weather station found at the nearby Svarteberget Research Station, and soil moisture was measured using samples from 10 stands in Krycklan. For a thorough description of the weather and soil moisture data, see [45].

Averaged, stand-wise backscatter data were extracted from the geocoded SAR images for each stand in both Remningstorp and Krycklan. A 50 m $\times$ 50 m DEM was used for geocoding and normalization. Although high-resolution lidar DEMs were also available for both test sites, they were not used because the evaluation scenario would be less realistic as comparable DEM resolutions are not available on global scale. All normalization procedures were performed before averaging, that is on high-resolution SAR data. A buffer zone of 10 m was also added to avoid border effects. In some cases, there were several geocoded SAR images acquired in the same scenario (same site, same imaging geometry, and same acquisition occasion). Also, not all stands were covered by all images, and thus the number of available stands was different for different scenarios. In Table III, the number of stands and the number of geocoded SAR images available for each scenario are shown.

Henceforth, the different data sets will in some cases be referred to using shorter notation:

- **Site ID (SID)**: Re for Remningstorp and Kr for Krycklan,
- **Group ID (GID)**: INS for *in-situ* based stand-wise data, and LID for lidar based stand-wise data,
- **Date ID (DID)**: *Mar*, *Apr*, and *May* for the acquisitions in Remningstorp in 2007, and *Oct* for the acquisitions in Krycklan in 2008.

III. BIOMASS RETRIEVAL MODELS

In the following section, the models evaluated in this paper will be described. A motivation for the selection of the models introduced in this paper will be given. The basic geometry is shown in Fig. 3.

In this paper, the following convention will be used:

$$[X]_{\mathrm{dB}} = 10\log_{10}(X) \tag{1}$$

where X is a power ratio. Also

$$\widehat{\mathcal{W}}_{\mathrm{Mn}} = \log_{10}(\widehat{\mathcal{B}}_{\mathrm{Mn}}) \tag{2}$$

where $\widehat{\mathcal{B}}_{\mathrm{Mn}}$ is a biomass estimate from model Mn in tons/ha.

The scattering coefficient σ^0 is the averaged radar cross section per unit area [52]. It can be defined as

$$\sigma^0_{\mathrm{PQ}} = \frac{4\pi \left\langle |S_{\mathrm{PQ}}|^2 \right\rangle}{A} \tag{3}$$

where S_{PQ} is the scattering matrix element for polarization PQ and A is the area of a resolution cell. It is common to choose A to be the projection of a slant range resolution cell to the ground [53]

$$A = \frac{A_0}{\cos\psi_i} \tag{4}$$

where A_0 is the area of the slant range resolution cell, and $\cos\psi_i$ is a projection factor

$$\cos\psi_i = \hat{n} \cdot (\hat{x} \times \hat{k}_i) \tag{5}$$

where $\hat{n}$ is the ground surface normal unit vector, $\hat{x}$ is the unit vector pointing in the flight direction, and $\hat{k}_i$ is the unit vector pointing in the propagation direction, see Fig. 3.

For a rough, forested surface, the normalization to σ^0 is not sufficient due to a residual dependence on the angle of incidence (caused by different penetration depths). A better normalization called γ^0 is used

$$\gamma^0_{\mathrm{PQ}} = \frac{\sigma^0_{\mathrm{PQ}}}{\cos\theta_i} \tag{6}$$

where θ_i is the local incidence angle (see Fig. 3).

A. *Topographic and Temporal Effects*

In Fig. 4, scattering coefficients for HH, HV, and VV, and the ratio HH/VV are plotted against biomass for all data from Remningstorp and Krycklan. The x-axes are the same for all four plots. The y-axes have the same scale (spacing between grid lines), but the values are shifted for better viewing. Color coding refers to the acquisition time. Running average curves are also plotted in order to simplify trend investigation.

Looking at the three polarizations HH, HV, and VV in Fig. 4, the following observations can be made:

1) VV backscatter is poorly correlated with biomass in all cases,
2) HH backscatter shows much higher variability in Krycklan than in Remningstorp,
3) backscatter at all polarizations is typically several dB lower in Krycklan than in Remningstorp,
4) reduced sensitivity can be observed in Krycklan at all polarizations above approximately 100 tons/ha,
5) an average backscatter shift by around 0–2 dB can be seen from March to May in the Remningstorp data.

Following point 1), it can be concluded that, of all polarizations, VV is least sensitive to biomass, making it a potential indicator of other properties, such as topography, moisture conditions, and forest structure. The observation from point 2) can be explained by the influence of topography. Krycklan data feature higher slopes and better directional representation for each stand (acquisitions from multiple headings). The backscatter shift referred to in 3) may be explained by different forest structure and moisture change. Also, the problem described in 4) is most certainly an effect of topography (most of the high-biomass LID-stands in Krycklan are located in topographic terrain, see Fig. 6 and Section III-C). Finally, the backscatter shift in 5) is most likely due to moisture change. Radiometric calibration has been carefully evaluated using trihedral corner reflectors (see [44]), and the maximal measured variation is only 0.8 dB. It is thus concluded that the measured backscatter shift cannot be explained by a radiometric calibration error.

When trying to define a model suitable for both Remningstorp and Krycklan, the five points mentioned above need to be taken into consideration. It is apparent that biomass retrieval from one curve fitted to all (or parts of) the data may often give very poor results when applied on (parts of) the rest of the data.

One possible way to avoid the aforementioned problems is by finding a biomass indicator less susceptible to temporal and topographic variations. This can be partly achieved by using the ratio of HH- and VV-backscatter, the co-polar ratio. This observable has been plotted against biomass in the bottom plot to the right in Fig. 4. By creating the HH- to VV-backscatter ratio, common factors are eliminated. Biophysical forest parameters such as forest structure, ground surface roughness, and moisture will to some degree have similar impact on both HH and VV, and their contribution in biomass estimation can be decreased by the use of the HH/VV ratio. Whereas the temporal and site-to-site change has been reduced, the variability is still high. Therefore, instead of using the ratio on its own, it will be combined with HV backscatter, which has previously shown the most consistent correlation with biomass [38], at least in areas with modest topographic variations.

As mentioned, the influence of topography has been decreased by the inclusion of the HH/VV ratio, but not fully suppressed. A complementary way of improving the retrieval is by finding a way to compensate for topographic variations using explicit functions, derived either from experimental data, from models, or from both.

An additional important factor to be considered is the number of regression parameters. With too many regression parameters (too many predictors), the risk of overfitting increases, and the model may lack generality. Moreover, the demand on training data increases as more points are needed for stable fitting. On the other hand, with too few regression parameters, the chosen predictors may not be sufficient for accurate modeling. It is thus important to optimize the number of model parameters.

B. *Basic Model*

The first approach for a biomass retrieval model is based on a linear function of backscatter in three polarization channels (based on [11], [14]–[16])

$$\widehat{\mathcal{W}}_{\mathrm{M1}} = a_0 + a_1 \left[\gamma^0_{\mathrm{HV}}\right]_{\mathrm{dB}} + a_2 \left[\gamma^0_{\mathrm{HH}}\right]_{\mathrm{dB}} + a_3 \left[\gamma^0_{\mathrm{VV}}\right]_{\mathrm{dB}} \tag{M1}$$

where a_0 to a_3 are model parameters and γ^0_{PQ} is the normalized scattering coefficient gamma nought for polarization PQ. The model (M1) makes use of three observables, and thus four parameters need to be estimated. The results show that a_3 has very high uncertainty making γ^0_{VV} not suitable for retrieval

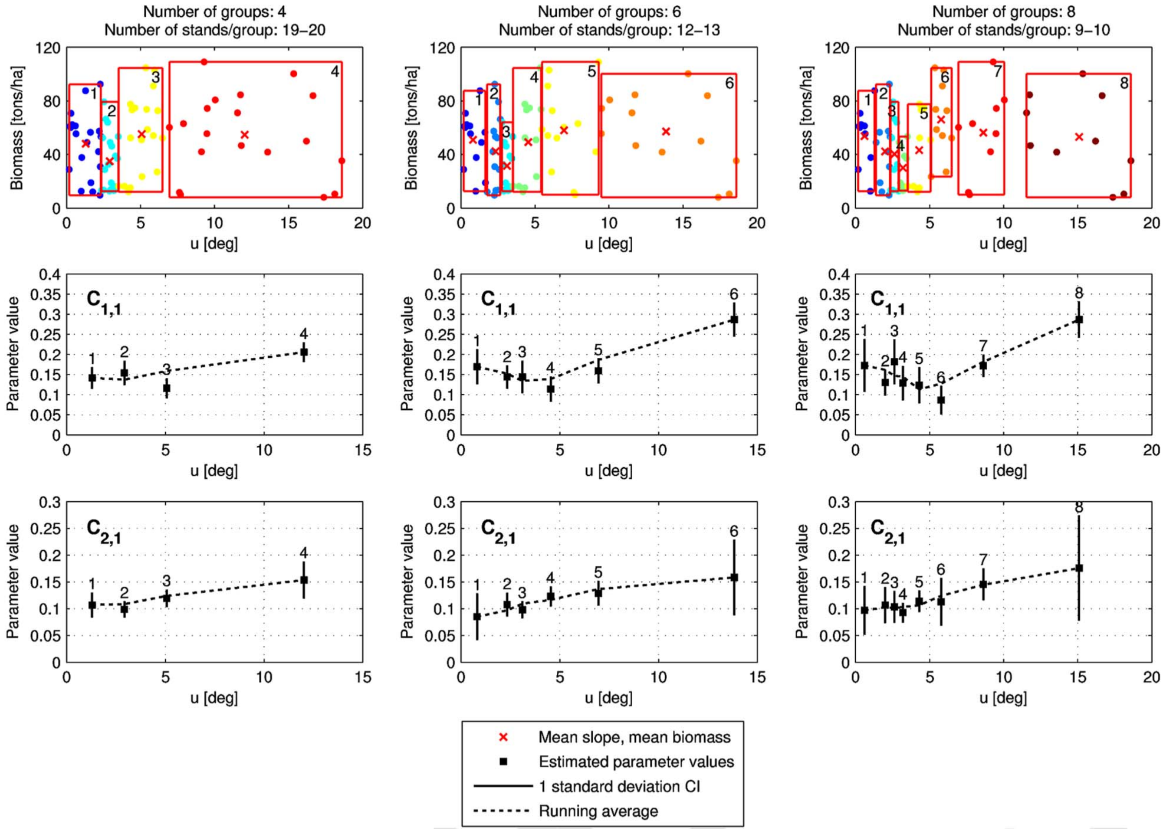


Fig. 5. Results of topographic investigation on LID-stands from Krycklan with upper biomass limit of 120 tons/ha. The topmost row of plots shows clustering of the data points in groups with similar u-angle for three grouping setups (four, six, and eight groups with similar number of data points). The groups are delineated with red bounding boxes showing the variability in u and biomass of each group. The red crosses represent the mean slope-mean biomass points for each group. Each group has a number appointed to it in the upper right corner of the corresponding bounding box. The second and third rows of plots show how the second parameter of the fitted models varies with u for two models. Running average curves are shown for easier trend investigation. One standard deviation confidence intervals for the estimated parameters are also shown.

(as already observed in Fig. 4). Furthermore, earlier studies indicate that a model based on both HH and HV may not be significantly better than one based on HV alone [15]. Thus, a simpler model using only one polarization will be evaluated (also used in [38])

$$\widehat{\mathcal{W}}_{\mathrm{M2}} = a_0 + a_1 \left[\gamma^0_{\mathrm{HV}}\right]_{\mathrm{dB}}. \tag{M2}$$

Following the observations about the co-polar ratio made in Fig. 4 and Section III-A, i.e., setting:

$$a_3 = -a_2$$

in (M1), a new model including the HH/VV ratio is constructed

$$\widehat{\mathcal{W}}_{\mathrm{M3}} = a_0 + a_1 \left[\gamma^0_{\mathrm{HV}}\right]_{\mathrm{dB}} + a_2 \left(\left[\gamma^0_{\mathrm{HH}}\right]_{\mathrm{dB}} - \left[\gamma^0_{\mathrm{VV}}\right]_{\mathrm{dB}}\right) \tag{M3}$$

which makes use of all three observables but only three parameters need to be estimated. A similar model was presented in [11].

C. New Model With Topographic Correction

Although the topographic correction introduced in [4] and [28] has shown good results at VHF-band, its functional form

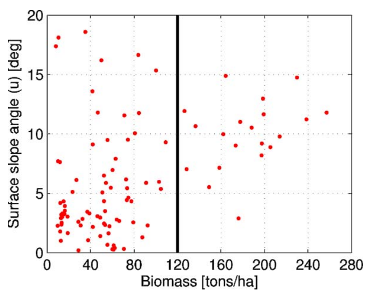


Fig. 6. Distribution of biomass and surface slope for all 97 LID-stands in Krycklan. Note that above approximately 120 tons/ha, most stands are located in sloping terrain. The black line indicates the upper biomass limit for the stands used during the parameter study described in Section III-C.

is too complicated for this work. Instead, a different approach is chosen. In order to find one single, most important topographic indicator, the following functions relating biomass to the two observables HV and HH/VV ratio were fitted to the experimental data:

$$\widehat{\mathcal{W}}_1 = C_{1,0} + C_{1,1} \left[\gamma^0_{\mathrm{HV}}\right]_{\mathrm{dB}} \tag{7}$$

$$\widehat{\mathcal{W}}_2 = C_{2,0} + C_{2,1} \left(\left[\gamma^0_{\mathrm{HH}}\right]_{\mathrm{dB}} - \left[\gamma^0_{\mathrm{VV}}\right]_{\mathrm{dB}}\right) \tag{8}$$

being the two main elements of (M3). $\widehat{\mathcal{W}}_1$ and $\widehat{\mathcal{W}}_2$ are related to biomass according to (2). The experimental data were divided into smaller groups with similar ground slope, and the fitting was done separately for each group. This way, each model parameter could be studied against the mean value of the topographic indicator for each group.

Four topographic indicators were considered in this study: the local incidence angle θ_i, the difference between local and nominal incidence angles $\theta_i - \theta_0$, the surface slope angle u, and the surface slope direction angle v. Although this study was done for all four indicators, the most conclusive results of this study, as well as the best biomass retrieval results, were achieved using u-based topographic correction. Therefore, only the results from that part of the study are presented.

In first row of plots in Fig. 5, the results from grouping by similar surface slope angle u are shown in three plots. The data points used here consisted of LID-stands from Krycklan with upper biomass limit set to 120 tons/ha. This limit was introduced to allow as uniform biomass-slope distribution as possible (see Fig. 6). The number of groups varies between 4 (to the left), 6 (in the middle), and 8 (to the right). Each group has approximately the same number of members. For each stand, the mean backscatter coefficient from four headings was used to reduce the variability due to different angles v.

In the second and third rows of plots in Fig. 5, the values of the second parameters $C_{1,1}$ and $C_{2,1}$ in (7) and (8) are plotted against u for three grouping setups. The constant parameters $C_{1,0}$ and $C_{2,0}$ depend not only on u, but also on other effects that cannot be predicted from the observables. They are thus not studied here. Whereas $C_{1,1}$ appears to be difficult to relate to u with a simple function, $C_{2,1}$ shows a more clear dependence on u. The first approximation of this dependence is a linear function, which suggests an additional term in (M3) consisting of the product of the surface slope u and the HH/VV ratio

$$\widehat{\mathcal{W}}_{\mathrm{M4}} = a_0 + a_1 \left[\gamma^0_{\mathrm{HV}}\right]_{\mathrm{dB}} + a_2 \left(\left[\gamma^0_{\mathrm{HH}}\right]_{\mathrm{dB}} - \left[\gamma^0_{\mathrm{VV}}\right]_{\mathrm{dB}}\right) + a_3 \cdot u \left(\left[\gamma^0_{\mathrm{HH}}\right]_{\mathrm{dB}} - \left[\gamma^0_{\mathrm{VV}}\right]_{\mathrm{dB}}\right). \quad \text{(M4)}$$

D. Reference Models

As reference, models presented in previous works by other researchers will be used. First, a single polarization model

$$\widehat{\mathcal{W}}_{\mathrm{R1}} = C_0 + C_1 \left(\left[\gamma^0_{\mathrm{HV}}\right]_{\mathrm{dB}} - b_0\right) \quad \text{(R1)}$$

with constants $C_0 = 3.8914$ and $C_1 = 0.1301$ as presented in [54]. The parameter b_0 is not explicitly included in [54], but is needed, and can be estimated from training data. Note, that (R1) is a simplified version of (M2) with constant slope ($a_1 = C_1$ and $a_0 = C_0 - C_1 \cdot b_0$).

Also, a seven-parameter model is used [14]

$$\widehat{\mathcal{W}}_{\mathrm{R2}} = a_0 + a_1 \left[\sigma^0_{\mathrm{HV}}\right]_{\mathrm{dB}} + a_2 \left[\sigma^0_{\mathrm{HV}}\right]^2_{\mathrm{dB}} + + a_3 \left[\sigma^0_{\mathrm{HH}}\right]_{\mathrm{dB}} + a_4 \left[\sigma^0_{\mathrm{HH}}\right]^2_{\mathrm{dB}} + a_5 \left[\sigma^0_{\mathrm{VV}}\right]_{\mathrm{dB}} + a_6 \left[\sigma^0_{\mathrm{VV}}\right]^2_{\mathrm{dB}}. \quad \text{(R2)}$$

In [14], a more advanced model including topographic corrections was also presented and proved suitable for biomass retrieval from P-band SAR data acquired with the AirSAR platform over the Yellowstone National Park. However, that model was not used in this study because a comparison with (R2) showed that the latter model was in fact more suitable for BioSAR data and also had fewer parameters (7 instead of 14). Note, that in (R2), σ^0 is used instead of γ^0.

IV. Model Validation and Discussion

In this section, the models presented in Section III will first be tested on data sets from Remningstorp to evaluate the influence of temporal change, mainly in terms of moisture conditions (Section IV-B). Thereafter, the models will be tested on data sets from Krycklan to evaluate the influence of topography (Section IV-C). In Section IV-D, the models will be evaluated across sites, i.e., models with parameters fitted to one test site will be used for biomass retrieval in the other test site. Next, in Section IV-E model errors will be studied against biomass for the three models that showed the best performance in the first three tests. Finally, in Section IV-F, biomass maps will be produced using the best model, and mapping errors will be pointed out and discussed.

Define the estimation error as

$$\widehat{R}(i) = \widehat{\mathcal{B}}(i) - \mathcal{B}_{\mathrm{ref}}(i) \quad (9)$$

where $\widehat{\mathcal{B}}(i)$ is the estimated biomass using SAR observation i, $\mathcal{B}_{\mathrm{ref}}(i)$ is the corresponding reference biomass. Note, that one single observation index i sweeps both through all stands *and* all acquisitions. The accuracy of the models will be evaluated using several quantitative measures.

- **Root-mean-square error** (RMSE) is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_i \widehat{R}(i)^2} \quad (10)$$

where N is the *total* number of observations.

- **Bias** is defined as the mean of the estimation error

$$\mathrm{bias} = \frac{1}{N}\sum_i \widehat{R}(i). \quad (11)$$

With this notation, positive bias means overestimation, and negative bias means underestimation.

- **Standard deviation of the estimation error** can be computed from (10) and (11) as

$$\text{standard deviation} = \sqrt{(\mathrm{RMSE})^2 - (\mathrm{bias})^2}. \quad (12)$$

- **Coefficient of determination** R^2 is a measure of how well a linear model fits the data in comparison with a simple average [55]. It is computed according to

$$R^2 = 1 - \frac{\sum_i \left(\mathcal{B}_{\mathrm{ref}}(i) - \widehat{\mathcal{B}}(i)\right)^2}{\sum_i \left(\mathcal{B}_{\mathrm{ref}}(i) - \overline{\mathcal{B}}_{\mathrm{ref}}\right)^2} \quad (13)$$

where

$$\overline{\mathcal{B}}_{\text{ref}} = \frac{1}{N}\sum_i \mathcal{B}_{\text{ref}}(i)$$

is the mean reference biomass. For accurate modeling, R^2 should be as close to one as possible. Values below zero indicate that better modeling results would be achieved with an average of the reference data.

- **Relative error** is defined as

$$\text{relative error} = 100\% \cdot \frac{\widehat{\mathcal{B}} - \mathcal{B}_{\text{ref}}}{\mathcal{B}_{\text{ref}}}. \tag{14}$$

A. *Data Selection and Model Training*

Since the model performance depends on the reliability of model parameter estimation (model training), the choice of the data used for training demands care.

First, the training data need to cover a large parameter range and have a reasonable accuracy. Lidar-based measurements present a good compromise between accuracy and coverage. Therefore, LID-stands presented in Table II will be used as training data.

The number of SAR measurements is not equal for all stands, and not all stands are always covered (see Table III). Also, in some cases, more than one geocoded SAR image is available for each scenario (same site, same imaging geometry, same acquisition date). A bias problem may thus occur. To minimize that problem, only one measurement per stand from each site, each date, and each heading was chosen to be used, and only the LID-stands covered by all images were used for training.

In Remningstorp, two geocoded images with zero nominal baseline were available for each acquisition date at the 200-degree heading. Since no systematic differences could be observed in the stand-wise data between the two acquisitions, the second acquisition specified in Table III was arbitrarily chosen. In case of the two headings 134° and 314° in Krycklan, for which multiple images were available, the choice was made to maximize the number of covered stands. The following images were therefore used:

- **Remningstorp**:
  — heading 179°: one image for each date (0110, 0206, and 0412),
  — heading 200°: one image for each date (0109, 0306, and 0411).
- **Krycklan**:
  — heading 43°: one image (0304),
  — heading 134°: one image (0104),
  — heading 314°: one image (0103),
  — heading 358°: one image (0301).

The underlined numbers in parentheses refer to the identification numbers of each image, as shown in Table III and in [44], [45].

In total, Remningstorp data suitable for training were limited to a maximum of 46 LID-stands (out of 58) and six acquisitions for each stand (out of nine, see Table III). For Krycklan, data suitable for training were limited to a maximum of 97 LID-stands and four acquisitions for each stand (out of seven, see Table III). Note, that in many cases, smaller subsets of these data sets were used for training. In cases when more than one acquisition per stand was used, different observations in the training data set were not entirely independent of each other, which might cause problems in the statistical analysis. In Section IV-E, this issue is pointed out and discussed.

Since all the models used in this text are linear, least-squares as implemented in Matlab function `regress` was used for parameter estimation.

For best quantitative validation, high-accuracy INS-stands were used. For temporal validation in Remningstorp, the same restrictions as for training data applied to validation data in order to be able to make fair comparison between headings. For the other validation scenarios, all available SAR acquisitions for each stand were used for biomass prediction, giving up to nine biomass values for some stands in Remningstorp and up to seven biomass values for some stands in Krycklan. This approach increases the influence of the well-represented stands during validation.

B. *Temporal Validation*

In this part, the models were trained using LID-stands in Remningstorp and validated using INS-stands from the same test site. Only the stands fully covered by both 179- and 200-degree acquisitions were used. Each combination of dates was examined, as well as the results of training and validation on all three dates. RMSE are presented in Table IV in tons/ha together with the coefficients of determination R^2. The mean biomass for validation data is 181 tons/ha. In this comparison, model (M4) was not included since topography is not significant in Remningstorp.

Looking at same date retrieval (training and validation on the same date), all models show reasonable performance with RMSE ranging between 35 and 60 tons/ha (19–33% of mean biomass). However, as the retrieval scenario becomes more difficult, and the training and validation dates are further apart, the single polarization models (R1) and (M2) often show significantly higher errors compared to models including all polarizations.

Comparing the two headings (and keeping in mind that the 179-degree heading features steeper incidence angles), it can be observed that for models (R1) and (M2), the retrieval is more stable across dates for the 179-degree heading (however, it gives in general worse results). Moreover, the data set used for training seems to affect the results much more for the 179-degree heading than for the 200-degree heading, for which only the temporal distance between training and validation data seems of an importance (the error is lowest on the diagonal and higher off-diagonal). This is clearly visible for models (M1) and (M3) at the 179-degree heading, where training on May data gives RMSE around 40 tons/ha, no matter which date is used for validation. For training on April data, the same values lie over 60 tons/ha.

Also, when trained and validated using all temporal acquisitions, full polarization models (R2), (M1), and (M3) show

TABLE IV

RESULTS OF **TEMPORAL VALIDATION** OF MODELS 1–5 IN TERMS OF RMSE (tons/ha, FIRST ROW) AND R^2 (SECOND ROW). COLOR CODING BY RMSE RELATIVE MEAN BIOMASS (181 tons/ha): WHITE FOR 20% AND BELOW, BLACK FOR 100% AND ABOVE

Heading: 179°

Model:	R1			R2			M1			M2			M3		
TRAINING	Mar	Apr	May	Mar	Apr	May	Mar	Apr	May	Mar	Apr	May	Mar	Apr	May
VALIDATION Mar	**50**	**61**	**71**	**42**	**46**	**82**	**55**	**84**	**41**	**55**	**89**	**85**	**51**	**78**	**38**
	0.43	**0.14**	**-0.16**	**0.6**	**0.53**	**-0.53**	**0.32**	**-0.59**	**0.61**	**0.31**	**-0.81**	**-0.67**	**0.4**	**-0.4**	**0.66**
VALIDATION Apr	**65**	**53**	**52**	**41**	**39**	**71**	**49**	**65**	**41**	**64**	**59**	**56**	**47**	**59**	**39**
	0.04	**0.37**	**0.39**	**0.62**	**0.65**	**-0.16**	**0.44**	**0.05**	**0.62**	**0.07**	**0.21**	**0.3**	**0.5**	**0.21**	**0.66**
VALIDATION May	**75**	**58**	**53**	**32**	**31**	**48**	**43**	**60**	**37**	**73**	**57**	**53**	**43**	**60**	**37**
	-0.27	**0.23**	**0.35**	**0.76**	**0.78**	**0.47**	**0.57**	**0.17**	**0.68**	**-0.22**	**0.26**	**0.35**	**0.58**	**0.17**	**0.69**
TRAINING	All			All			All			All			All		
All	**58**			**44**			**50**			**60**			**46**		
	0.24			**0.55**			**0.43**			**0.19**			**0.51**		

Heading: 200°

Model:	R1			R2			M1			M2			M3		
TRAINING	Mar	Apr	May	Mar	Apr	May	Mar	Apr	May	Mar	Apr	May	Mar	Apr	May
VALIDATION Mar	**35**	**49**	**69**	**45**	**54**	**63**	**45**	**55**	**73**	**42**	**90**	**151**	**38**	**53**	**68**
	0.72	**0.45**	**-0.09**	**0.54**	**0.32**	**0.1**	**0.55**	**0.31**	**-0.22**	**0.59**	**-0.85**	**-4.2**	**0.67**	**0.36**	**-0.06**
VALIDATION Apr	**54**	**40**	**42**	**39**	**45**	**50**	**42**	**42**	**54**	**50**	**51**	**85**	**39**	**41**	**51**
	0.34	**0.64**	**0.59**	**0.64**	**0.54**	**0.42**	**0.6**	**0.59**	**0.33**	**0.44**	**0.41**	**-0.63**	**0.65**	**0.61**	**0.41**
VALIDATION May	**75**	**55**	**46**	**43**	**40**	**42**	**51**	**40**	**45**	**71**	**51**	**53**	**49**	**40**	**43**
	-0.27	**0.3**	**0.51**	**0.58**	**0.64**	**0.61**	**0.41**	**0.63**	**0.55**	**-0.16**	**0.41**	**0.35**	**0.46**	**0.64**	**0.57**
TRAINING	All			All			All			All			All		
All	**49**			**45**			**41**			**55**			**39**		
	0.46			**0.53**			**0.61**			**0.3**			**0.65**		

Color coding by RMSE_{INS} [tons/ha]: ≤ 36 $\longleftrightarrow$ ≥ 181

TABLE V

RESULTS OF **TOPOGRAPHIC VALIDATION** OF MODELS 1–6 IN TERMS OF RMSE (tons/ha) AND R^2 (IN PARENTHESES). COLOR CODING BY RMSE RELATIVE MEAN BIOMASS (95 tons/ha): WHITE FOR 20% AND BELOW, BLACK FOR 100% AND ABOVE

Model:	R1	R2	M1	M2	M3	M4
TRAINING HEADING: 43°						
VALID.: 43°	**36 (0.37)**	**30 (0.56)**	**30 (0.58)**	**34 (0.45)**	**31 (0.55)**	**33 (0.48)**
134°	**35 (0.31)**	**31 (0.48)**	**29 (0.53)**	**37 (0.23)**	**29 (0.52)**	**33 (0.39)**
314°	**35 (0.26)**	**36 (0.22)**	**31 (0.42)**	**37 (0.19)**	**34 (0.34)**	**37 (0.2)**
358°	**40 (0.3)**	**32 (0.55)**	**32 (0.56)**	**40 (0.32)**	**38 (0.37)**	**39 (0.36)**
All	**36 (0.3)**	**33 (0.41)**	**30 (0.5)**	**37 (0.25)**	**32 (0.45)**	**35 (0.34)**
TRAINING HEADING: 134°						
VALID.: 43°	**43 (0.1)**	**37 (0.33)**	**30 (0.58)**	**43 (0.11)**	**30 (0.57)**	**30 (0.56)**
134°	**38 (0.18)**	**31 (0.47)**	**29 (0.52)**	**38 (0.18)**	**30 (0.51)**	**30 (0.48)**
314°	**40 (0.07)**	**35 (0.26)**	**34 (0.33)**	**39 (0.08)**	**35 (0.26)**	**34 (0.3)**
358°	**45 (0.11)**	**42 (0.25)**	**38 (0.36)**	**45 (0.11)**	**40 (0.3)**	**39 (0.35)**
All	**40 (0.13)**	**34 (0.36)**	**32 (0.45)**	**40 (0.14)**	**33 (0.41)**	**33 (0.42)**
TRAINING HEADING: 314°						
VALID.: 43°	**37 (0.35)**	**31 (0.55)**	**27 (0.67)**	**34 (0.46)**	**26 (0.67)**	**27 (0.64)**
134°	**35 (0.31)**	**35 (0.31)**	**31 (0.47)**	**44 (-0.08)**	**37 (0.25)**	**28 (0.57)**
314°	**36 (0.25)**	**29 (0.52)**	**28 (0.55)**	**44 (-0.14)**	**42 (-0.03)**	**30 (0.47)**
358°	**41 (0.29)**	**31 (0.59)**	**31 (0.57)**	**42 (0.25)**	**42 (0.24)**	**40 (0.31)**
All	**36 (0.29)**	**32 (0.44)**	**29 (0.53)**	**43 (0)**	**38 (0.2)**	**30 (0.51)**
TRAINING HEADING: 358°						
VALID.: 43°	**37 (0.35)**	**32 (0.52)**	**31 (0.53)**	**35 (0.43)**	**33 (0.47)**	**34 (0.43)**
134°	**35 (0.31)**	**32 (0.42)**	**31 (0.47)**	**37 (0.23)**	**32 (0.44)**	**33 (0.39)**
314°	**36 (0.25)**	**38 (0.16)**	**32 (0.38)**	**37 (0.18)**	**34 (0.31)**	**36 (0.22)**
358°	**41 (0.29)**	**33 (0.53)**	**32 (0.55)**	**40 (0.32)**	**39 (0.36)**	**39 (0.34)**
All	**36 (0.29)**	**34 (0.36)**	**32 (0.46)**	**37 (0.25)**	**33 (0.39)**	**35 (0.34)**
TRAINING HEADING: All						
VALID.: 43°	**38 (0.3)**	**31 (0.53)**	**30 (0.56)**	**36 (0.37)**	**31 (0.54)**	**32 (0.53)**
134°	**36 (0.29)**	**30 (0.51)**	**29 (0.52)**	**36 (0.27)**	**30 (0.5)**	**31 (0.47)**
314°	**36 (0.23)**	**35 (0.28)**	**32 (0.4)**	**36 (0.22)**	**34 (0.33)**	**34 (0.31)**
358°	**42 (0.25)**	**33 (0.52)**	**33 (0.53)**	**40 (0.29)**	**38 (0.37)**	**38 (0.38)**
All	**37 (0.27)**	**32 (0.44)**	**31 (0.49)**	**37 (0.27)**	**32 (0.44)**	**33 (0.42)**

Color coding by RMSE_{INS} [tons/ha]: ≤ 19 $\longleftrightarrow$ ≥ 95

TABLE VI

RESULTS OF **FLAT-TO-TOPOGRAPHIC VALIDATION** OF MODELS 1–6 IN TERMS OF RMSE (tons/ha) AND R^2 (IN PARENTHESES). COLOR CODING BY RMSE RELATIVE MEAN BIOMASS (95 tons/ha): WHITE FOR 20% AND BELOW, BLACK FOR 100% AND ABOVE

	Model:	R1	R2	M1	M2	M3	M4
	TRAINING DATA: March, both headings						
VALID.:	43°	68 (-1.19)	75 (-1.69)	62 (-0.81)	71 (-1.41)	60 (-0.69)	48 (-0.12)
	134°	63 (-1.21)	63 (-1.2)	57 (-0.79)	66 (-1.45)	55 (-0.68)	48 (-0.29)
	314°	66 (-1.55)	68 (-1.76)	58 (-0.99)	69 (-1.81)	57 (-0.91)	51 (-0.54)
	358°	69 (-1.05)	74 (-1.39)	61 (-0.63)	72 (-1.24)	60 (-0.53)	51 (-0.1)
	All	65 (-1.3)	67 (-1.45)	58 (-0.83)	68 (-1.53)	56 (-0.74)	49 (-0.33)
	TRAINING DATA: April, both headings						
VALID.:	43°	62 (-0.85)	73 (-1.51)	53 (-0.36)	67 (-1.15)	51 (-0.26)	41 (0.21)
	134°	57 (-0.8)	59 (-0.92)	48 (-0.29)	61 (-1.11)	46 (-0.21)	40 (0.11)
	314°	59 (-1.07)	64 (-1.42)	50 (-0.46)	64 (-1.43)	48 (-0.38)	43 (-0.1)
	358°	63 (-0.73)	72 (-1.23)	54 (-0.24)	68 (-1)	52 (-0.16)	45 (0.13)
	All	59 (-0.88)	63 (-1.18)	50 (-0.35)	64 (-1.2)	48 (-0.26)	42 (0.06)
	TRAINING DATA: May, both headings						
VALID.:	43°	59 (-0.68)	70 (-1.35)	50 (-0.18)	64 (-0.94)	46 (-0.02)	35 (0.43)
	134°	54 (-0.6)	55 (-0.67)	45 (-0.11)	58 (-0.85)	42 (0.02)	35 (0.32)
	314°	56 (-0.85)	61 (-1.2)	46 (-0.26)	60 (-1.14)	44 (-0.14)	38 (0.12)
	358°	60 (-0.57)	70 (-1.1)	51 (-0.1)	64 (-0.8)	48 (0.02)	42 (0.23)
	All	56 (-0.68)	60 (-0.96)	46 (-0.16)	60 (-0.94)	44 (-0.04)	37 (0.26)
	TRAINING DATA: All dates, both headings						
VALID.:	43°	63 (-0.91)	70 (-1.36)	53 (-0.33)	66 (-1.05)	50 (-0.2)	38 (0.3)
	134°	58 (-0.87)	57 (-0.8)	48 (-0.28)	60 (-1.02)	46 (-0.16)	38 (0.18)
	314°	60 (-1.16)	63 (-1.31)	49 (-0.43)	63 (-1.32)	47 (-0.33)	42 (-0.03)
	358°	64 (-0.79)	70 (-1.11)	53 (-0.22)	66 (-0.91)	51 (-0.11)	44 (0.18)
	All	60 (-0.95)	61 (-1.06)	49 (-0.32)	62 (-1.1)	47 (-0.21)	40 (0.12)

Color coding by RMSE_{INS} [tons/ha]: ≤ 19 $\longleftrightarrow$ ≥ 95

better results, particularly for the 200-degree heading with retrieval error as low as 39 tons/ha (22%).

It can be observed here that (R1) often performs better than (M2) in spite of the fact that it has one parameter instead of two, but otherwise the same structure. This is an indication of possible overfitting with (M2). Model (R1) was in fact developed from (M2) using data sets from several different test sites and campaigns (both tropical and boreal). As the estimates of the slope parameter in (M2) were found consistent for these data sets, the slope could be set to a constant. Almost all performance analysis in this paper is done using independent training and validation sets, which helps to detect overfitting.

C. Topographic Validation

In this part, the models were trained and validated using different heading combinations in Krycklan. The RMSE and R^2 are shown for all training-validation combinations in Table V. The mean biomass level for Krycklan INS-stands is 95 tons/ha. The models which include all three polarizations, (R2), (M1), (M3), and (M4), show slightly better performance than the two single polarized models (R1) and (M2), but the improvement is small. Perhaps surprisingly, the correction in (M4) does not improve the retrieval results in this case because the variability in backscatter from one stand is not reduced by the model (since only the slope angle u is included in the model and this angle is constant for all acquisition geometries).

In general, all models give errors higher than approximately 27% (26 tons/ha). Validation results are more conclusive for the two main headings (134° and 314°) because the number of validation points is 27 and 28, compared to 9 and 10 for the other two headings. Also, the distribution of slopes for different biomass levels is nonuniform in the training data. The high-biomass stands are situated in sloping terrain, see Fig. 6.

D. Across-Site Validation

The across-site test was done in two steps: training in Remningstorp and validation in Krycklan, and vice versa. These two tests will be evaluated separately.

1) Flat-to-Topographic: A problem occurs when the models are trained using Remningstorp data and validated using Krycklan data: Remningstorp data do not include enough topographic variations for reliable training; the retrieval models perform poorly if only Remningstorp data are used, see Table VI. Retrieval errors are at minimum 37% (35 tons/ha), but the variability of the data is large, and the coefficient of determination is low. In terms of RMSE, model (M4) performs best here. However, R^2-values are low.

In Fig. 7, scatter plots showing estimation results for all six models are shown. Acquisitions from all three dates and both headings in Remningstorp were used for training (model parameters as in Table VII). Retrieval results for all Krycklan data are shown in the plots, in red for LID-stands and in black for INS-stands. For all models except (M4), biomass in Krycklan is underestimated. For (M4), the variability in data is larger compared to the rest of the models, but the bias is reduced.

2) Topographic-to-Flat: Here, LID data from the topographic area of Krycklan, featuring a variety of stands in

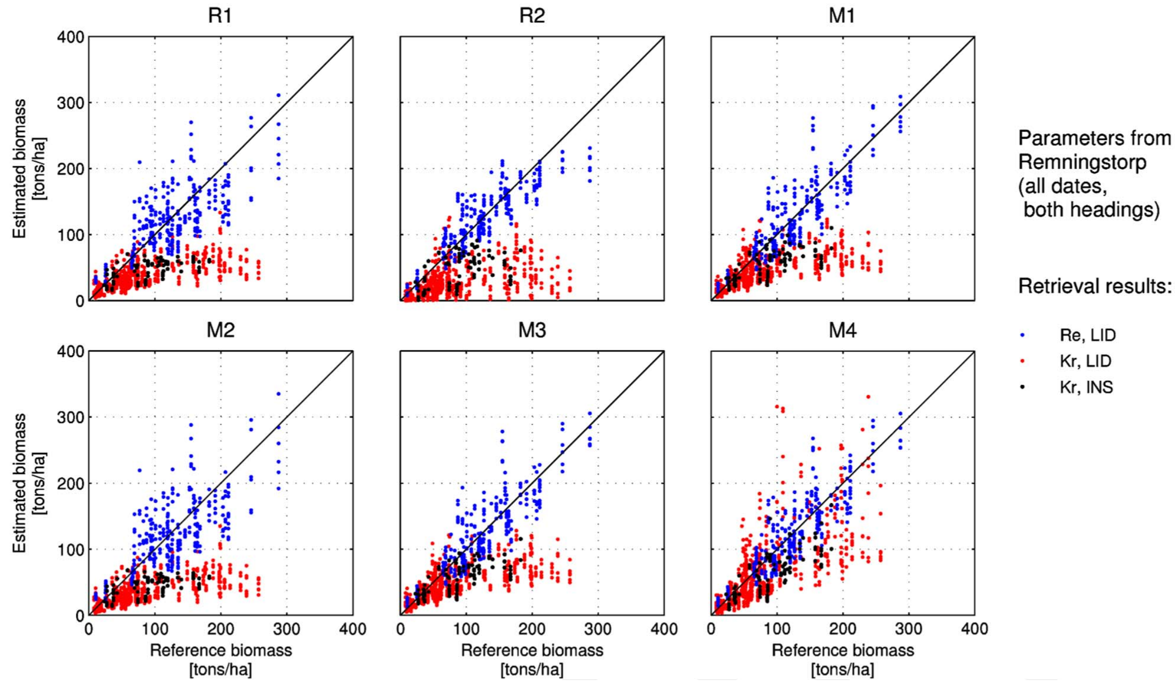


Fig. 7. Comparison of the six evaluated models: training on Remningstorp data.

TABLE VII
PARAMETER VALUES FOR THE SIX TESTED MODELS. TRAINING DATA CONSISTS OF ALL AVAILABLE LID-STANDS IN RESPECTIVE TEST SITE COVERED BY EXACTLY ONE IMAGE FROM EACH HEADING AND EACH DATE, SEE SECTION IV-A. THE PARAMETERS WRITTEN IN ITALICS WERE FOUND VERY UNCERTAIN (THEIR UNCERTAINTY INTERVALS INCLUDE ZERO)

Model:	Parameters estimated in Remningstorp	
(R1)	$b_0 = 2.827 \pm 0.138$	
(R2)	$a_0 = 0.842 \pm 0.650$	$a_1 = 0.065 \pm 0.022$
	$a_2 = -0.206 \pm 0.111$	$a_3 = -0.122 \pm 0.094$
	$\mathit{a_4 = -0.001 \pm 0.002}$	$a_5 = -0.010 \pm 0.004$
	$\mathit{a_6 = -0.006 \pm 0.007}$	
(M1)	$a_0 = 2.886 \pm 0.146$	$a_1 = 0.078 \pm 0.016$
	$a_2 = 0.072 \pm 0.010$	$a_3 = -0.056 \pm 0.015$
(M2)	$a_0 = 3.632 \pm 0.136$	$a_1 = 0.140 \pm 0.012$
(M3)	$a_0 = 2.933 \pm 0.138$	$a_1 = 0.089 \pm 0.011$
	$a_2 = 0.068 \pm 0.009$	
(M4)	$a_0 = 2.967 \pm 0.137$	$a_1 = 0.093 \pm 0.011$
	$a_2 = 0.056 \pm 0.011$	$a_3 = 0.713 \pm 0.411$

Model:	Parameters estimated in Krycklan	
(R1)	$b_0 = 0.766 \pm 0.190$	
(R2)	$a_0 = 2.507 \pm 1.246$	$\mathit{a_1 = 0.029 \pm 0.059}$
	$\mathit{a_2 = 0.061 \pm 0.144}$	$\mathit{a_3 = -0.105 \pm 0.115}$
	$\mathit{a_4 = -0.001 \pm 0.003}$	$\mathit{a_5 = -0.002 \pm 0.004}$
	$\mathit{a_6 = 0.001 \pm 0.005}$	
(M1)	$a_0 = 3.280 \pm 0.203$	$a_1 = 0.138 \pm 0.014$
	$a_2 = 0.049 \pm 0.012$	$a_3 = -0.113 \pm 0.016$
(M2)	$a_0 = 4.087 \pm 0.191$	$a_1 = 0.149 \pm 0.012$
(M3)	$a_0 = 3.402 \pm 0.222$	$a_1 = 0.109 \pm 0.013$
	$a_2 = 0.063 \pm 0.013$	
(M4)	$a_0 = 3.129 \pm 0.211$	$a_1 = 0.093 \pm 0.013$
	$a_2 = 0.020 \pm 0.015$	$a_3 = 0.605 \pm 0.134$

different slope conditions, were used for training of the models. In Table VIII, the resulting RMSE values are shown together with the coefficient of determination R^2. The mean biomass for Remningstorp INS-stands is 181 tons/ha. It can be observed that retrieval errors as low as 22% (40 tons/ha) can be achieved with (M4). Single-polarization models (R1) and (M2), and model (M3) show all extremely high errors going above 100% of mean biomass level. This validation scenario shows clearly the advantage of models (R2), (M1), and (M4). For (M4), the R^2-values are also high (see Table VIII).

In Fig. 8, scatter plots showing estimation results for all six models are shown. Acquisitions from all four headings in Krycklan were used for training (model parameters as in Table VII). Retrieval results for all Remningstorp data are shown in the plots, in blue for LID-stands and in black for INS-stands. For all models except (M4) and (R2), biomass in Remningstorp is overestimated. For (M4), the variability in data is larger compared to (R2), but the bias (underestimation) observed above 200 tons/ha is reduced.

E. Error Analysis

Looking at the results presented in the previous three sections, it can be observed that models (R2), (M1), and (M4) show best overall performance of the six studied models. Models (M1) and (M4) have the advantage of having less parameters and showing better results in flat-to-topographic retrieval. Although (R2) gives less variability (improved precision) in the higher biomass levels, a loss of sensitivity (reduced accuracy, higher bias) can be observed for biomass values above 200 tons/ha. Whereas the precision of a model can be improved using spatial averaging, it is difficult to improve the accuracy. Therefore, a limited increase in variability is an acceptable tradeoff for lower bias.

TABLE VIII
RESULTS OF **TOPOGRAPHIC-TO-FLAT VALIDATION** OF MODELS 1–6 IN TERMS OF RMSE (tons/ha) AND R^2 (IN PARENTHESES). COLOR CODING BY RMSE RELATIVE MEAN BIOMASS (181 tons/ha): WHITE FOR 20% AND BELOW, BLACK FOR 100% AND ABOVE

	Model:	R1	R2	M1	M2	M3	M4
		TRAINING DATA: All headings					
VAL. 179°:	Mar	>200 (-12.25)	56 (0.29)	100 (-1.28)	>200 (-36.78)	>200 (-12.37)	59 (0.22)
	Apr	157 (-4.61)	59 (0.2)	64 (0.05)	>200 (-15.8)	180 (-6.37)	41 (0.61)
	May	123 (-2.48)	48 (0.48)	86 (-0.7)	>200 (-9.66)	175 (-5.98)	41 (0.62)
	All	181 (-6.45)	55 (0.32)	85 (-0.64)	>200 (-20.75)	>200 (-8.24)	48 (0.48)
		TRAINING DATA: All headings					
VAL. 200°:	Mar	187 (-6.78)	74 (-0.22)	68 (-0.03)	>200 (-21.2)	198 (-7.77)	46 (0.54)
	Apr	137 (-3.2)	71 (-0.11)	58 (0.25)	>200 (-11.6)	178 (-6.06)	40 (0.64)
	May	88 (-0.73)	62 (0.13)	58 (0.26)	157 (-4.48)	147 (-3.78)	41 (0.63)
	All	143 (-3.57)	69 (-0.07)	61 (0.16)	>200 (-12.43)	176 (-5.87)	42 (0.6)
		TRAINING DATA: All headings					
VAL. both:	Mar	>200 (-8.42)	69 (-0.06)	79 (-0.41)	>200 (-25.88)	>200 (-9.15)	50 (0.44)
	Apr	144 (-3.62)	67 (-0.01)	60 (0.19)	>200 (-12.84)	179 (-6.13)	41 (0.63)
	May	100 (-1.26)	58 (0.24)	68 (-0.03)	177 (-6.04)	156 (-4.44)	41 (0.63)
	All	156 (-4.43)	65 (0.05)	70 (-0.08)	>200 (-14.92)	184 (-6.57)	44 (0.57)

Color coding by RMSE_{INS} [tons/ha]: ≤ 36 $\longleftrightarrow$ ≥ 181

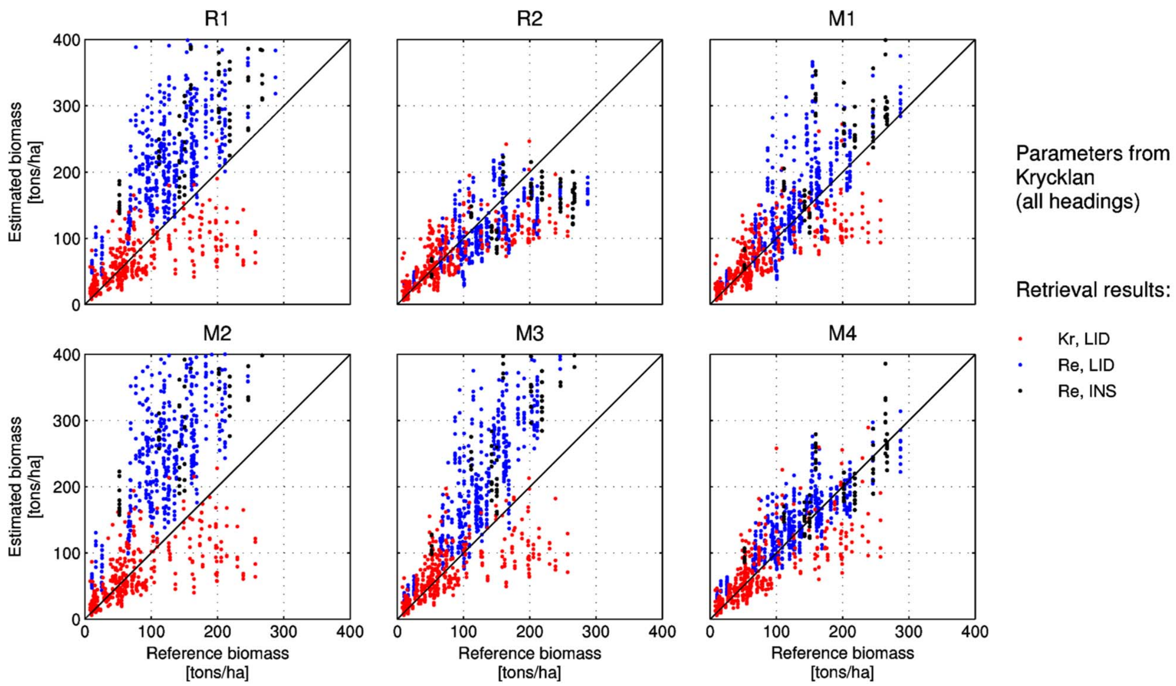


Fig. 8. Comparison of the six evaluated models: Training on Krycklan data.

As mentioned in Section IV-A, all observations used for training are not completely independent, since several observations from the same stand but with different imaging geometry and/or acquisition date are used simultaneously for parameter estimation. This breach of independence can be observed in Fig. 4 as clustering of observations from the same stand. This might induce slightly different parameter estimates compared to the estimates, which would be obtained if the full dependence structure of the observations was known. However, since the majority of pairs of observations are independent, these differences are likely to be small. Moreover, small differences in parameter estimates compared to "true" parameter values are not of concern in this study. The main focus of this paper is not the parameter estimation, but rather the performance analysis and the comparison of different models. The only real concern is the estimation of confidence intervals, which will be affected by the presence of unknown correlation between observations.

With the above discussion in mind, some conclusions can nevertheless be drawn from Table VII containing the estimated regression parameters. In particular, some of the coefficients for (R2) are not significantly different from zero (their confidence intervals include zero). This indicates that the model contains too many predictors. Note also, that the parameters of model (M4) are similar for both Remningstorp and Krycklan. This is

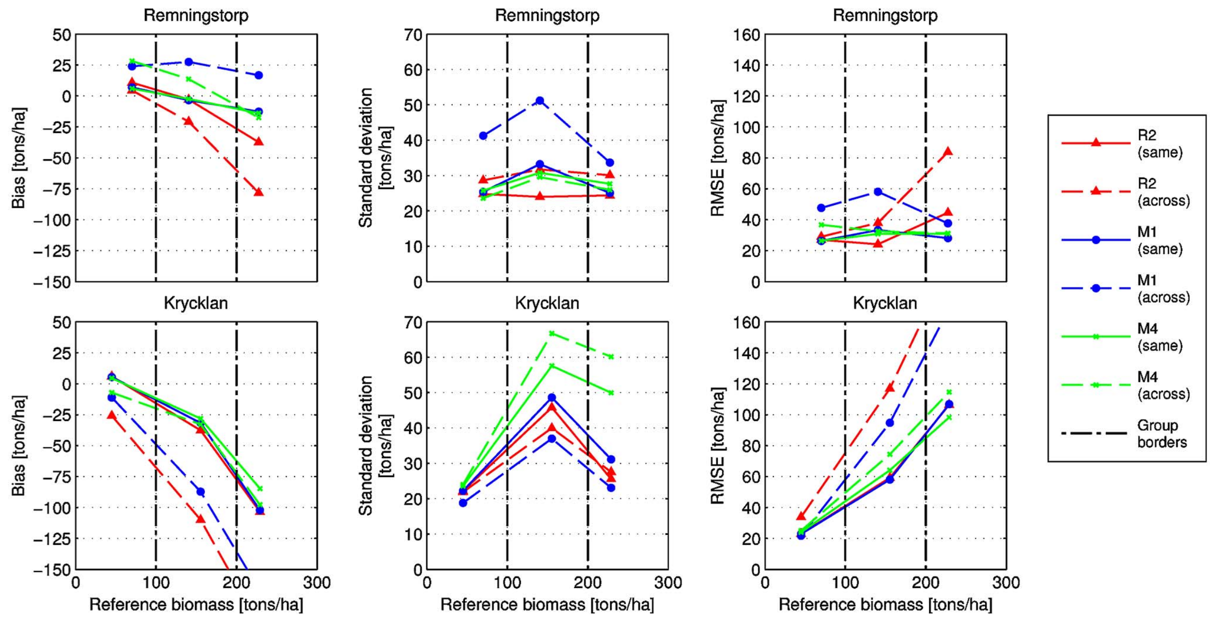


Fig. 9. Three different types of model errors are here plotted versus biomass: bias to the left, standard deviation of the estimation error in the middle, and RMSE to the right, as defined in (10)–(12). Only models (R2), (M1), and (M4) are compared. Model parameters as in Table VII were used. “Same” means that the model parameters estimated for the same site were used. “Across” means that the model parameters estimated for the other site were used. LID-stands were used, and averaging was done in three intervals: 0–100 tons/ha, 100–200 tons/ha, and 200–300 tons/ha. The interval borders are plotted in blue dashed lines. The number of data points in each group is at least 50. Note: some lines may cover each other.

an indication that the coefficients of this model are stable over a broad range of forest conditions.

In Fig. 9, bias (mean of the estimation error), standard deviation of the estimation error, and RMSE are plotted against biomass for models (R2), (M1), and (M4). These quantities have been defined in (10)–(12). For this study, the model parameters were those specified in Table VII. Statistics were computed for LID-stands in both Remningstorp and Krycklan, and the averaging was done in three intervals: low biomass (0–100 tons/ha), medium biomass (100–200 tons/ha), and high biomass (200–300 tons/ha).

It can be observed that all three models perform almost equally well when both trained and evaluated in Remningstorp (solid lines in the top three plots in Fig. 9). Model (R2) shows higher bias in the high-biomass group (underestimation with approximately 40 tons/ha), but the variability is quite small (standard deviation up to 30 tons/ha). When training and validation are both done in Krycklan (solid lines in the bottom three plots in Fig. 9), one can observe a strong underestimation occurring for stands with biomass above 100 tons/ha and a high variability. The origin of this bias can probably be related to the nonuniform biomass-slope distribution mentioned earlier and shown in Fig. 6, but a clear conclusion is difficult to be made as the number of independent data points is low. Also, the fact that none of the models compensates for variability with angle v contributes to the observed large variability. All three models perform similarly.

It is during across-site validation that (M4) proves itself better than the other two models. Lower bias is observed when training on Remningstorp and applying to Krycklan (dashed lines in the bottom three plots in Fig. 9). In the opposite case, (R2) shows lower bias for low-biomass stands, but higher in the two other groups (dashed lines in the top three plots in Fig. 9). Although (M4) shows in some cases slightly higher standard deviation of residuals, this effect can be reduced by spatial averaging. Bias is more difficult to reduce and should thus be kept as low as possible. Altogether, (M4) is observed as the best of the six models examined in this paper. Note, that in Krycklan, there is a lack of stands with high biomass and low slopes, whereas in Remningstorp these types of stands are common. An extrapolation is made for such stands when the model (M4) is trained in Krycklan and evaluated in Remningstorp. The exact influence of this effect on the retrieval is unclear.

F. Biomass Mapping Performance Analysis

In order to evaluate mapping performance of the new model, biomass maps were created from SAR images using (M4). In Fig. 10, a set of biomass maps is shown. To the left, biomass maps based on lidar scanning are shown. In the middle and to the right, two biomass maps extracted from SAR using (M4) are shown. For both Remningstorp and Krycklan, the same SAR images as used for training were used (those described in Section IV-A, six images for Remningstorp and four images for Krycklan). Geocoded images with pixel size 2 m $\times$ 2 m were first filtered using an average filter with a 5 $\times$ 5 window to match the resolution of the lidar-based biomass maps. Next, the filtered SAR images were re-sampled using linear interpolation to the same grid as the lidar-based biomass maps (10 m $\times$ 10 m). Thereafter, all biomass maps were filtered with a 7 $\times$ 7

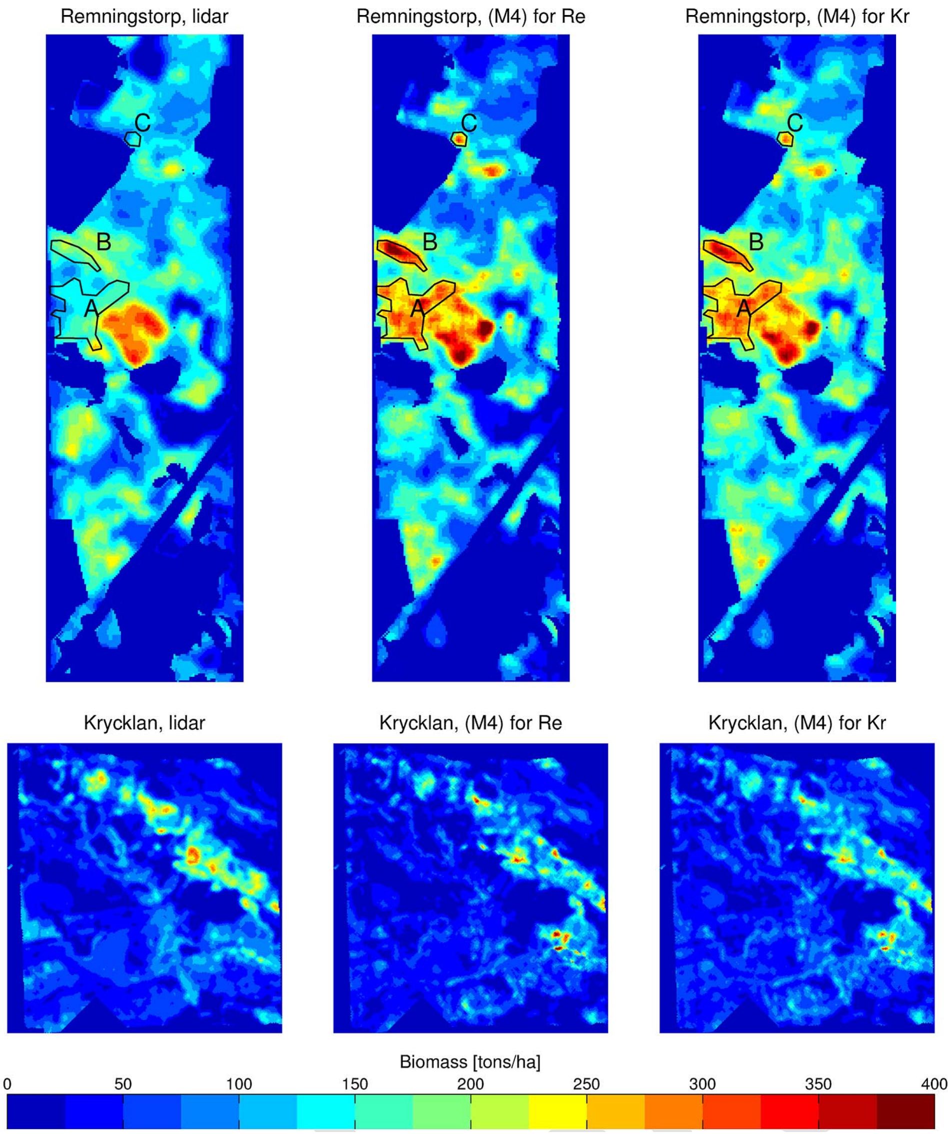


Fig. 10. Extracted biomass maps for Remningstorp and Krycklan. Biomass maps are quantized in intervals of 25 tons/ha. Model (M4) was used to create the maps. For Remningstorp, the north direction is upwards. All Krycklan maps have been rotated by 45° counterclockwise for better viewing and the north east direction is upwards. In all images, the resolution is 70 m × 70 m, and the pixel size is 10 m × 10 m. The size of the imaged region is 3700 m × 1130 m for Remningstorp and 3450 m × 3270 m for Krycklan. The scales are the same in both x- and y-direction. Three regions of interest discussed in Section IV-F are also marked.

average filter in order to reduce resolution to approximately 70 m × 70 m to match the size of the smallest stand in the data sets used for training (0.5 ha). Biomass maps were then produced from all SAR images and averaged. In Fig. 10, only the regions covered by all acquisitions in the respective test sites are shown. The parameters used for map creation can be found in Table VII.

The SAR-based biomass maps show good qualitative agreement with the lidar-based maps. However, in some regions, there are distinct differences between the maps. Three such examples are marked with black contours in Fig. 10.

In the large, irregular region "A" in the central-left part of the Remningstorp map, an overestimation with 100–150 tons/ha is observed. One INS-stand (here called #5, biomass: 167 tons/ha) is located within this region. A careful cross-check with reference *in-situ* and lidar data does not indicate any major issues related to the biomass map itself. However, according to Table 8.1 in [44], 50% of all trees in stand #5 are pines, which contributes to 95% of the total biomass in this stand. The remaining 5% is concentrated in a layer of understory vegetation. This fact has been observed during field visits, and it can also be seen in the lidar height data. The understory layer makes a large contribution to the HV backscatter through the increased number of vegetation scatterers. An investigation in the original SAR data shows that HV is more affected by this vegetation layer than HH.

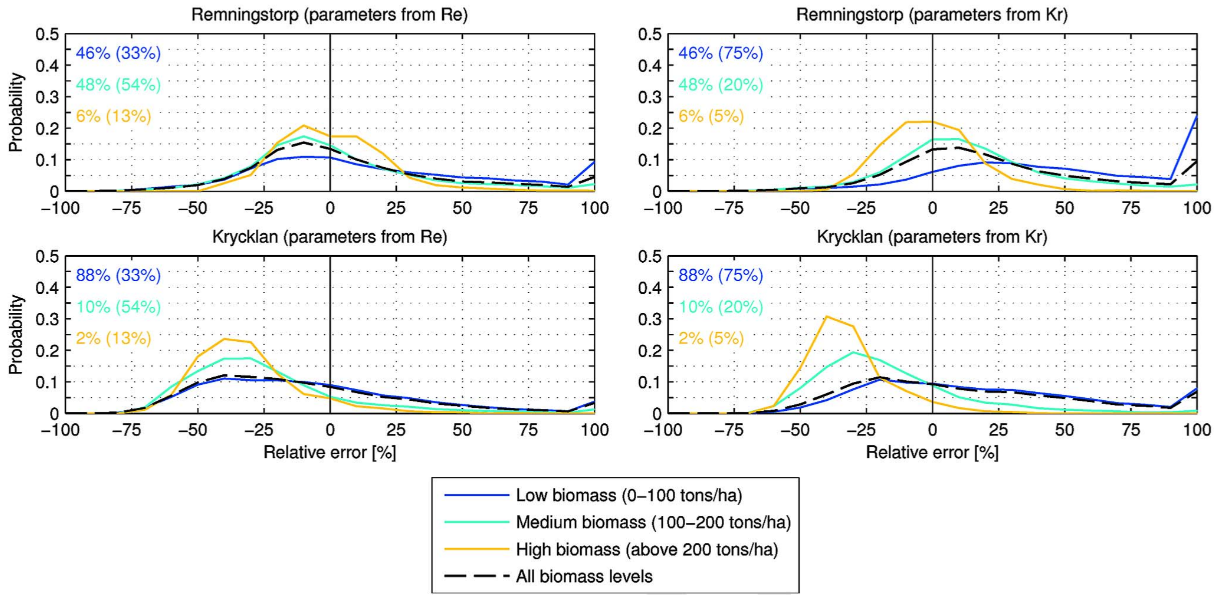


Fig. 11. Probability distributions for the relative difference between the lidar maps and the maps created using (M4) and SAR data are plotted here. Four curves are plotted, one for each biomass group (0–100 tons/ha, 100–200 tons/ha, and above 200 tons/ha), and one for all biomass levels. The distribution of the different groups among the image pixels is shown as percentage values in the upper left corner of each plot. In parentheses, the corresponding values for the training data are shown.

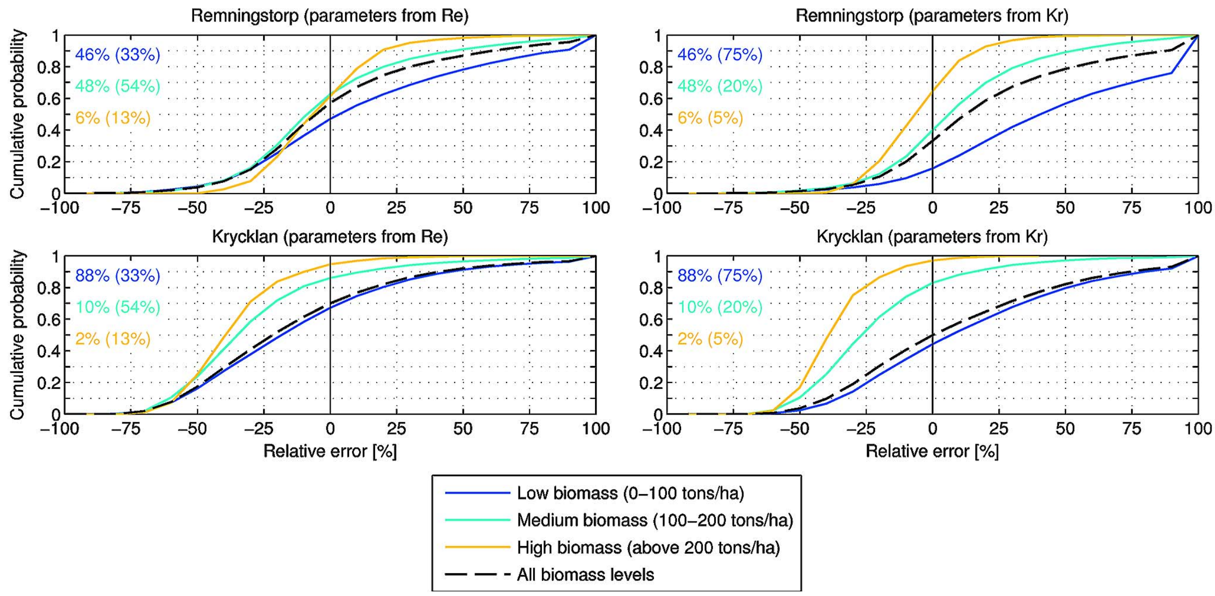


Fig. 12. Cumulative distributions for the relative difference between the lidar maps and the maps created using (M4) and SAR data are plotted here. Four curves are plotted, one for each biomass group (0–100 tons/ha, 100–200 tons/ha, and above 200 tons/ha), and one for all biomass levels. The distribution of the different groups among the image pixels are shown as percentage values in the upper left corner of each plot. In parentheses, the corresponding values for the training data are shown.

In the oblong region “B,” a disagreement of the order of 100–150 tons/ha between lidar and SAR biomass maps is observed in Fig. 10. One forest stand is located within region “B.” This stand is shown in [44, Fig. 6.17] as #11 (biomass: 273 tons/ha, not used in this study due to its small size, 20 m $\times$ 50 m). An investigation of the lidar height data shows that the high-biomass area containing stand #11 is small and surrounded by sparser forest with lower trees. Therefore, filtering of the lidar map will lead to underestimation of biomass around stand #11. Also, the DEM shows, that region “B” is located on a slope, which increases the HV backscatter. This leads to an overestimation of biomass in the SAR-based biomass map. Summarizing, the disagreement between lidar and SAR in region “B” is both due to an overestimation in the SAR map, and an underestimation in the lidar map.

Also, in region "C," another disagreement is observed. The region consists of a group of tall trees situated on plane ground, with virtually no forest between them and the SAR. This increases the difference between HH- and VV-backscatter through the double-bounce effect, thus increasing the ratio. Moreover, smoothing of biomass map decreases the reference biomass level in a similar way as in region "B."

In Figs. 11 and 12, histograms and cumulative distributions for the relative error defined in (14) are shown. Here, the lidar-based biomass map was used as $\mathcal{B}_{ref}$, and the estimated SAR biomass maps were used as $\widehat{\mathcal{B}}$. The data have been divided in three biomass groups: 0–100 tons/ha, 100–200 tons/ha, and 200 tons/ha and above. In the upper left corner of each subplot, the size of each group relative the total number of pixels in percent is shown (in parentheses, corresponding percentage of the training data in each group is shown). In black dashed lines, the corresponding distributions for the whole image are plotted.

In general, between 35 and 50% of all pixels are estimated with relative error smaller than 25%. In Remningstorp, particularly good estimation results are obtained for pixels with lidar biomass higher than 200 tons/ha (80–90% pixels showed relative error smaller than 25%). There is also a group of pixels with low lidar biomass, for which biomass is overestimated with more than 100%. However, in terms of biomass error (measured in tons per hectare), this overestimation is not large.

In Krycklan, a general underestimation is observed for pixels with biomass larger than 100 tons/ha, particularly when Remningstorp-based parameters are used. However, since only 12% of all pixels in the Krycklan map correspond to lidar biomass lower than 100 tons/ha, and the topography in Remningstorp is not strong, these results are less conclusive.

V. SUMMARY AND CONCLUSIONS

A new biomass retrieval model for boreal forest using polarimetric P-band SAR backscatter is presented in this paper. The model is based on two main SAR quantities: the HV backscatter and the HH/VV backscatter ratio, and it also includes a first-order topographic correction, the ground slope angle u.

The paper is based on analysis of data from two airborne P-band SAR campaigns, BioSAR 2007 and 2008, conducted in the two Swedish test sites Remningstorp and Krycklan, separated by 720 km. The examined stand-level biomass interval is 0–300 tons/ha and the surface slope goes up to 19°, measured on a 50 m $\times$ 50 m posting DEM. Only forest stands with areas greater than 0.5 ha are used in this work. An average difference between the data from Remningstorp and Krycklan is observed in all polarization channels, and more work is needed to fully understand and model it in terms of seasonal, topographic, and forest structure differences.

Compared to previously published models, the new model shows less bias induced by temporal change and topographic variability. Also, it gives reliable biomass retrieval results during across-site validation, that is when biomass estimation in one test site is evaluated using a model developed using data from the other test site.

First, all relevant models were tested on data sets coming from Remningstorp, acquired at three occasions during the spring of 2007, each separated by roughly one month. This test showed that the use of multiple polarizations significantly improves the performance. Also, the use of the HH/VV ratio instead of HH and VV channels separately simplifies the model without sacrificing any performance.

The models were also tested for bias due to topographic variability using SAR data acquired from different directions in topographic terrain in Krycklan. The new model gave errors of 27–40 tons/ha (corresponding to 28–42% of the the mean biomass in Krycklan, 95 tons/ha), whereas all the other models gave comparable or worse results. The results of this test were not conclusive, due to non-uniform biomass-slope distribution in the training data.

Thereafter, the across-site retrieval performance was evaluated. The test site used for training was thus distinctly different from the test site used for validation. With model parameters estimated on Krycklan data, biomass in Remningstorp could be estimated with RMSE of 40–59 tons/ha, or 22–33% of the mean biomass in Remningstorp (181 tons/ha) of the mean biomass. The other models produced errors that were at least 50% higher. In the inverse scenario, the Krycklan site was not well represented in the training data set (too small topographic variability in Remningstorp), and errors of 35–51 tons/ha were measured (37–54% of the mean biomass in Krycklan). In terms of RMSE, the new model showed better results than the other models. The coefficient of determination R^2 was, however, low, and it was concluded that the training set was not sufficiently representative in terms of ground surface slopes.

Last, biomass maps estimated using the new model with two parameter sets (one for each test site) were compared to lidar-based biomass maps. The biomass maps were created by averaging biomass estimates from six SAR images in Remningstorp and four SAR images in Krycklan. A good qualitative agreement was observed between the lidar-based biomass maps and the SAR-based biomass maps. However, in some regions biomass was overestimated by SAR, which could be explained based on basic scattering properties of forest in connection to observations made in field and in the lidar data. Between 35 and 45% of all pixels in the maps were estimated with relative difference between the maps smaller than 25%. In Remningstorp, particularly good agreement was obtained for pixels with lidar-estimated biomass higher than 200 tons/ha (80–90% pixels showed relative difference smaller than 25%). In Krycklan, a general underestimation was observed for pixels with biomass larger than 100 tons/ha, particularly when Remningstorp-based parameters were used. However, since only 12% of all pixels in the Krycklan map correspond to lidar biomass lower than 100 tons/ha, and the topography in Remningstorp is not strong, these results are not conclusive.

ACKNOWLEDGMENT

The authors would like to thank both the Swedish National Space Board and the ESA for funding this project, the German Aerospace Center (DLR) for the ESAR data, and Swedish University of Agricultural Sciences (SLU) for the field data.

REFERENCES

[1] S. Solomon, D. Qin, M. Manning, M. Marquis, K. Averyt, M. M. B. Tignor, and H. L. Miller, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge, U.K.: Cambridge Univ. Press, 2007. [Online]. Available: http://www.ipcc.ch/publications_and_data/publications_ipcc_fourth_assessment_report_wg1_report_the_physical_science_basis.htm

[2] C. Parker, A. Mitchell, M. Trivedi, and N. Mardas, "The little REDD book: A guide to governmental and non-governmental proposals for reducing emissions from deforestation and degradation," in *Global Canopy Programme*. Oxford, U.K.: John Krebs Field Station, Nov. 2008.

[3] E. Næsset, T. Gobakken, S. Solberg, T. G. Gregoire, R. Nelson, G. Ståhl, and D. Weydahl, "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.

[4] K. Folkesson, G. Smith-Jonforsen, and L. M. H. Ulander, "Model-based compensation of topographic effects for improved stem-volume retrieval from CARABAS-II VHF-band SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1045–1055, Apr. 2009.

[5] T. Le Toan, A. Beaudoin, J. Riom, and D. Guyon, "Relating forest biomass to SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 2, pp. 403–411, Mar. 1992.

[6] A. Beaudoin, T. Le Toan, S. Goze, E. Nezry, A. Lopes, E. Mougin, C. C. Hsu, H. C. Han, J. A. Kong, and R. T. Shin, "Retrieval of forest biomass from SAR data," *Int. J. Remote Sens.*, vol. 15, no. 14, pp. 2777–2796, Sep. 1994.

[7] H. Israelsson, J. Askne, and R. Sylvander, "Potential of SAR for forest bole volume estimation," *Int. J. Remote Sens.*, vol. 15, no. 14, pp. 2809–2826, Sep. 1994.

[8] Y. Rauste, T. Häma, J. Pulliainen, K. Heiska, and M. Hallikainen, "Radar-based forest biomass estimation," *Int. J. Remote Sens.*, vol. 15, pp. 1797–2808, Sep. 1994.

[9] K. J. Ranson and G. Sun, "Mapping biomass of a northern forest using multifrequency SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 2, pp. 388–396, Mar. 1994.

[10] M. L. Imhoff, "Radar backscatter and biomass saturation: Ramifications for global biomass inventory," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 2, pp. 511–518, Mar. 1995.

[11] E. J. Rignot, R. Zimmermann, and J. J. van Zyl, "Spaceborne applications of P band imaging radars for measuring forest biomass," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 5, pp. 1162–1169, Sep. 1995.

[12] D. H. Hoekman and M. J. Quiñones, "Land cover type and biomass classification using AirSAR data for evaluation of monitoring scenarios in the Colombian Amazon," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 685–696, Mar. 2000.

[13] J. R. Santos, C. C. Freitas, L. S. Araujo, L. V. Dutra, J. C. Mura, F. F. Gama, L. S. Soler, and S. J. S. Sant'Anna, "Airborne P-band SAR applied to the aboveground biomass studies in the Brazilian tropical rainforest," *Remote Sens. Environ.*, vol. 87, no. 4, pp. 482–493, Nov. 2003.

[14] S. Saatchi, K. Halligan, D. Despain, and R. Crabtree, "Estimation of forest fuel load from radar remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1726–1740, Jun. 2007.

[15] G. Sandberg, L. M. H. Ulander, J. E. S. Fransson, J. Holmgren, and T. Le Toan, "L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest," *Remote Sens. Environ.*, vol. 115, no. 11, pp. 2874–2886, Nov. 2011.

[16] S. Saatchi, M. Marlier, R. L. Chazdon, D. B. Clark, and A. E. Russell, "Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass," *Remote Sens. Environ.*, vol. 115, no. 11, pp. 2836–2849, Nov. 2011.

[17] M. Neumann, S. S. Saatchi, L. M. H. Ulander, and J. E. S. Fransson, "Assessing performance of L- and P-band polarimetric interferometric SAR data in estimating boreal forest above-ground biomass," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 714–726, Mar. 2012.

[18] S. R. Cloude and K. P. Papathanassiou, "Polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1551–1565, Sep. 1998.

[19] K. P. Papathanassiou and S. R. Cloude, "Single-baseline polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2352–2363, Nov. 2001.

[20] S. R. Cloude and K. P. Papathanassiou, "Three-stage inversion process for polarimetric SAR interferometry," *Proc. Inst. Elect. Eng.—Radar, Sonar Navigat.*, vol. 150, no. 3, pp. 125–134, Jun. 2003.

[21] I. H. Woodhouse, "Predicting backscatter-biomass and height-biomass trends using a macroecology model," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 4, pp. 871–877, Apr. 2006.

[22] M. Neumann, L. Ferro-Famil, and A. Reigber, "Multibaseline polarimetric SAR interferometry coherence optimization," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 1, pp. 93–97, Jan. 2008.

[23] M. Neumann, L. Ferro-Famil, and A. Reigber, "Estimation of forest structure, ground, and canopy layer characteristics from multibaseline polarimetric interferometric SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1086–1104, Mar. 2010.

[24] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.

[25] S. R. Cloude, "Polarization coherence tomography," *Radio Science*, vol. 41, no. 4, p. RS4 017, Aug. 2006.

[26] S. Tebaldini, "Single and multipolarimetric SAR tomography of forested areas: A parametric approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2375–2387, May 2010.

[27] B. Hallberg, G. Smith-Jonforsen, L. M. H. Ulander, and G. Sandberg, "A physical-optics model for double-bounce scattering from tree stems standing on an undulating ground surface," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 9, pp. 2607–2621, Sep. 2008.

[28] G. Smith-Jonforsen, L. M. H. Ulander, and X. Luo, "Low VHF-band backscatter from coniferous forests on sloping terrain," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 10, pp. 2246–2260, Oct. 2005.

[29] J. J. van Zyl, "The effect of topography on radar scattering from vegetated areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 1, pp. 153–160, Jan. 1993.

[30] Y.-C. Lin and K. Sarabandi, "Electromagnetic scattering model for a tree trunk above a tilted ground plane," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 4, pp. 1063–1070, Jul. 1995.

[31] Y. Dong and J. A. Richards, "Studies of the cylinder-ground double bounce scattering mechanism in forest backscatter models," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 1, pp. 229–231, Jan. 1995.

[32] J. Lopez-Sanchez, H. Esteban-Gonzalez, M. Baquero-Escudero, and J. Fortuny-Guasch, "An electromagnetic scattering model for multiple tree trunks above a tilted rough ground plane," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2, pp. 659–667, Mar. 1999.

[33] G. Smith and L. M. H. Ulander, "A model relating VHF-band backscatter to stem volume of coniferous boreal forest," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 728–740, Mar. 2000.

[34] M. L. Williams, PolSARproSim: A Coherent, Polarimetric SAR Simulation of Forests for PolSARPro; Design Document and Algorithm Specification (u1.0), Eur. Space Agency, Paris, France. [Online]. Available: http://envisat.esa.int/polsarpro/Manuals/PolSARproSim_Design.pdf

[35] Candidate Earth Explorer Core Mission: BIOMASS, Report for Assessment, Eur. Space Agency, Paris, France, Tech. Rep. SP-1212/2. [Online]. Available: http://esamultimedia.esa.int/docs/SP1313-2_BIOMASS.pdf

[36] M. Arcioni, P. Bensi, J.-L. Bézy, B. Carnicero, M. Davidson, M. Drinkwater, F. Fois, A. Gabriele, R. Haagmans, F. Hélière, P. Ingmann, V. Kangas, M. Kern, S. Kraft, J. Langen, A. Lecuyot, C.-C. Lin, R. Meynart, K. Scipal, and P. Silvestrin, "Biomass, CoReH2O, PREMIER: ESA's candidate 7th earth explorer missions," in *Proc. IEEE IGARSS*, Honolulu, HI, Jul. 2010, pp. 673–676.

[37] K. Scipal, M. Arcioni, J. Chave, J. Dall, F. Fois, T. Le Toan, C. Lin, K. Papathanassiou, S. Quegan, F. Rocca, S. Saatchi, H. Shugart, L. Ulander, and M. Williams, "The BIOMASS mission—An ESA earth explorer candidate to measure the biomass of the earth's forests," in *Proc. IEEE IGARSS*, Honolulu, HI, Jul. 2010, pp. 52–55.

[38] T. Le Toan, S. Quegan, M. W. J. Davidson, H. Balzter, P. Paillou, K. Papathanassiou, S. Plummer, F. Rocca, S. Saatchi, H. Shugart, and L. Ulander, "The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle," *Remote Sens. Environ.*, vol. 115, no. 11, pp. 2850–2860, Nov. 2011.

[39] S. Quegan, T. Le Toan, J. Chave, J. Dall, K. P. Papathanassiou, F. Rocca, L. Ulander, M. Williams, S. S. Saatchi, and H. Shugart, BIOMASS, Report for Mission Selection, Eur. Space Agency, Paris, France, Tech. Rep. SP-1324/1. [Online]. Available: http://esamultimedia.esa.int/docs/EarthObservation/SP1324-1_BIOMASSr.pdf

[40] M. J. Soja, G. Sandberg, and L. M. H. Ulander, "Topographic correction for biomass retrieval from P-band SAR data in boreal forests," in *Proc. IEEE IGARSS*, Honolulu, HI, 2010, pp. 4776–4779.

[41] G. Sandberg, M. J. Soja, and L. M. H. Ulander, "Impact and modeling of topographic effects on P-band SAR backscatter from boreal forests," in *Proc. IEEE IGARSS*, Vancouver, Canada, 2011, pp. 3522–3525.

[42] M. Neumann, S. S. Saatchi, L. M. H. Ulander, and J. E. S. Fransson, "Parametric and non-parametric forest biomass estimation from PolInSAR data," in *Proc. IEEE IGARSS*, Vancouver, Canada, Jul. 2011, pp. 420–423.

[43] L. M. H. Ulander, G. Sandberg, and M. J. Soja, "Biomass retrieval algorithm based on P-band BioSAR experiments of boreal forest," in *Proc. IEEE IGARSS*, Vancouver, Canada, Jul. 2011, pp. 4245–4248.

[44] I. Hajnsek, R. Scheiber, L. Ulander, A. Gustavsson, G. Sandberg, S. Tebaldini, A. M. Guarnieri, F. Rocca, F. Bombardini, and M. Pardini, BioSAR 2007 Technical Assistance for the Development of Airborne SAR and Geophysical Measurements During the BioSAR 2007 Experiment: Final Report Without Synthesis, Eur. Space Agency, Paris, France, Tech. Rep. ESA Contract No. 20755/07/NL/CB. [Online]. Available: http://earth.esa.int/campaigns/DOC/biosar_finalreports_nosynthesis.pdf

[45] I. Hajnsek, R. Scheiber, M. Keller, R. Horn, S. Lee, L. M. H. Ulander, A. Gustavsson, G. Sandberg, T. Le Toan, S. Tebaldini, A. M. Guarnieri, and F. Rocca, BioSAR 2008 Technical Assistance for the Development of Airborne SAR and Geophysical Measurements During the BioSAR 2008 Experiment: Final Report—BioSAR Campaign, Tech. Rep. ESA Contract No. 22052/08/NL/CT. [Online]. Available: http://earth.esa.int/campaigns/DOC/BIOSAR2_final_report.pdf

[46] L. M. H. Ulander, A. Gustavsson, P. Dubois-Fernandez, X. Dupuis, J. E. S. Fransson, J. Holmgren, J. Wallerman, L. E. B. Eriksson, G. Sandberg, and M. J. Soja, "BioSAR 2010—A SAR campaign in support to the BIOMASS mission," in *Proc. IEEE IGARSS*, Vancouver, BC, Canada, Jul. 2011, pp. 1528–1531.

[47] L. M. H. Ulander, A. Gustavsson, B. Flood, D. Murdin, P. Dubois-Fernandez, X. Depuis, G. Sandberg, M. J. Soja, L. E. B. Eriksson, J. E. S. Fransson, J. Holmgren, and J. Wallerman, BioSAR 2010 Technical Assistance for the Development of Airborne SAR and Geophysical Measurements During the BioSAR 2010 Experiment: Final Report, Eur. Space Agency, Paris, France, Tech. Rep. ESA Contract No. 4000102285/10/NL/JA/ef. [Online]. Available: http://earth.esa.int/campaigns/DOC/BioSAR_2010_final_report_v1.0.pdf

[48] J. E. S. Fransson, F. Walter, and L. M. H. Ulander, "Estimation of forest parameters using CARABAS-II VHF SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 720–727, Mar. 2000.

[49] L. G. Marklund, "Biomassafunktioner för Tall, Gran och Björk i Sverige," Institutionen för Skogstaxering, Sveriges Lantbruksuniversitet, Umeå, Sweden, Rapport 45, 1988.

[50] H. Petersson, "Biomassafunktioner for Trädfraktioner av Tall, Gran Och Björk i Sverige," SLU, Institutionen för skoglig resurshållning och geomatik, Umeå, Sweden, Tech. Rep. Arbetsrapport 59, 1999.

[51] J. E. S. Fransson, M. Magnusson, G. Sandberg, L. M. H. Ulander, and A. Gustavsson, "Forest Database for Single Trees in Remningstorp, Version 2.0," Västergötland, Sweden, 2008.

[52] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL: CRC Press, 2009.

[53] L. M. H. Ulander, "Radiometric slope correction of synthetic aperture radar images," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 5, pp. 1115–1122, Sep. 1996.

[54] T. Le Toan, L. Ulander, G. Sandberg, B. Petrucci, L. Villard, P. Borderies, P. Dubois-Fernandez, S. Angelliaume, M. Qui nones, D. Hoekman, K. Papathanassiou, F. Kugler, S.-K. Lee, F. Garestier, A. Couhert, and J. Valteau, "P-band SAR Wave Interaction and Information Retrieval: Analysis and Inversion of P-Band SAR Data for Forest Biomass and Height Mapping," Eur. Space Agency, Paris, France, Tech. Rep. ESA Contract No. 20290/06/NL/LvH, 2011.

[55] N. R. Draper and H. Smith, *Applied Regression Analysis,* 3rd ed. Hoboken, NJ: Wiley, 1998.

Maciej Jerzy Soja was born in 1985 in Warsaw, Poland. He received the B.Sc. and M.Sc. degrees in engineering physics from Chalmers University of Technology, Gothenburg, Sweden, in 2008 and 2009, respectively. Since December 2009, he has been working toward the Ph.D. degree in remote sensing in the Radar Remote Sensing Group at the Department of Earth and Space Sciences at Chalmers University of Technology, Gothenburg.

His main research topic is synthetic aperture radar in forestry.

Gustaf Sandberg received the M.Sc. degree in engineering physics from Chalmers University of Technology, Gothenburg, Sweden, in 2006, where he is currently working toward the Ph.D. degree in the Department of Earth and Space Sciences.

His main research interests lie in synthetic aperture radar analysis, with emphasis on ionospheric effects as well as L- and P-band measurements of forests.

Lars M. H. Ulander (S'86–M'90–SM'04) received the M.Sc. degree in engineering physics in 1985 and the Ph.D. degree in electrical and computer engineering in 1991, both from Chalmers University of Technology, Gothenburg, Sweden.

Since 1995, he has been with the Swedish Defense Research Agency (FOI) in Linköping where he is the Director of Research in radar signal processing. He is also an Adjunct Professor in radar remote sensing at Chalmers University of Technology. His research areas are synthetic aperture radar, electromagnetic scattering models, and remote sensing applications. He is the holder of five patents and the author or coauthor of over 250 professional publications, of which more than 50 are in peer-reviewed scientific journals.

Paper C

Model-Based Biomass Estimation of a Hemi-Boreal Forest from Multi-Temporal TanDEM-X Acquisitions

Authors:

J. I. H. Askne, J. E. S. Fransson, M. Santoro, **M. J. Soja**, and L. M. H. Ulander

Status:

Published in *Remote Sensing*, vol. 5, no. 11, pp. 5574–5597, October 2013

Remote Sens. **2013**, *5*, 5574-5597; doi:10.3390/rs5115574

OPEN ACCESS

Remote Sensing

ISSN 2072-4292

www.mdpi.com/journal/remotesensing

Article

Model-Based Biomass Estimation of a Hemi-Boreal Forest from Multitemporal TanDEM-X Acquisitions

Jan I.H. Askne [1,*], Johan E.S. Fransson [2], Maurizio Santoro [3], Maciej J. Soja [1] and Lars M.H. Ulander [1,4]

[1] Department of Earth and Space Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden; E-Mails: maciej.soja@chalmers.se (M.J.S.); lars.ulander@foi.se (L.M.H.U.)

[2] Department of Forest Resource Management, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden; E-Mail: johan.fransson@slu.se

[3] Gamma Remote Sensing, Worbstrasse 225, CH-3073 Gümligen, Switzerland; E-mail: santoro@gamma-rs.ch

[4] Department of Radar Systems, Swedish Defence Research Agency (FOI), SE-581 11 Linköping, Sweden.

* Author to whom correspondence should be addressed; E-Mail: jan.askne@chalmers.se; Tel.: +46-703-495-795; Fax: +46-31-772-1884

Received: 24 June 2013; in revised form: 22 October 2013 / Accepted: 23 October 2013 / Published: 29 October 2013

Abstract: Above-ground forest biomass is a significant variable in the terrestrial carbon budget, but is still estimated with relatively large uncertainty. Remote sensing methods can improve the characterization of the spatial distribution and estimation accuracy of biomass; in this respect, it is important to examine the potential offered by new sensors. To assess the contribution of the TanDEM-X mission, eighteen interferometric Synthetic Aperture Radar (SAR) image pairs acquired over the hemi-boreal test site of Remningstorp in Sweden were investigated. Three models were used for interpretation of TanDEM-X signatures and above-ground biomass retrieval: Interferometric Water Cloud Model (IWCM), Random Volume over Ground (RVoG) model, and a simple model based on penetration depth (PD). All use an allometric expression to relate above-ground biomass to forest height measured by TanDEM-X. The retrieval was assessed on 201 forest stands with a minimum size of 1 ha, and ranging from 6 to 267 Mg/ha (mean biomass of 105 Mg/ha) equally divided into a model training dataset and a validation test dataset. Biomass retrieved using the IWCM resulted in a Root Mean Square Error (RMSE) between 17%

and 33%, depending on acquisition date and image acquisition geometry (angle of incidence, interferometric baseline, and orbit type). The RMSE in the case of the RVoG and the PD models were slightly higher. A multitemporal estimate of the above-ground biomass using all eighteen acquisitions resulted in an RMSE of 16% with $R^2 = 0.93$. These results prove the capability of TanDEM-X interferometric data to estimate forest aboveground biomass in the boreal zone.

Keywords: TanDEM-X; InSAR; forestry; boreal; biomass estimation; model-based; allometry

1. Introduction

Forest above-ground dry biomass (AGB, herewith simply referred to as biomass) is an important variable for the global carbon budget, not only due to the uptake of carbon dioxide in the process of photosynthesis, but also because forests store huge amounts of carbon, which are eventually released into the atmosphere following a disturbance [1]. Accurate and timely mapping of forest AGB is therefore crucial to support carbon cycle modeling. Traditional methods based on forest inventories and aerial photography, and more recently, LiDAR campaigns, give accurate estimates of AGB; however, such methods are expensive and become inefficient whenever frequent and large-scale mapping is needed. Therefore, there is a need for development of alternative methods for frequent and large-scale biomass mapping [2].

One of the more promising techniques for above-ground dry biomass mapping is Synthetic Aperture Radar (SAR), *cf.* [3]. Being an active sensor, radar is independent of weather and external illumination. Spaceborne SAR missions currently in operation are characterized by an image resolution on the order of meters. In addition, interferometric SAR, InSAR, offers the possibility to exploit two further observables besides the radar backscatter, namely the coherence and the interferometric phase. These are affected by the forest structure and, thus, are related to forest variables such as tree height, and stem volume or AGB. In a single-pass acquisition scenario, the association between InSAR observables and forest variables is expected to be maximized because temporal decorrelation can be assumed to be negligible. Experimental evidence on the suitability of single-pass InSAR to estimate forest variables at X-band (wavelength of approximately 3 cm) was provided by data acquired by airborne sensors [4–6], and during the Shuttle Radar Topography Mission (SRTM) [7].

In June 2010, the TanDEM-X (TerraSAR-X add-on for Digital Elevation Measurement) satellite was launched. Together with the almost identical twin-satellite TerraSAR-X (launched in June 2007), the first satellite-based single-pass SAR interferometer was formed. In the bistatic mode of the TanDEM-X mission (consisting of the TanDEM-X and TerraSAR-X satellites), only one satellite is used for transmission while both satellites are used for reception. For simplicity, we will refer to this mission as the TDM mission. In TDM data, temporal decorrelation is limited to a minimum because of the small along-track baseline between the sensors. The primary objective of TDM is to obtain a global Digital Elevation Model (DEM) with an absolute height accuracy better than 10 m and an equatorial spatial resolution of 12 m [8]. Because of the limited penetration of microwaves into the canopy, X-band interferograms over forests are characterized by an elevation offset which is dependent on forest canopy height and density [9]. This offset suggests exploiting TDM imagery to estimate tree height although a reference for the ground elevation is needed. Since X-band microwaves do not significantly penetrate the closed canopy of a dense forest, a Digital Terrain Model (DTM) for the ground surface needs to be provided by some other, independent method, for example P-band SAR [10], or LiDAR [6,7]. Besides forest height estimation, retrieval of above-ground dry biomass was also investigated in some studies. In [10], a Root Mean Square Error (RMSE) of 46.1 Mg/ha (biomass range up to ≈ 360 Mg/ha) was obtained for biomass in a tropical forest using airborne SAR in X- and P-band, and in [7] RMSE = 19% was obtained using SRTM in X-band.

The objective of this study was to develop and assess estimation methods based on models linking X-band InSAR observations to forest biomass. For this, single polarized (VV), bistatic interferometric TanDEM-X data acquired between June 2011 and August 2012 over Remningstorp, a hemi-boreal test site situated in southern Sweden, were used. Three InSAR models were employed and evaluated: Interferometric Water Cloud Model (IWCM) [11–14], Random Volume over Ground model (RVoG) [15–17] and a simple model based on the penetration depth (PD) of X-band microwaves. As reference, biomass estimates derived from LiDAR scanning data acquired during the BioSAR 2010 campaign [18,19], performed within the BIOMASS phase-A study [20] were used. By means of 201 forest stands equally divided into a training and a validation dataset, properties of the model parameters were determined and biomass retrieval accuracy for the different models was quantified.

2. Test Site and Datasets

Remningstorp (58°30′N, 13°40′E) is an estate with 1,200 ha productive forest area in the hemi-boreal zone, which is the transition between the boreal and the temperate zone [21]. Forest species consist primarily of Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and birch (*Betula* spp.). The test site is fairly flat with elevations ranging from 120 m to 145 m above sea level.

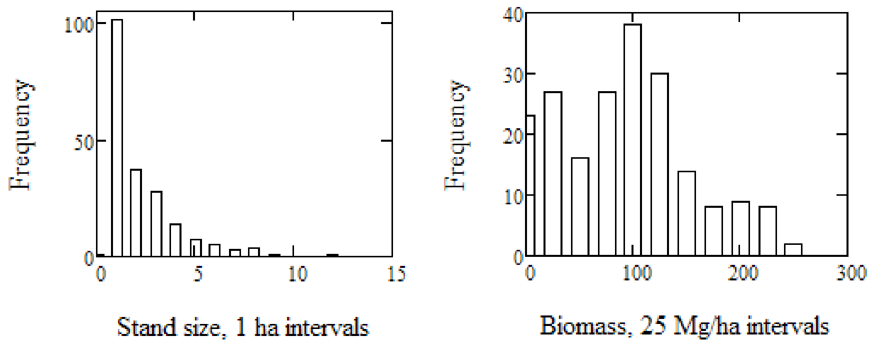
2.1. Field Observations

Field data used for this study were collected in 2010 [19] and consisted of 212 field plots with 10 m radius allocated over the estate. The survey assessed the stem volume, tree height, diameter at breast height (*i.e.*, 1.3 m above ground level), stem density, tree species composition based on proportion of total stem volume, and above-ground dry biomass, including stem, bark, branches and needles, but excluding stump and roots. Biomass was expressed in Mg of dry mass per hectare. The field survey and the biomass estimation for the field plots was carried out according to the Heureka forestry decision support system [22] with functions described in [23]. In addition, seven 80 m × 80 m field plots were inventoried *in situ* through single tree measurements (including all trees with diameter at breast height, DBH > 0.05 m). The measurements made on tree level included GPS position, DBH, species and height. Biomass and stem volume (including the stem above stump, and bark, and expressed in m^3/ha) were estimated using functions developed in [24] and [25], respectively.

The biomass for the 212 field plots was then related to LiDAR metrics by least-squares regression in line with similar studies, e.g., [26]. The LiDAR data were collected with a density of 69 returns/m^2 on average. The LiDAR metrics selected to establish the relationship with biomass were chosen based on studies of correlations and residual plots and included for example height percentiles of LiDAR returns, vegetation densities from proportions of LiDAR returns, and tree species stratification information. The final model was tested for overfitting, and the coefficient of determination, R^2, between the biomass estimated for the 212 field plots and the biomass derived from the corresponding LiDAR metrics, was estimated to 0.81. Spatially explicit estimates of biomass were then derived from the function linking the LiDAR metrics to biomass. The LiDAR-based biomass was obtained for the entire forest estate and represented the reference biomass dataset for this study. The accuracy of the LiDAR-based estimates of biomass was determined by comparing it to the biomass for the seven 80 m × 80 m field plots which gave a RMSE of 12.7% [19].

The Remningstorp estate was divided into 665 delineated polygons, of which 562 consisted of forest stands, *i.e.*, areas of homogeneous tree cover, species composition and canopy structure. The remaining 103 polygons consisted of open fields, pastures, private lots, water, *etc.* The 562 forest stands, of which 201 stands were at least 1 ha large, were characterized by full LiDAR coverage and did not experience major forest cover changes between the LiDAR and the TDM acquisitions. The digital map with the forest stand boundaries was eroded with a 10-m buffer zone around the boundary of each stand to reduce border effects on the evaluation of the TDM data. The biomass of the 201 stands was between 6 and 267 Mg/ha with a mean of 105 Mg/ha. Figure 1 illustrates the distribution of stand sizes and biomass for the 201 forest stands. More than 90% of the forest stands were smaller than 5 ha (Figure 1a). The biomass presented an almost uniform distribution up to 150 Mg/ha; several stands were characterized by biomass above 200 Mg/ha (Figure 1b).

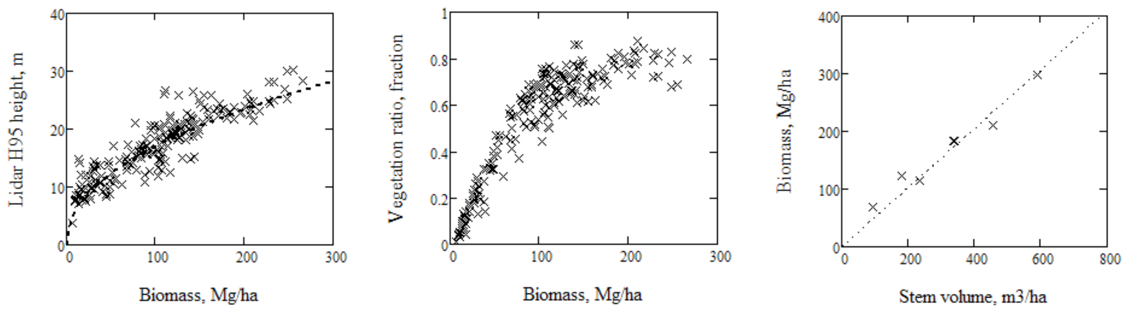
Figure 1. Distribution of forest stand size (**a**) and biomass (**b**) for the 201 forest stands ≥ 1 ha at the test site in Remningstorp.



(**a**) (**b**)

Some of the LiDAR metrics and the LiDAR-based biomass for the 201 forest stands were used to support the interpretation of the TDM interferometric signatures and to support the modeling phase relating the interferometric data to biomass (see Section 3). In Figure 2a, stand-level averages of the 95th percentile of LiDAR return values above a height threshold of 1.0 m or 10% of the maximum height (H95) have been plotted against the LiDAR-based biomass to assess the validity of the allometric function to be then used in the modeling phase, see Equation (3). Vegetation ratio derived from the LiDAR data [19], *i.e.*, the ratio of LiDAR return values above a height threshold of 5.0 m and the total number of returns, provide information on canopy closure and can be considered as proxy for a similar parameter used in modeling, namely the area-fill factor (see Section 3). To understand the relationship between canopy closure and biomass, Figure 2b shows the relationship between vegetation ratio and the LiDAR-based biomass for the 201 forest stands. Figure 2a,b shows strong correlations between the illustrated forest variables, indicating the suitability of empirical relationships to be implemented during the modeling phase of the interferometric signatures to express these solely as a function of forest biomass.

Figure 2. (**a**) Stand-wise values of LiDAR heights (H95) *versus* LiDAR-based biomass and an allometric equation relating basal-area-weighted mean height to biomass, see Equation (3); (**b**) LiDAR vegetation ratio *versus* LiDAR-based biomass, and (**c**) biomass *versus* stem volume for the seven 80 m × 80 m plots with single tree measurements.



(**a**) (**b**) (**c**)

The *in situ* information on biomass and stem volume from the seven 80 m × 80 m forest inventory plots was used to derive a linear equation linking the biomass, *B*, and the stem volume, *V*. In Equation (1), BEF represents a biomass expansion factor (expanding to include branches and needles as well as stem and bark) which was estimated to be 0.512 Mg/m^3 with $R^2 = 0.97$, *cf.* Figure 2c,

$$B = BEF \cdot V \tag{1}$$

Although the BEF was estimated using a small dataset of samples and the relationship between biomass and stem volume is in theory dependent on tree species, age and local conditions, *cf.* [27] and [28], it was assumed that a single and approximate value was sufficient for the purpose of this investigation.

Furthermore, an allometric relationship between basal-area-weighted mean height, *h*, and stem volume, *V*, was considered

$$h(V) = (2.44\ V)^{0.46} \tag{2}$$

Equation (2) was derived using measurements from test sites in Sweden and Finland and verified by means of 4,188 randomly chosen National Forest Inventory field plots located in different regions of Sweden [11,13]. Combining Equations (1) and (2), an allometric relationship between biomass and height was obtained:

$$B(h) = 0.21\ h^{2.17} \tag{3}$$

illustrated in Figure 2a. Although stem volume and biomass also depend on other forest variables, for example tree species and tree number density, it was shown in [29] using a plant structure model that biomass can generally be modeled from height using the same functional form as Equation (3), where the exponent is determined by a scaling parameter related to species and thinning practices.

The allometric relationship, Equation (3), is not entirely optimal for the Remningstorp data, *cf.* Figure 2a, but was still used due to its general nature. The aim of the allometric relationship is to decrease the number of forest variables in the models relating these to TDM observables and have them expressed purely as functions of biomass.

2.2. TanDEM-X SAR Dataset

A large number of TanDEM-X InSAR acquisitions were available for the Remningstorp test site (Table 1). In Table 1, the stand-level mean values of the Height of Ambiguity (HOA) and the along-track baseline (ATB) are shown. HOA is the height interval corresponding to a phase difference of 2π; it is a measure for the sensitivity of the InSAR phase to elevation and is inversely proportional to the perpendicular component of the across-track baseline. ATB is the distance between the satellites along track. Since each pixel in the monostatic active and the bistatic passive image is focused in azimuth at its Zero Doppler Time [30] the two images can be looked as monostatic and ATB provides a measure for temporal decorrelation [31].

Table 1. TanDEM-X InSAR acquisitions and weather conditions at the time of image acquisition (HOA = Height of Ambiguity, ATB = Along-Track Baseline). The sign of HOA depends on the satellite positions, but has no relevance for the results in this paper.

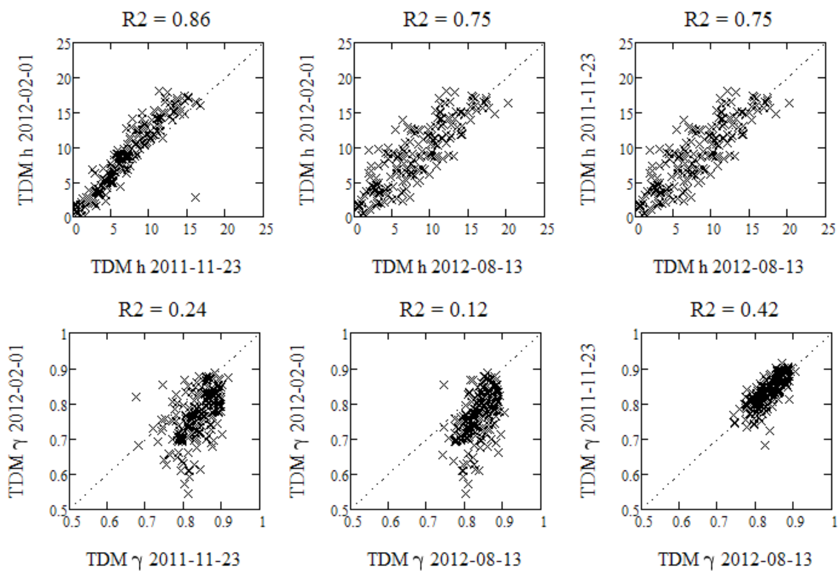
Date	HOA (m)	ATB (m)	Incidence Angle	Temperature (°C)	Wind Speed (m/s)	Precipitation (mm)
2011-06-04	49	110	41°	24	2	0
2011-11-23	−185	4	34°	5	3	1.2
2011-12-26	−178	64	34°	6	5	0
2012-01-17	−172	−1	34°	−2	1	0
2012-01-28	−182	7	34°	−3	2	0
2012-02-01	80	267	41°	−4	1	0.2
2012-02-08	−179	29	34°	−3	1	0
2012-02-12	−79	−244	41°	−3	1	0.2
2012-02-19	−186	−8	34°	1	6	3.4
2012-02-23	79	232	41°	3	3	0
2012-03-01	−186	−11	34°	5	4	0
2012-03-12	−187	−11	34°	4	2	0
2012-03-23	−183	14	34°	−1	0	0
2012-05-28	349	262	34°	15	2	1.3
2012-07-22	339	262	34°	14	3	4.1
2012-08-02	315	233	34°	15	3	0.6
2012-08-13	358	229	34°	14	0	0
2012-08-24	301	208	34°	13	3	0.2

TDM images were provided by German Aerospace Center (DLR) in a co-registered single-look complex (SLC) format with common spectral filtering applied during pre-processing. Interferometric processing of TanDEM-X data were done with a Matlab-based algorithm [32] developed specially for interferometric processing of TanDEM-X data and based on [30]. The first step of interferometric processing consists of interferogram flattening for curved Earth and surface elevation. For this, InSAR phase was simulated using an airborne LiDAR DTM with a spatial resolution of 2 m and mean height error lower than 0.5 m. The DTM was acquired by the Swedish National Land Survey (Lantmäteriet) within an ongoing nationwide LiDAR scanning campaign. The DTM was interpolated to range-azimuth coordinates (radar geometry) using satellite state vectors and look geometry information provided in the metadata of the images. For each stand, a complex average of all pixels within the stand was computed, and the corresponding phase was converted to elevation (in the following referred to as TDM height) by a multiplication with HOA/2π. Absolute height calibration was performed through the subtraction of the mean TDM height for non-forested stands (at least 0.5 ha in size and scattered over the entire estate) from the TDM heights. For a few stands, TDM heights were slightly below the mean TDM height for non-forest areas. In such cases, the absolute TDM height was set equal to zero. This applied also for one stand (2012-05-28) with an offset of −1.5 m, probably due to the large HOA, which was 349 m.

Among the stands, one stand (ID = 189, 1.42 ha, 25.1 Mg/ha) presented TDM heights from −11.7 m to 36.0 m. This stand also had a low coherence. High-resolution LiDAR data shows, that the stand is highly irregular and consists of several disjoint parts of high trees alternating with low vegetation or open ground. Such a stand is easily detected by comparing TDM observations in a multitemporal approach. In the following the TDM height of this stand has been put to zero, *i.e.*, the biomass will be assumed zero.

The temporal consistency of TDM heights was high (Figure 3, upper row), in particular between acquisitions with similar HOA (R^2 = 0.99 for 2012-02-01 and 2012-02-23, for example). When acquisitions with different HOA were compared, R^2 tended to decrease with increasing HOA for one or both of the acquisitions. Almost the opposite applied for coherence but then with much lower R^2. The dynamic range of coherence was found to decrease with increasing HOA. A large HOA corresponds to a short perpendicular baseline and thus to low volume decorrelation.

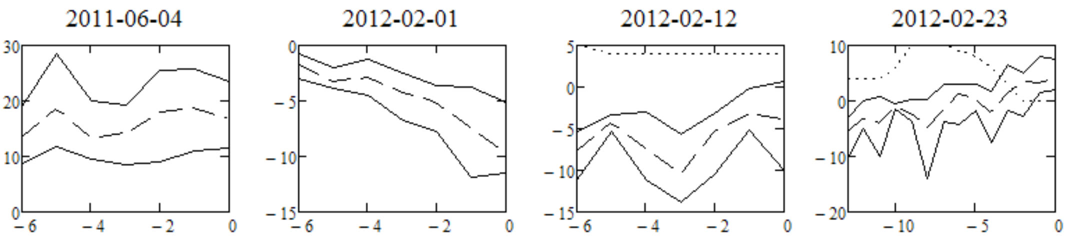
Figure 3. Temporal consistency visualized by means of scatterplots of stand-wise TDM height (TDM h, upper row) and TDM coherence (TDM γ, lower row). HOA = 80 m for 2012-02-01, HOA = −185 m for 2011-11-23, and HOA = 358 m for 2012-08-13.



2.3. Meteorological Data

Table 1 lists temperature and wind speed measured within one hour of the satellite overpass at the closest official meteorological station, Hällum, situated 23 km from Remningstorp. Precipitation in Table 1 was recorded in Remningstorp for the date of acquisition. For the four image pairs with |HOA| ≤ 80 m showing the strongest agreement between TDM heights, the history of temperature and snow depth are further illustrated in Figure 4 in order to support the interpretation of the measurements. The history of the maximal, mean, and minimal temperatures has been plotted. For three of the acquisitions in Figure 4, temperature change for the six preceding days is shown. For the fourth acquisition (2012-02-23), data from two preceding weeks are shown to illustrate the more complex situation.

Figure 4. History of temperature for the Remningstorp test site prior to four TDM acquisitions with |HOA| ≤ 80 m. The units on the x- and y-axis are days and temperature (°C), respectively. The snow layer in cm is dotted for the acquisition dates 2012-02-12 and 2012-02-23.



3. Interferometric Forest Models

The models, used for interpretation of TDM observations and to explain their relation to forest height or biomass, were selected to be simple enough in order to make inversion possible. This means, that the forest properties can only be described by a few parameters. The Interferometric Water Cloud Model, the Random Volume over Ground model, as well as the simple model based on the penetration depth presented below fulfilled these requirements.

3.1. Interferometric Water Cloud Model

The Interferometric Water Cloud Model, IWCM, is a model for the complex coherence of a forest. The IWCM was introduced to explain the coherence of forest at C-band [11,12,14]. The model assumes that the medium, characterized by a certain forest height, h, and stem volume, V, can be described by a random vegetation layer like the Water Cloud Model [33], with uniform scatterer density but generalized to include gaps. The proportion of the area with vegetation relative the total area is denoted as the area-fill, η. Recently, and in agreement with IWCM, it was shown using observations of the spatial scattering spectrum of TanDEM-X data over tropical forests that forests cannot be modeled as a layered medium, but by a model with gaps or random scattering "clouds" [34].

The attenuation factor in the vegetation layer is described by α, representing the mean attenuation for the idealized vegetation layer (either vegetation or no vegetation), and is a function of, for example, the dielectric constant of the scatterers. According to [35], the forest backscatter, σ_{for}, can also be described by an empirical, exponential stem volume dependence characterized by β, as demonstrated by scatterometer measurements at X- and C-band. There are then two alternative expressions for the backscatter:

$$\sigma_{for}^0 = \eta\left[\sigma_{gr}^0 e^{-\alpha h} + \sigma_{veg}^0(1 - e^{-\alpha h})\right] + (1-\eta)\sigma_{gr}^0 = \sigma_{gr}^0 e^{-\beta V} + \sigma_{veg}^0\left(1 - e^{-\beta V}\right) \tag{4}$$

In Equation (5), σ_{gr}^0 represents the ground backscatter and σ_{veg}^0 represents s the vegetation layer backscatter.

The complex expression for the coherence of this random volume with gaps is then described by each of the independent scattering parts weighted by system and temporal decorrelation, the product of which will be represented by the coherence parameters γ_{gr} and γ_{veg}, for the ground and vegetation parts, respectively. For the vegetated part there is also volume decorrelation, γ_{vol}, related to the forest height. The complex coherence (assuming the phase of the ground surface has been compensated for) is then given by:

$$\begin{aligned}\tilde{\gamma} &= \frac{\left\{\eta\left[\gamma_{gr}\sigma_{gr}^0 e^{-\alpha h} + \gamma_{veg}\tilde{\gamma}_{vol}\sigma_{veg}^0(1 - e^{-\alpha h})\right] + \gamma_{gr}(1-\eta)\sigma_{gr}^0\right\}}{\sigma_{for}^0} \\ &= \frac{\gamma_{gr}\sigma_{gr}^0 e^{-\beta V} + \gamma_{veg}\tilde{\gamma}_{vol}\sigma_{veg}^0\left(1 - e^{-\beta V}\right)}{\sigma_{for}^0}\end{aligned} \tag{5}$$

For the two expressions in Equations (4) and (5) to agree, a requirement on the area-fill factor η is given by a relation between α and β according to:

$$\eta(V) = \frac{1 - e^{-\beta V}}{1 - e^{-\alpha h(V)}} \tag{6}$$

with $\eta \to 1$ when $V \to \infty$. According to Equation (6), β can be described by the area-fill η, the attenuation α, and $h(V)$, *i.e.*, by the vegetation density and attenuation through the vegetation. h and V will, whenever needed, be related through the allometric relationship in Equation (2), *i.e.*, $h(V)$ represents the basal-area-weighted mean forest height. Equation (4) shows that β characterizes the transition from dominant ground scattering to dominant vegetation scattering.

If the variation of scattering with height is only determined by the attenuation, the volume decorrelation is determined by, [11]:

$$\tilde{\gamma}_{vol} = \frac{\int_0^h e^{-\alpha(h-z')} \cdot e^{-jK(B_n)z'} \mathrm{d}z'}{\int_0^h e^{-\alpha(h-z')}\, \mathrm{d}z'} = \frac{\alpha}{\alpha - jK(B_n)} \frac{e^{-jK(B_n)h} - e^{-\alpha h}}{1 - e^{-\alpha h}} \tag{7}$$

For the TanDEM-X bistatic mode $K(B_n) = 2\pi B_n/\lambda R \sin\theta = 2\pi/HOA$, where B_n is the component of the baseline perpendicular to the line of sight, λ is the wavelength, R is the slant range distance, and θ is the incidence angle.

Often an extinction coefficient, κ_{eff}, is used to define the attenuation through a homogeneous vegetation layer without gaps. From Equation (5), κ_{eff} is obtained by the definition $\frac{2\kappa_{eff}h(V)}{\cos\theta_i} = \beta V$. With h expressed by means of the allometric expression in Equation (2) and β expressed by Equation (6), it is obtained:

Remote Sens. **2013**, *5* **5583**

$$\kappa_{eff}(V) = -\frac{cos\theta_i}{2h(V)} ln\,[1 - \eta(V)\big(1 - e^{-\alpha h(V)}\big)] \tag{8}$$

The expression for κ_{eff} in Equation (8) illustrates how the extinction coefficient is dependent on stem volume, forest height, area-fill, as well as temperature, humidity, *etc.*, through the attenuation coefficient α. For dense vegetation layers, $\eta(V) = 1$ and $\kappa_{eff}(V)$ tends to $\alpha\, cos\theta/2$.

Equation (4) can be rewritten as

$$\tilde{\gamma} = \frac{\gamma_{veg}\tilde{\gamma}_{vol} + \gamma_{gr}m}{1 + m} \tag{9}$$

where

$$m = \frac{\sigma_{gr}^0 e^{-\beta V}}{\sigma_{veg}^0 (1 - e^{-\beta V})} \tag{10}$$

is the ratio describing the relative amount of ground scattering compared to volume scattering (ground-to-volume ratio).

The TDM coherence, γ, and the estimated height, z_{est} are defined as

$$\gamma = |\tilde{\gamma}| \quad z_{est} = -\frac{arg[\tilde{\gamma}]}{2\pi} HOA \tag{11}$$

where arg stands for argument of the complex-valued term within brackets.

The IWCM has been used for C-band data in order to derive stem volume from coherence, and it was found to be suitable for the retrieval of stem volume at several test sites in Sweden and Finland. For the one-day repeat-pass interval of the European Remote Sensing ERS-1/2 mission, the coherence for stable winter conditions was found to be useful for stem volume estimation [13,14,36,37], while the interferometric phase height was found unstable [38]. As will be illustrated below, it is instead primarily the interferometric phase height which has the highest sensitivity to biomass for TanDEM-X.

3.2. Random Volume over Ground Model

The Random Volume over Ground, RVoG, was introduced for studies of polarimetric SAR interferometry, PolInSAR, [17]. The model can be obtained by excluding gaps in the analysis of IWCM, which means one less unknown parameter. This means that $\eta \equiv 1, \beta V = \alpha h$, and $\kappa_{eff} = \alpha cos\theta/2$ is the extinction coefficient. The ground-to-volume ratio m then changes to

$$m = \frac{\sigma_{gr}^0 e^{-\alpha h}}{\sigma_{veg}^0 (1 - e^{-\alpha h})} \tag{11}$$

and the expression in Equation (9) for the complex coherence is used for cases when the temporal decorrelation can be neglected. RVoG has shown to be useful in PolInSAR height estimation without the need of training stands as long as the extinction coefficient can be assumed polarization-independent and a polarization combination with no ground contribution can be found [5,16]. RVoG was also used for the single polarized case, e.g., [5,6]. When m can be assumed negligible and a lidar DTM is available, the tree height and the extinction coefficient can be estimated without training stands. Biomass can then be estimated by means of an allometric relationship [39,40]. With training stands, the RVoG model parameters and model properties can be studied in detail.

3.3. Penetration Depth Model

If a medium is dense and attenuation is such that ground scattering is negligible, *i.e.*, $exp[-\alpha h] << 1$, then $m \approx 0$ and only volume decorrelation remains, if $\gamma_{veg} = 1$, see Equations (7), (9), and (12):

$$\gamma_{vol} = \frac{\alpha}{\alpha - jK(B_n)} \frac{e^{-jK(B_n)h} - e^{-\alpha h}}{1 - e^{-\alpha h}} \approx e^{-jK(B_n)\left[h - \frac{1}{\alpha}\right]} \tag{12}$$

If $K(B_n)/\alpha << 1$ then the estimated interferometric phase height is approximated as $h - 1/\alpha$, while if $K(B_n)/\alpha \approx 1$ a correction has to be introduced depending on α and HOA, and $h - 1/\alpha_{eff}$ represents the interferometric phase height. Setting $h(V) - 1/\alpha_{eff}$ equal to the TDM height gives a simple model for the TDM height, H_{TDM}. This model is denoted as PD. Neglecting that the approximations made are not valid for small biomass, the biomass could be estimated from

$$B = 0.21(H_{TDM} + \alpha_{eff}^{-1})^{2.17} \tag{13}$$

3.4. Estimation of Model Parameters

The IWCM contains six unknown parameters ($\sigma_{gr}°$, $\sigma_{veg}°$, γ_{gr}, γ_{veg}, α, and β) that need to be determined (Table 2). The traditional estimation approach consists of a least-squares regression to reference data of the forest variable in the model and corresponding observations of backscatter and complex coherence. This is referred to as "training" the model. The models in Equations (4) and (5) assumed to be formulated as dependent on stem volume, V, have been transformed in this study to a dependence on biomass by means of Equation (1). The allometric expression in Equation (2), $h(V)$, and the BEF are assumed to be known *a priori*. In addition, there is a need for knowledge of the ground phase or, in an equivalent manner, of the TDM height for non-forest areas nearby the test site. All model parameters are estimated by fitting the IWCM to the sets of backscatter, coherence and InSAR height observations from the TanDEM-X dataset and the corresponding LiDAR-based values of biomass forming a training dataset. In addition, the estimates of vegetation ratio, which mimic the area-fill factor are used. It should be noted that the area-fill does not correspond exactly to the LiDAR measured vegetation ratio. 1-η represents the fraction of gaps in the vegetation, and the gaps have to be larger for microwaves with longer wavelengths than for the LiDAR to propagate freely. The dielectric properties of the surrounding vegetation will also affect the wave propagation and transmission. Consequently, the area-fill can be expected to be higher than the LiDAR-based vegetation ratio and also vary with environmental conditions.

Table 2. Input and output of the model training phase for the three models.

Model	Input	Output
IWCM	Parameters of satellite look geometry, DTM, $h(B)$, BEF, biomass , $\sigma°_{TDM}$, γ_{TDM} and h_{TDM} of training stands	$\sigma_{gr}^{o}, \sigma_{veg}^{o}$, γ_{gr}, γ_{veg}, α, β,
RVoG	Parameters of satellite look geometry, DTM, $h(B)$, BEF, biomass , $\sigma°_{TDM}$, γ_{TDM} and h_{TDM} of training stands	$\sigma_{gr}^{o}, \sigma_{veg}^{o}$, γ_{gr}, γ_{veg}, α,
PD	DTM, $h(B)$, biomass and h_{TDM} of training stands	α_{eff},

In the RVoG model, vegetation gaps are excluded, which reduces the number of unknowns to five since the β-parameter is eliminated (Table 2). The unknown model parameters were estimated in a similar manner as in the case of the IWCM.

In the PD model, only the α_{eff}-parameter is unknown. It is noted that this model is a simple function of the TDM height. For estimating biomass, the dependence on the allometric relationship, which is indirect for the other models, is very clear in this case. The estimate of α_{eff} is obtained by fitting the model to measurements of InSAR height for the training stands (Table 2).

3.5. Model Training and Inversion Procedure

The 201 forest stands with a size of at least 1 ha were divided in two datasets, one for training and one for validation. Then the datasets were interchanged and training and validation was repeated.

As a first-order approximation σ_{gr}^0 and σ_{veg}^0 in Equation (5) were estimated from the TDM backscatter measurements of the 20 stands with the smallest and largest biomass. Because of the large scatter of the measured backscatter with respect to biomass, it was assumed that a constant $\beta = 0.007$ would return a realistic approximation of the two backscatter model parameters [41]. To correct for errors in the regression, σ_{gr}^0 was manually adapted such that the model curve of Equation (4) would have been within the range of observations. The estimates of the coherence model parameters in Equation (5) were then obtained. γ_{gr} was determined as the mean of the TDM coherence for the ten stands with highest coherence and with biomass close to zero, while γ_{veg} was assumed equal to γ_{gr}.

In the case of IWCM, the values of the two remaining unknown parameters α and β were estimated by means of least-squares fitting of the model to the observations in the training dataset using the Levenberg-Marquardt method together with the constraint that the area-fill (expressed by the LiDAR vegetation ratio) is <1. Since the TDM height was the most accurate SAR observation (Figure 3), the two unknown model parameters were estimated by minimizing the quadratic difference between the InSAR height predicted by the model (z_{est}) and the corresponding observed TDM heights ($H_{TDM,i}$):

$$min \sum_i (z_{est}(B_i, \alpha, \beta) - H_{TDM,i})^2 \tag{15}$$

In Equation (15), B_i represents the LiDAR-based value of biomass of training stand i. A minimization similar to Equation (15) was also done for the coherence to determine a correction to γ_{veg}. The effect of the fine tuning of this parameters did not have any significant effect on the estimate of z_{est} and therefore on the estimation of the remaining model parameters. Once the model parameters α and β were estimated, the model could be inverted to obtain estimates of biomass B_j for each forest stand j in the validation dataset by estimating the roots of the expression $z_{est}(B_j, \alpha, \beta) - H_{TDM,j} = 0$.

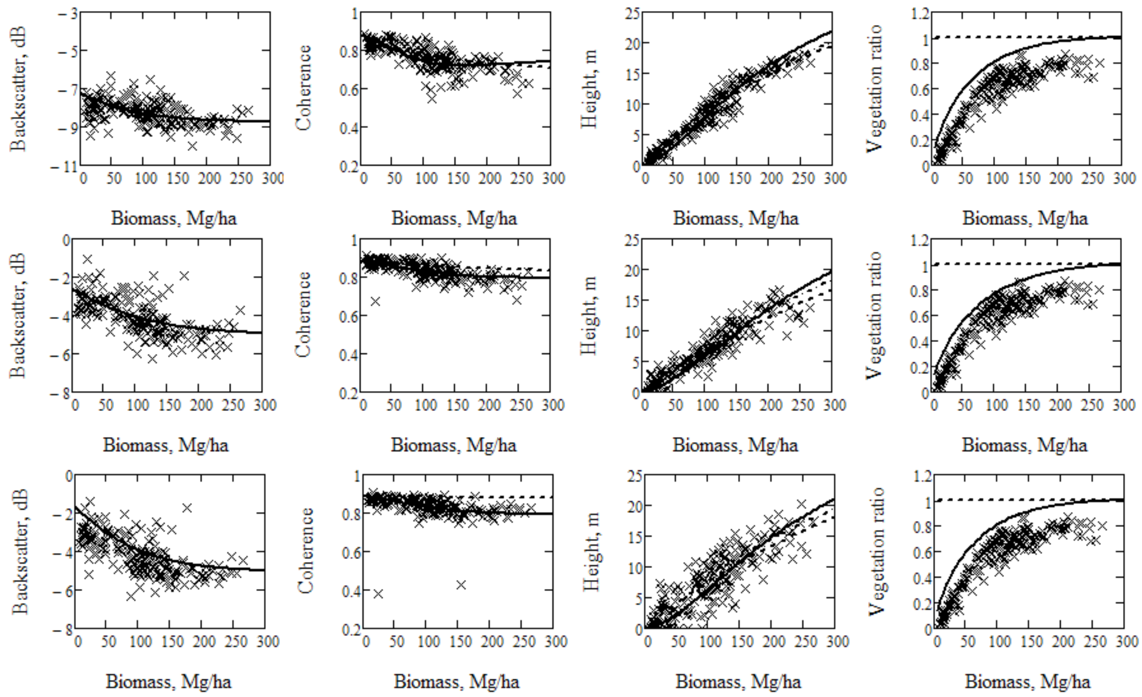
In the case of the RVoG model, m can sometimes be assumed negligible, but here m was included in the analysis since the contribution from ground is important for forest heights up to ≈15 m (depending on the attenuation observed in summer and in winter). The parameter α was determined by fitting the model to TDM height by least-squares fitting. Biomass was then estimated in the same way as for the IWCM.

The simplest model, PD, is a model for the TDM height expressed by the allometric expression for the forest height and an attenuation coefficient for the penetration, $h(B) - \alpha_{eff}^{-1}$. From the training stands an estimate for α_{eff}, was obtained. The biomass for the validation stands was then estimated from Equation (14).

4. Modeling and Biomass Retrieval Results

Results concerning the estimation of the model parameters are presented first (Section 4.1). Then, retrieval results are presented in the form of RMSE values of biomass estimation (Section 4.2). RMSE can be considered a measure of the usefulness of the different models as well as a measure of the biomass estimation accuracy.

Figure 5. Scatter plots and model results for three TDM acquisitions, in order from top to bottom 2012-02-01 (HOA = 80 m), 2011-11-23 (HOA = −185 m), and 2012-08-13 (HOA = 358 m). IWCM results for backscatter, coherence, TDM height and vegetation ratio (area-fill): solid line, RVoG results for coherence and TDM height: dashed line, and PD results for TDM height: dash-dotted line.



4.1. Model Properties and Parameters Estimates

Figure 5 illustrates model-based backscatter, coherence, height and area-fill (vegetation ratio) as a function of biomass with respect to observations from the TanDEM-X dataset and corresponding LiDAR-based biomass. Three examples are shown representing a winter (2012-02-01), a fall (2011-11-23), and a summer (2012-08-13) acquisition as well as different HOA: 80 m, −185 m and 358 m, respectively. Backscatter and coherence presented weak sensitivity with respect to biomass. The backscatter was characterized by a dynamic range of 2–3 dB, with a clear decreasing trend for increasing biomass which can be connected with relatively rough and wet conditions of the soil.

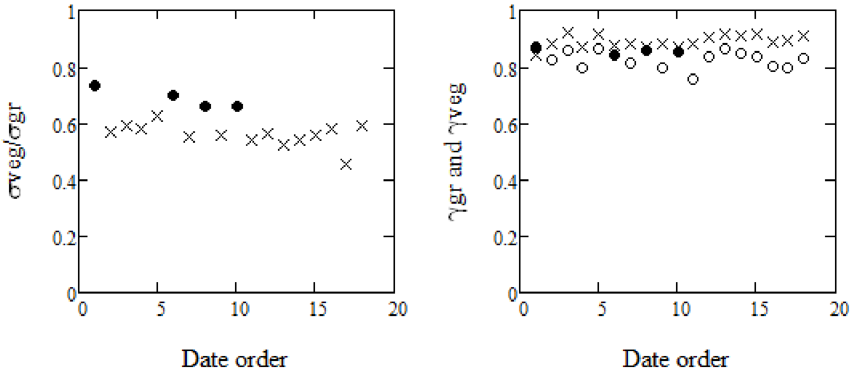
Coherence was mostly above 0.7 and showed a decreasing dynamic range for increasing HOA (The two extreme coherence values in the bottom line are associated with stands crossed by a high voltage power line). Compared to backscatter and coherence, the InSAR heights showed the strongest sensitivity to biomass. The association strength between TDM heights and biomass however decreased for increasing HOA. The association between vegetation ratio and biomass was also strong. Saturation of the vegetation ratio slightly above 100 Mg/ha could be observed. The three models could reproduce the trend in the observations regardless of the observable and the acquisition date (Figure 5). Some discrepancies occurred in the range of the highest biomass values.

Estimates of the model parameters for each TDM acquisition are presented in Table 3 and illustrated in Figures 6, 7 and 8. Figure 6 illustrates the IWCM estimates of the two backscatter and the two coherence parameters for each acquisition. The ratio $\sigma^o_{veg}/\sigma^o_{gr}$ indicates stronger sensitivity of the backscatter to biomass for shallow incidence angles (filled circles in Figure 6a). The backscatter ratio and β determine the biomass for which the backscatter from ground and vegetation are similar, and from this criterion as well as $\exp[-\alpha h(V)] \approx 0.15$ it is found that the ground has an influence up to a biomass of 50–100 Mg/ha and forest heights of 12–17 m. The four cases with shallow incidence angle of 41° were also those with |HOA| ≤ 80 m. For these cases, the coherence was relatively noisy, in particular on 2011-06-04, whereas the coherence parameters γ_{veg} and γ_{gr} were almost equal, as expected for cases without temporal decorrelation (crosses and filled circles in Figure 6b). Images with |HOA| ≥ 172 m presented instead a slight difference with $\gamma_{veg} < \gamma_{gr}$. This difference could not be explained as an effect of ATB length or specific environmental conditions though.

Table 3. TDM acquisition date, model parameters α, β, and modeled TDM height in m at 150 Mg/ha (H150) for IWCM, and α for RVoG and PD.

#	Date	α IWCM	β IWCM	H150 IWCM	α RVoG	α_{eff} PD
1	2011-06-04	0.20	0.0093	15.0	0.14	0.17
2	2011-11-23	0.12	0.0056	9.4	0.09	0.10
3	2011-12-26	0.11	0.0053	8.9	0.08	0.10
4	2012-01-17	0.15	0.0070	11.2	0.11	0.12
5	2012-01-28	0.10	0.0049	8.5	0.07	0.09
6	2012-02-01	0.15	0.0070	12.0	0.11	0.12
7	2012-02-08	0.12	0.0056	9.2	0.09	0.10
8	2012-02-12	0.17	0.0080	13.1	0.12	0.14
9	2012-02-19	0.14	0.0068	10.9	0.10	0.11
10	2012-02-23	0.16	0.0078	12.8	0.12	0.13
11	2012-03-01	0.14	0.0066	10.4	0.10	0.11
12	2012-03-12	0.12	0.0058	9.6	0.09	0.10
13	2012-03-23	0.13	0.0061	9.8	0.09	0.10
14	2012-05-28	0.16	0.0076	11.8	0.12	0.12
15	2012-07-22	0.14	0.0068	10.8	0.10	0.11
16	2012-08-02	0.14	0.0068	10.9	0.10	0.11
17	2012-08-13	0.14	0.0066	10.3	0.10	0.11
18	2012-08-24	0.14	0.0065	10.6	0.10	0.11

Figure 6. (**a**) Estimates of $\sigma^o_{veg}/\sigma^o_{gr}$ (*cf.* Equation (5)) (× and • for incidence angles of 34° and 41°, respectively) and (**b**) estimates of γ_{gr} (×) and γ_{veg} (*cf.* Equation (4)) (o and • for |HOA| ≤ 80 m or above, respectively) *versus* TDM acquisition date (date order).



(**a**) (**b**)

Figure 7 illustrates the range of modeled TDM heights with respect to biomass being delimited by the curves corresponding to the maximum (2011-06-04) and minimum (2012-01-28) model estimated TDM heights. For biomass of 150 Mg/ha Table 3 lists the corresponding modeled TDM height, which varied between 8.5 m and 15.0 m. Such variability must be compensated for by a model-based approach to retrieve biomass in order to correctly interpret the dependence of TDM height on biomass.

Figure 7. Range of model-based estimates of TDM heights *versus* biomass. The range is delimited by the curves (solid for IWCM, dashed for PD model) corresponding to the maximum and minimum of model estimated TDM heights *versus* biomass.

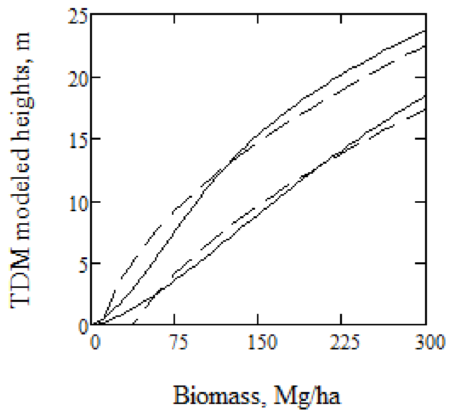
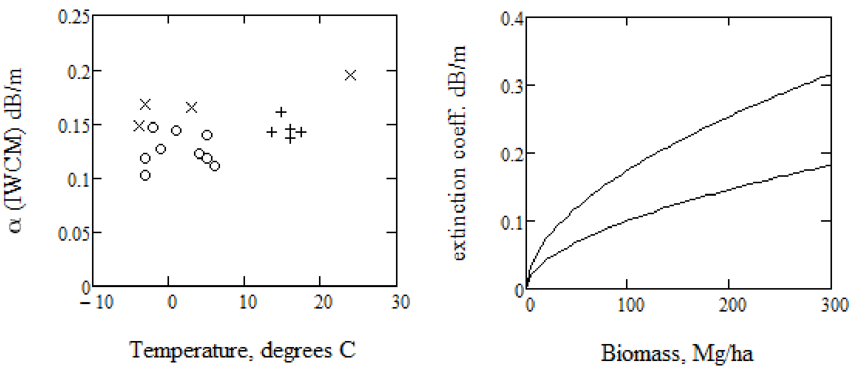


Figure 8 illustrates the estimates of the α parameter in the case of the IWCM with respect to temperature and the corresponding range of the extinction coefficient κ_{eff} delimited by the two curves for the smallest and largest estimates of κ_{eff} using Equation (8). The attenuation in the vegetated fraction (α) did not present clear dependence on temperature (Figure 8), nor we could identify any dependence on HOA. This result is in contrast with the assumption that α should be lower in case of sub-zero temperatures. However, since the few acquisitions characterized by frozen environmental conditions took place when temperature was close to 0 °C, it is not possible to conclude that such an assumption is incorrect. The increase of the extinction coefficient with biomass shown in Figure 8b also means an increase with area-fill (*i.e.*, LiDAR vegetation ratio). The range 0.1–0.3 dB/m confirm previous estimates derived in [5,6].

Figure 8. (**a**) Estimates of α in the case of the IWCM with respect to temperature in °C (× for |HOA| ≤ 80 m, o for |HOA| ≈ 180 m, and + for |HOA| ≈ 330 m) and (**b**) corresponding range of the extinction coefficient with respect to biomass as delimited by the curves corresponding to the maximum and minimum α (acquisition dates 2011-06-04 and 2012-01-28, respectively).



(**a**) (**b**)

4.2. Biomass Estimation

The biomass of each stand in the validation dataset was estimated based on the trained models and compared with the LiDAR-based estimates of biomass. The biomass retrieval accuracy expressed in the form of the RMSE between the TDM biomass and the reference biomass is given in Table 4 for the different models. The estimates obtained with the IWCM presented slightly better accuracy compared to the retrieval based on the other models. Nonetheless, the difference between the models is relatively small, in particular between IWCM and PD.

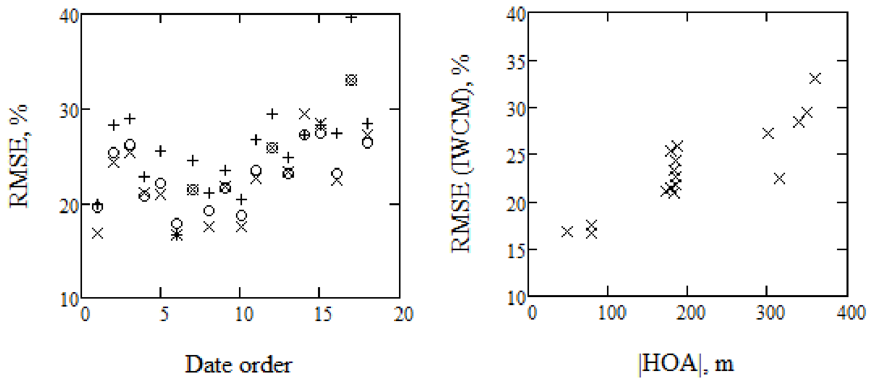
Table 4 indicates that winter-time data with long-lasting frozen conditions (February–March 2012, see also Figure 4) was more suitable for retrieval compared to other acquisitions. In Figure 9 the RMSEs are plotted with respect to date and to |HOA|. Figure 9b shows a clear difference depending on whether |HOA| was ≤ 80 m or above; in the former cases the retrieval RMSE was much lower. Since the TDM height is determined by the InSAR phase, and 2π corresponds to a TDM height = HOA, a certain phase error will have increased effect on TDM height for increasing HOA.

The RMSE reported in Table 4 was obtained for the 201 forest stands larger than 1 ha. Taking into account smaller forest stands as well, resulted in larger retrieval errors as shown for the case of the IWCM-based retrieval in Table 5 for the two cases with lowest and highest RMSE, *i.e.*, for the TDM acquisitions on 2012-02-01 (HOA = 80 m) and 2012-08-13 (HOA = 358 m) respectively. The number of forest stands in the validation dataset increased to 315 when using a threshold on stand size of 0.5 ha (mean biomass of 110 Mg/ha). It further increased to 403 when the threshold was 0.25 ha (mean biomass of 112 Mg/ha).

Table 4. Single-image relative RMSE for the biomass estimated using the models IWCM, RVoG, and PD.

#	Date	RMSE % IWCM	RMSE % RVoG	RMSE % PD
1	2011-06-04	16.8	19.9	19.5
2	2011-11-23	24.3	28.2	25.4
3	2011-12-26	25.4	28.9	26.2
4	2012-01-17	21.1	22.8	20.7
5	2012-01-28	20.8	25.4	22.0
6	2012-02-01	16.7	16.7	17.9
7	2012-02-08	21.3	24.5	21.4
8	2012-02-12	17.5	21.1	19.1
9	2012-02-19	21.7	23.5	21.7
10	2012-02-23	17.5	20.5	18.7
11	2012-03-01	22.5	26.8	23.5
12	2012-03-12	25.9	29.3	25.8
13	2012-03-23	23.3	24.8	23.0
14	2012-05-28	29.5	27.2	27.2
15	2012-07-22	28.4	28.3	27.4
16	2012-08-02	22.5	27.4	23.1
17	2012-08-13	33.0	39.7	33.1
18	2012-08-24	27.2	28.3	26.4
	Mean RMSE	23.1	25.7	23.4

Figure 9. (**a**) Illustrating relative RMSE variation with TDM acquisition date (date order) × for IWCM, + for RVoG, o for PD. (**b**) Illustrating relative RMSE IWCM *versus* HOA.



(**a**) (**b**)

Table 5. Relative RMSE in the case of the IWCM as a function of minimum forest stand size. The two acquisitions with the smallest and largest RMSE are shown for simplicity.

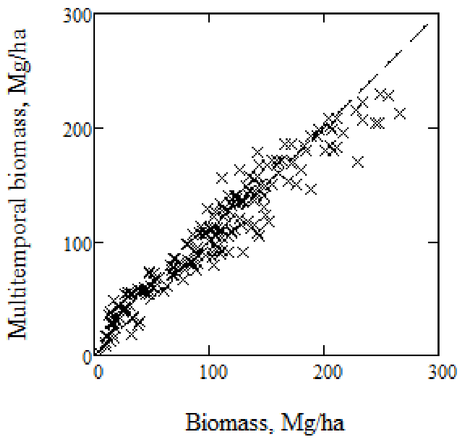
Date	RMSE % Stands ≥ 1 ha	RMSE % Stands ≥ 0.5 ha	RMSE % Stands ≥ 0.25 ha
2012-02-01	16.7	20.2	22.6
2012-08-13	33.0	35.2	38.3

It was previously observed (Figure 7) that the TDM height varied considerably between different acquisitions; however, for each acquisition the biomass of each stand, *i*, was rather constant thanks to the compensation embedded in the models. An important aspect of satellite observations is the possibility to further exploit repeated observations to reduce the uncertainties in the single-image estimates of the biomass [13,14,41]. A multitemporal combination of the different biomass estimates, *l*, is here proposed in which the weighting factor are based on the HOA since it has been shown above how the noise in the TDM height is increasing with HOA, *cf.* Figures 3, 5, and 9. The multitemporal estimate of the biomass of each stand, Bmt_i, is defined as

$$Bmt_i = \sum_{l=1}^{18} \frac{HOA_l^{-2}}{\sum_1^{18} HOA_l^{-2}} B_{l,i} \tag{14}$$

The multitemporal biomass determined by IWCM resulted in a RMSE of 16.5% or 17.3 Mg/ha and $R^2 = 0.93$, for stands > 1 ha (Figure 10).

Figure 10. Scatterplot of biomass derived from a multi-temporal combination of 18 TanDEM-X InSAR pairs with respect to LiDAR-based biomass for 201 stands larger than 1 ha.



5. Discussion

The study on above-ground dry biomass retrieval with TanDEM-X interferometry follows a number of investigations on remote sensing data and retrieval techniques at the Remningstorp test site. Several of these studies have dealt with the use of multitemporal spaceborne and airborne SAR data. Interferometric SAR datasets acquired at C-band with one-day temporal separation were evaluated to retrieve forest stem volume in [13] achieving RMSE of 27%. In [27] L- and P-band SAR backscatter was used for biomass estimation with RMSE between 31% and 46% for L-band and 18% and 27% for P-band. In [42] a biomass model for P-band with training data from Krycklan, a test site in northern Sweden, was used, and validation data from Remningstorp resulted in RMSE 22–33%. For CARABAS VHF-band SAR single image estimates from different flight directions resulted in RMSE 11–25% [43]. However, it should be noted that RMSE is not the only way to compare different methods. In a first report on the use of TanDEM-X data from Remningstorp with the goal of biomass estimation [40], two monostatic acquisitions were studied with a delay of 3 s, which complicates the analysis due to temporal decorrelation. For a 95% confidence interval four biomass classes up to 250 Mg/ha could be distinguished.

The RMSE obtained in the present study (16%) demonstrates the significant contribution of multitemporal TanDEM-X interferometric data to the quantification of forest aboveground biomass. Similar accuracy has earlier been found by means of other X-band satellite InSAR measurements; namely with SRTM [7], and TanDEM-X [44]. Since a dependence on tree species is expected [7,45], a knowledge about the species based on other data could improve RMSE. A shorter time difference between LiDAR measurements (August 2010) and TDM measurements (June 2011–August 2012) also could improve the RMSE, but the time difference is relatively short and the growth (approximately 5 Mg/ha/year) has not been compensated for. A local value of *BEF i.e.*, 0.512 was used, due to lack of a more general value. In the literature somewhat higher values can be seen, see e.g., [46], and when testing *BEF = 0.58* (determined from proportions of tree species of the seven 80 m × 80 m stands and BEF-factors for pine, spruce and deciduous according to [27]) the RMSE values were slightly changed (varying from 16.4% to 31.1%) but the mean value of RMSE for IWCM was unchanged at 23.1%. Since BEF is used for training as well as validation stands, the effect of BEF is small.

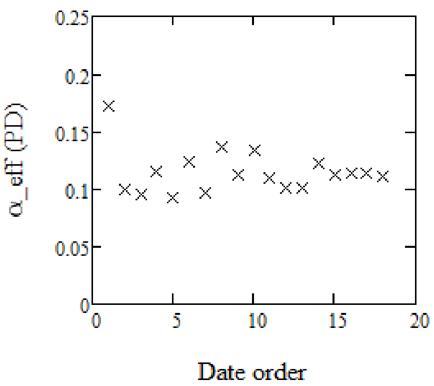
There is a close relationship between biomass and TDM height, which is determined by the forest height and the penetration depth. The latter is related to the vegetation density. The low extinction values, *i.e.*, <0.3 dB/m, which can be assumed to be related to gaps in the vegetation down to different levels, make the X-band microwaves to propagate up to 6–10 m into the vegetation according to the Penetration Depth model in the studied cases.

In the present study data the tendency to "saturate" at biomass > 200 Mg/ha in Figure 10 is probably not caused by the saturation effect observed in backscatter studies, since the measurement is based on the volume decorrelation and penetration depth of the upper vegetation layer, but may instead be caused by changes in the forest density, *i.e.*, number and dimension of gaps (Figure 2). This should be further investigated. The deviation from the reference line below 50 Mg/ha, could be related to the sensitivity of TDM data to specific conditions of the ground. In particular, the measured backscatter showed some deviations from the model in this region (Figure 5). However, it should also be noted that the reference biomass was obtained from LiDAR measurements which in turn were characterized by their own set of uncertainties and errors. The forest height varied more than 10 m in height in the biomass range < 50 Mg/ha (Figure 2) indicating certain complexity of the TDM and LiDAR metrics in this interval of biomass.

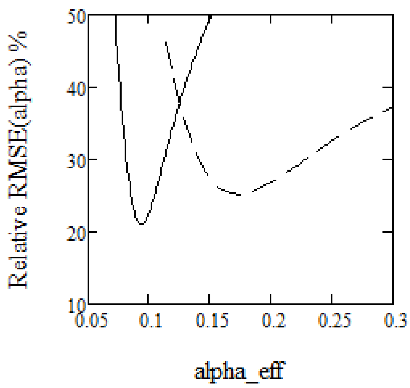
The number of training stands used in this study was half of the total number of stands, which in practical terms might be difficult to achieve when the aim is to map larger areas. With fewer training stands, it is assumed that the representativeness of the model parameters decreases in a manner related to forest homogeneity and measurement noise. To verify the impact of the number of stands used for model training on the retrieval, an extreme case was considered of a training dataset formed by only five stands chosen with approximate intervals of 50 Mg/ha. The RMSE for the best performing acquisition (2012-02-01) increased from 16.7% to 18.1% for stands larger than 1 ha, illustrating the possibility to limit the training dataset to a small number of training stands.

So far the accuracy of biomass estimation has been analyzed using models with parameters trained by stands having known properties. It would be very valuable if training stands could be avoided. The PD model is appealing since it contains only one unknown and the retrieval RMSE was close to the result obtained with the IWCM (Table 4), in spite of the model approximation developed with emphasis on mature stands. Figure 11a shows that the estimate of the α_{eff} parameter in the PD model obtained with training stands was relatively constant with one exception (see also Table 3). If α_{eff} can be estimated from known conditions it may be possible to give a first-order estimate of forest height and biomass depending on how sensitive the estimates are for the correct α_{eff}. A preliminary sensitivity analysis of the retrieval RMSE of biomass with respect to α_{eff} is illustrated in Figure 11b and should be further investigated, when a wider range of environmental conditions have been studied.

Figure 11. (**a**) Estimates of α_{eff} (PD model) *versus* acquisition date in date order and (**b**) sensitivity of the relative RMSE using PD to α_{eff} for the two acquisitions with highest (dashed line, 2011-06-04) and lowest (solid line, 2012-01-28) α_{eff}-values.



(**a**)



(**b**)

An analysis of two TanDEM-X acquisitions from a spruce dominated area in southeast Norway used a linear function [44] between biomass and TDM height without intercept, $B \propto h_{TDM}$. The relative RMSE at the stand level was 19% using a biomass increase of 14 Mg/ha per m increase of TDM height. Such a linear relation is in line with our results, *cf.* Figure 5. However, as shown by the variation of H150 (TDM height at 150 Mg/ha) there is a variation between the different acquisitions. Assuming a linear relation, $B = \chi\, h_{TDM}$, results in χ varying between 10.3 and 16.4, with a mean of 13.3 Mg/ha per m TDM height.

The result in [44], and the analysis of PD show the importance of extending the TanDEM-X analysis to a wider range of environmental conditions and to investigate if a fixed value of a single parameter model results in a sufficiently high accuracy over a wide range of conditions. If so, the use of training stands can be avoided.

6. Conclusions

Eighteen interferometric TanDEM-X bistatic image pairs (VV-polarization) acquired between June 2011 and August 2012 over the test site of Remningstorp, situated in southern Sweden, have been studied in order to determine the potential of model-based above-ground dry biomass estimation. LiDAR-based estimates of biomass and vegetation ratio, acquired in August 2010 [19], were used as reference data. In order to interpret the TanDEM-X observations, the Interferometric Water Cloud

Model (IWCM), was primarily used. The Random Volume over Ground model (RVoG), was also studied, and a new model based on the penetration depth (PD) concept, was introduced. All the used models are based on physical principles, but in a significantly simplified form. Therefore, it should be stressed, that the parameters represent simplifications of complex phenomena. However, the physical relevance makes it possible to relate the variation of the parameters to other measurements and to environmental influence.

The relative RMSE of biomass associated with a retrieval based on the IWCM for forest stands ≥ 1 ha varied between 17% and 33% (relative to the mean value of 105 Mg/ha), with the best estimates obtained for small HOA. The relative RMSE for biomass retrieval based on the RVoG model varied between 17% and 40%. The relative RMSE for biomass retrieval based on the simple PD model was between 18% and 33%. Taking the mean of all 18 TDM estimates of stand biomass weighted inversely proportional to HOA^2 resulted in an RMSE of 16% in the IWCM case for forest stands larger than 1 ha. The environmental influence (temperature, humidity, rain *etc.*) on the TDM height resulted in a variation from 8.9 m to 15.2 m at a biomass of 150 Mg/ha, and this variation has to be taken care of by the model analysis.

The presented analysis demonstrates that TanDEM-X InSAR data together with an accurate high-resolution DTM, a fairly straightforward allometric expression, and forest stands for training model parameters, have a potential to estimate above-ground dry biomass with high accuracy in the case of forest conditions like those in Remningstorp. The results obtained by means of the bistatic TanDEM-X (VV-pol) are among the best remote sensing estimates of biomass obtained so far from Remningstorp. In a more general perspective, these results indicate the suitability of TanDEM-X data to retrieve boreal forest biomass with accuracy and spatial resolution as required by forest inventories, *cf.* [2].

Acknowledgments

ESA and the BIOSAR2010 team are gratefully acknowledged for supporting the collection of forest inventory data. Leif Eriksson, Gustaf Sandberg, and Henrik Persson are acknowledged for support and discussions, and Heather Reese for proofreading the manuscript. TanDEM-X data were distributed through proposal XTI_VEG0376 3D Forest Parameter Retrieval from TanDEM-X Interferometry. This work was partly supported by the Swedish National Space Board.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Houghton, R.A. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* **2005**, *11*, 945–958.
2. Houghton, R.A.; Hall, F.; Goetz, S.J. Importance of biomass in the global carbon cycle. *J. Geophys. Res.: Biogeosci.* **2009**, *114*, doi:10.1029/2009JG000935.
3. Ouchi, K. Recent trend and advance of synthetic aperture radar with selected topics. *Remote Sens*. **2013**, *5*, 716–807.

4. Balzter, H.; Luckman, A.; Skinner, L.; Rowland, C.; Dawson, T. Observations of forest stand top height and mean height from interferometric SAR and LiDAR over a conifer plantation at Thetford Forest, UK. *Int. J. Remote Sens*. **2007**, *28*, 1173–1197.
5. Hajnsek, I.; Kugler, F.; Lee, S.; Papathanassiou, K.P. Tropical-forest-parameter estimation by means of Pol-InSAR: The INDREX-II campaign. *IEEE Trans. Geosci. Remote Sens*. **2009**, *47*, 481–493.
6. Praks, J.; Antropov, O.; Hallikainen, M.T. LIDAR-Aided SAR interferometry studies in boreal forest: Scattering phase center and extinction coefficient at X- and L-Band. *IEEE Trans. Geosci. Remote Sens*. **2012**, *50*, 3831–3843.
7. Solberg, S.; Astrup, R.; Gobakken, T.; Næsset, E.; Weydahl, D.J. Estimating spruce and pine biomass with interferometric X-band SAR. *Remote Sens. Environ.* **2010**, *114*, 2353–2360.
8. Krieger, G.; Moreira, A.; Fiedler, H.; Hajnsek, I.; Werner, M.; Younis, M.; Zink, M. TanDEM-X: A satellite formation for high-resolution SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3317–3341.
9. Weydahl, D.J.; Sagstuen, J.; Dick, Ø.B.; Rønning, H. SRTM DEM accuracy assessment over vegetated areas in Norway. *Int. J. Remote Sens.* **2007**, *28*, 3513−3527.
10. Neeff, T.; Dutra, L.V.; Santos, J.R.; Freitas, C.C.; Araújo, L.S. Tropical forest measurement by interferometric height modeling and P-band backscatter. *For. Sci.* **2005**, *51*, 585−594.
11. Askne, J.; Dammert, P.; Ulander, L.M.H.; Smith, G. C-band repeat-pass interferometric SAR observations of forest. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 25–35.
12. Askne, J.; Dammert, P.; Ulander, L.M.H.; Smith, G. Multi-temporal repeat pass SAR interferometry of boreal forests. *IEEE Trans. Geosci. Remote Sens*. **2003**, *41*, 1540–1550.
13. Askne, J.; Santoro, M. Experiences in Boreal Forest Stem Volume Estimation from Multitemporal C-Band InSAR. In *Recent Interferometry Applications in Topography and Astronomy*; Padron, I., Ed.; InTech: Morn Hill, Winchester, UK, 2012; 169–194. Available online: http://cdn.intechopen.com/pdfs/33100/InTech-Experiences_in_boreal_forest_stem_volume_estimation_from_multitemporal_c_band_insar.pdf (accessed on 25 October 2013).
14. Santoro, M.; Askne, J.; Smith, G.; Fransson, J.E.S. Stem volume retrieval in boreal forests from ERS-1/2 interferometry. *Remote Sens. Environ.* **2002**, *81*, 19–35.
15. Cloude, S.R.; Papathanassiou, K.P. Polarimetric SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1551–1565.
16. Papathanassiou, K.P.; Cloude, S.R. Single-baseline polarimetric SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 2352–2363.
17. Treuhaft, R.N.; Madsen, S.N.; Moghaddam, M.; van Zyl, J.J. Vegetation characteristics and underlaying topography from interferometric data. *Radio Sci.* **1996**, *31*, 1449–1495.
18. Ulander, L.M.H.; Gustavsson, A.; Dubois-Fernandez, P.; Dupuis, X.; Fransson, J.E.S.; Holmgren, J.; Wallerman, J.; Eriksson, L.; Sandberg, G.; Soja, M. BIOSAR 2010–A SAR Campaign in Support to the BIOMASS Mission. In Proceedings of 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vancouver, BC, Canada, 24–29 July 2011; pp. 1528–1531.

19. Ulander, L.M.H.; Gustavsson, A.; Flood, B.; Murdin, D.; Dubois-Fernandez, P.; Depuis, X.; Sandberg, G.; Soja, M.J.; Eriksson, L.E.B.; Fransson, J.E.S.; *et al. BIOSAR 2010 Technical Assistance for the Development of Airborne SAR and Geophysical Measurements during the BioSAR 2010 Experiment; Final Report*; ESA contract no. 4000102285/10/NL/JA/ef; ESA: Noordwijk, The Netherlands, 2011; p. 212. Available online: http://earth.esa.int/campaigns/DOC/BioSAR_2010_final_report_v1.0.pdf (accessed on 25 October 2013).
20. ESA. *Report for Mission Selection: Biomass*; European Space Agency: Noordwijk, The Netherlands, 2012.
21. Ahti, T.; Hämet-Ahti, L.; Jalas, J. Vegetationzones and their sections innorthwestern Europe. *Annales Botanici Fennici* **1968**, *5*, 169–211.
22. Wikström, P.; Edenius, L.; Elfving, B.; Eriksson, L.O.; Lämås, T.; Sonesson, J.; Öhman, K.; Wallerman, J.; Waller, C.; Klintebäck, F. The Heureka forestry decision support system: An overview. *Math. Comput. For. Nat.-Resour. Sci.* **2011**, *3*, 87–94.
23. Petersson, H. *Biomassafunktioner för Trädfraktioner av Tall, Gran och Björk i Sverige*; Department of Forest Resource Management, Swedish University of Agricultural Sciences: Umeå, Sweden, 1999.
24. Marklund, L.G. *Biomassafunktioner för Tall, Gran och Björk i Sverige*; Institutionen för Skogstaxering, Sveriges Lantbruksuniversitet: Umeå, Sweden, 1988.
25. Näslund, M. *Funktioner och Tabeller för Kubering av Stående Träd. Tall, Gran och Björk i Södra Sverige samt Hela Landet*; Statens Skogsforskningsinstitut: Stockholm, Sweden, 1947.
26. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **2002**, *80*, 88–99.
27. Sandberg, G.; Ulander, L.M.H.; Fransson, J.E.S.; Holmgren, J.; Toan, T.L. L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest. *Remote Sens. Environ.* **2011**, *115*, 2874–2886.
28. Jalkanen, A.; Mäkipää, R.; Ståhl, G.; Lehtonen, A.; Petersson, H. Estimation of the biomass stock of trees in Sweden: comparison of biomass equations and age-dependent biomass expansion factors. *Ann. For. Sci.* **2005**, *62*, 845–851.
29. Woodhouse, I.H. Predicting backscatter-biomass and height-biomass trends using a macroecology model. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 871–877.
30. Balss, U.; Breit, H.; Duque, S.; Fritz, T.; Rossi, C. TanDEM-X Payload Ground Segment: CoSSC Generation and Interferometric Considerations. Available online: https://tandemx-science.dlr.de/pdfs/TD-PGS-TN-3129_CoSSCGenerationInterferometricConsiderations_1.0.pdf (accessed on 25 October 2013).
31. Huber, S.; Fiedler, H.; Krieger, G.; Zink, M. TanDEM-X Performance Optimization. In Proceedings of International Radar Symposium, Cologne, Germany, 5–7 September 2007.
32. Soja, M.; Ulander, L. Digital Canopy Model Estimation from TanDEM-X Interferometry Using High-resolution Lidar DEM. In Proceedings of 2013 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2013, Melbourne, VIC, Australia, 22–26 July 2013.
33. Attema, E.P.W.; Ulaby, F.T. Vegetation modelled as a water cloud. *Radio Sci.* **1978**, *13*, 357-364.
34. De Zan, F.; Krieger, G.; López-Dekker, P. On some spectral properties of TanDEM-X Interferograms over forested areas. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 71–75.

35. Pulliainen, J.T.; Heiska, K.; Hyyppä, J.; Hallikainen, M.T. Backscattering properties of boreal forests at the C-and X-band. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 1041–1050.
36. Fransson, J.E.S.; Smith, G.; Askne, J.; Olsson, H. Stem volume estimation in boreal forests using ERS-1/2 coherence and SPOT XS optical data. *Int. J. Remote Sens.* **2001** *22*, 2777–2791.
37. Santoro, M.; Shvidenko, A.; McCallum, I.; Askne, J.; Schmullius, C. Properties of ERS-1/2 coherence in the Siberian boreal forest and implications for stem volume retrieval. *Remote Sens. Environ.* **2007**, *106*, 154–172.
38. Santoro, M.; Askne, J.; Dammert, P. Tree height retrieval from ERS interferometric phase in boreal forest. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 207–217.
39. Mette, T.; Papathanassiou, K.P.; Hajnsek, I.; Zimmermann, R. Forest Biomass Estimation Using Polarimetric SAR Interferometry. In Proceedings of Workshop on POLinSAR—Applications of SAR Polarimetry and Polarimetric Interferometry, Frascati, Italy, 14–16 January 2003.
40. Caicoya, A.T.; Kugler, F.; Hajnsek, I.; Papathanassiou, K. Boreal Forest Biomass Classification with TanDEM-X. In Proceedings of 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012.
41. Santoro, M.; Beer, C.; Cartus, O.; Schmullius, C.; Shvidenko, A.; McCallum, I.; Wegmüller, U.; Wiesmann, A. Retrieval of growing stock volume in boreal forest using hyper-temporal series of Envisat ASAR ScanSAR backscatter measurements. *Remote Sens. Environ.* **2011**, *115*, 490–507.
42. Soja, M.J.; Sandberg, G.; Ulander, L.M.H. Regression-based retrieval of boreal forest biomass in sloping terrain using P-band SAR backscatter intensity data. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2646–2665.
43. Folkesson, K.; Smith-Jonforsen, G.; Ulander, L.M.H. Validating backscatter models for CARABAS SAR images of coniferous forests. *Can. J. Remote Sens.* **2008**, *34*, 480–495.
44. Solberg, S.; Astrup, R.; Breidenbach, J.; Nilsen, B.; Weydahl, D. Monitoring spruce volume and biomass with InSAR data from TanDEM-X. *Remote Sens. Environ.* **2013**, *139*, 60–67.
45. Ulander, L.M.H.; Askne, J.I.H.; Eriksson, L.E.B.; Franssson, J.E.S.; Persson, H.; Soja, M. Effects of Tree Species and Season on Boreal Forest Biomass Estimates from TanDEM-X. In Proceedings of 4th TanDEM-X Science Team Meeting, Oberpfaffenhofen, Germany, 12–14 June 2013; p. 15. Available online https://tandemx-science.dlr.de/pdfs/ScienceMeeting2013_OralPresentations/TDX_8 _Forest_Applications/3_Ulander.pdf (accessed on 25 October 2013).
46. Häme, T.; Salli, A.; Lahti, K. Estimation of Carbon Storage in Boreal Forests Using Remote Sensing Data, Pilot Study. In *The Finnish Programme on Climate Change, Progress Report*; Kanninen, M., Anttila, P., Eds.; Publications of the Academy of Finland; VAPK Publishing: Helsinki, Finland, 1992; Vol. 3/92; pp. 250–255.

Paper D

Estimation of Forest Height and Canopy Density from a Single Complex Correlation Coefficient

Authors:

M. J. Soja, H. Persson, and L. M. H. Ulander

Status:

Accepted for publication in *IEEE Geoscience and Remote Sensing Letters*, August 2014

Estimation of Forest Height and Canopy Density From a Single InSAR Correlation Coefficient

Maciej Jerzy Soja, Henrik Persson, and Lars M. H. Ulander, *Senior Member, IEEE*

***Abstract*—A two-level model (TLM) is introduced and investigated for the estimation of forest height and canopy density from a single ground-corrected InSAR complex correlation coefficient. The TLM models forest as two scattering levels, namely, ground and vegetation, separated by a distance Δh and with area-weighted backscatter ratio μ. The model is evaluated using eight VV-polarized bistatic–interferometric TanDEM-X image pairs acquired in the summers of 2011, 2012, and 2013 over the managed hemi-boreal test site Remningstorp, which is situated in southern Sweden. Ground phase is removed using a high-resolution digital terrain model. Inverted TLM parameters for thirty-two 0.5-ha plots of four different types (regular plots, sparse plots, seed trees, and clear-cuts) are studied against reference lidar data. It is concluded that the level distance Δh can be used as an estimate of the 50th percentile forest height estimated from lidar (for regular plots: $r > 0.95$ and root-mean-square difference (σ) $< 10\%$, or 1.8 m). Moreover, the uncorrected area fill factor $\eta_0 = 1/(1 + \mu)$ can be used as an estimate of the vegetation ratio, which is a canopy density estimate defined as the fraction of lidar returns coming from the canopy to all lidar returns (for regular plots: $r > 0.59$ and $\sigma \approx 10\%$, or 0.07).**

***Index Terms*—Canopy density, forest height, interferometric model, interferometry, synthetic aperture radar (SAR), TanDEM-X, two-level model (TLM).**

I. Introduction

THERE is a great need for a tool suitable for frequent mapping of large forest areas. Global forest biomass is one of the largest uncertainties in the current climate models [1]. An efficient tool for deforestation detection is needed for the implementation of international agreements [2]. Remote assessment of forest quality is also needed for biodiversity studies and commercial forestry.

Two important forest parameters are forest height and canopy density. In airborne lidar scanning, these parameters are often highly correlated with biomass [3]. Moreover, they can be also used for the assessment of forest quality, as well as deforestation detection.

Manuscript received April 3, 2014; revised August 11, 2014; accepted August 27, 2014. This work was supported by the Swedish National Space Board.

M. J. Soja is with the Department of Earth and Space Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden (e-mail: maciej.soja@chalmers.se).

L. M. H. Ulander is with the Department of Earth and Space Sciences, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden, and also with the Radar Systems Unit, Swedish Defence Research Agency (FOI), SE-581 11 Linköping, Sweden.

H. Persson is with the Department of Forest Resource Management, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden.

Digital Object Identifier 10.1109/LGRS.2014.2354551

Across-track interferometric synthetic aperture radar (InSAR) is a technique in which the scattering center elevation is measured from small phase differences between two SAR acquisitions made at slightly different incidence angles [4]. If a high-resolution digital terrain model (DTM) is available, which is the case in Sweden and many other European countries, the phase introduced by ground topography can be removed, and the remaining phase term is related to the scattering center elevation above ground, which is related to forest height and canopy density [5]. However, the inversion of these parameters from InSAR data is not trivial.

It has been shown that forest height can be estimated from C-, L-, and P-band fully polarimetric InSAR data using random volume over ground (RVoG) model inversion [6]–[8]. For single-polarized InSAR data, simplified versions of the RVoG can also provide estimates of forest height [9]–[11]. The RVoG models forest as a horizontally homogeneous volume, and there is no parameter directly related to the horizontal structure.

In the interferometric water cloud model (IWCM) [12]–[14], the horizontal structure is modeled using canopy gaps. The IWCM has been mainly investigated for stem volume (and later biomass) retrieval, although forest height retrieval has also been studied [15].

The scope of this letter is to introduce a model suitable for the estimation of both forest height and canopy density from a single ground-corrected InSAR complex correlation coefficient. The model is evaluated using single-pass X-band InSAR data acquired with the TanDEM-X twin-satellite system over a hemi-boreal forest in Sweden.

II. Modeling

The complex correlation coefficient $\tilde{\gamma}$ is the main observable in an interferometric SAR system. For the two images s_1 and s_2, it is defined as

$$\tilde{\gamma} = \frac{\mathsf{E}[s_1 s_2^*]}{\sqrt{\mathsf{E}\left[|s_1|^2\right] \mathsf{E}\left[|s_2|^2\right]}} \tag{1}$$

where $\mathsf{E}[\bullet]$ is the expectation value operator, and * is the complex conjugate operator.

In the case of TanDEM-X InSAR data, the temporal decorrelation over forests is negligible due to the almost simultaneous acquisition scenario [16]. Furthermore, assume that the signal-to-noise ratio (SNR) and system decorrelation effects are small, images s_1 and s_2 have been filtered to the same 2-D frequency spectrum [17], absolute phase calibration has been applied to remove phase offset, and complex multilooking has been applied to reduce phase and coherence estimation errors (see

(a) IWCM (b) RVoG (c) TLM

Fig. 1. Comparison between the different models described in Section II.

Section III-B). In that case, the main decorrelation effect is volume decorrelation, which can be modeled from the vertical backscattering profile $\sigma_v(z)$ using

$$\tilde{\gamma} \approx \tilde{\gamma}_{\text{vol}} = \frac{\int_{-\infty}^{\infty} \sigma_v(z) e^{jk_z z}\, dz}{\int_{-\infty}^{\infty} \sigma_v(z)\, dz} \tag{2}$$

where z is the vertical coordinate, and k_z is the vertical wavenumber, which, for a bistatic system, is

$$k_z = 2\pi B_\perp / (\lambda R \sin\theta) \tag{3}$$

where $B_\perp$ is the perpendicular baseline, λ is the wavelength, R is the average range to the satellites, and θ is the average angle of incidence.

A. Vegetation as Random Volume

A common way to solve the integrals in (2) is through the assumption of a known vertical backscattering profile.

In the IWCM [12]–[14], vegetation is modeled as a volume of randomly oriented scatterers located above a ground plane and covering a certain fraction of the total area, which is called area fill factor [see Fig. 1(a)]. The vertical backscattering profile is then an exponential function described by an attenuation coefficient. The total number of parameters needed for correlation coefficient modeling is five: volume height, ground height, area fill factor, attenuation coefficient, and a ground-to-volume backscatter ratio. Note that, in its full formulation, originally developed for stem volume estimation from repeat-pass ERS-1/2 interferometry [12], [13], the IWCM includes additional empirical functions and temporal decorrelation modeling.

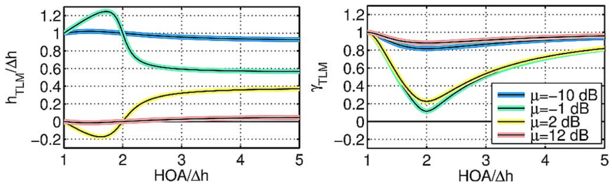
In the RVoG model [6], the gaps modeled by the IWCM are neglected [see Fig. 1(b)], which limits the number of parameters to four. Possible further simplifications used in the past include neglecting the ground component, assuming known ground phase, or fixing extinction coefficient [9]–[11]. The main advantage of the RVoG is that it requires fewer parameters than the IWCM, but the canopy gaps are not modeled. There is, thus, no natural parameter, which can be used as an estimate for canopy density.

B. Vegetation as Scattering Levels

A different way to solve the integral in (2) is by simplifying it to a sum of a finite number of discrete scattering levels.

In the special case of two scattering levels, i.e., ground and vegetation, located at $z = z_0$ and $z = z_0 + \Delta h$, with respective backscattering coefficients σ^0_{gr} and σ^0_{veg}, and covering area fractions $1 - \eta$ and η, respectively, (2) simplifies to a two-level model (TLM) [see Fig. 1(c)], i.e.,

$$\tilde{\gamma}_{\text{TLM}} = e^{i\Phi_0}(\mu + e^{ik_z \Delta h})/(\mu + 1) \tag{4}$$



(a) Relative interferometric height (b) Coherence

Fig. 2. Interferometric height and coherence modeled with TLM.

where $\Phi_0 = k_z z_0$ is the ground phase, $\mu = \rho(1-\eta)/\eta$ is the area-weighted (ground-to-vegetation) backscatter ratio, and $\rho = \sigma^0_{\text{gr}}/\sigma^0_{\text{veg}}$ is the backscatter ratio. In the case of known Φ_0, the TLM requires only two independent parameters describing the scene (μ and Δh) to model one ground-corrected complex correlation coefficient, and the inversion can be carried out without any additional data.

C. Influence of Baseline on Height and Coherence

Assuming $\Phi_0 = 0$, the interferometric height is computed from the phase of (4) as

$$h_{\text{TLM}} = \frac{\text{HOA}}{2\pi}\left[\tan^{-1}\left(\frac{\sin\left(2\pi\frac{\Delta h}{\text{HOA}}\right)}{\cos\left(2\pi\frac{\Delta h}{\text{HOA}}\right) + \mu}\right) + \pi n\right] \tag{5}$$

where n is an integer describing the phase ambiguity, and $\text{HOA} = 2\pi/k_z$ is the height of ambiguity for a bistatic system (maximal height that can be unambiguously resolved by the interferometric system). Likewise, the coherence is computed from the magnitude of (4) as

$$\gamma_{\text{TLM}} = |\tilde{\gamma}_{\text{TLM}}| = \frac{\sqrt{1 + \mu^2 + 2\mu\cos\left(2\pi\frac{\Delta h}{\text{HOA}}\right)}}{1+\mu}. \tag{6}$$

In Fig. 2(a), the ratio $h_{\text{TLM}}/\Delta h$ is plotted against $\text{HOA}/\Delta h$ for four different values of μ. Phase unwrapping has been performed to ensure continuity and that $0 \le h_{\text{TLM}} \le \Delta h$ is satisfied for large HOA. The interferometric height is fairly independent of HOA if scattering at one level is significantly larger than at the other level (large or small μ). If μ is close to unity, scattering at both levels becomes equally significant, and an interference effect is observed for low HOA, i.e., when the phase difference between the two levels is large. In the case of dominant ground-level scattering, negative interferometric heights are obtained. In the case of dominant vegetation-level scattering, interferometric height exceeding Δh is observed. In both cases, the interference effect becomes weaker for $\text{HOA} > 2\Delta h$.

Coherence dependence on HOA is illustrated in Fig. 2(b). Coherence is maximized when Δh is either close to HOA or much smaller than HOA. In the first case, the phase is similar at both levels during the integration in the numerator of (2) due to constructive interference of two adjacent periods. In the second case, the phase is similar because the phase of $e^{ik_z z}$ changes very little between 0 and Δh, and $e^{ik_z z}$ is virtually

TABLE I
FIELD PLOT GROUPING USED IN THIS STUDY

Group:	Description:
Regular plots (RP)	mature, homogeneous, dense plots, no undervegetation
Sparse plots (SP)	plots with sparse seed trees from past harvesting, regrown undervegetation and young trees
Seed trees (ST)	recently harvested plots with sparse seed trees, no undervegetation (two in 2013)
Clear-cuts (CC)	recently harvested plots, bare soil, no undervegetation (one in 2012, additional two in 2013)

TABLE II
LIDAR METRICS USED AS REFERENCE IN THIS STUDY

Metric:	Property:	Description:
vegetation ratio (VR)	canopy density	ratio between lidar returns above 1 m or 10% of maximal height and all lidar returns
H50	forest height	50th percentile (median) height for all returns above 1 m or 10% of maximal height
H95	forest height	95th percentile height for all returns above 1 m or 10% of maximal height

constant, giving maximal coherence. Coherence is minimized for $\Delta h = \text{HOA}/2$, that is, when the interference between the two levels is perfectly destructive. Note that TLM coherence is the same for μ and its reciprocal $1/\mu$.

III. EXPERIMENTAL DATA

In this letter, data acquired over the boreal test site of Remningstorp (58° 28′ N, 13° 38′ E), which is situated in southern Sweden, are used. Remningstorp features fairly flat topography with ground height varying between 120 and 145 m above sea level. The forest consists primarily of Norway spruce, Scots pine, and different birch species. The annual growth rate of the forest is about 10–20 cm, but it is neglected in this study as the study period covers only three growing seasons.

A. Field Plots and Lidar Data

A set of 32 circular 40-m radius plots is available for Remningstorp. Species-specific field data on biomass, tree number density, and tree height have been used together with field observations, optical SPOT-5 images, and lists with forest management procedures to assess the state of each plot at the time of each SAR acquisition. As a result of these procedures, a time line for the observed changes has been established, and four groups with significant difference in forest structure have been created. These groups are presented in Table I.

Three lidar metrics have been extracted from maps with 10 m $\times$ 10 m pixels provided within the BioSAR 2010 campaign [18]. These lidar metric are presented in Table II. As the lidar data have been collected in August 2010, which is before any harvesting procedures have been conducted within the plots, the reference data are not valid for the plots with seed trees and clear-cuts.

B. SAR and DTM Data

TanDEM-X (TDM) is a twin-satellite X-band (9.65 GHz) SAR interferometer in which acquisitions are made almost simultaneously [16]. Eight bistatic–interferometric VV-polarized

TABLE III
SUMMARY FOR THE TDM DATA USED IN THIS STUDY. MEAN VALUES FOR ALL PLOTS ARE GIVEN. BACKGROUND SHADING IS BY HOA GROUP

Nr:	Date:	$B_\perp$ [m]:	HOA [m]:	Coherence:
1	20110604	281	49	0.65
2	20110809	266	52	0.66
3	20110820	257	54	0.65
4	20120601	433	32	0.73
5	20120828	371	37	0.74
6	20130702	270	51	0.66
7	20130724	226	61	0.73
8	20130804	221	63	0.73

TDM acquisitions made at low HOA in the ascending mode are available and used in this study (see Table III). The nominal incidence angle is within the interval 41.2°–41.7°. The data have been divided in three groups according to the approximate HOA level: 35 m, 50 m, and 60 m (images 4 and 5, 1–3 and 6, and 7 and 8, respectively).

As ground reference, a DTM with a 2 m $\times$ 2 m grid posting and a mean height error lower than 0.5 m has been used [19]. Four plots are not covered by the two images from August 2011.

TDM data have been interferometrically processed according to [17]. The raw interferograms have been flattened in radar geometry using a linearly interpolated DTM and taking into consideration the quasi-bistatic acquisition geometry and satellite displacement between signal transmission and reception. A 5-m buffer zone has been added prior to plot-level averaging of the ground-corrected interferograms. The total number of looks has been estimated to 330 for image 1 (from 2011-06-04) and 430 for the remaining seven images. Absolute phase calibration has been done using ground reference points derived from a nonforest mask. No unwrapping has been found necessary due to the limited height variations in the flattened interferograms. Geocoding error and height measurement errors have been estimated using two 5-m trihedral corner reflectors situated within the test site. The geocoding offset has been found lower than 2 m, and the standard height estimation error has been found lower than 10 cm.

IV. MODEL INVERSION

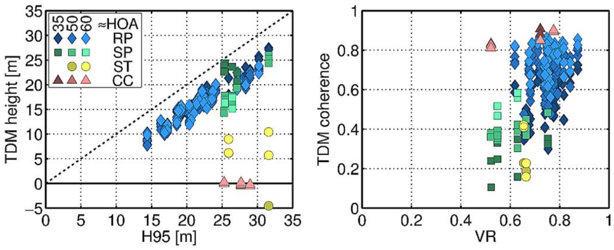
For a fixed μ and $\Phi_0 = 0$, the TLM is a circle in the complex plane, with its center in $\mu/(1+\mu)$, radius $1/(1+\mu)$, and passing through unity. Area-weighted backscatter ratio μ can be obtained from the complex correlation coefficient $\tilde{\gamma}$ as

$$\mu = \frac{1 - |\tilde{\gamma}|^2}{1 - 2\mathrm{Re}[\tilde{\gamma}] + |\tilde{\gamma}|^2} \tag{7}$$

whereas level distance Δh can be found using (7) in (4), i.e.,

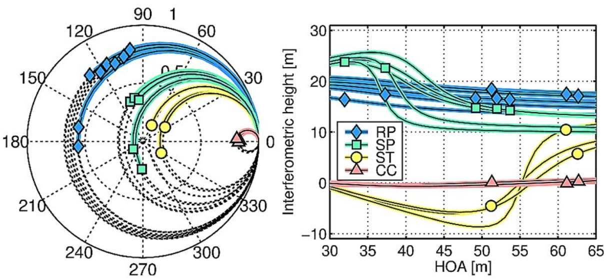
$$\Delta h = \frac{1}{k_z}\left[\tan^{-1}\left(\frac{2\mathrm{Im}[\tilde{\gamma}](1-\mathrm{Re}[\tilde{\gamma}])}{2\mathrm{Re}[\tilde{\gamma}](1-\mathrm{Re}[\tilde{\gamma}])-(1-|\tilde{\gamma}|^2)}\right)+\pi n\right] \tag{8}$$

where n is an integer describing the ambiguity of the inversion; and $\mathrm{Re}[\bullet]$ and $\mathrm{Im}[\bullet]$ are the real and imaginary part operators, respectively. The 2π-ambiguity can be manually resolved by checking the most probable values in relation to the type of studied forest. For the data used in this study, $n = 0$ has been chosen in all cases.



(a) TDM height vs. H95 (b) TDM coherence vs. VR

Fig. 3. Scatter plots for TDM observables against lidar reference. Color coding according to plot type has been applied, with different shades for different HOA groups.



(a) Complex plane (b) Interferometric height vs. HOA

Fig. 4. Sample fitting results for the different plot types.

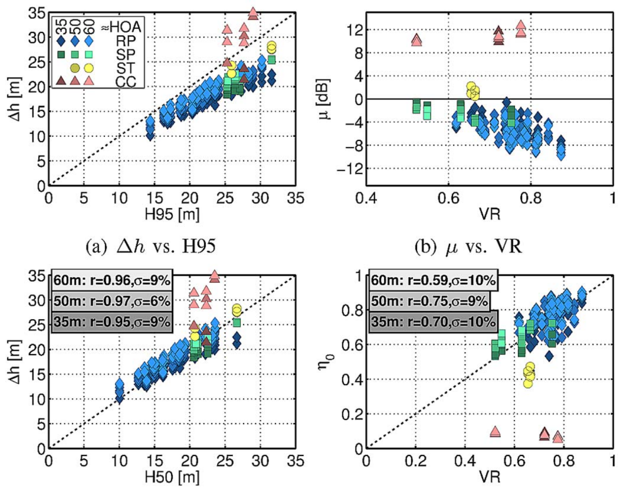
If μ and ρ are known, the area fill factor is $\eta = \rho/(\rho + \mu)$. Normally, the backscatter ratio ρ is not known. Assuming $\rho = 1$, i.e., equal ground and canopy backscattering coefficients, the uncorrected area fill factor η_0 can be obtained, i.e.,

$$\eta_0 = 1/(1 + \mu). \tag{9}$$

V. RESULTS

TDM interferometric height and coherence are plotted against H95 and VR, respectively, for all eight images and all available plots in Fig. 3. Color coding has been applied for the different plot types and color shading for different HOA groups. A large variance of the interferometric height and a generally lower coherence can be observed for the sparse plots with H95 around 25 m and for plots with seed trees. This agrees well with the TLM predictions for area-weighted backscatter ratios μ close to unity (see Fig. 2). Both regular plots and clear-cuts show more predictable behavior with stable height estimates and a higher coherence. Note that the tallest sparse plot is also the densest of all sparse plots, and it shows similar behavior as the regular plots. For the regular plots, the correlation between TDM height and H95 is high, but an underestimation of 5–10 m is observed. Note also that a large coherence variability is introduced by the differences in the interferometric baseline.

As shown in Fig. 4(a), finding Δh and μ limits to finding a circle in the complex plane. The ratio between the circle center position and its radius is equal to μ. The highest area-weighted backscatter ratio μ is thus obtained for clear-cuts, and the lowest



(a) Δh vs. H95 (b) μ vs. VR

(c) Δh vs. H50 (d) η_0 vs. VR

Fig. 5. Scatter plots for inverted TLM parameters against lidar reference. Pearson correlation coefficients r and root-mean-square differences sigma (σ), computed separately for the different HOA groups, are shown for the unaltered plots (both regular and sparse plots).

is obtained for the regular dense plots. In Fig. 4(b), TDM interferometric heights and TLM model curves are plotted against HOA. It can clearly be seen that, in cases when ground- and vegetation-level contributions are comparable, large variance of the interferometric height is to be expected, as in the cases of sparse pine plots and plots with seed trees.

Parameters Δh and μ inverted using (7) and (8) are plotted against H95 and VR in Fig. 5(a) and (b), respectively, for all plots and acquisitions. It is concluded that Δh is a biased estimate of H95, but a better estimate of H50 ($r \geq 0.96$, and root-mean-square difference σ is around 10% of the mean H50) [see Fig. 5(c)]. Note that there is a HOA-dependent offset between the acquisitions. The area-weighted backscatter ratio μ does not measure the same property as VR, which is clearly shown in Fig. 5(b). The uncorrected area fill factor η_0 is a better estimate of VR ($r \geq 0.59$ and σ around 10% of the mean VR) [see Fig. 5(d)]. However, a significant variance can be observed.

As observed earlier in Fig. 3(a), plots with seed trees show low and highly variable interferometric heights. In Fig. 5, it can be observed that a harvesting procedure in which sparse seed trees are left does not affect forest height inversion, while the inverted canopy density is lower. For clear-cuts, forest height inversion produces significantly biased results [see Fig. 5(a) and (c)].

Since phase calibration has been done using open fields, the interferometric height for clear-cuts is close to zero, giving an almost real-valued correlation coefficient. However, coherence is between 0.8 and 0.9, mainly due to SNR and system decorrelation effects. TLM inversion does not provide reliable height estimates because the TLM cannot predict a real-valued correlation coefficient with coherence lower than 1 and low height. This is probably the main cause of the inflated Δh estimates for clear-cuts and the HOA-dependent offset.

VI. Discussion

The interference effect modeled by the TLM and observed in the data for sparse plots and seed trees occurs when ground- and vegetation-level contributions are similar in strength. Whether this high ground-level contribution at X-band is an effect of dielectric penetration through the scatterers or penetration through canopy gaps has been discussed earlier, most recently in [20], but it is still an open question. The results shown here hint that the inclusion of canopy gaps in modeling is useful, as the inversion provides a parameter that is related to canopy density.

The estimation of canopy density using the uncorrected area fill factor η_0 is based on the assumption that $\rho = 1$. Although this assumption appears to be valid for the studied X-band data, the influence of wavelength on ρ is expected to be strong, and the assumption will not hold at other frequencies. If an estimate of the area fill factor η is available, e.g., from lidar VR, ρ can be estimated from μ at scene or plot level and studied against parameters such as weather, season, and ground surface roughness. A better knowledge of ρ can then improve estimation of canopy density from μ.

The TLM in the presented form has been developed for VV-polarized data, and polarization dependence has not been studied. It has been shown in [11] that a difference in the interferometric height can be observed between HH and VV. Therefore, TLM inversion of HH-polarized data requires a separate study.

This study has been limited to InSAR data with low HOA (smaller than 65 m), for which there is a large variation of the interferometric phase between the two levels. However, TLM inversion is expected to work at larger HOA values as well, as long as volume decorrelation is the most significant decorrelation effect.

VII. Conclusion

In this letter, a two-level model (TLM) has been introduced and used for the estimation of forest height and canopy density from bistatic–interferometric VV-polarized TanDEM-X (TDM) data. With an access to the global TDM data, the presented approach can be used on large scale in countries where national high-resolution DTMs exist. Since the DTM is temporally stable in most forested regions, the presented approach requires only one DTM acquisition, and frequent mapping of forest height and canopy density can thereafter be carried out using a spaceborne SAR system such as TDM. The approach is therefore suitable for cost-effective mapping and monitoring of national forest resources. The HOA-dependent height estimation bias observed primarily for open fields and clear-cuts can be avoided through the inclusion of a coherence calibration step during InSAR processing, in which decorrelation effects such as SNR and system decorrelation are compensated for.

Acknowledgment

The authors would like to thank the German Aerospace Center (DLR) for the TDM data (proposal XTI_VEGE0306); the Swedish Land Survey for the DTM and SPOT-5 data; the European Space Agency for the BioSAR 2010 lidar data; and G. Eriksen, H. Kjellin, J. Askne, A. Berg, and L. Eriksson for their respective contributions to this work.

References

[1] IPCC Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge, U.K.: Cambridge Univ. Press 2007, IPCC.

[2] C. Parker, A. Mitchell, M. Trivedi, and N. Mardas, *The Little REDD Book: A Guide to Governmental and Non-Governmental Proposals for Reducing Emissions from Deforestation and Degradation*. Oxford, U.K.: Global Canopy Programme, John Krebs Field Station, Nov. 2008.

[3] J. Hyyppä *et al.*, "Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests," *Int. J. Remote Sens.*, vol. 29, no. 5, pp. 1339–1366, 2008.

[4] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse Problems*, vol. 14, no. 4, pp. R1–R54, Aug. 1998.

[5] S. Solberg, R. Astrup, T. Gobakken, E. Naesset, and D. J. Weydahl, "Estimating spruce and pine biomass with interferometric X-band SAR," *Remote Sens. Environ.*, vol. 114, no. 10, pp. 2353–2360, Oct. 2010.

[6] S. R. Cloude and K. P. Papathanassiou, "Polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1551–1565, Sep. 1998.

[7] K. P. Papathanassiou and S. R. Cloude, "Single-baseline polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2352–2363, Nov. 2001.

[8] F. Garestier, P. C. Dubois-Fernandez, and I. Champion, "Forest height inversion using high-resolution P-band Pol-InSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3544–3559, Nov. 2008.

[9] F. Garestier, P. C. Dubois-Fernandez, and K. P. Papathanassiou, "Pine forest height inversion using single-pass X-band PolInSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 59–68, Jan. 2008.

[10] J. Praks, O. Antropov, and M. T. Hallikainen, "LIDAR-aided SAR interferometry studies in boreal forest: Scattering phase center and extinction coefficient at X- and L-band," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, pp. 3831–3843, Oct. 2012.

[11] 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, Oct. 2014.

[12] J. I. H. Askne, P. Dammert, L. M. H. Ulander, and G. Smith, "C-band repeat-pass interferometric SAR observations of the forest," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 25–35, Jan. 1997.

[13] J. I. H. Askne, M. Santoro, G. Smith, and J. E. S. Fransson, "Multitemporal repeat-pass SAR interferometry of boreal forests," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 7, pp. 1540–1550, Jul. 2003.

[14] 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.*, vol. 5, no. 11, pp. 5574–5597, Oct. 2013.

[15] M. Santoro, J. I. H. Askne, and P. Dammert, "Tree height retrieval from ERS interferometric phase in boreal forest," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 2, pp. 207–217, Feb. 2005.

[16] G. Krieger *et al.*, "TanDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3317–3341, Nov. 2007.

[17] S. Duque, U. Balss, C. Rossi, T. Fritz, and W. Balzer, "TanDEM-X payload ground segment, CoSSC generation and interferometric considerations," Remote Sens. Technol. Inst., German Aerospace Center (DLR), Cologne, Germany, Tech. Rep., 2012.

[18] L. M. H. Ulander *et al.*, "BioSAR 2010: Technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2010 experiment: Final report," Eur. Space Agency (ESA), Paris, France, Tech. Rep. Contract No. 4000102285/10/NL/JA/ef, 2011.

[19] "Produktbeskrivning: GSD-Höjddata, grid 2 +," Lantmäteriet, Gävle, Sweden, Tech. Rep., 2013.

[20] F. De Zan, G. Krieger, and P. Lopez-Dekker, "On some spectral properties of TanDEM-X interferograms over forested areas," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 1, pp. 71–75, Jan. 2013.

Paper E

Estimation of Forest Biomass from Two-Level Model Inversion of Single-Pass InSAR Data

Authors:

M. J. Soja, H. Persson, and L. M. H. Ulander

Status:

Submitted to *IEEE Transactions on Geoscience and Remote Sensing*, July 2014

Estimation of Forest Biomass from Two-Level Model Inversion of Single-Pass InSAR Data

Maciej Jerzy Soja, Henrik Persson, and Lars M. H. Ulander

Abstract

A model for above-ground biomass estimation from single-pass interferometric synthetic-aperture radar data is presented. Forest height and canopy density estimates Δh and η_0, obtained from two-level model inversion, are used in a power function with slope K and exponents α and β. The model is compared to a linear, zero-intercept model, scaling the interferometric height to biomass.

Eighteen bistatic, VV-polarized TanDEM-X (TDM) acquisitions made over two test sites in the summers of 2011, 2012, and 2013 are used. Remningstorp is a hemi-boreal forest in southern Sweden, with flat topography and 32 circular plots (area: 0.5 ha, biomass: 42–242 t/ha, height: 14–32 m). Krycklan is a boreal forest in northern Sweden, 720 km north-north-east from Remningstorp, with significant topography and 31 stands (area: 2.4–26.3 ha, biomass: 23–183 t/ha, height: 7–21 m). For all acquisitions, the nominal incidence angle is 41° and the height-of-ambiguity is in the interval 32–63 m. High-resolution digital terrain model has been used for ground correction during InSAR processing.

The proposed model explains 65–89% of the variance observed in the data, with a residual root-mean-square error (RMSE) 12–19% (median: 15%). If model training and validation are carried out on different acquisitions or between test sites, the prediction RMSE increases (12–80%, median: 30%). With α fixed and β a site-dependent constant, the prediction RMSE is lower (12–56%, median: 17%), while the residual RMSE is similar (12–29%, median: 16%). The linear, zero-intercept model shows similar residual and prediction performance for the Krycklan data, whereas for the Remningstorp data and across-site retrieval, the performance is poorer.

Index Terms

above-ground biomass (AGB), forest height, canopy density, interferometric model, two-level model (TLM), interferometric synthetic-aperture radar (InSAR), TanDEM-X

I. Introduction

Forest are important natural resources because of their economic value and their crucial role in the local and global ecosystems [1]. Efficient and sustainable management procedures are required to maintain healthy and productive forests.

One of the key elements in forest management is to have reliable information for short- and long-term planning. The above-ground dry biomass, here shortly called biomass or the AGB, is especially important for carbon cycle studies, while other parameters such as forest height and canopy density can both aid biomass estimation and provide additional information on the forests. However, current methods of collecting forest information are expensive, and therefore more cost-effective methods need to be developed. Remote sensing in combination with field inventories has the potential to meet these requirements, and provide frequent and high-resolution mapping of forest variables. Large-scale mapping is also needed for natural disaster management, so that the damages caused by, e.g., storms, can be minimized.

Aerial photography has traditionally been used for forest mapping [2], [3]. This technique has the advantage of being relatively easy to implement and interpret, but it requires cloud-free acquisitions and good flying weather. Moreover, it is less efficient on large scale and whenever frequent updates are needed. Spaceborne photography is more efficient in terms of coverage and acquisition rate, but it has lower resolution. More advanced optical techniques, such as photogrammetry [4], [5], can provide additional information on forests, but the even stricter requirements on the acquired data make their use more difficult on an operational scale.

In recent years, airborne lidar scanning (ALS) has become popular. The technique uses laser pulses transmitted downwards from an airborne platform, which are used to sample height at high vertical and horizontal resolutions [4], [6]–[10]. Due to the high resolutions and the penetration of laser pulses through canopy gaps, ALS can provide information on both horizontal and vertical forest structure, and many important forest parameters can be derived from the data. ALS is today considered the most accurate remote sensing technique in forestry [10]. However, the technique is relatively expensive, and thus inefficient for frequent and large scale mapping. Spaceborne lidar, on the other hand, has yet unresolved resolution, coverage, and technology limitations.

Synthetic-aperture radar (SAR) is an active remote sensing technique in which radio- or microwave-frequency pulses are used to probe the environment. Spaceborne SAR sensors can provide weather- and daylight-independent imagery of the Earth with resolutions down to a couple of meters. Through the choice of the center frequency, SAR systems can be optimized to fit different needs [11]. In forestry, low frequency bands, like the VHF-band (30–300 MHz) and the lower UHF-band (300–1000 MHz, according to the IEEE standard) are more suitable for imaging of tree trunks and ground surface, while the high frequency bands, like the X-band (8–12 GHz), are more suitable for the imaging of tree canopies. SAR is one of the most promising tools for forest remote sensing and many past and ongoing studies are dedicated to the retrieval of forest parameters from SAR data [12].

The TanDEM-X system consists of two, almost identical X-band SAR satellites flying in a tight tandem formation, at a distance of a few hundred meters during the operational phase. Using the principles of interferometric SAR (InSAR), small phase differences between the two acquired SAR images are used to measure the position of the scattering center [13], i.e. to create a digital elevation model (DEM). With the tight tandem formation and bistatic-mode acquisitions of TanDEM-X, the temporal changes between the two SAR acquisitions are minimal, and the acquired height measurements are very precise.

The acquired DEM can be corrected for ground topography if a high-resolution digital terrain model (DTM) is available, and a map of the scattering center elevation above ground can be obtained. In Sweden, there is a national, lidar-scanned DTM with a grid posting of $2\,\mathrm{m} \times 2\,\mathrm{m}$ and a height accuracy better than $0.5\,\mathrm{m}$ [14]. Similar DTMs exist or are being created in many other countries. Since the changes of the ground surface are very slow in most forested regions, only one lidar scanning is required to obtain a high-resolution DTM, and after that mapping of forest canopy can be done with the TanDEM-X system.

The exact position of the scattering center above ground in forests is related to the structure of the forest and it depends on forest properties such as forest height and canopy density. In several studies, this relation has been investigated. In [15], [16], random volume over ground (RVoG) model inversion has been applied to estimate forest height from single-pass X-band InSAR data. In [17], a linear relation between biomass and the measured elevation of the scattering center above ground has been observed. This study has been based on an approach developed earlier for the SRTM X-SAR data [18]. In [19], biomass estimates have been obtained from ground-corrected TDM interferograms using the inversion of the interferometric water cloud model (IWCM), which includes an allometric relation between forest height and biomass, as well as temporal decorrelation. Both interferometric coherence and phase, and backscatter intensity data have been used in the inversion process. In [20], a multiple regression approach using interferometric height, coherence, and their transformed versions has been used to estimate biomass, separately for two test sites in Sweden.

It has been shown in [21], [22] that the inversion of a two-level model (TLM) can provide estimates of forest height and canopy density in a hemi-boreal forest in Sweden. In this study, the inverted parameters will be used to estimate biomass. Data from two boreal test sites in Sweden, separated by 720 km, will be used. The new model will be evaluated both for its explanatory and predictive values.

II. Method

In this section, the basic models used in this study will be described. First, it will be shown how biomass can be estimated from forest height and canopy density. Next, it will be shown how forest height and canopy density can be estimated from InSAR data. Thereafter, the results from the first two sections will be used together, and a new model for biomass estimation from InSAR data will be presented. Finally, the evaluation method used in this study will be described.

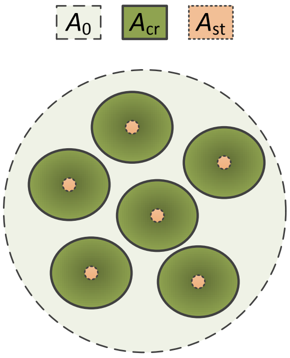


Fig. 1. Geometrical visualization of a forest plot.

A. AGB from Forest Height and Canopy Density

Above-ground biomass (AGB) is defined as the total dry mass of all above-ground forest, most commonly measured in terms of biomass density, i.e., as mass per area unit. As a large part of the AGB is confined to the stem (around 3/4 for spruce and pine in Sweden, according to [23]), a simple geometrical argument [24], [25] suggests that the AGB is a function of forest height, basal area, a taper factor accounting for the non-cylindrical trunk shape, the oven-dry wood density, and an expansion factor for the conversion of stem biomass to total aboveground biomass. After merging the last three factors into a single, forest type-dependent constant C, the AGB can be estimated from:

$$\widehat{\mathrm{AGB}} = C \cdot h \cdot \frac{A_{st}}{A_0}, \tag{1}$$

where A_{st} is the total basal area for all trees, A_0 is the ground area of the plot, see Figure 1, and h is the basal area-weighted forest height.

In field inventories, the total basal area is estimated from stem diameter measurements. The measurement of forest height is more time-consuming, and many allometric equations for biomass computation require only stem diameter measurements [25]–[27].

In remote sensing, there are several techniques for forest height estimation, including lidar scanning [7], [28], [29], polarimetric SAR interferometry [30]–[32], photogrammetry [4], [5], and radargrammetry [33], [34], but the estimation of the basal area is difficult, due to canopy closure, shadowing, and too low resolution. On the other hand, the size of tree crowns can be estimated from aerial photography [35]–[37], lidar [9], [38], [39], or SAR interferometry [21], [22]. Since several studies show a reasonable correlation between canopy diameter and stem diameter for many tree species [40]–[42], the total crown area A_{cr} will in the following be used as a predictor of the total basal area:

$$\widehat{\mathrm{AGB}} = C' \cdot h \cdot \eta_{cr}, \tag{2}$$

where C' is a forest type-dependent constant, which is the product of C and the ratio between the total basal area and the total canopy area, and

$$\eta_{cr} = \frac{A_{cr}}{A_0} \tag{3}$$

is the fractional canopy coverage, which is a measure of canopy density.

The main purpose of the argument above is to show that a multiplicative model is appropriate for biomass estimation from forest height and canopy density. However, this argument is based on several simplifying assumptions regarding the shape of the trees, their intrinsic wood properties, their spatial distribution, tree parameter distribution within a plot, etc. In reality, the dependence of the AGB on the two forest parameters is expected to be more complicated. For instance, there will be a residual dependence of C' on height and basal area, which may affect the dependence of the AGB estimate on h and η_{cr} in (2). Exponents α and β are therefore introduced to create an improved model, based on the experience from field inventories [25], [26]:

$$\widehat{\mathrm{AGB}} = C'' \cdot h^{\alpha} \cdot \eta_{cr}^{\beta}, \tag{4}$$

where C'' is a new, forest type-dependent constant.

B. Forest Height and Canopy Density from InSAR

In synthetic-aperture radar interferometry [13], the complex correlation coefficient is the main observable and it is defined as:

$$\widetilde{\gamma} = \frac{\mathsf{E}\left[s_1 s_2^*\right]}{\sqrt{\mathsf{E}\left[|s_1|^2\right]\mathsf{E}\left[|s_2|^2\right]}}, \tag{5}$$

where s_1 and s_2 are the two interferometric images, * is the complex conjugate operator, and $\mathsf{E}[\bullet]$ is the expectation value operator.

Coherence is the magnitude of the complex correlation coefficient and it is a measure of similarity between two images. The phase of the correlation coefficient carries information about the vertical distribution of the scatterers. In applications, the complex correlation coefficient is estimated from a finite number of samples, and the interferometric phase is affected by noise. The total noise level will increase with decreasing number of independent samples and/or decreasing coherence [43].

The loss of coherence (decorrelation) can be caused by up to four different effects: temporal changes in the scene, geometric differences between the two images, thermal noise, and system imperfections [44], [45].

Volume decorrelation is a geometric effect caused by the distribution of scatterers in the vertical direction z. It can be modeled from the vertical backscattering profile $\sigma(z)$ using [46], [47]:

$$\widetilde{\gamma}_{vol} = \frac{\int_{-\infty}^{\infty} \sigma(z) e^{ik_z z} \mathrm{d}z}{\int_{-\infty}^{\infty} \sigma(z) \mathrm{d}z}, \tag{6}$$

with k_z being the vertical wavenumber, which for a bistatic acquisition geometry is:

$$k_z = \frac{2\pi}{\mathrm{HOA}} = \frac{2\pi B_\perp}{\lambda R \sin\theta}, \tag{7}$$

where HOA is the height-of-ambiguity, $B_\perp$ is the perpendicular baseline, λ is the wavelength, R is the average range, and θ is the average angle of incidence. HOA is the height corresponding to a 2π-phase shift in the interferogram, and it is the maximal height difference, which can be unambiguously resolved by the interferometric system.

In the two-level model (TLM) [21], [22], [48], forest is modeled as two scattering levels, ground and vegetation, at the respective elevations z_0 and $z_0 + \Delta h$, and with the respective backscattering coefficients σ_{gr}^0 and σ_{veg}^0. The vertical backscattering profile $\sigma(z)$ is therefore:

$$\sigma(z) = (1-\eta)\sigma_{gr}^0 \delta(z - z_0) + \eta \sigma_{veg}^0 \delta(z - (z_0 + \Delta h)), \tag{8}$$

where η is the area-fill factor (the fraction of the total area covered by the vegetation level) and $\delta(\bullet)$ is the Dirac delta function.

Inserting (8) in (6) yields:

$$\widetilde{\gamma}_{vol} = e^{ik_z z_0} \cdot \frac{\mu + e^{ik_z \Delta h}}{\mu + 1}, \tag{9}$$

where

$$\mu = \rho \cdot \frac{1-\eta}{\eta} \tag{10}$$

is the area-weighted backscatter ratio with the ground-to-vegetation backscatter ratio ρ defined as:

$$\rho = \frac{\sigma^0_{gr}}{\sigma^0_{veg}}. \tag{11}$$

The first exponential term in the expression for the TLM in (9) introduces a phase term related to ground topography z_0. This phase term can also be observed in the measured complex correlation coefficient $\widetilde{\gamma}$. If ground topography is known, for example from an external digital terrain model (DTM), then this exponential term can be used to compensate the complex correlation coefficient for ground topography, so that the ground-corrected complex correlation coefficient can be obtained:

$$\widetilde{\gamma}_{gc} = \frac{\mathsf{E}\left[s_1 s_2^* e^{-ik_z z_0}\right]}{\sqrt{\mathsf{E}\left[|s_1|^2\right]\mathsf{E}\left[|s_2|^2\right]}}. \tag{12}$$

From the phase of the ground-corrected complex correlation coefficient $\widetilde{\gamma}_{gc}$, the interferometric height (scattering center position above ground) can be obtained using:

$$h_{gc} = \frac{\arg\left(\widetilde{\gamma}_{gc}\right) + 2\pi n}{k_z} = \text{HOA}\left(\frac{\arg\left(\widetilde{\gamma}_{gc}\right)}{2\pi} + n\right), \tag{13}$$

where $\arg(\bullet)$ is the argument operator and the integer n describes the ambiguity of the phase computation.

In the absence of decorrelation effects other than volume decorrelation, the measured ground-corrected complex correlation coefficient $\widetilde{\gamma}_{gc}$ can be modeled by the TLM expression in (9) with $z_0 = 0$:

$$\widetilde{\gamma}_{gc} = \frac{\mu + e^{ik_z \Delta h}}{\mu + 1}. \tag{14}$$

Since this equation has two unknowns (Δh and μ) and two observables (the real and imaginary parts of $\widetilde{\gamma}_{gc}$), it can be solved without the need for multiple acquisitions. The TLM describes a circle in the complex plane with its center on the positive x-axis and passing through $\widetilde{\gamma}_{gc}$ and 1. The solutions for μ and Δh are:

$$\mu = \frac{1 - \gamma_{gc}^2}{1 - 2\mathsf{Re}\left[\widetilde{\gamma}_{gc}\right] + \gamma_{gc}^2}, \tag{15}$$

$$\Delta h = \frac{\tan^{-1}\left[\frac{2\mathsf{Im}[\widetilde{\gamma}_{gc}](1-\mathsf{Re}[\widetilde{\gamma}_{gc}])}{2\mathsf{Re}[\widetilde{\gamma}_{gc}](1-\mathsf{Re}[\widetilde{\gamma}_{gc}])+\gamma_{gc}^2-1}\right] + \pi n}{k_z}, \tag{16}$$

where $\gamma_{gc} = |\widetilde{\gamma}_{gc}|$ is the ground-corrected coherence, $\mathsf{Re}[\bullet]$ and $\mathsf{Im}[\bullet]$ are the real and imaginary part operators, respectively, and n is an integer describing the ambiguity of the inversion. The lowest positive Δh is chosen in cases when HOA is larger than forest height.

In [21], [22], it has been shown that Δh estimated from VV-polarized TanDEM-X data is correlated with H95, which is a lidar metric for forest height (see Section III-E for the definition). It has also been shown in [21], [22] that the uncorrected area-fill factor defined as:

$$\eta_0 = \frac{1}{1+\mu} = \frac{\frac{1}{2}\left(1+\gamma_{gc}^2\right) - \mathsf{Re}\left[\widetilde{\gamma}_{gc}\right]}{1 - \mathsf{Re}\left[\widetilde{\gamma}_{gc}\right]}. \tag{17}$$

is correlated with vegetation ratio, which is a lidar metric for canopy density (see Section III-E for a definition). The uncorrected area-fill factor can be obtained by solving (10) for η under the assumption that $\rho = 1$. The validity of this assumption has been discussed in [21], and it has been concluded that at high frequencies, such as for the X-band data used in [21], [22], the ground- and vegetation-level scattering coefficients are similar as the wavelength is short compared to the size of the scatterers and the orientation of the scatterers can be considered random.

As mentioned earlier, the derivation of (15), (16), and (17) is based on the assumption that the total decorrelation is only caused by the volume effect. In the TanDEM-X system used in this study, the near-simultaneous, bistatic acquisition scenario minimizes the temporal decorrelation [49]. Common-band filtering of both master and slave images deals with most of the spatial decorrelation effect caused by different range and Doppler frequency bands [50]. Therefore, the two most significant decorrelation effects other than volume decorrelation are caused by the finite SNR and the system imperfections. In [49], the total coherence for soil and rock for VV-polarized TanDEM-X acquisitions in mid-swath and at a 41-degree incidence angle has been modeled to approximately 0.88 for an occurrence level of 50% and 0.82 for an occurrence level of 90%. In this study, however, all decorrelation effects other than volume decorrelation will be neglected for practical reasons, and the validity of this assumption will be discussed in Section V.

Due to the high resolution of TanDEM-X data and relatively large regions of interests, a large number of independent samples can be used during the estimation of the complex correlation coefficient, and the errors in coherence and phase estimation are negligible. An estimate of the number of looks used during the computation of the complex correlation coefficient will be given in Section III-G.

C. Biomass Models

The following biomass model, called the TLM biomass model or shortly the TBM, is introduced:

$$\widehat{\mathrm{AGB}} = K \cdot \Delta h^{\alpha} \cdot \eta_0^{\beta}, \tag{18}$$

where K, α, and β are unknown model parameters, and Δh and η_0 are forest height and canopy density estimates obtained from the TLM inversion using (16) and (17), respectively.

Fig. 2. Location of the two test sites used in this study.

The TBM will be compared to a linear, zero-intercept model, which scales the ground-corrected interferometric height to biomass:

$$\widehat{\mathrm{AGB}} = D \cdot h_{gc}, \tag{19}$$

where D is a scaling factor, which needs to be estimated from the training data. This model has been proposed in [17]. In the following, it will be referred to as the scaling model, or simply the SM.

D. Evaluation Strategy

The models will be evaluated using multiple TanDEM-X acquisitions acquired over two, geographically separated test sites in Sweden, during three consecutive summers and at different HOAs. The models will be tested both for their explanatory values (that is how well they can be fitted to the data), and their predictive values (that is how well they can predict biomass from other data). The models will thus be tested for their robustness to the change of test site, acquisition year, and acquisition HOA. They will also be used to produce biomass maps, to see how well the spatial structures can be reproduced.

III. Data

A. Test Sites

Remningstorp is a hemi-boreal test site situated in southern Sweden (58° 28' N, 13° 38' E), see Figure 2. It is fairly flat with ground slopes at stand level lower than 5° (computed from a 50 m × 50 m digital terrain model, DTM). The test site covers approximately 1200 ha of productive forest land, and the forest consists primarily of Norway spruce (*Picea abies* (L.) Karst.), Scots pine (*Pinus sylvestris* L.), and birch (*Betula* spp.). For a description of the test site, see [51].

Krycklan is a boreal test site located in northern Sweden (64° 14' N, 19° 46' E), see Figure 2. Krycklan is situated 720 km north-north-east of Remningstorp. Unlike Remningstorp, Krycklan has a strongly undulating topography with ground slopes on stand level up to 19° (again, computed from a 50 m $\times$ 50 m DTM). The forest is dominated by Norway spruce and Scots pine. For a description of the test site, see [52].

B. In-Situ Data

A set of 32 circular, 40-meter radius plots is available for Remningstorp. Field inventories were conducted during the autumn of 2010 and spring of 2011. For each plot, all trees with a diameter at breast height (dbh) higher than 5 cm were calipered and tree species were determined. Height was measured for a subset of roughly 10% of the trees. Out of the thirty-two plots, twenty-one are spruce-dominated (more than 2/3 of biomass), five are pine-dominated, and two are birch-dominated. Three plots consist of a mixed spruce and pine forest and one plot consists of a mixed forest with all three tree species.

In Krycklan, a set of 31 stands of irregular shape and sizes between 2.4 and 26.3 hectares were inventoried in the summer of 2008. Systematic grids of circular field plots (radius 10 m) were laid out in each stand. The spacing of each grid was selected to give between 8 and 13 field plots per stand. For each field plot, all trees with a dbh higher than 4 cm were calipered and the species were determined. Tree height and age were also measured for 1–2 randomly chosen sample trees in each field plot. Of the thirty-one stands, five are spruce-dominated, thirteen are pine-dominated, three are mixed coniferous, and the remaining ten are mixed forest stands.

C. Biomass Estimates

For both test sites, estimates of above-ground dry biomass have been made from the in-situ data using the Heureka system [53], which implements the allometric functions described in [23]. The allometric functions have been derived using multiple regression analysis of data from 1286 trees (Norway spruce, Scots pine, and birch) from 131 stands located across Sweden and described in [54], [55].

Stem volume growth has been modeled in Heureka using the radial growth functions described in [23]. Although SAR acquisitions have been made in the summer, which is in the middle of a growth season, biomass estimates for the end of the preceding growth season will be used throughout this study. The performance of the volume growth model used in Heureka has been evaluated in [56] using 1711 permanent plots from the National Forest Inventory (NFI) database. The prediction error (RMSE) for the stem volume has been found to be around 15%, and a small underestimation (bias) of 2% has been observed for spruce. Due to the close relation between forest volume and above-ground dry biomass, similar errors are also expected for the AGB.

A realistic estimate of the uncertainty in the reference biomass data used in this study is 15%, primarily based on the results presented in [56] and the errors presented in [23]. Although the sampling procedures in Remningstorp and Krycklan include the measurement of the dbh for a large set of trees (all trees with dbh larger than 5 cm in Remningstorp), height has only been measured for a subset of trees, and thereafter extrapolated to the other trees using regression from the dbh. Since both the dbh and height are used for biomass estimation, the input variables to the allometric equations are correlated, which increases the uncertainty of the aggregated estimates. A possible bias will also occur when the models presented in [23] are used locally, on data which may deviate from the data used for the derivation of these models. Additional uncertainties, such as in-situ measurement errors and errors introduced during the determination of plot areas also contribute to the total error.

D. Forest Change Detection

After field measurements, several plots/stands have been altered through clearing, thinning, or clear-cutting. In Remningstorp, the altered plots have been identified using lists of management procedures provided by the managing company, SPOT-5 image analysis, and field visits. Three plots have been altered between the SAR acquisitions from 2011 and from 2012, and additional eight between the SAR acquisitions from 2012 and from 2013. In Krycklan, only SPOT-5 image analysis has been used. Two stands have been altered already before the first SAR acquisition in 2011, but no changes have been detected after that. Altered plots/stands have been disregarded in this study.

E. Lidar Data

Two lidar metrics have been extracted from $10\,\mathrm{m} \times 10\,\mathrm{m}$ maps provided within the BioSAR 2008 and 2010 campaigns [51], [52]. The 95th-percentile forest height, called H95, has been computed as the 95th percentile of all lidar returns above a threshold of 1 m or 10% of the maximal return within a $10\,\mathrm{m} \times 10\,\mathrm{m}$ cell. The lidar vegetation ratio, called VR, has been computed as the ratio between the number of returns from above that threshold to all returns. The VR is thus a measure of canopy density.

Biomass maps with a $10\,\mathrm{m} \times 10\,\mathrm{m}$ resolution have been derived from multiple regression analysis of different lidar metrics and species stratification maps, see [51], [52]. In Remningstorp, 212 circular field plots with a radius of 10 m and distributed in a systematic grid over the entire test site, have been used for model training. In Krycklan, the previously mentioned field plots located within the 31 stands, together with additional 110 circular field plots surveyed with the same methodology and positioned within the central part of the test site have been used for model training. The uncertainty in the biomass maps is estimated to 20%, based on the uncertainties reported in [51], [52]. The maps will only be used in a qualitative, side-by-side comparison.

Note that the lidar data have not been corrected for growth. Since the lidar data are only used in qualitative comparisons, this does not affect the quantitative results presented in this study.

F. Digital Terrain Model

As ground level reference, the national, lidar-scanned digital terrain model (DTM) acquired by the Swedish Land Survey is used [14]. The DTM has a 2 m $\times$ 2 m grid, with a mean height error lower than 0.5 meters. Lidar scanning has been performed from an airplane flying at an altitude between 1700 and 2300 meters. Point density on the ground is between 0.5 and 1 point per square meter. In the southern part of the country, lidar scanning has primarily been performed during non-vegetative periods to minimize the contribution of leaves, grass, crops, etc.

G. InSAR Data

TanDEM-X is a twin-satellite, X-band (9.65 GHz) SAR interferometer in which acquisitions are made almost simultaneously [49]. Bistatic-interferometric, VV-polarized, stripmap-mode TanDEM-X (TDM) acquisitions made at low HOA in the ascending mode are used in this study. The choice of the low-HOA data is motivated by the better sensitivity to forest height [57]. A summary of the data can be found in Table I, where background color coding by HOA has been applied. Note that the data from 2012 feature lower HOAs than the data from 2011 and 2013. The nominal angle of incidence varies between 41.2 and 41.7 degrees for Remningstorp and between 40.4 and 41.0 degrees for Krycklan. For images 1 and 9 in Table I (the first acquisitions for each test site), the scene center resolutions provided by the DLR in the meta files are: 1.8 m in ground range and 6.6 m in azimuth. For the rest of the images, the ground range resolution is 2.7 m and the azimuth resolution is 3.3 m.

The InSAR data have been interferometrically processed using an in-house developed algorithm based on [50]. The raw interferograms have been ground-corrected in radar geometry using a linearly interpolated DTM and taking into consideration the quasi-bistatic acquisition geometry and satellite displacement between transmission and reception of the signals. A 5-meter buffer zone has been added prior to plot/stand-level averaging of the ground-corrected interferograms. The lowest number of looks has been estimated to 320, computed as the ratio between the area of the smallest plot/stand in the data set (excluding the buffer zone), and the ground range and azimuth resolutions for the image with the lowest resolution. Absolute phase calibration has been done using ground reference points derived from a non-forest mask. No unwrapping has been found necessary due to the limited height variations in the flattened interferogram. Geocoding error and height measurement errors have been estimated using two 5-meter trihedral corner reflectors situated within the Remningstorp site. The geocoding offset has been found lower than 2 m and the standard deviation of the measured elevation of the scattering center has been found lower than 10 cm. For the creation of the ground-corrected coherence and height images, a 5×5 averaging window has been used.

TABLE I

SUMMARY FOR THE EXPERIMENTAL DATA USED IN THIS STUDY. MEAN VALUES FOR ALL PLOTS ARE GIVEN. BACKGROUND SHADING HAS BEEN APPLIED ACCORDING TO HOA. N IS THE NUMBER OF AVAILABLE PLOTS/STANDS FOR EACH ACQUISITION.

			InSAR data			In-situ & lidar data						
						N	Biomass [t/ha]			H95 [m]		
Nr	Site	Date	$B_\perp$ [m]	HOA [m]	Coherence		min	mean	max	min	mean	max
1	Remningstorp	20110604	282	49	0.65	32	42	148	242	14	23	32
2		20110809	266	52	0.67	28	42	150	242	14	23	32
3		20110820	258	54	0.66							
4		20120601	432	32	0.54	29	42	145	242	14	22	30
5		20120828	370	37	0.54							
6		20130702	270	51	0.66	21	42	143	242	14	22	30
7		20130724	226	61	0.73							
8		20130804	220	63	0.73							
9	Krycklan	20110617	258	52	0.71	29	23	94	183	7	16	21
10		20110720	250	54	0.75							
11		20110811	242	55	0.76							
12		20110822	240	56	0.78							
13		20120717	374	36	0.59							
14		20120808	360	37	0.61							
15		20120819	350	39	0.62							
16		20130601	270	50	0.73							
17		20130623	260	52	0.71							
18		20130726	216	62	0.79							

Color coding by HOA:	30 m	40 m	50 m	60 m

IV. RESULTS

In this section, the two models will be fitted to the experimental data, and biomass will be estimated.

The significance of each model parameter will be studied using the Student's t-test. This test evaluates the hypothesis that the expectation value of the normally distributed parameter estimate $\hat{\beta}$ is β_0. The t-statistic is computed as:

$$t = \frac{\hat{\beta} - \beta_0}{\hat{\sigma}_\beta}, \tag{20}$$

where $\hat{\sigma}_\beta$ is the estimated standard deviation of $\hat{\beta}$. The Student's t-test will here be used to test the hypothesis that $\beta_0 = 0$. For a known number of degrees of freedom, the probability p of obtaining a certain t-statistic can be computed from the t-distribution. A low p-value means that β is a significant parameter.

The goodness-of-fit of each model will be evaluated using the coefficient of determination R^2, which describes the fraction of the total variability observed in the data that can be explained by the model:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}, \tag{21}$$

where Y_i are the observed values, $\hat{Y}_i$ are the corresponding modeled values, and $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ is the average observed value.

The model error will be evaluated using the root-mean-square error (RMSE), which is computed as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}. \tag{22}$$

The fitting of the models has been done using non-linear least squares, as implemented in the `nls` function provided within the R-package [58]. Although the use of linear least squares is possible for the TBM model after a logarithmic transform, non-linear regression has in many cases provided lower RMSE and larger R^2, and it is therefore used throughout this study.

A. *Interferometric Height and Coherence*

In Figure 3(a), TanDEM-X interferometric height h_{gc} is plotted against lidar height H95, separately for each year. A good correlation can be observed, but the interferometric height is approximately 5–10 m lower for almost all stands.

For some Remningstorp plots with H95 just above 25 m, the interferometric height is approximately 5 m higher for the 2012 acquisitions (with HOAs equal to 32 m and 37 m) than for the 2011 and 2013 acquisitions (with HOAs around and above 50 m). This effect has previously been discussed in [21], where it has been concluded that it is caused by an interference effect occurring when ground- and vegetation-level scattering is of similar strength and when the distance between the respective scattering centers is around HOA/2. The affected plots consist of former seed trees with new understorey vegetation. The trees are sparse, and allow for a significant penetration through the gaps, and the understorey vegetation layer boosts the ground-level scattering.

In Figure 3(b), ground-corrected TanDEM-X coherence γ_{gc} is plotted against lidar vegetation ratio, separately for each year. It can be observed that coherence is consistently lower for the acquisitions from 2012 than for the other two acquisitions, due to the larger baseline, and for some plots, the coherence falls below 0.3. It is noted that these plots are the same, sparse plots with former seed threes and rich understorey vegetation, and that this low coherence occurs due to the aforementioned interference effect.

In Figure 3(c), TanDEM-X interferometric height is plotted against reference biomass. For Krycklan, there is a good correlation between the interferometric height and biomass. For Remningstorp, however, there is a large height variance, especially in the case of the data from 2012, with low HOA. It is noted that the plots with relatively low biomass but high interferometric height are the same sparse plots that have been discussed in the previous paragraphs. Note that one of these plots have been altered in 2012 and additional two in 2013, as described in [21], and they are not included in the scatter plots from 2012 and 2013.

B. TLM Inversion Products

In Figure 4(a), the level distance Δh inverted using (16) is plotted against lidar height H95. It can be observed that the correlation between Δh and H95 is better than between the interferometric height and H95, but the bias is different, as observed in [21], [22]. Note that the slope of the inverted level distance Δh changes at low H95. The reason for this will be discussed in Section V.

In Figure 4(b), the area-weighted backscatter ratio μ inverted using (15) is plotted in decibels against lidar vegetation ratio. Although these two parameters measure different properties, a good correlation can be observed.

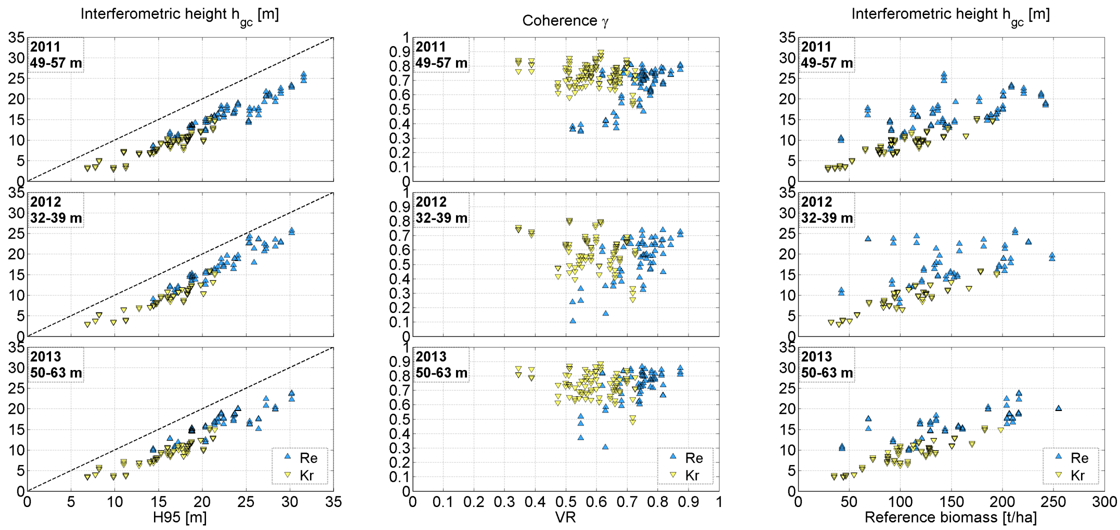
In Figure 4(c), the uncorrected area-fill factor η_0 inverted using (17) is plotted against lidar vegetation ratio. The correlation is good for most stands.

C. Parameter Estimation Results

In Table II, estimates of K, α, and β are shown for the TLM biomass model (TBM). The TBM is able to explain between 65% and 89% of the variance observed in the data. It can be observed that α is similar for both test sites and for most acquisitions, with most values close to one. The other exponent, β, is similar for all acquisitions made over the same test site, but it changes between the test sites. For acquisitions made in Krycklan, it is close to one, whereas for those made in Remningstorp, it is closer to three. The third parameter, K, shows a large variance with values between 0.4 and 27.1.

Based on these observations, it is reasonable to let the exponents become constants. The exponent α can be fixed to the same value for both test sites, while the exponent β must change between test sites. In Table III(a), regression results for the TBM with α fixed to 1.25, and β fixed to 2.64 for Remningstorp and 1.16 for Krycklan are shown. The chosen values are all average values for the estimates presented in Table II. The estimated values of the slope constant K', are more stable than the estimates of K, between 6.6 and 10.2, without any significant difference between the two test sites. Note that the lowest R^2 is obtained for the image from Remningstorp with the lowest HOA. Since the choice of the fixed parameters is based on the mean of all values, it is biased towards acquisitions with large HOA, as they are more frequent.

In Table III(b), regression results for the scaling model (SM) are shown. It can be observed that the slope is very stable for both Remningstorp and Krycklan, but it changes between the two sites. For Remningstorp, it is between 8.3 and 9.6, whereas for Krycklan, it is between 11.3 and 12.2. The two lowest values are obtained for the images

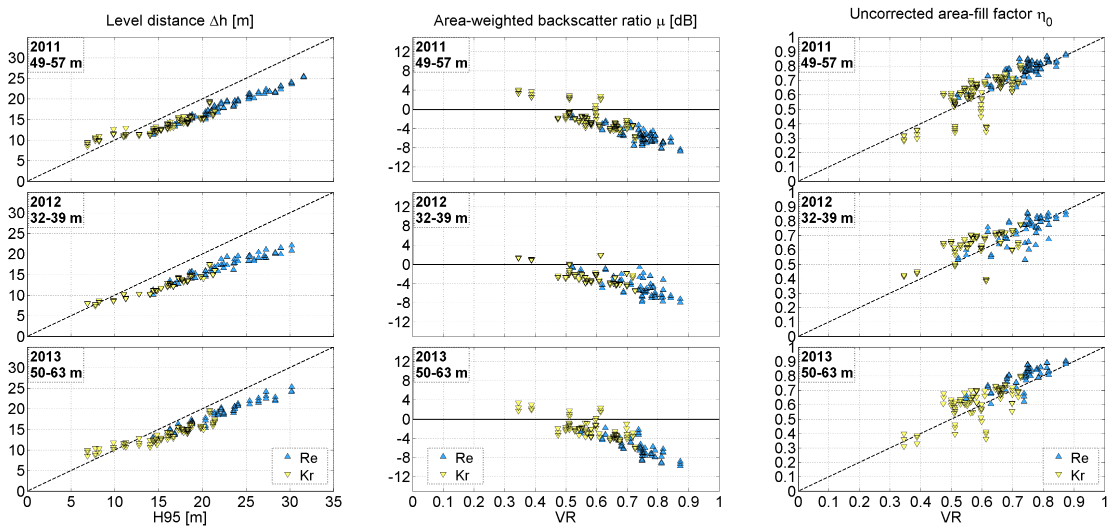


(a) TDM interferometric height vs. lidar H95

(b) TDM ground-corrected coherence vs. lidar VR

(c) TDM height vs. reference biomass

Fig. 3. TanDEM-X interferometric height and ground-corrected coherence are plotted against reference data. Acquisition year and HOA intervals are shown for each subplot. Note: several points may overlap. In Remningstorp, there are 88 points in 2011, 58 points in 2012, and 63 points in 2013. In Krycklan, there are 116 points in 2011 and 87 points in both 2012 and 2013.



(a) Level distance vs. lidar H95

(b) Area-weighted backscatter ratio vs. lidar vegetation ratio

(c) Uncorrected area-fill factor vs. lidar vegetation ratio

Fig. 4. Inverted TLM parameters are plotted against reference data. Acquisition year and HOA intervals are shown for each subplot. Note: several points may overlap. In Remningstorp, there are 88 points in 2011, 58 points in 2012, and 63 points in 2013. In Krycklan, there are 116 points in 2011 and 87 points in both 2012 and 2013.

over Remningstorp acquired at the lowest HOA values. For these, the coefficient of determination is -0.07 and 0.13, and the SM is not able to explain the variance for these acquisitions.

Note that the fixed parameters chosen above as averages of the estimated parameters shown in Table II can also be estimated from regression of all relevant data but this approach has not been chosen here. The main purpose of this part is to show that fixed exponents provide a more stable model. The choice of exact parameter values is not of primary interest to this study.

D. Biomass Estimation Results

Residual scatter plots for the models are shown together with two-sigma error bars in Figure 5. It can be observed that only for the Remningstorp acquisitions from 2012, fixing of the exponents in the TBM significantly decreases model performance, as it also can be observed by comparing the R^2-values in Table II and Table III(a). The SM performs poorer in Remningstorp than in Krycklan.

Both residual and prediction root-mean-square error (RMSE) values are shown in Table IV for the TBM, in Table V for the TBM with fixed exponents, and in Table VI for the SM. As expected, the performance of the TBM is poorer in across-site evaluation and for large difference in HOA in Remningstorp, primarily due to the differences in the exponent β. It can also be observed that the TBM with fixed exponents gives a much lower and more stable prediction RMSE without increasing the residual RMSE significantly. In the case of the SM, both residual and prediction RMSEs are low in Krycklan, but higher in Remningstorp and across-sites.

In Figure 6, the dependence of the RMSE on the parameters K, α, and β is studied. The default values of the parameters are: $K = 7.42$, $\alpha = 1.25$, $\beta = 2.64$ for Remningstorp and $\beta = 1.16$ for Krycklan, and they are marked with vertical lines. Each parameter is varied around its default value and the RMSE is computed. While one parameter is varied, the other two are held constant at their default values. Remningstorp and Krycklan are shown separately, and color coding according to HOA has been applied. Note the significant difference between Remningstorp and Krycklan in model sensitivity to different parameter settings at different HOAs.

E. Biomass Mapping

In Figure 7, biomass maps obtained using the TBM and SM are shown for both Remningstorp and Krycklan, and compared to lidar-derived biomass maps. For both test sites, the first acquisition from 2011 has been used: nr 1 for Remningstorp and nr 9 for Krycklan), together with the respective parameters presented in Table II and Table III(b). Note that forest management procedures may have been conducted between the acquisition of lidar and TanDEM-X data. Note also that, in Remningstorp, regions not covered by lidar scanning have been masked out.

As observed earlier during the residual study, the TBM performs well both in Remningstorp and in Krycklan, if the parameters obtained from acquisitions made at similar HOA and within the same test site are used. Although the SM performs well in Krycklan, it can be observed that it cannot reproduce the variance of biomass in Remningstorp.

V. DISCUSSION

The TLM biomass model (TBM) is able to explain 65–89% of the variance observed in data, with a residual RMSE of 12-19% of the mean biomass, for 18 TanDEM-X images acquired over two test sites in Sweden. In cases when different data are used for training and validation, the model shows poorer results, with a prediction RMSE often exceeding 30% in across-site scenarios or when the difference in HOA is large. However, the TBM can be stabilized by fixing α and by letting β be a site-dependent constant. In that case, the prediction RMSE is below 20% for most of the acquisitions.

The TBM is here compared to a linear, zero-intercept model, which scales the interferometric height to biomass. This scaling model (SM) has earlier been used in [17], where a stand-level residual RMSE of 19% has been obtained for a Norwegian test site using two images, one from the ascending orbit and one from the descending orbit, with the respective HOAs 23 m and 122 m. In [17], the scaling factor has been estimated to 14 t/ha/m, while in this study, the same factor is between 8.3 and 9.6 for Remningstorp and between 11.3 and 12.2 for Krycklan. A method for biomass change detection has also been proposed in [17], based on direct scaling of the change in the interferometric height to change in biomass, without the need for a high-resolution DTM. From the results of this study, it can be concluded that whereas Krycklan shows many similarities with the Norwegian test site, Remningstorp appears to be significantly different. The dependence of the interferometric height on HOA and the horizontal forest structure is especially large in Remningstorp, and the model presented in [17] does not function well in this test site.

In [19], an approach based on the interferometric water cloud model (IWCM) and multi-temporal averaging of stand-level biomass estimates from eighteen acquisitions made over Remningstorp at HOAs between 49 m and 358 m, in both summer and winter, gives an RMSE of 16%, whereas multi-temporal averaging of seven images acquired at temperatures below 3°C gives an RMSE of 14%. In the case of a single image, the RMSE is in the interval 17–33%. Additionally, a penetration depth (PD) model is introduced in [19], which uses a height-to-biomass allometric equation to compute biomass from the sum of the interferometric height and the penetration depth. The PD-based approach gives RMSE values in the interval 18–33%. The penetration depth is the only parameter that needs to be estimated (the allometric relation is assumed to be known). However, the model does not account for the horizontal structure of the forest, which can be problematic in the case of a commercial forest, where management activities such as thinnings and clearings affect the denseness of the forest and its biomass, but not necessarily its height.

In [20], an approach based on multiple regression of the interferometric height and the ground-corrected coherence, and their transformed versions gives an RMSE of 14% for one TanDEM-X image acquired over Krycklan (nr 9 in Table I) and 17% for one TanDEM-X image acquired over Remningstorp (nr 1 in Table I). The approach presented in [20] is based on multiple regression of interferometric observables, and the results may not be representative, especially considering the limited extent of the data used in the study.

A. TLM Inversion Requirements

As biomass predictors, the TBM uses two parameters obtained from the inversion of a two-level model (TLM). The inversion of the TLM requires a high-resolution digital terrain model (DTM), and it is based on the assumption that volume decorrelation is the dominant decorrelation effect.

In Sweden, there is a lidar DTM covering the whole country, and similar products are or will soon be available in many other countries. Therefore, the presented approach can already be used in many places and on a large scale, as the global TanDEM-X data used for DEM generation have been made available by the DLR for scientific use. Since ground surface is temporally stable in most forested areas, the availability of high-resolution DTMs will only increase with time. The exact requirements on the DTM have not been studied here, but it is possible that a coarser DTM than can be sufficient, as exemplified in [17]. Other techniques, such as low-frequency SAR, may also be used as ground reference.

In the presented approach, volume decorrelation has been assumed to be the dominant decorrelation effect. In the case of the single-pass interferometric, bistatic-mode TanDEM-X data used here, the most significant decorrelation sources other than volume effects are the thermal noise and system imperfections. These effects have been ignored throughout this study to keep the inversion process simple.

However, it has been observed in [21] that the TLM inversion provides unrealistically high Δh for clear-cuts, and that the inverted level distance Δh is affected by a HOA-dependent offset. In Figure 4(a), a change of slope has been observed for low H95. Additionally, a HOA dependence has been observed in the estimated model parameters α, β, and K presented in Table II, and that this dependence is stronger in Remningstorp, where the forest is taller and the relative HOA is lower.

A probable cause for these effects is that the SNR and system decorrelation effects have not been considered in the TLM inversion process. Since the interferometric phase has been calibrated using non-forested areas, the complex correlation coefficient for low forest and open areas has a high, yet non-unitary real part and low imaginary part. The TLM cannot model a complex correlation coefficient with high coherence and low phase without making Δh close to HOA/2. This introduces a HOA-dependent offset in the estimated Δh.

A solution for this issue can be obtained through the modeling of a real-valued system and SNR decorrelation term γ_{sys}, and replacing $\tilde{\gamma}_{gc}$ by $\frac{\tilde{\gamma}_{gc}}{\gamma_{sys}}$ in (15), (16), and (17). If γ_{sys} can be estimated, e.g., from the data, then an improvement of the TLM inversion performance can be expected.

A second probable cause for these effects is that at low HOA relative forest height, the modeling of the exact vertical distribution of scatterers becomes more important, as the phase change with height is larger. The assumption of two scattering levels may then become too simplistic.

Nevertheless, the issues related to HOA dependence can be avoided in practical use by a sensible choice of HOA. Moreover, η_0 can often compensate for the large Δh, as observed for Remningstorp in Figure 7, where biomass mapping in open fields is accurate.

B. Influence of Forest Structure

An interesting observation can be made about the estimated values for the exponent β, associated with the canopy density estimate η_0, which changes significantly between Remningstorp and Krycklan. This can be explained by different structure of the trees in Remningstorp and in Krycklan. It can be estimated using the Heureka system, that 68% of the total biomass in the 32 plots in Remningstorp is confined to the stem, while for the 31 stands in Krycklan, the same number is 76%. In [59], it is concluded that the trees in northern Sweden generally have smaller crowns than in southern Sweden. A larger β in Remningstorp will reduce the contribution of the canopy density estimate η_0 to the total biomass, thus compensating for the fact that the forest in Remningstorp has in general denser canopies, at similar biomass and height.

The TLM can be used to study temporal change of canopy density from multi-temporal, single-pass InSAR acquisitions. By keeping Δh constant for all acquisitions and letting η_0 vary, an over-determined equation system is obtained. A time series study of η_0 can provide information on the change of the canopy density, due to seasonal variations, management procedures such as clearing, thinning, and clear-cutting, or natural disasters, and, eventually, biomass change can be estimated as well.

C. Future Development

There is an ongoing debate about the mechanisms of microwave penetration into forest canopy, and whether the significant penetration of X-band SAR into the canopy is primarily due to the dielectric penetration through the scatterers or penetration through the canopy gaps [60]. In this study, it is shown that the inclusion of canopy gaps in an interferometric model can be beneficial for model inversion, but the dielectric penetration has been disregarded, and further discussion on the penetration mechanisms is left for future studies.

The presented approach has been evaluated on VV-polarized, X-band SAR data. An evaluation of this approach on other frequencies, for instance C-band, as well as other polarizations is of large interest. The presented approach is principally not restricted to the used frequency, although the TLM inversion process may need to be revisited at other frequencies, where a different choice of ρ may be motivated. Also, as shown in [16], the difference between the interferometric height at HH- and VV-polarization can be several meters. Although the exact relation between the inverted parameters Δh and η_0 and the lidar estimates of forest height and canopy density, and biomass will most likely be different at other polarizations and frequencies, the presented approach may still be useful.

This study has been restricted to data acquired at a 41-degree nominal angle of incidence. The influence of the incidence angle requires a separate study. An evaluation of the presented approach on tropical forest is also of interest. As the tropical forest is, in general, taller and denser, the penetration through canopy gaps is expected to be lower, which certainly will affect TLM inversion.

VI. Conclusions

A new biomass model is proposed, in which biomass is estimated from forest height and canopy density estimates obtained from the inversion of a two-level model (TLM) using single-pass interferometric SAR data. In this study, bistatic-mode, VV-polarized TanDEM-X data acquired at a 41-degree nominal incidence angle over two Swedish test sites separated by 720 km are used together with the national, digital terrain model (DTM) with a grid posting of $2\,\mathrm{m} \times 2\,\mathrm{m}$ and a vertical accuracy better than 0.5 m. Compared to other studies, the presented approach provides similar or better results in terms of biomass retrieval, a larger data set has been used for the evaluation, and across-site and across-acquisition biomass retrieval scenarios have been studied.

It is here concluded that the two test sites used in this study feature quite different forest, and regional training of the new model is required in operational use. However, only one of the three model parameters has been found significantly dependent on the test site, and the regional model training can be done using only a few data points, e.g., from the National Forest Inventory database. The HOA dependence most likely caused primarily by the lack of system and SNR decorrelation modeling can be suppressed either by choosing HOAs larger than approximately twice the forest height, which is the case for most of the global TanDEM-X acquisitions over boreal forests, or by the modeling of a real-valued system and SNR decorrelation term.

Since a high-resolution digital terrain model is required for TLM inversion, the presented approach is suitable for frequent mapping of large areas of forest in regions with known topography. However, the ground surface is most often temporally stable, and only one DTM acquisition is required. Thereafter, forest height, canopy density, and biomass mapping can be done using spaceborne SAR with large coverage, high resolution, and frequent acquisitions. Therefore, the presented approach is useful for the monitoring of national forest resources, and for improved forest management. With an access to the global TanDEM-X data, national maps of forest height, canopy density, and biomass can be created.

ACKNOWLEDGMENTS

The authors would like to thank the German Aerospace Center (DLR) for the TanDEM-X data and support provided within proposal XTI_VEGE0376, the Swedish Land Survey for the DTM and SPOT-5 data, Skogssällskapet for the information on forest management activities, and the European Space Agency (ESA) for lidar data from the BioSAR 2008 and 2010 campaigns.

REFERENCES

[1] IPCC, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, M. M. Dahe Qin, M. Marquis, K. Averyt, M. M. B. Tignor, and H. L. Miller, Eds. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2007.

[2] S. H. Spurr, *Aerial photographs in forestry*. Ronal Press Company, 1948.

[3] J. Hyyppä, H. Hyyppä, M. Inkinen, M. Engdahl, S. Linko, and Y.-H. Zhu, "Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes," *Forest Ecology and Management*, vol. 128, no. 1–2, pp. 109–120, 2000.

[4] B. St-Onge, Y. Hu, and C. Vega, "Mapping the height and above-ground biomass of a mixed forest using lidar and stereo Ikonos images," *International Journal of Remote Sensing*, vol. 29, pp. 1277–1294, 2008.

[5] H. Persson, J. Wallerman, H. Olsson, and J. E. Fransson, "Estimating forest biomass and height using optical stereo satellite data and a DTM from laser scanning data," *Canadian Journal of Remote Sensing*, vol. 39, no. 03, pp. 251–262, 2013. [Online]. Available: http://pubs.casi.ca/doi/abs/10.5589/m13-032

[6] M. Nilsson, "Estimation of tree heights and stand volume using an airborne lidar system," *Remote Sensing of Environment*, vol. 56, pp. 1–7, 1996.

[7] J. Hyyppä, H. Hyyppä, D. Leckie, F. Gougeon, X. Yu, and M. Maltamo, "Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests," *International Journal of Remote Sensing*, vol. 29, pp. 1339–1366, 2008.

[8] E. Næsset, T. Gobakken, J. Holmgren, H. Hyyppä, J. Hyyppä, M. Maltamo, M. Nilsson, H. Olsson, Å. Persson, and U. Söderman, "Laser scanning of forest resources: The nordic experience," *Scandinavian Journal of Forest Research*, vol. 19, pp. 482–499, 2004.

[9] S. C. Popescu, "Estimating biomass of individual pine trees using airborne lidar," *Biomass and Bioenergy*, vol. 31, pp. 646–655, 2007.

[10] E. Næsset, T. Gobakken, S. Solberg, T. G. Gregoire, R. Nelson, G. Ståhl, and D. Weydahl, "Model-assisted regional forest biomass estimation using lidar and InSAR as auxiliary data: A case study from a boreal forest area," *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3599–3614, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0034425711003257

[11] K. J. Ranson, G. Sun, R. H. Lang, N. S. Chauhan, R. J. Cacciola, and O. Kilic, "Mapping of boreal forest biomass from spaceborne synthetic aperture radar," *Journal of Geophysical Research*, vol. 102, pp. 29 599–29 610, 1997.

[12] B. Koch, "Status and future of laser scanning, synthetic aperture radar and hyperspectral remote sensing data for forest biomass assessment," {*ISPRS*} *Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 6, pp. 581–590, 2010, {ISPRS} Centenary Celebration Issue. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924271610000845

[13] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse Problems*, vol. 14, pp. R1–R54, 1998.

[14] Swedish Land Survey (Lantmäteriet), "Produktbeskrivning: GSD-Höjddata, grid 2+," Tech. Rep., 2013, (in Swedish).

[15] J. Praks, O. Antropov, and M. T. Hallikainen, "LIDAR-aided SAR interferomety studies in boreal forest: Scattering phase center and extinction coefficient ant X- and L-band," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3831–3843, 2012.

[16] F. Kugler, D. Schulze, I. Hajnsek, H. Pretzsch, and K. P. Papathanassiou, "TanDEM-X Pol-InSAR performance for forest height estimation," *IEEE Transactions on Geoscience and Remote Sensing*, 2014, (published on-line, in press).

[17] S. Solberg, R. Astrup, J. Breidenbach, B. Nilsen, and D. Weydahl, "Monitoring spruce volume and biomass with InSAR data from TanDEM-X," *Remote Sensing of Environment*, vol. 139, pp. 60–67, 2013.

[18] S. Solberg, R. Astrup, T. Gobakken, E. Naesset, and D. J. Weydahl, "Estimating spruce and pine biomass with interferometric X-band SAR," *Remote Sensing of Environment*, vol. 114, pp. 2353–2360, 2010.

[19] 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 Sensing*, vol. 5, pp. 5574–5597, October 2013.

[20] H. Persson and J. E. S. Fransson, "Relation between interferometric TanDEM-X data and boreal forest variables," *Remote Sensing of Environment*, 2014, (submitted).

[21] M. J. Soja, 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, (submitted).

[22] M. J. Soja and L. M. H. Ulander, "Two-level forest model inversion of interferometric TanDEM-X data," in *10th European Conference on Synthetic Aperture Radar (EUSAR)*, 2014, pp. 1137–1140.

[23] H. Petersson, "Biomassafunktioner for trädfraktioner av tall, gran och björk i sverige," SLU, Institutionen för skoglig resurshållning och geomatik, Arbetsrapport 59, Umeå, Tech. Rep., 1999, (in Swedish).

[24] M. Cannell, "Woody biomass of forest stands," *Forest Ecology and Management*, vol. 8, no. 3–4, pp. 299–312, 1984. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0378112784900628

[25] J. Chave, C. Andalo, S. Brown, M. A. Cairns, J. Q. Chambers, D. Eamus, H. Fölster, F. Fromard, N. Higuchi, T. Kira, J.-P. Lescure, B. W. Nelson, H. Ogawa, H. Puig, B. Rièra, and T. Yamakur, "Tree allometry and improved estimation of carbon stocks and balance in tropical forest," *Oecologia*, vol. 145, pp. 87–99, 2005.

[26] S. Brown, A. J. R. Gillespie, and A. E. Lugo, "Biomass estimation methods for tropical forests with applications to forest inventory data," *Forest Science*, vol. 35, pp. 881–902, 1989.

[27] D. Zianis, P. Muukkonen, R. Mäkipää, and M. Mencuccini, *Biomass and Stem Volume Equations for Tree Species in Europe*, ser. Silva Fennica. The Finnish Society of Forest Science, 2005.

[28] E. Næsset, "Determination of mean tree height of forest stands using airborne laser scanner data," {*ISPRS*} *Journal of Photogrammetry and Remote Sensing*, vol. 52, no. 2, pp. 49–56, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924271697830006

[29] J. Holmgren, *Estimation of forest variables using airborne laser scanning*. Department of Forest Economics, Swedish University of Agricultural Sciences, 2003, vol. 278, no. 2. [Online]. Available: http://diss-epsilon.slu.se/archive/00000334/

[30] S. R. Cloude and K. P. Papathanassiou, "Polarimetric SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 5, pp. 1551–1565, 1998. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=718859

[31] K. P. Papathanassiou and S. R. Cloude, "Single-baseline polarimetric SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 2352–2363, 2001. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=964971

[32] S. R. Cloude and K. P. Papathanassiou, "Three-stage inversion process for polarimetric SAR interferometry," *IEE Proceedings on Radar, Sonar & Navigation*, vol. 150, no. 3, pp. 125–134, 2003. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1222372

[33] H. Raggam, K. Gutjahr, R. Perko, and M. Schardt, "Assessment of the stereo-radargrammetric mapping potential of TerraSAR-X multibeam spotlight data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 2, pp. 971–977, 2010.

[34] H. Persson and J. Fransson, "Forest variable estimation using radargrammetric processing of TerraSAR-X images in boreal forests," *Remote Sensing*, vol. 6, pp. 2084–2107, 2014.

[35] Y. Wu and A. H. Strahler, "Remote estimation of crown size, stand density, and biomass on the Oregon transect," *Ecological Applications*, vol. 4, pp. 299–312, 1994.

[36] J. Zhou, C. Proisy, X. Descombes, I. Hedhli, N. Barbier, J. Zerubia, J. P. Gastellu-Etchegorry, and P. Couteron, "Tree crown detection in high resolution optical and lidar images of tropical forest," in *Proc. SPIE*, vol. 7824, 2010. [Online]. Available: http://dx.doi.org/10.1117/12.865068

[37] S. A. Soenen, D. R. Peddle, R. J. Hall, C. A. Coburn, and F. G. Hall, "Estimating aboveground forest biomass from canopy reflectance

model inversion in mountainous terrain," *Remote Sensing of Environment*, vol. 114, no. 7, pp. 1325–1337, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S003442570900368X

[38] S. C. Popescu, R. H. Wynne, and R. F. Nelson, "Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass," *Canadian Journal of Remote Sensing*, vol. 29, pp. 564–577, 2003.

[39] S. C. Popescu and R. H. Wynne, "Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height," *Photogrammetric Engineering and Remote Sensing*, vol. 50, pp. 589–604, 2004.

[40] Y. Illvessalo, "On the correlation between the crown diameter and the stem of trees," *Metsatieteellisen tutkimuslaitoksen julkaisuja*, vol. 38, pp. 1–32, 1950.

[41] G. E. Hemery, P. S. Savill, and S. N. Pryor, "Applications of the crown diameter-stem diameter relationship for different species of broadleaved trees," *Forest Ecology and Management*, vol. 215, no. 13, pp. 285–294, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378112705003373

[42] J. Kalliovirta and T. Tokola, "Functions for estimating stem diameter and tree age using tree height, crown width and existing stand database information," *Silva Fennica*, vol. 39, pp. 227–248, 2005.

[43] J.-S. Lee, K. W. Hoppel, S. A. Mango, and A. R. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 5, pp. 1017–1028, September 1994.

[44] E. Rodriguez and J. Martin, "Theory and design of interferometric synthetic aperture radars," *IEE Proceedings F Radar and Signal Processing*, vol. 139, no. 2, pp. 147–159, 1992.

[45] H. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 5, pp. 950–959, 1992.

[46] J. Hagberg, L. Ulander, and J. Askne, "Repeat-pass SAR interferometry over forested terrain," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 2, pp. 331–340, March 1995.

[47] J. Askne, P. Dammert, L. M. H. Ulander, and G. Smith, "C-band repeat-pass interferometric SAR observations of the forest," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 1, pp. 25–35, 1997.

[48] M. J. Soja, H. Persson, and L. M. H. Ulander, "Estimation of boreal forest biomass from two-level model inversion of interferometric TanDEM-X data," in *IEEE Geoscience and Remote Sensing Symposium (IGARSS)*, Quebec, Canada, 2014, (accepted).

[49] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, and M. Zink, "TanDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 11, pp. 3317–3341, 2007.

[50] S. Duque, U. Balss, C. Rossi, T. Fritz, and W. Balzer, "TanDEM-X payload ground segment, CoSSC generation and interferometric considerations," Remote Sensing Technology Institute, German Aerospace Center (DLR), Tech. Rep., 2012.

[51] L. M. H. Ulander, A. Gustavsson, B. Flood, D. Murdin, P. Dubois-Fernandez, X. Depuis, G. Sandberg, M. J. Soja, L. E. B. Eriksson, J. E. S. Fransson, J. Holmgren, and J. Wallerman, "BioSAR 2010 Technical Assistance for the Development of Airborne SAR and Geophysical Measurements during the BioSAR 2010 Experiment: Final Report," ESA contract no. 4000102285/10/NL/JA/ef, Tech. Rep., 2011. [Online]. Available: http://earth.esa.int/campaigns/DOC/BioSAR_2010_final_report_v1.0.pdf

[52] I. Hajnsek, R. Scheiber, M. Keller, R. Horn, S. Lee, L. M. H. Ulander, A. Gustavsson, G. Sandberg, T. Le Toan, S. Tebaldini, A. M. Guarnieri, and F. Rocca, "BioSAR 2008 technical assistance for the development of airborne SAR and geophysical measurements during the BioSAR 2008 experiment: Final report – BioSAR campaign," ESA contract no. 22052/08/NL/CT, Tech. Rep., 2009. [Online]. Available: http://earth.esa.int/campaigns/DOC/BIOSAR2_final_report.pdf

[53] P. Wikström, L. Edenius, B. Elfving, L. O. Eriksson, T. Lämäs, J. Sonesson, K. Öhman, J. Wallerman, C. Waller, and F. Klintebäck, "The Heureka forestry decision support system: An overview," *Mathematical and Computational Forestry & Natural-Resource Sciences*, vol. 3, pp. 87–94, 2011.

[54] L. G. Marklund, "Biomassafunktioner för gran i Sverige," Institutionen för skogstaxering, Sveriges lantbruksuniversitet, Umeå, Sweden, Rapport 43, 1987.

[55] ——, "Biomassafunktioner för tall, gran och björk i Sverige," Institutionen för skogstaxering, Sveriges lantbruksuniversitet, Umeå, Sweden, Rapport 45, 1988.

[56] N. Fahlvik, B. Elfving, and P. Wikström, "Evaluation of growth models used in the Swedish forest planning system Heureka," *Silva Fennica*, vol. 48, pp. 1–17, 2014.

[57] M. J. Soja 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)*, Melbourne, VIC, Australia, July 2013, pp. 165–168.

[58] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2013. [Online]. Available: http://www.R-project.org/

[59] A. Jakobsons, "The relationship between crown diameter and other tree factors, diameter at breast height in particular: Analysis based on the sample tree material of the national forest inventory," Stockholms skoghögskolan, insitutionen för skogstaxering, Tech. Rep., 1970.

[60] F. De Zan, G. Krieger, and P. Lopez-Dekker, "On some spectral properties of TanDEM-X interferograms over forested areas," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 1, pp. 71–75, January 2013.

Maciej Jerzy Soja was born in 1985 in Warsaw, Poland. He received the B.Sc. and M.Sc. degree in engineering physics from Chalmers University of Technology, Gothenburg, Sweden, in 2008 and 2009, respectively. Since December 2009, he is pursuing his Ph.D. degree in remote sensing in the Radar Remote Sensing Group at the Department of Earth and Space Sciences at Chalmers University of Technology. His main research topic is synthetic aperture radar in forestry.

Henrik Persson was born in 1983. He received the M.Sc. degree in engineering physics from Luleå University of Technology in 2009. After two years of working at the ETH Zürich in Switzerland he started pursuing his Ph.D. degree in forestry remote sensing at the Swedish University of Agricultural Sciences in September 2011. His main research topic is modelling of forest variables from different satellite sensors.

Lars M. H. Ulander (S'86-M'90-SM'04) received the M.Sc. degree in engineering physics in 1985 and the Ph.D. degree in electrical and computer engineering in 1991, both from Chalmers University of Technology. Since 1995 he has been with the Swedish Defence Research Agency (FOI) in Linköping where he is Director of Research in radar signal processing. He is also Adjunct Professor in radar remote sensing at Chalmers University of Technology. His research areas are synthetic aperture radar (SAR), electromagnetic scattering models and remote sensing applications. He is the author or co-author of over 250 professional publications, of which more than 50 are in peer-reviewed scientific journals. He is the holder of five patents.

TABLE II

RESULTS FOR THE TLM BIOMASS MODEL (TBM). FOR EACH PARAMETER, THE ESTIMATED STANDARD DEVIATIONS σ, t-STATISTICS, AND p-VALUES ARE SHOWN. FOR THE WHOLE MODEL, THE COEFFICIENTS OF DETERMINATION R^2 ARE SHOWN.

TBM: $\widehat{\mathrm{AGB}} = K \Delta h^{\alpha} \eta_0^{\beta}$

Nr	Site	K	σ	t	p	α	σ	t	p	β	σ	t	p	R^2
1	Remningstorp	**9.4**	4.87	1.9	6.4e-02	**1.2**	0.18	6.7	2.6e-07	**2.7**	0.37	7.3	4.5e-08	**0.78**
2		**11.6**	6.49	1.8	8.6e-02	**1.1**	0.19	6.0	3.3e-06	**3.0**	0.39	7.6	5.5e-08	**0.81**
3		**27.1**	15.39	1.8	9.1e-02	**0.9**	0.19	4.6	9.6e-05	**3.0**	0.38	7.9	3.0e-08	**0.82**
4		**0.4**	0.37	1.0	3.5e-01	**2.4**	0.39	6.0	2.2e-06	**1.8**	0.34	5.3	1.4e-05	**0.65**
5		**1.7**	1.13	1.5	1.5e-01	**1.8**	0.24	7.3	9.2e-08	**2.0**	0.30	6.5	7.4e-07	**0.76**
6		**7.4**	5.63	1.3	2.0e-01	**1.2**	0.25	4.8	1.3e-04	**2.6**	0.44	5.8	1.8e-05	**0.81**
7		**6.5**	3.71	1.7	9.8e-02	**1.3**	0.19	6.8	2.4e-06	**2.9**	0.36	8.0	2.4e-07	**0.89**
8		**20.1**	15.34	1.3	2.1e-01	**0.9**	0.25	3.8	1.2e-03	**3.2**	0.55	5.8	1.7e-05	**0.81**
9	Krycklan	**15.0**	11.34	1.3	2.0e-01	**1.0**	0.25	3.9	5.7e-04	**1.4**	0.26	5.5	9.4e-06	**0.85**
10		**9.6**	6.76	1.4	1.7e-01	**1.1**	0.23	4.8	5.8e-05	**1.3**	0.28	4.5	1.3e-04	**0.84**
11		**13.7**	9.25	1.5	1.5e-01	**1.0**	0.22	4.6	1.1e-04	**1.4**	0.27	5.2	2.2e-05	**0.86**
12		**8.5**	5.09	1.7	1.1e-01	**1.2**	0.21	5.7	5.8e-06	**1.2**	0.24	5.0	3.3e-05	**0.87**
13		**3.9**	2.37	1.6	1.2e-01	**1.5**	0.21	6.9	2.4e-07	**0.9**	0.34	2.6	1.6e-02	**0.82**
14		**4.6**	2.96	1.5	1.3e-01	**1.4**	0.22	6.4	8.3e-07	**0.9**	0.35	2.6	1.5e-02	**0.82**
15		**6.7**	4.53	1.5	1.5e-01	**1.3**	0.22	5.7	4.9e-06	**1.2**	0.38	3.0	5.4e-03	**0.82**
16		**8.9**	6.47	1.4	1.8e-01	**1.1**	0.24	4.8	5.4e-05	**1.1**	0.35	3.1	4.8e-03	**0.81**
17		**20.5**	16.41	1.2	2.2e-01	**0.9**	0.26	3.4	2.1e-03	**1.4**	0.28	4.9	4.9e-05	**0.84**
18		**8.8**	6.21	1.4	1.7e-01	**1.1**	0.24	4.8	5.5e-05	**1.0**	0.26	4.0	5.0e-04	**0.83**

Color coding for acquisition number by HOA: | 30 m | 40 m | 50 m | 60 m |

TABLE III

Results for the TLM biomass model (TBM) with fixed exponents and the scaling model (SM). For each slope parameter, the estimated standard deviations σ, t-statistics, and p-values are shown. The coefficients of determination R^2 are also shown.

(a) TBM with fixed exponents: $\widehat{\mathrm{AGB}}_{Re} = K'\Delta h^{1.25}\eta_0^{2.64}$, $\widehat{\mathrm{AGB}}_{Kr} = K'\Delta h^{1.25}\eta_0^{1.16}$

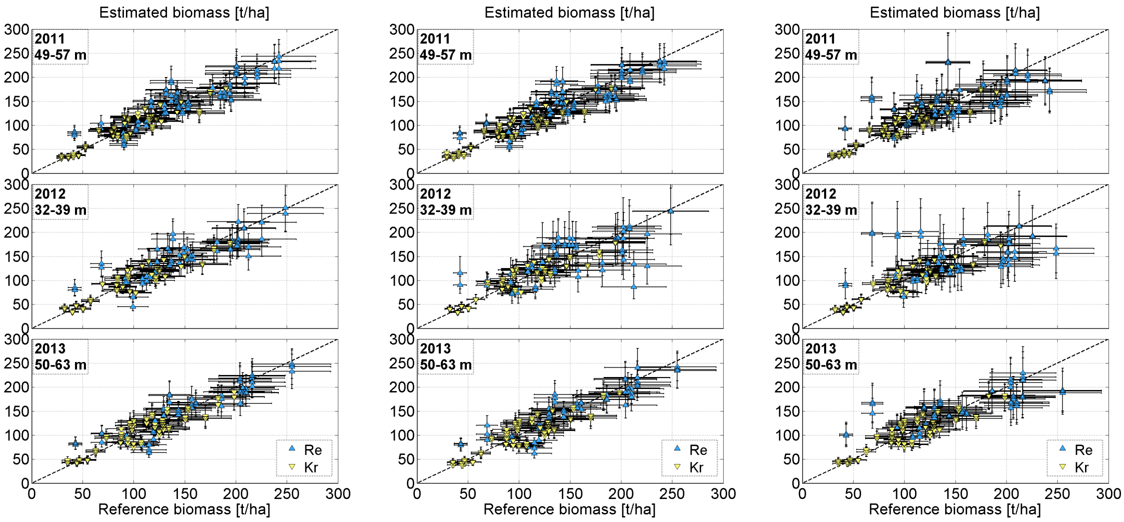
Nr	Site	K'	σ	t	p	R^2
1	Remningstorp	**7.6**	0.20	37.2	2.7e-27	**0.78**
2		**7.5**	0.21	35.5	3.4e-24	**0.80**
3		**8.0**	0.23	34.2	9.1e-24	**0.79**
4		**10.2**	0.55	18.5	3.0e-17	**0.22**
5		**8.8**	0.32	27.8	5.9e-22	**0.64**
6		**7.1**	0.23	30.3	3.4e-18	**0.81**
7		**7.1**	0.18	39.3	2.1e-20	**0.88**
8		**7.2**	0.25	28.5	1.1e-17	**0.78**
9	Krycklan	**6.7**	0.19	36.3	4.4e-25	**0.84**
10		**6.6**	0.18	36.1	5.1e-25	**0.84**
11		**6.8**	0.18	37.4	1.9e-25	**0.85**
12		**6.7**	0.17	39.5	4.1e-26	**0.87**
13		**7.5**	0.22	34.0	2.6e-24	**0.81**
14		**7.4**	0.22	34.4	1.8e-24	**0.81**
15		**7.3**	0.21	35.2	9.8e-25	**0.82**
16		**7.0**	0.20	34.9	1.2e-24	**0.81**
17		**7.2**	0.19	37.7	1.5e-25	**0.83**
18		**6.8**	0.19	35.9	5.7e-25	**0.82**

(b) SM: $\widehat{\mathrm{AGB}} = Dh_{gc}$

Nr	Site	D	σ	t	p	R^2
1	Remningstorp	**9.0**	0.40	22.2	1.4e-20	**0.40**
2		**9.1**	0.45	20.2	7.8e-18	**0.41**
3		**9.4**	0.44	21.2	2.3e-18	**0.46**
4		**8.3**	0.53	15.6	2.4e-15	**-0.07**
5		**8.4**	0.49	17.4	1.5e-16	**0.13**
6		**9.4**	0.48	19.3	2.1e-14	**0.54**
7		**9.6**	0.45	21.6	2.5e-15	**0.63**
8		**9.6**	0.50	19.4	1.9e-14	**0.54**
9	Krycklan	**11.5**	0.31	36.6	3.6e-25	**0.85**
10		**11.4**	0.32	36.0	5.3e-25	**0.84**
11		**11.7**	0.31	37.6	1.6e-25	**0.85**
12		**11.6**	0.29	39.4	4.7e-26	**0.87**
13		**11.3**	0.32	35.4	8.7e-25	**0.82**
14		**11.4**	0.32	35.1	1.1e-24	**0.82**
15		**11.4**	0.32	35.5	8.2e-25	**0.83**
16		**11.8**	0.34	35.2	9.8e-25	**0.81**
17		**12.2**	0.32	38.3	9.8e-26	**0.84**
18		**11.9**	0.32	36.6	3.3e-25	**0.83**

Color coding for acquisition number by HOA:

30 m	40 m	50 m	60 m



(a) TBM: $\widehat{\mathrm{AGB}} = K\Delta h^{\alpha}\eta_0^{\beta}$

(b) TBM with fixed exponents: $\widehat{\mathrm{AGB}}_{Re} = K'\Delta h^{1.25}\eta_0^{2.64}$, $\widehat{\mathrm{AGB}}_{Kr} = K'\Delta h^{1.25}\eta_0^{1.16}$

(c) SM: $\widehat{\mathrm{AGB}} = Dh_{gc}$

Fig. 5. Residual scatter plots are shown. Acquisition year and HOA intervals are shown for each subplot. Horizontal error bars show the 15%-uncertainty of the reference biomass estimates. Vertical error bars represent the residual RMSE values presented on the diagonal of Tables IV–VI. Note: several points may overlap. In Remningstorp, there are 88 points in 2011, 58 points in 2012, and 63 points in 2013. In Krycklan, there are 116 points in 2011 and 87 points in both 2012 and 2013.

TABLE IV
RESIDUAL AND PREDICTION RMSE VALUES (IN PERCENT OF THE AVERAGE BIOMASS, WHICH CAN BE FOUND IN TABLE I) FOR THE TBM. RESIDUAL RMSE VALUES ARE MARKED IN BOLDFACE CHARACTERS AND SHOWN ON THE DIAGONAL. OFF-DIAGONAL, PREDICTION RMSE VALUES ARE SHOWN.

TBM: $\widehat{\mathrm{AGB}} = K\Delta h^{\alpha}\eta_0^{\beta}$ (K, α, and β as in Table II)

Training data		Validation data																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Nr	Site	Remningstorp								Krycklan									
1	Remningstorp	**16**	15	16	39	25	17	14	17	46	43	45	41	48	47	46	43	50	44
2		16	**15**	16	41	26	17	14	16	48	45	48	43	50	49	48	45	52	46
3		18	17	**15**	38	25	23	21	20	41	37	40	35	41	40	39	37	45	39
4		68	72	68	**19**	33	73	77	80	29	29	31	30	39	38	37	33	33	30
5		29	31	29	27	**16**	32	33	36	35	33	36	33	43	42	41	36	39	34
6		17	16	19	41	27	**15**	12	17	48	45	48	44	51	50	49	46	52	46
7		17	17	20	44	30	16	**12**	16	53	50	53	49	55	55	54	51	56	51
8		17	16	19	44	30	17	14	**15**	51	47	50	45	51	51	49	47	55	48
9	Krycklan	32	33	29	24	23	33	34	35	**15**	16	15	15	19	18	17	16	16	16
10		31	32	29	23	21	31	32	33	15	**16**	15	14	20	19	18	17	17	16
11		32	33	30	24	23	34	34	35	15	16	**15**	15	19	18	17	16	16	16
12		32	33	31	23	22	32	33	35	15	16	15	**14**	19	19	18	16	16	15
13		53	55	52	26	35	52	55	56	23	23	21	20	**16**	16	16	18	18	21
14		50	51	49	25	33	49	51	52	22	22	20	19	16	**16**	16	17	17	20
15		46	47	45	24	30	47	49	50	19	20	18	18	16	16	**16**	17	15	18
16		35	36	34	23	24	35	36	37	16	17	16	15	18	17	17	**16**	15	16
17		36	37	33	26	27	38	38	39	17	19	16	18	18	18	17	17	**14**	17
18		31	32	30	23	22	31	32	33	16	16	15	14	19	18	18	16	16	**15**

Main color coding by RMSE value:	20%	40%	60%	80%
Color coding for the acquisition number by HOA:	30 m	40 m	50 m	60 m

TABLE V
RESIDUAL AND PREDICTION RMSE VALUES (IN PERCENT OF THE AVERAGE BIOMASS, WHICH CAN BE FOUND IN TABLE I) FOR THE TBM WITH FIXED EXPONENTS. RESIDUAL RMSE VALUES ARE MARKED IN BOLDFACE CHARACTERS AND SHOWN ON THE DIAGONAL. OFF-DIAGONAL, PREDICTION RMSE VALUES ARE SHOWN.

TBM with fixed exponents: $\widehat{\mathrm{AGB}}_{Re} = K'\Delta h^{1.25}\eta_0^{2.64}$, $\widehat{\mathrm{AGB}}_{Kr} = K'\Delta h^{1.25}\eta_0^{1.16}$ (K' as in Table III(a))

Training data		Validation data																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Nr	Site	Remningstorp								Krycklan									
1	Remningstorp	**16**	15	17	39	24	17	14	17	21	22	19	20	17	16	16	18	16	20
2		16	**15**	17	39	25	17	14	17	20	21	19	19	16	16	16	18	16	19
3		17	16	**16**	36	22	20	18	20	25	27	24	24	18	18	18	21	19	24
4		40	40	33	**29**	26	48	48	47	57	60	56	57	42	43	44	51	47	56
5		23	23	19	32	**20**	29	28	29	36	38	35	36	25	26	26	31	28	35
6		17	16	19	42	28	**15**	12	16	17	17	16	16	17	17	16	16	15	16
7		17	17	20	42	28	15	**12**	17	17	17	16	15	17	17	16	16	15	16
8		16	16	19	42	27	15	12	**16**	17	18	16	16	17	16	16	16	15	17
9	Krycklan	19	19	23	45	31	17	13	18	**15**	16	15	14	19	19	18	17	16	16
10		20	20	24	46	32	17	14	18	16	**16**	15	14	20	20	19	17	17	16
11		19	18	22	45	31	16	13	17	16	16	**15**	14	19	18	18	16	16	16
12		19	19	23	45	31	17	13	18	15	16	15	**14**	19	19	18	17	16	16
13		16	15	17	40	25	16	13	17	19	20	18	18	**16**	16	16	17	15	19
14		16	15	18	40	26	16	13	17	19	20	18	18	16	**16**	16	17	15	18
15		16	16	18	41	26	16	13	17	18	19	17	17	17	16	**16**	17	15	18
16		17	17	20	43	29	16	12	17	16	17	15	15	18	17	17	**16**	15	16
17		16	16	19	42	27	15	12	16	17	18	16	16	17	17	16	16	**15**	17
18		19	19	22	45	31	16	13	18	16	16	15	14	19	19	18	16	16	**16**

Main color coding by RMSE value:	20%	40%	60%	80%
Color coding for the acquisition number by HOA:	30 m	40 m	50 m	60 m

TABLE VI
RESIDUAL AND PREDICTION RMSE VALUES (IN PERCENT OF THE AVERAGE BIOMASS, WHICH CAN BE FOUND IN TABLE I) FOR THE SM. RESIDUAL RMSE VALUES ARE MARKED IN BOLDFACE CHARACTERS AND SHOWN ON THE DIAGONAL. OFF-DIAGONAL, PREDICTION RMSE VALUES ARE SHOWN.

SM: $\widehat{\mathrm{AGB}} = Dh_{gc}$ (D' as in Table III(b))

Training data		Validation data																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Nr	Site	Remningstorp								Krycklan									
1	Remningstorp	**26**	26	26	35	31	24	23	25	28	27	29	28	27	28	28	30	32	30
2		26	**26**	25	35	32	24	22	24	26	26	28	27	26	26	27	29	30	29
3		26	27	**25**	36	33	24	22	24	24	24	25	25	24	24	24	27	28	27
4		27	28	28	**34**	31	27	26	28	33	33	34	34	32	33	33	35	37	35
5		26	27	27	34	**31**	26	25	27	32	31	33	32	31	32	32	34	36	34
6		26	26	25	36	32	**24**	22	24	25	24	26	25	24	25	25	27	29	27
7		27	27	25	37	34	24	**21**	24	23	23	24	23	22	23	23	25	27	25
8		27	27	25	37	34	24	21	**24**	23	22	24	23	22	23	23	25	27	25
9	Krycklan	38	37	34	51	47	33	29	31	**15**	16	15	14	16	16	16	16	16	16
10		37	36	33	50	46	33	29	30	15	**16**	15	14	16	16	16	16	16	16
11		40	39	35	53	49	35	31	32	16	16	**15**	14	16	16	16	16	15	15
12		39	38	35	52	48	34	30	32	15	16	15	**14**	16	16	16	16	15	15
13		37	36	32	49	46	32	28	30	15	16	15	15	**16**	16	16	17	17	16
14		38	37	33	50	47	33	29	30	15	16	15	14	16	**16**	16	16	16	16
15		38	37	33	51	47	33	29	31	15	16	15	14	16	16	**16**	16	16	16
16		42	40	36	54	51	36	32	34	16	16	15	14	17	16	16	**16**	15	15
17		45	43	39	58	54	40	35	36	17	18	16	15	18	18	17	16	**15**	15
18		42	41	37	55	51	37	33	34	16	16	15	15	17	17	16	16	15	**15**

Main color coding by RMSE value:	20%	40%	60%	80%
Color coding for the acquisition number by HOA:	30 m	40 m	50 m	60 m

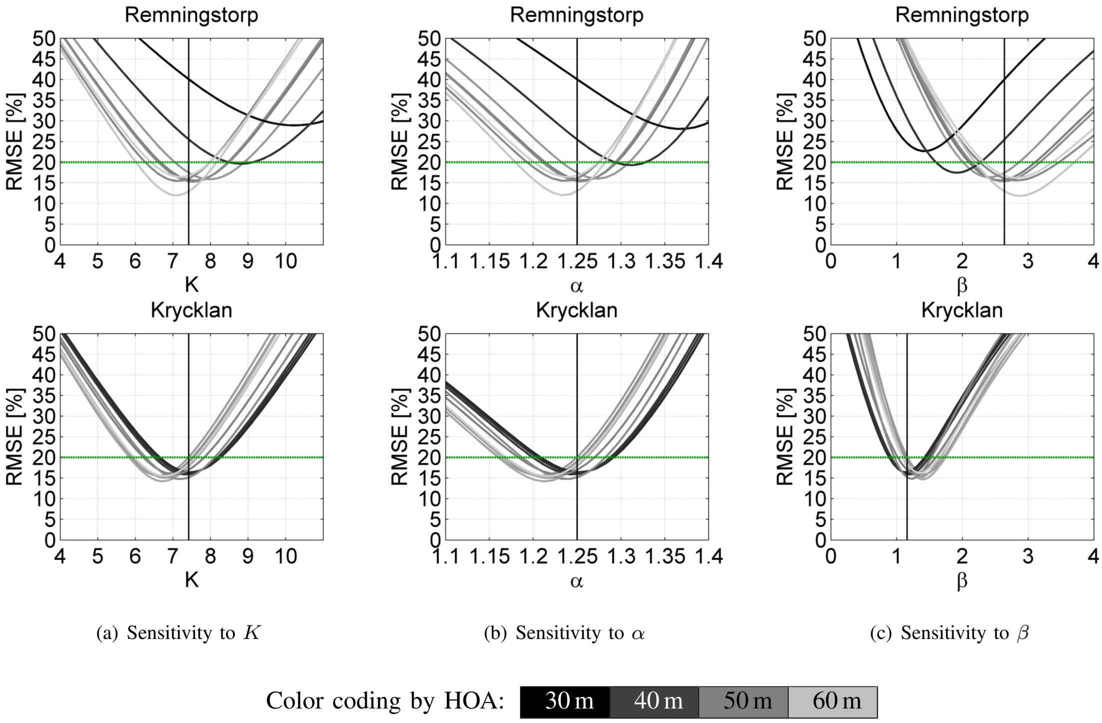
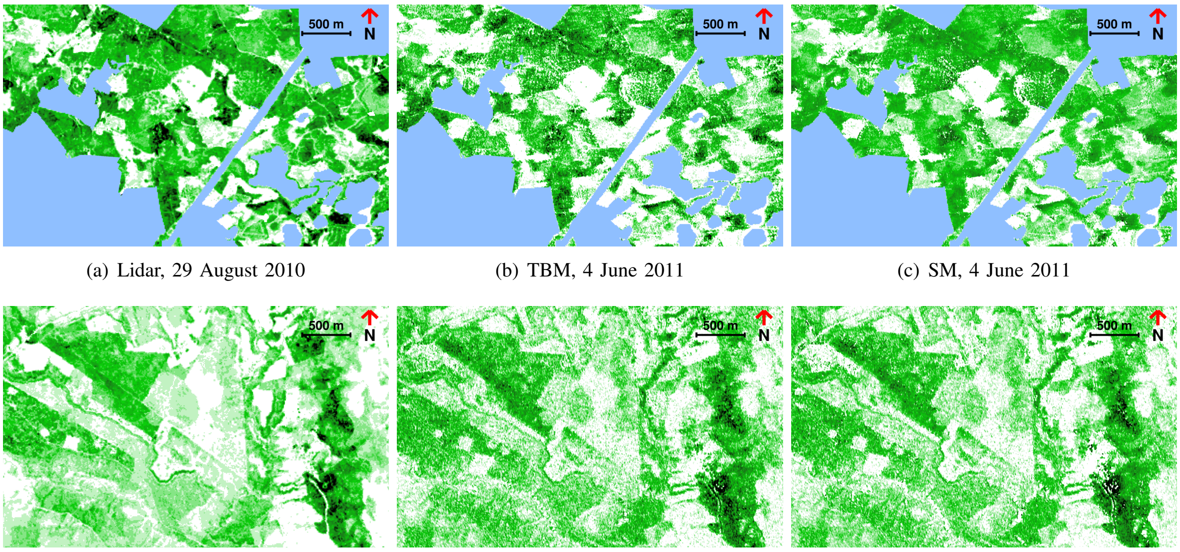


Fig. 6. Sensitivity of the TBM to small changes of parameters K, α, and β around a default setup with $\alpha = 1.25$, $\beta = 2.64$ for Remningstorp, $\beta = 1.16$ for Krycklan, and $K = 7.42$ (being the mean of all K'-values in Table III(a)). The vertical lines show the default values and the green, horizontal lines mark the 20% error level.



(a) Lidar, 29 August 2010 (b) TBM, 4 June 2011 (c) SM, 4 June 2011

(d) Lidar, 5-6 August 2008 (e) TBM, 17 June 2011 (f) SM, 17 June 2011

Color coding by biomass in t/ha:

N/A	0	25	50	75	100	125	150	175	200	225	250	275	300

Fig. 7. Mapping results for (a)–(c) Remningstorp and (d)–(f) Krycklan. The lidar biomass maps are compared to biomass maps obtained using the TLM biomass model (TBM) and the scaling model (SM) with single TanDEM-X image pairs from the corresponding dates. Model parameters estimated using the corresponding plot/stand-level estimates data are used. Note, that growth has not been modeled in the lidar maps.