

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

GENERATION OF FOREST HEIGHT AND BIOMASS MAPS OF A
BOREAL FOREST USING REMOTE SENSING DATA FROM
TANDEM-X, SRTM, LANDSAT AND AIRBORNE LASER
SCANNING

THESIS
PRESENTED
AS A PARTIAL REQUIREMENT
FOR ENVIRONMENTAL SCIENCES PH. D.

BY
YASER SADEGHI

OCTOBER 2017

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.07-2011). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

GÉNÉRATION DE CARTES DE LA HAUTEUR ET DE LA
BIOMASSE D'UNE FORÊT BORÉALE À L'AIDE DE DONNÉES DE
TÉLÉDÉTECTION DE TANDEM-X, SRTM, LANDSAT ET DE
BALAYAGE LASER AÉROPORTÉ

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DU DOCTORAT EN SCIENCES DE L'ENVIRONNEMENT

PAR
YASER SADEGHI

OCTOBRE 2017

ACKNOWLEDGEMENTS

First I would like to express my appreciation for my research director, Prof. Benoît St-Onge, for his availability, understanding and continued support. Working with Benoît and people at the Laboratoire de Cartographie des Dynamiques Forestières (LCDF) at UQAM was a fantastic experience. I would also like to thank my research co-director, Prof. Brigitte Leblon and my thesis' advisors, Dr. Marc Simard (Senior Engineer in Jet Propulsion Laboratory (JPL), NASA) and Dr. Kostas Papathanassiou (Senior Scientist in Microwaves and Radar Institute, German Aerospace Center (DLR)), for their help and great comments during my research. My gratitude also goes to the Natural Sciences and Engineering Research Council of Canada (NSERC) for financial support, the German Aerospace Centre (DLR) for providing us with access to the TanDEM-X data, as well as the Applied Geomatics Research Group (AGRG, Nova Scotia Community College, Middleton, NS), Dr. Christopher Hopkinson (University of Lethbridge, AB), and the Canadian Consortium for lidar Environmental Applications Research (C-CLEAR, Centre of Geographic Sciences, Lawrencetown, NS) for providing the lidar data. I thank the Département des Sciences du Bois et de la forêt (Université Laval) for granting access to Montmorency Forest, and the Quebec Ministère des Forêts, de la Faune et des Parcs (MFFP). I acknowledge that the Jet Propulsion Laboratory, California Institute of Technology (sponsored by the National Aeronautics and Space Administration US Participating Investigator program) has supported Marc Simard, which translated into support for our project. I would also like to express my gratitude towards First Resource Management Group Inc., and Philip E. J. Green, for their financial support in the last stages of this project. I am also grateful to the Centre d'étude de la forêt (CEF) for the financial support they have provided for my conference expenses and would like to thank the people of the Department of Geography and the Institut des Sciences de L Environnement of UQAM. Finally, I want to express my deepest love to my family and friends for all these years.

CO-AUTHORSHIP

This dissertation is a synthesis of the research carried out in the following four manuscripts:

Chapter II Canopy height model (CHM) derived from a TanDEM-X InSAR DSM and an airborne lidar DTM in boreal forest.
 Authors: Yaser Sadeghi, Benoît St-Onge, Brigitte Leblon, Marc Simard. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 9, No. 1, January 2016.

Chapter III Effects of TanDEM-X acquisition parameters on the accuracy of digital surface models of a boreal forest canopy.
 Authors: Yaser Sadeghi, Benoît St-Onge, Brigitte Leblon, Marc Simard. (In press), Canadian Journal of Remote Sensing. It has been accepted on 15 Dec 2016.

Chapter IV Mapping boreal forest biomass from a SRTM and TanDEM-X based canopy height model and Landsat spectral indices.
 Authors: Yaser Sadeghi, Benoît St-Onge, Brigitte Leblon, Jean-Francois Prieur, Marc Simard.
 Submitted to the International Journal of Applied Earth Observation and Geoinformation.

The methods described in this paper are covered in whole or part by a patent pending, US Patent Application No. 15/037,894 (USPTO Publication No. 20160292626). Co-inventors are Benoit St-Onge and Philip E. J. Green. The assignee is First Resource Management Group Inc.

Appendix Mapping forest canopy height using TanDEM-X DSM and airborne LiDAR DTM.
 Authors: Yaser Sadeghi, Benoît St-Onge, Brigitte Leblon, Marc Simard, Kostas Papathanassiou.
 Published as a conference paper in the Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International, INSPEC Accession Number: 14715356.

Some of the results of this project have been presented in their early version in conferences such as:

IGARSS (2012, Munich)
IGARSS (2014, Quebec)

ASAR (2013, Montreal)
ASAR (2015, Montreal)

PollnSAR and BIOMASS (2015, Frascati)

Table of Contents

| | |
|--|-------|
| LIST OF FIGURES..... | x |
| LIST OF TABLES..... | xii |
| RÉSUMÉ | xvi |
| ABSTRACT | xviii |
| CHAPTER I | |
| GENERAL INTRODUCTION..... | 1 |
| 1.1 Forest biomass mapping importance and requirements..... | 1 |
| 1.1.1 Optimal spatial resolution of biomass maps | 2 |
| 1.1.2 Optimal temporal resolution of biomass maps..... | 3 |
| 1.1.3 Optimal accuracy of biomass maps..... | 4 |
| 1.2 Forest biomass retrieval using spaceborne remote sensing..... | 6 |
| 1.2.1 Reflectance-biomass models..... | 7 |
| 1.2.2 Forest height-biomass models..... | 9 |
| 1.3 The need for combining data from different spaceborne sensors | 15 |
| 1.4 Optimal approach for producing a global biomass map | 16 |
| 1.5 General objective of the research project..... | 20 |
| 1.6 Specific objectives and thesis organization..... | 20 |
| CHAPTER II | |
| CANOPY HEIGHT MODEL (CHM) DERIVED FROM A TANDEM-X INSAR DSM AND AN AIRBORNE LIDAR DTM IN BOREAL FORES | 22 |
| 2.1 Résumé..... | 22 |
| 2.2 Abstract | 23 |
| 2.3 Introduction | 24 |

| | | |
|-------|--|----|
| 2.4 | Study area and data sources | 30 |
| 2.4.1 | Study area..... | 30 |
| 2.4.2 | Field and forest inventory data..... | 31 |
| 2.4.3 | Remote sensing data..... | 32 |
| 2.5 | Methods..... | 33 |
| 2.5.1 | Creating the TanDEM-X DSM | 33 |
| 2.5.2 | Creating the Lidar DSM..... | 34 |
| 2.5.3 | Creating the DTM and CHMs..... | 35 |
| 2.5.4 | Comparing InSAR and Lidar CHMs..... | 36 |
| 2.5.5 | Relationships between local incidence angle, TanDEM-X..... | 39 |
| 2.5.6 | Effects of forest structure on coherence and dominant ΔH | 39 |
| 2.6 | Results..... | 40 |
| 2.7 | Discussion | 58 |
| 2.8 | Conclusions..... | 61 |

CHAPTER III

| | | |
|-------|--|----|
| | EFFECTS OF TANDEM-X ACQUISITION PARAMETERS ON THE ACCURACY OF DIGITAL SURFACE MODELS OF A BOREAL FOREST CANOPY | 65 |
| 3.1 | Résumé..... | 65 |
| 3.2 | Abstract | 66 |
| 3.3 | Introduction | 67 |
| 3.4 | Materials..... | 71 |
| 3.4.1 | Study area and field data | 71 |
| 3.4.2 | TanDEM-X images | 73 |
| 3.4.3 | Lidar data | 74 |

| | | |
|-------|--|----|
| 3.5 | Methods..... | 75 |
| 3.5.1 | Creating the DTM, DSMs, and CHMs | 75 |
| 3.5.2 | Accuracy assessment of the InSAR CHMs..... | 76 |
| 3.5.3 | Consistency experiment | 76 |
| 3.5.4 | Baseline variation experiment..... | 77 |
| 3.5.5 | Incidence angle variation experiment | 77 |
| 3.5.6 | Phenology variation experiment | 77 |
| 3.5.7 | Analysis of the impact of acquisition conditions on canopy height predictions..... | 80 |
| 3.6 | Results..... | 81 |
| 3.6.1 | Coherence maps and CHMs..... | 81 |
| 3.6.2 | General prediction models | 87 |
| 3.7 | Discussion | 89 |
| 3.8 | Conclusions..... | 93 |

CHAPTER IV

| | | |
|-------|---|-----|
| | MAPPING BOREAL FOREST BIOMASS FROM A SRTM AND TANDEM-X BASED CANOPY HEIGHT MODEL AND LANDSAT SPECTRAL INDICES.... | 99 |
| 4.1 | Résumé..... | 99 |
| 4.2 | Abstract | 96 |
| 4.3 | Introduction | 96 |
| 4.4 | Study area and data | 101 |
| 4.4.1 | Study area..... | 101 |
| 4.4.2 | Reference data..... | 102 |
| 4.4.3 | Satellite data used to generate predictive variables..... | 103 |

| | | |
|------------|---|-----|
| 4.5 | Methods..... | 105 |
| 4.5.1 | Calculation of field plot biomass | 105 |
| 4.5.2 | Creation of lidar biomass maps..... | 106 |
| 4.5.3 | SRTM DEM correction..... | 106 |
| 4.5.4 | Creation and validation of the SAR CHM | 110 |
| 4.5.5 | Biomass prediction from remote sensing variables | 111 |
| 4.6 | Results..... | 111 |
| 4.7 | Discussion and conclusion | 120 |
| CHAPTER V | | |
| | SYNTHESIS, CONCLUSIONS AND FUTURE RESEARCH | 124 |
| 5.1 | Scientific contribution | 125 |
| 5.2 | Generalization and limitations | 131 |
| 5.3 | Future research directions | 133 |
| 5.4 | Concluding remarks | 136 |
| APPENDIX A | | |
| | MAPPING FOREST CANOPY HEIGHT USING TANDEM-X DSM AND AIRBORNE LIDAR DTM | 138 |
| A.1 | Abstract..... | 138 |
| A.2 | Introduction | 139 |
| A.3 | Study Site and Data | 139 |
| A.4 | Methodology..... | 140 |
| A.5 | Results | 142 |
| A.6 | Conclusions..... | 145 |
| | REFERENCES..... | 146 |

LIST OF FIGURES

| | |
|--|-----|
| Figure 1.1 Interferometric SAR acquisition geometry..... | 12 |
| Figure 2.1 location of the study area..... | 31 |
| Figure 2.2 Representation of the two definitions of canopy height..... | 35 |
| Figure 2.3 lidar CHM, InSAR CHM, coherence map, and map of the differences between the CHMs..... | 42 |
| Figure 2.4 The sample profiles of the InSAR DSM, the InSAR CHM..... | 43 |
| Figure 2.5 lidar vs. field-measured individual tree heights..... | 45 |
| Figure 2.6 Relationships between InSAR and lidar canopy surface height (CSH) ... | 47 |
| Figure 2.7 Relationships between InSAR DH and lidar DH..... | 50 |
| Figure 2.8 Relationships between InSAR CSH and lidar DH..... | 52 |
| Figure 2.9 Relationship between a) coherence and local incidence angle, b) Δh and coherence, and c) Δh and local incidence angle..... | 54 |
| Figure 2.10 Relationships of stem density (a & b), dominant height (DH) (c & d), canopy surface height (CSH) (e & f), and gap volume (g & h) with coherence, and dominant ΔH | 57 |
| Figure 3.1 Location of the Montmorency Research Forest..... | 73 |
| Figure 3.2 Coherence maps..... | 80 |
| Figure 3.3 InSAR CHM maps..... | 81 |
| Figure 3.4 Relationships between InSAR CHMs and the lidar CHM..... | 83 |
| Figure 3.5 Relationships between the InSAR CHMs..... | 86 |
| Figure 4.1 Location of the study site..... | 102 |
| Figure 4.2 SRTM correction process for landform..... | 109 |
| Figure 4.3 Lidar-predicted vs. Observed biomass for 191 field plots..... | 114 |
| Figure 4.4 The height maps for lidar and SAR (20 m resolution). | 117 |
| Figure 4.5 Scatterplots of the lidar CHM and the SAR..... | 119 |
| Figure 4.6 The relationship between lidar-predicted and SAR-Landsat biomass..... | 119 |
| Figure 4.7 The random forest regression..... | 119 |

| | |
|--|-----|
| Figure A.1 Lidar CHM..... | 141 |
| Figure A.2 TanDEM-X-lidar CHM (TanDEM-X DSM minus lidar DTM)..... | 142 |
| Figure A.3 TanDEM-X-lidar CHM vs. Lidar CHM..... | 143 |
| Figure A.4 Relationship between TanDEM-X InSAR coherence and LIA..... | 143 |
| Figure A.5 Relationship between TanDEM-X InSAR coherence and lidar CHM... | 144 |
| Figure A.6 Relationship between InSAR coherence and forest basal area..... | 145 |

LIST OF TABLES

| | |
|--|-----|
| Table 2.1 DBH, height and stem density of field plots..... | 32 |
| Table 2.2 Relationships between lidar and field-measured individual tree heights.... | 44 |
| Table 2.3 InSAR CSH relationship with the lidar CSH..... | 47 |
| Table 2.4 InSAR DH relationship with the lidar DH..... | 52 |
| Table 2.5 InSAR CSH and lidar CSH respective relationships to DH | 53 |
| Table 2.6 Relationships between LIA, coherence, and ΔH | 55 |
| Table 2.7 Relationship between stem density, DH, CSH, gap volume, coherence, .. | 58 |
| Table 3.1 The characteristics of tandem-x acquisitions | 74 |
| Table 3.2 Statistics of the regression between InSAR CHM and lidar CHM..... | 84 |
| Table 3.3 Statistics of the comparison experiments..... | 87 |
| Table 3.4 General model to predict lidar CHM | 88 |
| Table 3.5 Statistics of the comparison between the predicted and reference CHM using general models (*)...... | 89 |
| Table 4.1 General statistics of the main structural attributes of the 200 field plots of the Montmorency Forest | 103 |
| Table 4.2 Summary characteristics of the remote sensing data | 105 |
| Table 4.3 Plot level biomass statistics | 112 |
| Table 4.4 The statistical summary of the SRTM DEMS | 114 |
| Table 4.5 Statistical between the lidar CHM and the SAR CHM | 117 |
| Table 4.6 Statistical relationships between SAR-Landsat and lidar predictions of biomass using linear regression and random forest (RF)..... | 118 |

ACRONYMS

| | |
|-----------------------|---|
| AGB | Above Ground Biomass |
| AGL | Above Ground Level |
| ATLAS | Advanced Topographic Laser Altimeter System |
| ALS | Airborne Laser Scanning |
| AGRG | Applied Geomatics Research Group |
| CO₂ | Carbon Dioxide |
| C-CLEAR | Canadian Consortium for lidar Environmental Applications Research |
| CHM | Canopy Height Model |
| CSH | Canopy Surface Height |
| EVI | Enhanced Vegetation Index |
| DBH | Diameter at Breast Height |
| DVI | Difference Vegetation Index |
| DSM | Digital Surface Model |
| DTM | Digital Terrain Model |
| DH | Dominant Height |
| GLAS | Geoscience Laser Altimeter System |
| DLR | German Aerospace Centre |
| GEDI | Global Ecosystem Dynamics Investigation |
| GNDVI | Green Normalized Difference Vegetation Index |

| | |
|------------------|---|
| GRVI | Green-Red Vegetation Index |
| GVI | Green Vegetation Index |
| HoA | Height of Ambiguity |
| ICESat | Ice, Cloud, and land Elevation Satellite |
| InSAR | Interferometric Synthetic Aperture Radar |
| IDW | Inverse Distance Weighting |
| JPL | Jet Propulsion Laboratory |
| LVIS | Laser Vegetation Imaging Sensor |
| LAI | Leaf Area Index |
| Lidar | Light Detection and Ranging |
| LIA | Local Incidence Angle |
| MIMICS | Michigan Microwave Canopy Scattering Model |
| MODIS | Moderate Resolution Imaging Spectroradiometer |
| NGA | National Geospatial-Intelligence Agency |
| NSERC | Natural Sciences and Engineering Research Council of Canada |
| NDSI | Normalized Difference Snow Index |
| NDVI | Normalized Difference Vegetation Index |
| Pol-InSAR | Polarimetric Interferometry SAR |
| MFFP | Québec Ministère des Forêts, de la Faune et des Parcs |
| RF | Random Forest |
| RVoG | Random Volume over Ground |

| | |
|-----------------|--|
| RMSE | Root-Mean-Square Error |
| SPC | Scattering Phase Centre |
| SRTM | Shuttle Radar Topographic Mission |
| SNR | Signal to Noise Ratio |
| SSC | Single-look Slant range Complex |
| SAVI | Soil Adjusted Vegetation Index |
| SAR | Synthetic Aperture Radar |
| TanDEM-X | TerraSAR-X add-on for Digital Elevation Measurements |
| 3D | Three-dimensional |
| TIN | Triangulated Irregular Network |

RÉSUMÉ

Malgré des efforts de recherche considérables consacrés à estimer la biomasse forestière sur de vastes zones, une incertitude significative demeure dans les quantités de biomasse et de carbone piégé dans les forêts de la Terre. Dans cette thèse, nous présentons premièrement des arguments pour préconiser une approche de prédiction de la biomasse à partir de la hauteur de la forêt, plutôt qu'à partir de la réflectance, et nous proposons une logique qui mène à considérer que l'interférométrie radar (InSAR) spatiale réalisée en une passe est l'une des meilleures approches de cartographie de la hauteur de la forêt à l'échelle planétaire. Pour cette raison, nous avons utilisé les données des deux seules missions InSAR planétaires à une passe, TanDEM-X et SRTM (Shuttle Radar Topographic Mission). TanDEM-X, la première et la seule mission interférométrique globale en bande X, fut amorcée en 2010 et a généré un modèle numérique d'altitude (MNA) planétaire à haute résolution. À travers une comparaison détaillée à des données concomitantes de balayage laser aéroporté (BLA), nous avons d'abord démontré que ce MNA est en fait un modèle numérique de surface (MNS). En soustrayant de ce MNS un modèle numérique de terrain (MNT) obtenu par BLA, nous avons pu générer un modèle de hauteur de couvert (MHC). La résolution et l'exactitude de ce MHC InSAR-BLA ont été évaluées à des résolutions allant de 5 m à 25 m, et à l'échelle du peuplement forestier. Il fut constaté que ce type de MHC a une résolution inhérente plus grossière que celle d'un MHC correspondant créé par BLA. Son exactitude variait de 2.7 m (EMQ) à 5 m de résolution, jusqu'à 1.5 m à l'échelle du peuplement. Étant donné que les MNS de TanDEM-X sont générés à travers le monde selon des configurations et en des saisons différentes, nous suspectons que cette variabilité dans les conditions d'acquisition pourrait affecter l'exactitude des MNS de TanDEM-X. Nous avons testé cette hypothèse en évaluant l'exactitude de cinq jeux de données TanDEM-X acquis selon des différentes conditions géométrique et phénologique pour une forêt boréale en majeure partie sempervirente. Les résultats montrent des biais allant de 0.77 m à 1.56 m comparativement à des données de BLA, des r^2 allant de 0.68 à 0.38, et des EMQ allant de 2.06 m à 3.67 m. Parmi ces cinq jeux de données TanDEM-X, deux qui furent acquis dans des conditions quasi identiques différaient par 1.27 m (EMQ), alors que l'effet le plus prononcé provenait d'une large différence de ligne de base interférométrique, menant à une EMQ de 3.27 m entre les DSM générés respectivement avec une courte et longue ligne de base interférométrique. L'effet des changements phénologiques sur les estimations de la hauteur forestière était plus faible que ceux résultant des différences de lignes de base, avec une EMQ de 2.30 m entre les jeux de données acquis sans, et avec feuilles (dans le cas des arbres décidus). Ces résultats indiquent que, malgré des variations dans les conditions d'acquisition, une mosaïque TanDEM-X continue acquise avec des lignes de base appropriées pourrait servir à produire des estimations fiables et suffisamment homogènes des altitudes des surfaces des couverts forestiers dans le cas des forêts boréales fermées sempervirentes. Finalement, nous désirions proposer une solution de télédétection pour créer des MHC qui ne dépendrait pas de données de BLA, mais qui

serait plutôt entièrement fondée sur des données acquises par des capteurs satellitaires. Pour cela, nous avons utilisé une méthode de correction des MNA SRTM (acquis en bande C) permettant d'en faire des quasi-MNT. Un MHC InSAR a ensuite été produit par soustraction de ce MNT du MNS TanDEM-X, ce qui a résulté en une EMQ de 2.45 m, un r^2 de 0.43 et un biais de 0.07 m, lorsque comparé aux hauteurs obtenues par BLA à l'échelle du peuplement. Ensuite, un modèle de prédiction de la biomasse basé sur ce MHC ainsi que sur des indices de végétation fut développé. La biomasse forestière a pu ainsi être cartographiée complètement depuis l'espace avec une EMQ de 26 Mg ha⁻¹, et un r^2 de 0.62, comparativement à une carte très exacte de la biomasse à l'échelle du peuplement réalisée par BLA. Dans un avenir rapproché, les données SRTM pourraient être remplacées par celles d'une mission en bande L, TanDEM-L, ce qui mènerait à des MNT et MHC améliorés, et ainsi, à de meilleures cartes de biomasse forestière.

Mots-clés: TanDEM-X, modèle de hauteur de couvert (MHC), biomasse, radar interférométrique (InSAR), balayage lidar aéroporté (BLA).

ABSTRACT

Despite considerable research efforts devoted to estimating forest biomass over large areas, significant uncertainty remains in the quantities of biomass and carbon stored in the Earth's forests. In this thesis, we first present arguments for favouring a biomass-from-height approach to reflectance-based approaches, and propose a rationale for considering spaceborne single-pass interferometric synthetic aperture radar (InSAR) as being one of the best forest height global mapping approach. For this reason, we have used data from the only two global single-pass InSAR missions, TanDEM-X and SRTM (Shuttle Radar Topographic Mission). TanDEM-X, the first and only global X-band spaceborne single-pass interferometer mission, was launched in 2010 and generated a global digital elevation model (DEM) at high-resolution. Through a detailed comparison with concomitant airborne laser scanning (ALS) data, we first demonstrated that this DEM is, in fact, a digital surface model (DSM). By subtracting an ALS DTM (digital terrain model) from this surface model, we were able to generate a canopy height model (CHM). The inherent resolution and accuracy of such InSAR-ALS CHMs were assessed at spatial resolutions ranging from 5 m to 25 m, and at forest stand level. It was found that this type of CHM has a coarser inherent resolution compared to a corresponding ALS CHM. Its accuracy varied from 2.7 m (RMSE) at a 5 m resolution, to 1.5 m at stand level. Because TanDEM-X DSMs are generated worldwide using various sensor configurations, and at different seasons, it was suspected that these variable acquisition conditions may affect the accuracy of TanDEM-X DSMs. We tested this hypothesis by assessing the accuracy of five TanDEM-X datasets acquired under various geometrical and phenological conditions over a mostly evergreen boreal forest. The results show biases from 0.77 m to 1.56 m compared to ALS data, r^2 s from 0.68 to 0.38, and RMSEs from 2.06 m to 3.67 m. Among these five TanDEM-X datasets, two that were acquired in nearly identical conditions differed by 1.27 m (RMSE) while the strongest effect came from a big difference in the interferometric baseline, leading to a RMSE of 3.27 m between DSMs generated respectively with short and long baselines. The effect of phenological changes on forest height estimations was found weaker than baseline effects, with a RMSE of 2.30 m between leaf-on and leaf-off datasets (in the case of deciduous trees). These results indicate that, despite variations in the acquisition conditions, a continuous TanDEM-X mosaic acquired with proper baselines could produce a reliable and sufficiently homogeneous estimate of canopy surface elevations of evergreen closed-canopy boreal forests. Finally, we wanted to propose a remote sensing solution for creating CHMs that would not rely on ALS data, but rather on data acquired from satellite sensors. For this, we have used a method for correcting the SRTM DEMs (acquired in C-band) to create quasi-DTMs. An InSAR CHM was then produced by subtracting this DTM from a TanDEM-X DSM, resulting in a RMSE of 2.45 m, a r^2 of 0.43, and a bias of 0.07 m, when compared to ALS heights at stand level. Then, a biomass prediction model based on this CHM and Landsat vegetation indices was developed. Forest biomass could thus be mapped entirely from space with a RMSE of

26 Mg ha⁻¹, and a r^2 of 0.62, compared to a highly accurate ALS biomass map at stand level. In the near future, the SRTM data could be replaced by that of a planned InSAR satellite mission in L-band, TanDEM-L, leading to better DTMs and CHMs, and hence, improved forest biomass maps.

Key Words: TanDEM-X, canopy height model (CHM), biomass, interferometric synthetic aperture radar (InSAR), airborne laser scanning (ALS).

CHAPTER I

GENERAL INTRODUCTION

1.1 Forest biomass mapping importance and requirements

While the role of carbon dioxide (CO₂) in global warming is now ascertained, the fluxes of carbon between the biosphere and the other components of the overall terrestrial system are still being actively studied. Notably, the amount and spatial distribution of carbon stocks stored in forests at any time remain uncertain (Bustamante et al., 2016; Rodríguez-Veiga et al., 2016), and are constantly changing due to growth and disturbances. For example, great amounts of carbon are released into the atmosphere from forested areas as a result of intensive human activities and land cover changes (Gibbs et al., 2007). A number of methods have been developed to monitor forest carbon stocks changes on large scales (Hansen et al., 2003; Lefsky 2010; Baccini et al., 2012; Saatchi et al., 2011.b; Simard et al., 2011; Thurner et al., 2014), based on the fact that carbon represents half of the biomass of a tree (IPCC 2003). These models rely on spatially explicit data on the dynamics of the above ground biomass (AGB) of the forest, expressed as tons of forest biomass per hectare (Mg ha⁻¹). Therefore, the reduction of the uncertainty of global carbon stocks and fluxes could be achieved through a better assessment of the forest biomass (IPCC 2003).

A general estimation (Houghton et al., 2009), used here as an example, indicates that forest biomass averages 390 Mg ha⁻¹ in tropical forests, 270 Mg ha⁻¹ in temperate forests, and 83 Mg ha⁻¹ in boreal forests. However, large-scale biomass maps of tropical forests, generated with cell sizes ranging from 500 m to 1000 m by different authors, diverge rather strongly with regards to biomass levels with a standard deviation of the error between 11 to 108 Mg ha⁻¹ (Avitabile et al., 2016). The cause of this divergence

stems from differences in the methods used, whether at the level of field measurement interpolation or scaling up methods, modeling of environmental parameters used as predictors, or remote sensing techniques (Ometto et al., 2014; Houghton et al., 2001). Therefore, improving the accuracy of forest biomass estimation methods is essential. Furthermore, forests cover very wide areas of the Earth's surface and are rapidly changing, thus using satellite-based remote sensing with global coverage capability and short revisit time is preferable to field or airborne surveys. Fortunately, the number of satellites with the appropriate characteristics for monitoring forests at large scales, such as wavelength, spatial resolution, revisit time, and in some cases, interferometric capability, is increasing. However, this brings the question of selecting the best remote sensing data regarding optimal levels of spatial resolution, temporal resolution, and accuracy for producing timely and accurate forest biomass maps. The “optimal” term in this section is directly related to the possibility of validating these maps at specific resolutions. Availability of precise reference data is the baseline for determining the optimal resolution for the biomass maps.

1.1.1 Optimal spatial resolution of biomass maps

The spatial scales at which forest biomass changes varies between the fine grain level, such as the fall or regrowth of single trees to much coarser levels, such as those at which large fire events occur (Houghton et al., 2012; Houghton 2010). Dynamics such as deforestation, degradation, damage from disease or afforestation occur at various scales. Each of these biomass disturbance factors, separately or in combination with others, can modify biomass from very fine to broad scales (Vanderwel et al., 2013, a&b). Therefore, defining the optimal scale for capturing biomass changes is not immediately evident. In one instance, a cell size of 100 m for mapping biomass was suggested, for example, as the optimal spatial resolution because it roughly corresponds to the size of forest stands, which is in an average of one hectare (Houghton et al., 2012). It was also suggested that: "Global monitoring of forest aboveground biomass

(AGB) with sub-hectare ($1 \text{ ha} = 10^4 \text{ m}^2$) resolution will facilitate the understanding of carbon storage and its flux between forests and the atmosphere (Treuhaft et al., 2015)."

Furthermore, the size of field plots used for the calibration of the statistical models employed in biomass mapping must have a certain size relation to the pixel size of the remote sensing images used. Large differences in the respective sizes of field plots and the corresponding remote sensing pixels increase calibration and validation uncertainties. This problem is complicated by the fact that due to the intensive labor and important costs of establishing each field plot, the need for having a large number of well-distributed plots, the size of these needs to be kept small. This, in turn, puts a constraint on the cell size of the remote sensing images used for mapping biomass, especially as a large number of permanent small field plots (a few hundreds of m^2 per plot) are already established worldwide and regularly remeasured. Therefore, a rather high spatial resolution (small cell size) should be preferred (Hall et al., 2011; Le Toan et al., 2011; Hurtt et al., 2010; Bergen et al., 2009; Houghton et al., 2009). It follows from this argument that a sensor such as e.g. MODIS, with a pixel size of 250 m ($62\,500 \text{ m}^2$), cannot be said to have an optimal resolution. However, these MODIS maps, in the case of Canada with its very large areas of forest, could be said to be economically optimal for biomass mapping.

1.1.2 Optimal temporal resolution of biomass maps

Natural disturbances, such as fire, windthrow, insect disease/damage, or human activities causing, or inducing, deforestation as well as afforestation, affect between 0.2% to 3% of world forests annually, i.e. occur at a rather high temporal rate (Hall et al., 2011; Houghton 2005). On the other hand, forest growth, especially in boreal zones, is rather slow, lessening the need for high temporal resolution (short revisit time).

However, inter-seasonal biomass maps would be required to reduce the variability related to forest phenology (leaf-on vs. leaf-off), which can potentially affect the accuracy of forest height or biomass prediction (White et al., 2015; Anderson and Bolstad 2013). Due to the close relationship between image features or some of their derivatives, e.g. the normalized difference vegetation index (NDVI) and vegetation phenology (White et al., 2003), image acquisition is often limited to an optimal period. For example, in the northern mixed forests, deciduous trees may bear full, non-senescent leaves during only two to three months per year, while in the tropics; the leaf-on time is between six to nine months (Hall et al., 2011). If an optical sensor (instead of a radar sensor) is being considered, frequent cloudiness brings an additional and severe constraint regarding revisit rate. Moreover, each scene must be sufficiently lit by the sun. In the case of Landsat 8 for example, day scenes are acquired on descending orbits while night scenes unusable for biomass studies are taken on ascending orbits. This limits the revisit rate compared to radar sensors which can create scenes of the same area both on ascending and descending orbits. The temporal resolution should, therefore, be high enough in the case of optical imagery to ensure that every location of an area of interest can be seen at least once in the day time, without clouds, per leaf-on season, which translates into a quasi-daily revisit rate (Asner 2009; Olander et al., 2008).

1.1.3 Optimal accuracy of biomass maps

In a very general way, a biomass prediction model can be expressed as:

$$\hat{B} = f(P_i) + \varepsilon \quad [1.1]$$

Where \hat{B} is the biomass predicted using predictors P_i , and ε is the random error (unexplained variance). For any resolution cell of a biomass map, the error is the difference between the "true" biomass and \hat{B} .

The first desirable quality of a prediction model is that it should not be biased. A significant bias would affect the biomass totals over regions, or biomes, and has, therefore, a great impact on our capacity to assess the real amounts of stored C. In principle, if the model is properly calibrated using a representative sample, and well adapted to the distribution of values (linear, non-linear, etc.), it should not produce biased estimates. However, since the a priori distribution of biomass values, in their overall frequencies, or in space, is not known, designing a proper sampling plan over a large area is difficult. Furthermore, access to calibration plots is often arduous in remote areas, leading to a situation where plots are established near access roads (Maltamo et al., 2011), which are themselves not randomly distributed in space. In addition, sampling teams may avoid difficult terrains, such as windthrow areas or swamps, and introduce a certain bias (Fisher et al., 2008, Dalponte et al., 2011).

The model itself may also be inefficient at producing unbiased prediction over the full range of biomass values. It is well known (see next section) that models that predict biomass from image reflectance tend to saturate (Lu et al., 2016) at relatively low biomass levels (100-300 Mg ha⁻¹). In this case, the predictions over high biomass areas systematically underestimate the true value.

Random errors, often expressed as root mean square error (RMSE), or %RMSE, have less severe consequences in the case of an unbiased model because if these errors are normally distributed around 0, they will cancel out as resolution cells become larger. In general, the smaller the resolution cell is, the noisier the input signal will be (from, say, remote sensing images), leading to greater RMSE for smaller pixels. This brings us back to the optimal spatial resolution, which can be set such that the highest spatial resolution is preserved, while the RMSE values remain at an acceptable level.

This poses the question of the desired levels of random errors in spatially explicit biomass predictions. First, it should be remembered that field data used to calibrate biomass prediction models are themselves quite uncertain. Indeed, in most cases, the

biomass of trees falling within a plot is obtained by measuring trees in the field, and then using allometric equations to predict their biomass (e.g. Lambert et al., 2005). Many models predict single tree biomass using diameter at breast height (DBH) and height, within species-specific models (Baldasso et al., 2012; Chave et al., 2005; Henry et al., 2013). Often, the field height measurements are not carried out in full, and are uncertain for tall measured trees, leading to rather uncertain predictions of "field" values that are thereafter used as "reference" for calibration and verification of remote sensing models. Houghton et al. (2009) report that field-based biomass estimation may have an error of approximately 20% when using allometric equations. We have not found clear accuracy targets in the scientific literature concerning forest biomass estimates. In other carbon modelling domains, researchers report for example an uncertainty of 18% for carbon net uptake by oceans (Canadell et al., 2007). So setting a maximum, or optimal RMSE remains an open question at global scale for biomass estimation.

In summary, the optimal biomass prediction model should be designed such that it requires a small number of calibration plots, be unbiased, produce undependable predictions over the full range of biomass, and have an acceptable level of random prediction error at the set resolution. Considering the current state of the art, a 20% RMSE error should not in our opinion be deemed excessive.

1.2 Forest biomass retrieval using spaceborne remote sensing

Satellite remote sensing can be used to estimate forest biomass and monitor its changes on a global scale. Many studies have proposed biomass mapping methods using data from Earth observation satellites, but there is yet no agreement on the best approach. In general, two general families of methods can be identified: 1) reflectance-biomass models, based on a direct relationship between biomass and image reflectance or

backscatter, and 2) forest height-biomass models, where the forest height is used as the main predictor to estimate biomass.

1.2.1 Reflectance-biomass models

Reflectance is the proportion of incident energy that is returned by the forest elements back to the sensor. In the case of optical sensors, the energy is said to be "reflected," while for radar sensors, it is "backscattered." The proportion of returned energy is determined by geometric or physical parameters such as vegetation structure, pigments, water contents or its dielectric properties in the case of radar remote sensing. Canopy structure (element size and shape, the orientation of leaves or woody stems, surface roughness) will affect the returned energy in both the optical and radar domains. Leaf pigment, namely chlorophyll, and water contents, will affect reflectance in the optical domain. The volumetric moisture content of canopies will determine their dielectric properties and thus the amount of backscatter microwave energy. Reflectance or backscatter varies depending on wavelength. The short wavelengths of optical sensors (sun's energy in the visible to short wave infrared regions) react to a canopy's small components such as leaves and twigs. They cannot penetrate deeply into the canopy and only return a signal from the surface. They usually carry a signal on the percent cover of canopies, not on their height or biomass. Percent cover is the proportion of ground covered by vegetation. For this reason, biomass prediction models based on reflected energy measured using optical sensors tend to saturate at very low levels of biomass (Lu 2006).

Landsat vegetation indices were used for example to estimate forest biomass in the boreal and mixed-conifer forest with a RMSE of 49% (Frazier et al., 2014) and 27% (Pflugmacher et al., 2014), respectively. Optical ASTER spectral information was also used to predict the boreal forest biomass with a RMSE of 41% to 44.7% (Muukkonen

and Heiskanen 2005). By adding textural indices, Nichol and Sarker (2011) obtained a RMSE of 32 Mg ha⁻¹ using ALOS-AVNIR-2 for estimating forest biomass.

Unlike optical sensors, active radar sensors are not being affected by weather or cloud cover and hence are a good alternative to optical sensors for mapping forest biomass. In this case, due to the physical relationship between backscatter and volumetric density of canopy elements, the backscatter amplitude information obtained using various polarizations (H and V) or different radar bands are statistically linked to forest biomass. Radar sensors with shorter wavelengths, such as those of the X-band (approximately 3 cm) and C-band (approximately 5 cm) are sensitive to small components of the canopy. The microwave energy at such wavelengths is not able to penetrate deeply into the canopy to return a signal relative to biomass from the larger elements of the forest. Longer wavelength microwave in L-band (approximately 24 cm) and P-band (approximately 70 cm) interact with larger forest components, such as branches and stem (Lucas et al., 2010; Imhoff 1995.a; Imhoff 1995.b; Wang et al., 1995). It has been shown in different studies that the radar backscatter also saturates at certain biomass levels, depending on the wavelength. For example, P-band backscatter saturates at biomass levels of 100 to 300 Mg ha⁻¹ in tropical forest, and 200 Mg ha⁻¹ in the boreal and temperate forests (Saatchi et al., 2011.a; Imhoff 1995.b; Rignot et al., 1995). L-band saturates at 40 Mg ha⁻¹ to 272 Mg ha⁻¹ (Ahmed et al., 2014; Saatchi et al., 2011.a; Lucas et al., 2010; Imhoff 1995.b). A biomass saturation level of 20 Mg ha⁻¹ was observed for the tropical forest using C-band (Imhoff 1995.b). Therefore, models based on reflectance or backscatters are insufficient to cover the full range of the biomass values due to the weak sensitivity to biomass at higher levels. In the best case, biomass levels up to 300 Mg ha⁻¹ can be estimated, while there are many situations where AGB reaches 600 Mg ha⁻¹ or more in tropical forests. It was even reported that in a Canadian boreal forest field plot (Abitibi region, Quebec), a 400 m² plot had a biomass more than 600 Mg ha⁻¹ due to the chance concentration of several very large

trembling aspens (Benoît St-Onge, pers. comm. 2016). For these reasons, other approaches to biomass estimation from space must be sought.

1.2.2 Forest height-biomass models

Biomass can be conceived as the product of tree volume and tree density. Because the density of the woody parts of trees is not highly variable among different species of a given ecoregion, volume estimations can easily be translated into biomass. The volume of a single tree can, in turn, be simply viewed as being proportional to the product of tree width (say, DBH), and tree height. The biomass of a plot, therefore, depends on the number of tree per hectare, their DBH, and height. In closed canopies, during a large part of the growing stage of trees, there indeed exists a very close relationship between canopy height, volume ($\text{m}^3 \text{ha}^{-1}$) and biomass. Such relationships were well illustrated in numerous studies based on airborne lidar data (St-Onge et al., 2008.b; Næsset 2002). Moreover, using remotely sensed height from space as the sole predictor, Solberg et al., (2014) estimated boreal forest biomass with a RMSE of 43%. Lefsky et al. (2002) showed a single equation based on forest height explains 84% biomass variations for forests in North America and Canada. For this reason, measuring the height of forest canopies theoretically provides a very useful, albeit imperfect, predictor of biomass.

Forest height as the main predictor of forest biomass can be extracted to some extent using satellite-based 3D (three-dimensional) remote sensing (Lindberg et al., 2012; Hall et al., 2011; Le Toan et al., 2011), and can then be used to estimate wood volume, which is then converted to biomass, or biomass directly. Spaceborne sensors that acquired 3D data of potential use for forest structural mapping either produce data on sparse samples of forest canopies or offer continuous mapping capabilities. The GLAS (Geoscience Laser Altimeter System) sensor is an example of the first category. It has been used to produce forest vertical profiles within large footprints. Examples of the

other category are radar sensors with interferometry capability. They have been used to generate continuous height maps of either the surface of canopies or some intermediate level between the canopy surface and ground level. Radar sensors have here again an advantage over short wavelength sensors (i.e. using a laser) as they are not impeded by clouds. In the following subsections, we provide details on these different families of 3D spaceborne sensors, and also illustrate how they can be combined, amongst themselves or with optical image data.

1.2.2.1 Estimating forest height and biomass using the GLAS sensor

The only spaceborne lidar sensor for measuring forest attributes was GLAS. This instrument was mounted on the ICESat platform (Ice, Cloud, and land Elevation Satellite) as the part of the NASA Earth Observing System (Schutz et al., 2005). Its main mission objective was to help quantify changes in glacier elevations, but it was also used for other applications, such as the extraction of forest height and biomass. GLAS covered most of the Earth's surface from 86°N to 86°S latitudes from 2003 to 2009. It acquired vertical laser profiles in footprints with a diameter of 65 m separated by 172 m along the track, and tracks were ± 1 km apart at the equator. Of course, the acquisition could only be made in clear sky conditions. In the simplest forest cases, i.e. a single-storey dense canopy over the flat horizontal ground, the height of the canopy could be estimated based on the elevation difference between the first intensity peak (canopy surface), and the last one (ground level). However, because of their large footprint, these signals become very ambiguous in the case of sloping terrain or the presence of multi-storey forests. Extracting forest height from GLAS vertical profiles is said to be limited to areas having a slope less than 10% (Hilbert and Schmillius 2012). The ICESat platform carrying the GLAS sensor was decommissioned in 2010.

Despite its inherent limitations, GLAS data was used in attempts to characterize forest height or biomass. Information extracted from the GLAS footprints were extrapolated

using ancillary datasets such as MODIS (Moderate Resolution Imaging Spectroradiometer) and SRTM (Shuttle Radar Topographic Mission) to produce forest height maps (Lefsky 2010; Simard et al., 2011). Then, these forest height maps were converted to global biomass maps using height-biomass equations adapted to the different ecoregions. Saatchi et al. (2011.b) estimated the uncertainty of these maps as being of 30%. The main source of error stems from the uncertainty of the forest height derived from GLAS. One global forest height map based on GLAS was evaluated using field height measurements, which resulted in a RMSE of 6.1 m, and a r^2 of 0.5 (Simard et al., 2011).

ICESat-2, carrying the Advanced Topographic Laser Altimeter System (ATLAS) sensor, is set to be launched in 2018. Its laser sensor will use a photon counting technique for reconstructing return waveforms (Markus et al., 2016; Abdalati et al., 2010; Brunt et al., 2016). Although the footprint diameter is predicted to be much smaller, at 10 m, the use of a green laser (532 nm) instead of near infrared in the case of GLAS (1064 nm), will generate a much weaker signal over forests due to the markedly lower reflectance of vegetation in the green relatively to the near infrared. Applications to forest characterization are not expected to be highly successful based on the sensor's specifications (Herzfeld et al., 2013).

1.2.2.2 Estimating forest height and biomass using InSAR

The second type of sensors with spatially continuous mapping ability includes optical imagers with the stereo capability or radar systems with interferometry capability. These sensors can provide spatially continuous forest surface elevations or above ground forest heights mapping and can theoretically yield more accurate results than the sampling method (Houghton et al., 2007). Optical stereo images, however, suffer from the same limitations as any other optical imagery sensor, i.e. they cannot acquire data in cloudy conditions or nighttime. For this reason, this section will only present

radar-based solutions. Satellite-based interferometric radar is the only remote sensing technology that can guarantee fast acquisition of global and wall-to-wall 3D coverage without being impeded by cloudiness or limited to daytime. Graham (1974) introduced synthetic aperture radar for topographic mapping. InSAR (interferometric synthetic aperture radar) was later developed and used on airborne and spaceborne sensors in the 1980s and 1990s to generate elevation maps. In InSAR, the phase difference between two SAR images is calculated and converted into surface elevation differences. These relative elevations can then be converted to absolute values using control points of known elevation. Figure 1.1 presents the interferometric SAR acquisition geometry. Two sensors separated by a spatial baseline receive signals from which the phase difference can be measured to produce topographical data.

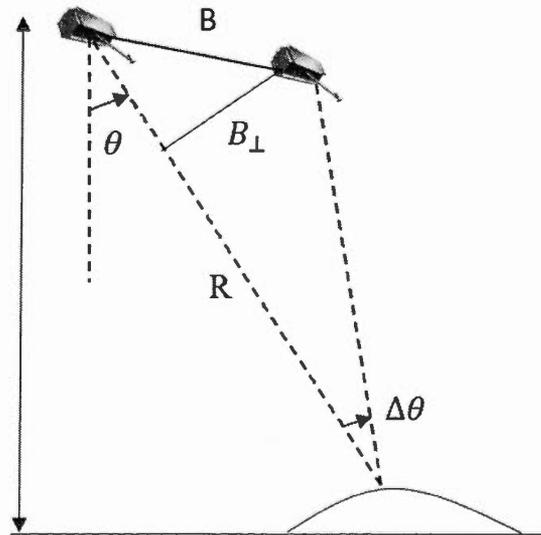


Figure 1.1 Interferometric SAR acquisition geometry. $\Delta\theta$ is the difference between the two incidence angles for two sensors, θ is incidence angle of sensor, B is baseline distance between two sensors, B_{\perp} is the perpendicular baseline, and R is the slant range distance.

The performance in converting phase to height critically depends on the baseline. The common term to express the effective spatial baseline is the vertical wavenumber, K_z . This term can be approximated by equation 1.2. This vertical wavenumber term relates the interferometric phase (φ) to the terrain height (h) by equation 1.3. Finally, the terrain height corresponding to the 2π interferometric phase change is called the height of ambiguity, or HoA.

$$K_z = \frac{2\pi B_{\perp}}{R\lambda \sin \theta} \quad [1.2]$$

$$h = \frac{\varphi}{K_z} \quad [1.3]$$

Interferometric image pairs can be acquired either simultaneously (single-pass) or in two passes (repeat-pass). The repeat-pass mode is the most common one since it requires just one sensor. However, it is limited to the reconstruction of the topography of bare lands, or studies of elevation changes, such as land subsidence caused by ground water extraction (Simard et al., 2012).

Minute changes in the forest between two passes suffice to create coherence problems between the two images that preclude interferometric reconstruction of the canopy elevations. Temporal decorrelation is the main unknown in repeat-pass InSAR systems. It is caused by the motion of forest elements (e.g. branches) due to the wind, as well as changes in moisture content or phenological state (fallen leaves, growth, or senescence) (Simard et al., 2012). Temporal decorrelation problems can be solved by using single-pass interferometers, a configuration in which two antennae separated by a certain distance (baseline) are operating simultaneously.

SRTM and TanDEM-X are the only two satellite missions that have acquired data in this mode over the world forests. The SRTM mission (using the C-band principally, with a wavelength of 5.6 cm, and to a lesser extent X-band, at 3.1 cm) operated in

February 2000 with the purpose of generating a global DEM (digital elevation Model) of the Earth. The single-pass TanDEM-X mission in X-band (3.1 cm) was put into operation in 2010 with the goal of producing a global high-resolution DEM (12 m) with a relative error of 2 m (completed in 2016).

The instantaneous phase difference allows the calculation of the height of the SPC (Scattering Phase Centre), which in turn is related to the canopy elevation. The vertical location of the SPC relative to the ground surface depends on canopy properties (species, density, foliage, etc.) and radar wavelength. The location of the SPC relatively to the canopy surface is often referred to as “penetration depth” (Liu et al., 2009; Thirion-Lefevre and Colin-Koeniguer 2007; Balzter et al., 2007.b; Izzawati et al., 2006). For a given forest structure, SPC is closer to the canopy surface at short wavelengths (e.g., in X-band), and deeper within the canopy at longer wavelengths (e.g., in C- or L-bands). In some studies, a combination of short and long wavelengths was used to estimate forest height by subtracting the InSAR ground height, based on L- or P-band, from the InSAR canopy surface height obtained in X-band (Balzter et al., 2007.a; Balzter et al., 2007.b; Neeff et al., 2005). Using single polarization, interferometric SAR can indeed only produce a single height surface such as a DTM or a DSM or something intermediary. Therefore, using two InSAR systems with short and long wavelengths or a combination of one radar system with the ancillary data such as lidar DTM elevations is needed.

In the case of multi-polarization interferometry however, i.e. Pol-InSAR (Polarimetric Interferometric SAR), techniques can be developed such that data from only one radar system may suffice to estimate forest height (Kugler et al., 2014; Soja et al., 2014; Askne et al., 2013; Praks et al., 2012; Hajnsek et al., 2009). Height can be estimated, for example, by inversion of the RVoG (Random Volume over Ground) model (Soja et al., 2014; Kugler et al., 2014; Lee and Fatoyinbo 2015). Again, a single-pass system,

with multi-polarimetric capabilities, is required for the RVoG or other inversion techniques.

An InSAR height map extracted from TanDEM-X was used to estimate forest volume and biomass by Solberg et al. (2013). It was found that in the studied region a 1 m forest height increment corresponds to a volume increase of $23 \text{ m}^3 \text{ ha}^{-1}$ and a biomass increase of 14 Mg ha^{-1} in the boreal forest. The accuracy of this estimation is 43–44% (RMSE) at the plot level, and 19–20% at the stand level. A similar process was applied in managed boreal forests using TanDEM-X InSAR information, and a RMSE of 32% for stem volume, and 20% for Lorey's height was achieved (Karila et al., 2015). Moreover, the height of mangroves was estimated with an accuracy of 10% using a PolInSAR approach by Lee and Fatoyinbo (2015).

1.3 The need for combining data from different spaceborne sensors

Using a single spaceborne sensor for mapping biomass is so far not sufficient to extract all the needed information for producing a global map of forest biomass. Single-pass InSAR sensors offer much promise, but there currently exists only one operational single-pass InSAR mission (TanDEM-X), providing data in dual polarization (HH, VV) only. As will be demonstrated in the next chapter, TanDEM-X interferograms can be used to generate DSMs, not DTMs. The most accurate source of DTMs is data from ALS, but world coverage is still very patchy. Moreover, InSAR cannot generate data that allows useful species identification, even in very broad classes (such as coniferous vs. deciduous trees). Species information is needed to select the proper height to biomass relationships in the presence of multispecific forests. The combination of InSAR to other remote sensing data therefore necessarily has to be considered.

As explained in section 1.2.2.1, the current existing large-scale forest height and biomass maps have been produced using a combination of GLAS with InSAR and

optical datasets. There is, however, an important limitation to these maps. Forest height was estimated only within the GLAS footprints, and only in those that did not fall on steeper slopes. These heights were then averaged per region, each region being defined by homogeneity criteria, for example in terms of MODIS spectra. The GLAS-derived heights are therefore taken as being representative of the average height of each region. However, there is no guarantee that this is the case. The forest height estimates are then used as the main predictor of biomass. There is a risk that the extrapolation of the GLAS-derived height to regions creates important errors, including regional or even global biomass biases.

Ideally, data combination should involve full spatial coverage data, available from space. While multispectral optical sensors able to provide some information on species abundance, for example, Landsat 8 or Sentinel 2, there is no current spaceborne mission that could provide a DTM. The closest is the SRTM data, but as we will show with more details in Chapter IV, the elevations of this dataset are located somewhere between the ground and the top of the canopy, depending on vegetational characteristics. As a very promising mission, TanDEM-L should provide a continuous quasi-DTM worldwide due to its much longer wavelength of 23.6 cm that will allow deep penetration in vegetation (Krieger et al., 2009). This mission from DLR is however still in its funding stage (Irena Hajnsek, DLR, pers. comm. 2016), and the earliest launch date is 2023 (DLR-TanDEM-L, 2016).

1.4 Optimal approach for producing a global biomass map

Based on the considerations presented in the preceding sections, we here define what is in our opinion the optimal approach for producing a global biomass map. In summary, the characteristics of this approach should be:

- to rely on the height-biomass relationship, and not on a reflectance- or backscatter-biomass relationship;
- full continuous spatial coverage of the world's forests, i.e. approximately from 72°N to 72°S;
- a spatial resolution fine enough (i.e. approx. 20x20 m or finer) to enable linking field plot data unambiguously to pixels of the remote sensing dataset for calibration and verification purposes;
- the capacity to acquire height data over an entire biome during one full leaf-on period;
- to achieve biomass predictions with a %RMSE of 20% or better, at the set resolution (e.g. approx. 20x20m);

Knowing that no spaceborne laser scanner is currently planned, and that stereo imagery is highly impeded by the presence of clouds, the above criteria point towards a single-pass InSAR solution. Ideally, the InSAR sensors should operate at multiple wavelengths, allowing for the generation of a DSM, and a DTM, from which a global CHM could be developed. They would also be able to acquire data in full polarimetric mode (HH, VV, HV, and VH) at a relatively long wavelength (L or P) to allow the inversion of models such as the RVoG model.

The current spaceborne system with the closest characteristics is TanDEM-X. However, its main, and quite important, limitation is that it cannot provide a DTM (and no other spaceborne sensor can). Interferograms derived from the X-band should represent the elevations of the surface of the canopy because backscattering originates from small components, such as leaves, needles or fine branches (Mougin et al., 1993). The scattering phase center should, therefore, be at, or just a bit below the top of the canopy. At the onset of this research project (2011), no one knew exactly where the TanDEM-X SPC would be located vertically, relative to the top of the canopy. It was

also unclear if the polarization mode would affect the location of the SPC. At the IGARSS (International Geoscience and Remote Sensing Symposium) in 2014, the forum where we first presented our results (see Appendix A), the location of the TanDEM-X SPC was still debated and the subject of many questions. These questions are answered in the next chapter, i.e., the first paper composing this thesis.

To map forest height as a predictor, and assuming for the time being that TanDEM-X can provide a global DSM, an accurate DTM is needed. This can only come from an active sensor, such as a lidar or radar sensor, because only those can send pulses that penetrate vegetation canopies, and then obtain a signal from the ground level. For the purpose of the demonstrations that we set out to make within this research project, we first had to recourse to an ALS DTM in order to assess the quality of the TanDEM-X data without introducing uncertainty about ground elevation. It is possible that large land masses will be covered by ALS soon. A current project ("Elevation project," from Natural Resources Canada, Geomatics Canada 2016) exists for creating a unified ALS DTM for the entire southern portion of Canada (approx. south of the 50th parallel). Similar projects exist for the United States, Scandinavia, large parts of Europe, etc. However, there does not seem to be any such endeavour in Russia, which contains a large part of boreal forests. The same applies to large parts of Africa and South America, where most tropical forests are found.

For this reason, we have explored, in Chapter IV (3rd thesis paper) the possibility of correcting DEMs from the SRTM mission to create "quasi-DTMs." The penetration of SRTM C-band in vegetation canopies is greater than in X-band but pulses do not reach the ground level in dense and tall forest stands. For example, a 12 m average height difference between SRTM DEM and a lidar DTM was reported by Su et al. (2015). There is, however, evidence that SRTM DEMs can be corrected to a certain degree (Su et al. in 2014 and 2015). Combining the corrected SRTM DEMs with TanDEM-X DSMs could, therefore, be a solution to procure a global CHM at a high-resolution.

From this CHM, forest biomass could be predicted, but the level of uncertainty could be higher and remains to be investigated. The DTM would not require frequent updates as the topography does not change rapidly, except following local catastrophic events (e.g. landslides). The DSM would itself require a frequent update, a task that the TanDEM-X mission can accomplish. Fortunately, DLR is already planning a follow-up mission to TanDEM-X (Irena Hajnsek pers. comm. 2016).

1.5 General objective of the research project

Ideally, we would have sought to make the demonstration that the new methods proposed in this thesis for mapping biomass could be applied for any forest in the world, therefore reaching the goal of providing a method for global biomass assessment. This would, however, require current DSMs (from TanDEM-X), DTMs (from lidar) and field datasets in the boreal, temperate, and tropical biomes, as well as specific datasets in the sparse savannah or subarctic forests. On the one hand, these resources are expensive and difficult to obtain. On the other, such a broad project exceeds the normal scope of a Ph.D. project and would entail an enormous amount of work, better suited for an entire team of researchers.

We have therefore limited the scope of this project to closed-canopy boreal forests due to location SPC which in these conditions will stay near the top of canopy. Our general objective was to show how TanDEM-X interferograms, when combined to other datasets, can be used to generate height maps which in turn are used to predict and map forest above ground biomass. Throughout this thesis, we test our methods at different resolutions, from 5 m to stand level, and assess accuracy at these different resolutions. It is hoped that our results can be generalized to most closed-canopy boreal forests of the world.

1.6 Specific objectives and thesis organization

This research thesis is mainly comprised of three papers presenting detailed methods and results on all aspects of our doctoral project. The accuracy of digital surface models extracted from TanDEM-X interferograms, and transformed into CHM by subtracting a lidar DTM, was evaluated in Chapter II. Due to availability of only HH polarization of TanDEM-X at global scale, we continued our process based on producing DSM from HH interferograms. However, the DSM based on VV polarization was extracted and compared to the HH version, and the differences were found to be negligible. This assessment was performed at different resolutions from 5 m to 25 m pixels and stand level. Then, the effects of various parameters such as forest height, forest density, gap volume, and LIA (Local Incidence Angle) on the CHM accuracy were studied.

Because the final DSMs products released by DLR (and distributed commercially by Airbus Defense and Space) are assembled from interferograms acquired at different moments in time, under various satellite configurations, we were concerned that they may have variable accuracy in terms of RMSE and even bias. For this reason, we have studied five DSMs extracted under variable TanDEM-X acquisitions parameters with regard to phenological conditions, interferometric baseline, and incidence angle. The effect of meteorological conditions was also assessed to a certain degree. The results are presented in the second research paper, i.e. Chapter III.

In the following chapter and last thesis paper, we present a method for correcting SRTM elevations and use the resulting quasi-DTM to generate a CHM by subtracting its elevations from a TanDEM-X DSM, thus creating a DSM entirely from satellite sensors. Using this CHM, TanDEM-X interferometric coherence, and Landsat 8 vegetation indices, we have created a biomass map.

Finally, Chapter V presents a summary of the results of three thesis papers and discusses the potential and limitations of the height and biomass maps created, in light of results obtained by other researchers, and presents future research directions. Our first thesis results were presented in a short paper published in the Proceedings of the IGARSS 2014, in Quebec City. This paper was placed in the appendix section. It remains of interest despite being largely superseded by the papers presented in Chapters II to IV, as it contains early and original results, some of which do not appear in the journal papers. These results have guided and influenced our research at this early stage.

CHAPTER II

CANOPY HEIGHT MODEL (CHM) DERIVED FROM A TANDEM-X INSAR DSM AND AN AIRBORNE LIDAR DTM IN BOREAL FOREST

This chapter has been published as:

Y. Sadeghi, B. St-Onge, B. Leblon, M. Simard, "Canopy height model (CHM) derived from a TanDEM-X InSAR DSM and an airborne lidar DTM in boreal forest," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., Vol. 9, issue. 1, pp. 381-397, January 2016.

2.1 Résumé

La mission TanDEM-X, première mission globale d'interferométrie RADAR à passage unique, permet de générer une cartographie spatialement continue des élévations de la canopée. Dans cet article, nous évaluons la possibilité d'utiliser les données de TanDEM-X, en combinaison avec un modèle externe numérique de terrain (MNT), pour cartographier les hauteurs de la canopée en forêt boréale. Une comparaison et la validation d'un modèle de hauteur de la canopée (MHC) issue des données TanDEM-X à un MHC de référence lidar a été réalisée sur deux bases comparatives soit: la hauteur de surface de la canopée (HSC) et la hauteur dominante (HD). Ces comparatifs ont été réalisés à l'échelle du peuplement à des résolutions spatiales allant de 5 m à 25 m. La validation du MHC TanDEM-X a permis de démontrer une RMSE de 2.7 m à

une résolution de 5 m et de 1.9 m à une résolution de 25 m (à l'échelle du peuplement or pour la hauteur moyenne à l'échelle du peuplement). Ces résultats nous ont permis de démontrer que le MHC TanDEM-X présentait une résolution plus grossière que le MHC lidar correspondant. La différence de la hauteur InSAR par rapport à la hauteur du lidar varie entre 1.2 m et 1.5 m, mais les hauteurs InSAR sous-estimaient la hauteur des arbres dominants de 4.6 m à 7.5 m. Des différences similaires ont été observées pour les variables HSC et HD TanDEM-X en comparaison au lidar (respectivement 6.04 m, 8.98 m, et 8.05 m pour différentes résolutions). Les résultats nous démontrent que les hauteurs interférométriques TanDEM-X sont très proches de la hauteur de référence de lidar et que la pénétration sous les hauteurs dominantes est causée par la propagation du signal des microondes entre les sommets des arbres et la surface principale du feuillage en forêt boréale. Enfin, la précision des estimations de hauteur InSAR n'a pas été sensible aux effets de la densité des arbres, mais a été légèrement affectée par des angles d'incidence locaux, le volume des trouées et la hauteur de la canopée.

Mot-clés: TanDEM-X modèle de hauteur de la canopée (MHC), lidar MHC, la hauteur de surface de la canopée (HSC), hauteur dominante (HD)

2.2 Abstract

The first global X-band spaceborne single-pass interferometer mission, TanDEM-X, provides a spatially-continuous map of global canopy elevations. In this paper we assess the use of TanDEM-X data, in combination with an external DTM, to map boreal canopy heights. A comparison of the TanDEM-X CHM to a validated reference lidar CHM was performed based on two definitions of canopy height: canopy surface height (CSH), and dominant height (DH) at spatial resolutions ranging from 5 m to 25 m, and at the stand level. We found the TanDEM-X CHM to have a coarser resolution than the corresponding lidar CHM. This was apparent in the height validation of the

TanDEM-X CHM, which had a RMSE of 2.7 m at the 5 m resolution, 1.9 m at the 25 m resolution, and 1.5 m at the stand level. The difference of the InSAR height compared to the lidar height varied between 1.2 m and 1.5 m, but InSAR heights were below the height of dominant trees by 4.6 m to 7.5 m. Similar discrepancies were observed for the lidar canopy surface height relatively to dominant height (respectively 6.04 m, 8.98 m, and 8.05 m). The results show that the TanDEM-X interferometric heights are very close to the lidar reference height and that penetration below the dominant height is caused by propagation of the microwave signal between the tree apices and the main foliage surface in the boreal forest. Finally, the accuracy of InSAR height estimates was not sensitive to tree density effects but was moderately affected by local incidence angles, gap volume, and canopy height.

Keywords: TanDEM-X canopy height model (CHM), lidar CHM, canopy surface height (CSH), dominant height (DH)

2.3 Introduction

Despite considerable research efforts that have been devoted to developing remote sensing methods for measuring global carbon stocks and fluxes, significant uncertainty remains in the estimates of carbon that is stored in the Earth's forests (Gonzalez et al., 2010; Neigh et al., 2013; Hall et al., 2011). Because canopy height is a key parameter for estimating biomass (Pflugmacher et al., 2008; Dubayah et al., 2010; Molto et al., 2014; Feldpausch et al., 2012; Kellner et al., 2009), satellite-based three-dimensional (3D) remote sensing is likely the most efficient and accurate method for quantifying forest biomass globally (Hall et al., 2011; Le Toan et al., 2011; Lindberg et al., 2012). Two broad types of 3D spaceborne sensors exist a) low point density, large-footprint lidar (Light Detection and Ranging); and b) imagers such as stereo optical sensors or synthetic aperture radar (SAR). They use two different remote sensing approaches for biomass mapping, which are respectively a) sampling followed by spatialization (e.g.,

interpolation) and b) spatially continuous mapping. The first type of sensor includes GLAS (Geoscience Laser Altimeter System), which was mounted on the ICESat platform (Ice, Cloud, and land Elevation Satellite) as part of the NASA Earth Observing System (Schutz et al., 2005). GLAS covered most of Earth's surface and was used for forest height estimation between 2003 and 2009; it was decommissioned in 2010. In the near future, similar sampling lidar sensors will become available, such as the ATLAS (Advanced Topographic Laser Altimeter System) sensor of ICESat-II (Moussavi et al., 2014; Abdalati et al., 2010), GEDI (Global Ecosystem Dynamics Investigation) (WWW.Science.nasa.gov), and the Japanese ISS/JEM-borne vegetation lidar sensors that will be carried by the International Space Station (Murooka et al., 2013). With respect to sampling approaches, height profile metrics are used to estimate forest height from the lidar waveform at the footprint locations (e.g., Simard et al. (2011), Lefsky (2010)), which is then converted to biomass using allometric equations (Pflugmacher et al., 2008; Lefsky 2010; Baccini et al., 2012; Saatchi et al., 2011.b). However, sampling methods that are used to estimate biomass are based on sparse samples and are generally spatialized using inference methods and ancillary data, such as MODIS (Moderate Resolution Imaging Spectroradiometer) imagery. These lidar sensors will have coverage that is limited to about 50°N-50°S, thereby omitting an important part of the boreal forest.

The second type of sensor includes optical imagers with stereo capability such as WorldView-3, or SAR systems that enable stereo-radargrammetry or interferometry. They can provide spatially continuous maps of forest surface elevations or forest heights above ground and can theoretically yield more accurate results than the sampling method (Houghton et al., 2007). The remote sensing images that are generated by these sensors can also be directly linked to existing field plot networks, provided that the remote sensing data completely overlap these plots. Spaceborne continuous 3D remote sensing methods can be implemented through stereo-photogrammetry (St-Onge et al., 2008.a), and with SAR images, through stereo-

radargrammetry (Vastaranta et al., 2014; Raggam et al., 2009), InSAR (Interferometric Synthetic Aperture Radar) (Solberg et al., 2013; Treuhaft et al., 2015; Solberg et al., 2010), or Pol-InSAR (Polarimetric Interferometry SAR) (Kugler et al., 2014; Neumann et al., 2012). Only SAR-based techniques can guarantee the acquisition of usable high-resolution images worldwide because they are unhindered by cloud coverage. InSAR and Pol-InSAR techniques generally produce more accurate results than does stereoradargrammetry (Vastaranta et al., 2014; Toutin et al., 2010; Persson 2014.a). With the InSAR technique, the phase difference between two SAR images that have been acquired either simultaneously or in two passes is used to estimate the height of the Scattering Phase Centre (SPC), which in turn is related to the elevation of the canopy. The vertical location of SPC relative to the ground surface depends upon canopy properties (species, density, etc.) and radar wavelength. The location of the SPC relative to the canopy surface is itself often referred to as “penetration” (Liu et al., 2009; Izzawati et al., 2006; Balzter et al., 2007.a,b; Thirion-Lefevre and Colin-Koeniguer 2007). For a given forest structure, SPC is closer to the canopy surface at short wavelengths (e.g., in X-band, (Praks et al., 2012)) and deeper within the canopy at longer wavelengths (e.g., in L-band, (Walker et al., 2007; Krieger et al., 2009).

Using single polarization interferometry, only one measurement of elevation per resolution cell can be obtained to produce a DSM (Digital Surface Model) that lies somewhere between the canopy surface and ground level. Therefore, ancillary data on ground topography are required to estimate canopy height. In the case of Pol-InSAR techniques, the estimation of canopy height can be performed without external information on bare terrain elevations (Kugler et al., 2014; Praks et al., 2012; Hajnsek et al., 2009; Askne et al., 2013; Soja et al., 2014), because both the canopy and ground surface levels are extracted from the polarimetric InSAR data. Canopy height estimation can be attempted by inversion of the RVoG (Random Volume over Ground) model (Treuhaft et al., 1996; Treuhaft et al., 2000). The inversion technique, however, is sensitive to the effective spatial baseline, which is expressed by the vertical

wavenumber (k_z). The vertical wavenumber is itself a function of the height of ambiguity, viz., a height difference that is proportional to the 2π phase changes. Therefore, inversion is only reliable for k_z values ranging from 0.05 to 0.2 (Kugler et al., 2014). Outside this range, the relationship between coherence and height saturates, and the inversion results become largely inaccurate, i.e. well over the 10% deviation mark which is considered as being an acceptable level of error (Kugler et al., 2015). For both Pol-InSAR and InSAR techniques, small changes in the canopy (moisture, position of scatterers, wind, etc.) can occur when the two images forming the interferometric pair are acquired at different times (i.e., different repeat passes, which generally occur on different dates); this leads to temporal decorrelation, which may hinder successful implementation of Pol-InSAR inversion techniques or InSAR methods (Askne et al., 2013; Lavalley et al., 2012; Lee et al., 2013; Simard et al., 2012). Furthermore, the ground signal in densely vegetated areas must remain sufficiently strong for the RVoG method to produce a reliable estimate of the bare earth elevations.

Temporal decorrelation problems can be solved by the use of single-pass interferometers, in which two antennae that are separated by a certain distance (baseline) are operated simultaneously. Only two satellite missions have acquired data in such a mode over world forests: SRTM (Shuttle Radar Topographic Mission) and TanDEM-X. The SRTM mission, which operated in the C-band (5.6 cm) and X-band (3.1 cm) in February 2000, had a coverage limited to 60°N-56°S and was discontinuous in the X-band. The TanDEM-X mission has operated in the X-band (3.1 cm) since June 2010 (Krieger et al., 2010), providing global and continuous coverage. SRTM interferometry data were used to produce a DSM having a 30 m pixel size, but its likely true resolution is coarser than 30 m (Farr et al., 2007). In contrast, the objective of the TanDEM-X mission is to acquire at least one interferometric pair in single polarization (HH) over all terrestrial areas at a resolution of 12 m. Dual-polarization data will only be acquired over selected sites for research purposes (Irena Hajnsek, *personal communication*; TanDEM-X Science Coordination, Deutsches Zentrum für Luft- und

Raumfahrt – DLR). For the foreseeable future, single InSAR analysis of TanDEM-X data will be the only means by which a global and complete forest DSM can be created. Provided that high-quality ground elevation data can be acquired (e.g., airborne scanning lidar or other high-quality topographical datasets), it is also the only way for eventually creating a global canopy height model from which biomass could be derived.

DSMs that are based on single-pass spaceborne SAR interferometry of the TanDEM-X (Solberg et al., 2013; Solberg et al., 2015.a), SRTM-X (although having incomplete coverage (Solberg et al., 2010; Walker et al., 2007)), and SRTM-C (Simard et al., 2008; Kellndorfer et al., 2004) systems can be used to create CHMs by subtracting an external DTM. In principle, interferograms that are derived from the X-band should represent elevations that are close to the surface of the canopy because backscattering originates from small components, such as leaves, needles or fine branches (Mougin et al., 1993). This leads to an expected low degree of penetration within the top most foliage for the TanDEM-X SPC. A few studies have indicated that the degree of X-band penetration within forest canopies is much greater (Kugler et al., 2014). When comparing SPC height to canopy height, it was found to be several meters below the average height of the dominant trees, e.g., 7 m using airborne RAMSES system developed by ONERA (Garestier et al., 2006), 9 m by DLR's E-SAR system (Hajnsek et al., 2009), 20-30 % of forest height in (Balzter et al., 2007.a; Balzter et al., 2007.b; Praks et al., 2012;) using E-SAR, or 1/4 to 1/3 of the way into the canopy height when employing GeoSAR (Hensley et al., 2001.a; Hensley et al., 2001.b). However, the manner in which penetration is defined and calculated varied across these studies. Furthermore, some researchers have shown that when forest density decreases, the apparent penetration depth difference between the polarized channels increases, due to the occurrence of canopy gaps (Garestier et al., 2008). Results were sometimes evaluated at high resolution (Askne et al., 2013), but were more often obtained at the stand level (Kugler et al., 2014, which is a very coarse spatial unit in which an unknown number of large

canopy gaps (several m^2 and more) can influence the results. Furthermore, the degree of canopy penetration is often evaluated relative to a rough surrogate of the reference dominant height of trees, such as the maximum lidar height within small local windows. Therefore, it is still unclear as to where exactly the TanDEM-X SPC is located within the canopy and what factors determine how the corresponding interferometric DSMs differ from high-resolution lidar DSMs that are often deemed as a reliable reference (Balzter et al., 2007.a,b; Praks et al., 2012; Hajnsek et al., 2009; Garestier et al., 2006; Hensley et al., 2001).

The objectives of this study were, therefore, to characterize the reliability of single polarization TanDEM-X interferograms for creating representative DSMs, and their usefulness in creating CHMs by subtracting an external DTM from the interferometric elevations. We hypothesize that the InSAR height should follow quite closely the corresponding lidar height that is used as a reference, but that the relationship between these two height surfaces will vary depending upon the spatial resolution at which the comparison is performed because of the presence of canopy gaps. We also posit that InSAR imaging factors such as the local incidence angle and InSAR coherence, together with forest structural attributes such as height, density, and gap volume should affect the location of the InSAR height in the canopy. These hypotheses and objectives should allow us to estimate the degree of similarity between the SPC height of TanDEM-X interferograms and a reference canopy height that was obtained from an airborne lidar dataset. We present a general comparison of TanDEM-X and lidar height surfaces, and then study the effects of InSAR imaging factors (local incidence angle and coherence) and forest attributes (height, density and gap volume of trees) on the discrepancies between InSAR and lidar heights. Most our experiments were performed at resolutions ranging from 5 x 5 m to 25 x 25 m, and at the stand scale, the latter only being included to allow comparisons with previous studies. Further, these experiments use two definitions of canopy height: the height of the highest lidar return in the canopy

(canopy surface height, CSH), and the average height of dominant trees (dominant height, DH).

2.4 Study area and data sources

2.4.1 Study area

The study was conducted in the Montmorency Research Forest, which is a 6600 ha site that is located about 70 km north of Quebec City, Canada (47°18' N, 71°08' W) (Figure 2.1). This area falls within the boreal shield ecozone zone (Leblanc and Bélanger 2000). It is mainly populated by balsam fir (*Abies balsamea* (L.) Miller), and paper birch (*Betula papyrifera* Marshall). There are also a few spruce species (*Picea mariana* (Miller) BSP, *P. glauca* (Moench) Voss) and trembling aspen (*Populus tremuloides* Michaux). The forest was actively harvested during the 1932-1944 period but has since largely regrown. Ecosystem-based management was implemented from 1992 onwards in the largest portion of the forest, while 13 % of the forest was declared protected and left untouched. The terrain within the study area ranges in elevation between 600 m and 1000 m, with locally strong slopes.

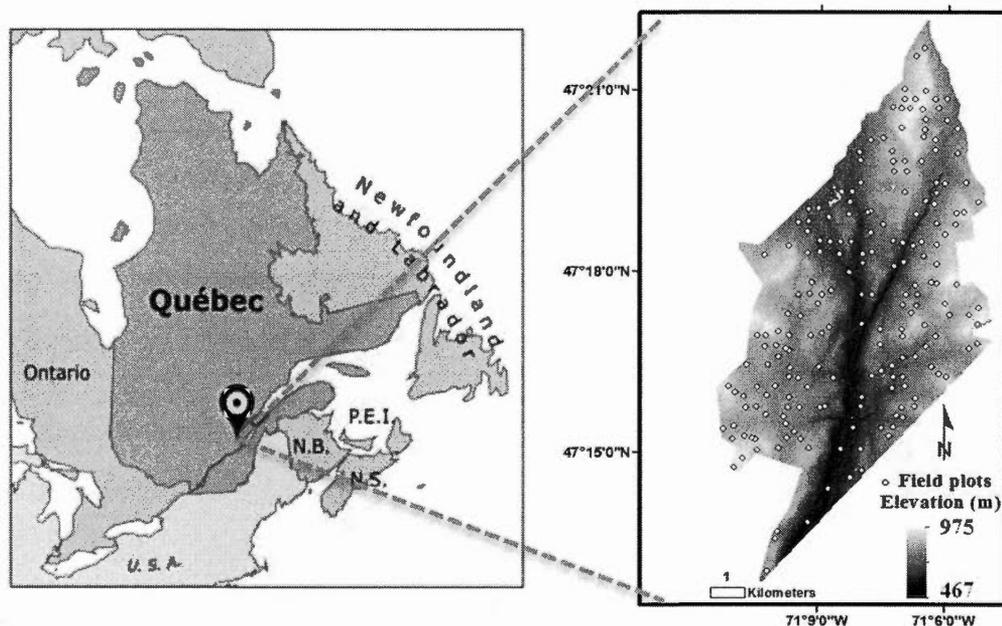


Figure 2.1 Location of the study area (left panel) showing the lidar DTM and the field plots (right panel)

2.4.2 Field and forest inventory data

The field data are comprised of measurements of individual tree heights and species, together with standard inventory dendrometric data for 200 permanent plots, each covering 400 m². Height measurements were performed on two or three trees that were located on each plot. These measurements were taken during the summers of 2012 (about 90 % of trees) and 2013 (about 10 % of trees) using a Vertex III clinometer (Haglöf Sweden AB). The diameter at breast height (DBH) of all trees having a diameter of 5 cm or more was measured. Tree density was defined in this study as the number of stems of DBH > 5 cm per hectare. The permanent plots are re-measured every five years as part of ongoing data collection activities at the Montmorency Forest, so the density values were current. Tree and plot statistics are presented in Table 2.1, while plot locations are shown in figure 2.1. We have used a forest stand map of the

study area to compile results on a stand-wise basis. The average area of the 3413 stands that were delineated on this map was about 2 ha.

Table 2.1 DBH, height and stem density of field plots

| | Trees ($n = 431$) | | Plots ($n = 200$) |
|---------|---------------------|---------------|-----------------------|
| | DBH (cm) | Height (m) | Density (stems/ha) |
| Minimum | 5 | 3.2 | 25 |
| Maximum | 116 | 26.4 | 2900 |
| Mean | 23 | 14.9 | 1298 |

2.4.3 Remote sensing data

A dual-polarization (HH/VV) high-resolution strip-map mode dataset was acquired on 15 July 2013 in bistatic mode by TanDEM-X on its ascending (right-looking) path. The images were acquired with range and azimuth resolutions of 1.2 m and 6.6 m, respectively, and obtained in the slant range (CoSSC) format for interferometric processing. The effective 107.54 m baseline resulted in the height of ambiguity of 43.56 m and an effective vertical interferometric wavenumber (k_z) of 0.145. The incidence angle at the image centre was 30.5°.

Airborne lidar data were collected on 6 and 9 August 2011, using an Optech ALTM 3100 laser (Optech Inc., Vaughan, ON, Canada) having a wavelength of 1046 nm, a divergence of 0.25 rad, a scan rate of 46 Hz and 56 Hz, and a maximum scan angle of 17°. The sensor was flown at 1000 m above ground level (AGL) with a pulse repetition frequency of 100 kHz, leading to a median first return density of 5 hits/m². The average distance between a lidar first return and its closest neighbour was 0.34 m, and the average distance between any random point (for example, representing a tree apex) and

the closest first return was 0.19 m. The sensor recorded up to 4 returns per pulse, which were classified as ground/non-ground by the lidar data provider.

2.5 Methods

2.5.1 Creating the TanDEM-X DSM

The TanDEM-X interferometric data were processed using the ENVI SARscape 5.0 Processor to produce co-polarized interferograms, complex coherence images (γ) (Costantini 1998) and DSMs that were co-registered with sub-pixel accuracy. Interferograms were flattened using a 90 m SRTM DEM. Phase unwrapping was performed using the minimum cost flow method (Costantini 1998), which was implemented within SARscape. The final products, which include coherence images and DSMs for HH acquisition with a 5 m spatial resolution, were georeferenced to a WGS-84 datum and UTM 19N projection using the SAR sensor orbit parameters.

To ensure that accuracy assessment of the InSAR DSM was not biased by gross errors, we masked out image regions where the following conditions were observed: lakes, forest blocks that were harvested during the time interval between the respective acquisitions of lidar and TanDEM-X data, layover areas (causing failure of interferometry), and a high tension electricity transportation corridor where pylons and suspended power cables were present. Polygons that were used for masking were drawn manually, based on visual analysis and ancillary information. In addition, areas, where the coherence value was lower than 0.3, were removed from the analysis because the corresponding interferometric results are unreliable (Martone et al., 2012; De Zan et al., 2013). These areas had a low signal to noise ratio (SNR) due to slope effects or low backscattering which led to poor coherence. Further, places, where localized phase unwrapping problems occurred, were also masked. These latter areas corresponded to terrain with very steep slopes, edges of clear-cuts, or sharp localized coherence changes

(e.g., between lakes and land). The total coverage of masked out pixels represents only 0.7 % of the area of the study region.

2.5.2 Creating the Lidar DSM

The canopy surface height was taken as being equivalent to the high-resolution lidar DSM created from the highest lidar return (first return) within each resolution cell. Due to the high sensitivity of the lidar sensor, this surface was postulated to be located very near the highest vegetation material within a resolution cell, possibly somewhat lower than the highest leaf or twig due to the minimum amount of laser energy interception required to trigger a first return (as demonstrated for example in figure 2.4 of (Doneus et al., 2010)). The lidar surface thus represented the micro-topography of the outer canopy layer, closely following the protuberances and gaps at the chosen resolution (Figure 2.2). The lidar DSM included treetops, tree sides, canopy gaps, shrubs, and had a height of zero in the absence of vegetation. To create a continuous lidar raster DSM from the lidar returns at 0.25 m resolution, all cells in which at least one first return had fallen received the Z -value of this return. If more than one first return fell within a given cell, only the highest Z -value was retained. Empty cells were filled with values that were obtained from a triangulated irregular network (TIN) built using neighbouring non-empty cells. To correct for small cavities (St-Onge 2008; Ben-Arie et al., 2009) that were caused by a laser pulse travelling obliquely very close to the side of a crown and hitting an object beneath it, a Laplacian filter was used to detect these voids, which were then filled by interpolating the values of the surrounding pixels.

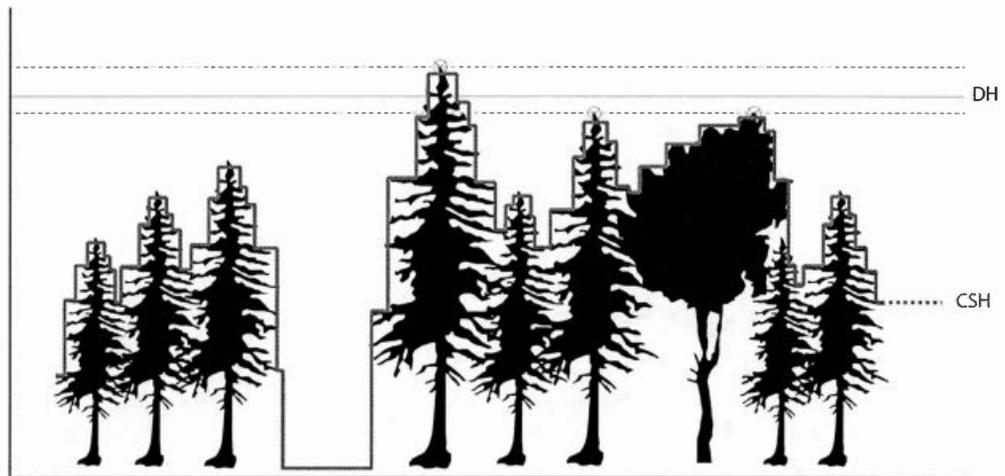


Figure 2.2 Representation of the two definitions of canopy height; a) canopy surface height (CSH, red line), and b) dominant height (DH, horizontal solid black line). Crossed circles indicate the height of the selected dominant trees

2.5.3 Creating the DTM and CHMs

A lidar DTM was necessary to convert the InSAR and lidar DSMs to CHMs. A procedure similar to that used for the creation of the DSM was applied. However, if more than one ground return fell into a given pixel, the lowest DTM value was retained. Empty pixels received an interpolated value, and no cavity filling was necessary. The respective CHMs were then created for the lidar DSM and the TanDEM-X DSM (hereafter, termed “InSAR CHM” for the sake of simplicity) by subtracting the lidar DTM values.

2.5.4 Comparing InSAR and Lidar CHMs

2.5.4.1 Accuracy of the Lidar CHM

Verifying the absolute accuracy of a lidar CHM would require comparing numerous field canopy surface height measurements to the corresponding lidar CHM heights. For this reason, such verification is rarely done. As a partial accuracy assessment, the lidar heights of 431 sample trees were compared to their corresponding field heights. Lidar height was defined as the highest valued CHM pixel within a given sample tree crown. Mean difference (bias), root-mean-square error (RMSE), and coefficient of determination (r^2) of the regression between the raw lidar heights and field heights were computed.

2.5.4.2 Height differences between the InSAR and Lidar canopy heights (ΔH)

When comparing InSAR and lidar canopy heights, we used two different definitions of height, which lead to two distinct methods of data preparation. First, canopy height was considered as the height of the *canopy surface* (CSH), within a given resolution cell (defined in section ‘Methods’ and Figure 2.2). Second, canopy height was defined as the *average height of dominant trees* (DH), which is routinely reported in standard forest inventories. This was computed as the average of the n tallest stems per unit surface, with n depending upon the various levels of spatial aggregation that was used in the study (Figure 2.2). For example, at the stand level, we used the common H_{100} metric (Mette et al., 2004). H_{100} is defined as the mean height of the 100 tallest trees per ha. Its use permits comparisons with other studies that have employed the same metric (e.g., (Solberg et al., 2013; Kugler et al., 2014; Mette et al., 2004)). We hereafter refer to these two definitions of canopy height, respectively, as “canopy surface height” (CSH) and “dominant height” (DH). CSH and DH were used in the following experiments. In all cases, the forest height differences (ΔH) were obtained by subtracting lidar from InSAR ($\Delta H = \text{InSAR} - \text{lidar}$).

In the first experiment, the InSAR CSH was compared to the corresponding lidar CSH values; a) at the initial InSAR resolution of 5 m, b) at various aggregated resolutions of 10 m, 15 m, 20 m and 25 m that resulted from 2x2, 3x3, 4x4, and 5x5 pixel aggregations, and c) at the forest stand level. Aggregation was applied to study the potential effect of surface generalization, both at a relatively high resolution (5x5 m to 25x25 m) and at the standard spatial unit used in forest management (stand). This allowed us to determine whether gains could be made by lowering image spatial resolution, e.g., by decreasing the effect of localized spurious InSAR heights if these were present, or by decreasing the effects of canopy gaps and isolated protruding trees. In all cases, comparisons between InSAR and lidar CSHs excluded masked regions. To bring the lidar CSH to the initial InSAR CSH resolution of 5 m, we aggregated the 0.25 m lidar pixels by calculating their mean value within a 5x5 m window. Any further lidar or InSAR data aggregation was performed by calculating their respective means of the 5 m pixels within each aggregation cell.

In the second experiment, we evaluated the capacity of the TanDEM-X data to represent variation in dominant tree height by comparing the InSAR DH to a reference map of dominant heights that were derived from the lidar data. The reference map was generated by identifying individual trees and extracting their heights using an in-house software application (SEGMA v. 9.0). Its processing steps are described as follows. First, the 0.25 m resolution lidar CHM is filtered using a Gaussian filter, in which the σ value varies proportionally to the local CHM height. Second, local maxima are detected on the filtered CHM and regions are grown around these maxima. The regions stop growing when certain criteria are met (e.g., strong valleys are formed between trees; the lowest crown height threshold is reached, and so on). Third, the maximum height of the non-filtered 0.25 m resolution CHM is extracted from each segmented crown. Its value corresponds exactly to that found in the initial lidar point file, minus the underlying DTM elevation. Fourth, because this height may often be underestimated due to the highest lidar return in a crown missing the apex, a correction

function was applied. This function was calibrated by regressing the field heights against the corresponding lidar heights for the 431 field-measured trees. A list results, which describes the horizontal position of each tree and its corrected height. From this list, we created maps of dominant height by calculating the H_{100} metric at various cell resolutions. In the field, these trees are selected by their diameter at breast height (DBH), but we here used their height to select them.

The InSAR dominant height was validated at the same resolutions that were used as in the previous canopy surface experiment (5x5 m up to 25x25 m, together with the stand level). To remain consistent with the concept underlying H_{100} measurement (100 trees/ha = 0.01 tree/m²), we retained only the maximum tree height at the 5x5 m and 10x10 m aggregation levels, together with the 2, 4 and 6 tallest trees, respectively, at the 15x15 m, 20x20 m and 25x25 m levels. At the stand level, mean dominant height was calculated from the n tallest trees, where $n = \text{stand area (in ha)} * 100$. In parallel, the highest InSAR CHM value was kept at the 10x10 m resolution, and the mean of the 2, 4, and 6 highest pixels, respectively, at the three aforementioned aggregation levels. Cells that were devoid of trees were excluded from analysis.

A third experiment was designed to isolate micro-topographic effects (e.g. the presence of gaps) on the canopy surface under the dominant height level from other factors that may affect X-beam penetration. Here, we compared both InSAR CSH and lidar CSH with the lidar DH. For conciseness, these calculations have been performed only at resolutions of 5 m and 25 m, and at the stand level. Our goal was to show how a generalized canopy surface lies systematically below dominant heights when gaps exist, regardless of whether it has been estimated from InSAR or lidar data.

In all three experiments, the discrepancy between the InSAR and corresponding lidar estimates is expressed in terms of the mean of the height differences (InSAR minus lidar), the root-mean-square error (RMSE, when a height variable is predicted), and the

coefficient of determination (r^2). Given a large number of cells at the 5x5 m to 25x25 m resolutions, only a systematic sample was used. Values were taken every 100 m in both X and Y directions, resulting in about 2000 samples. At the stand level, all forest inventory polygons were used.

2.5.5 Relationships between local incidence angle, TanDEM-X coherence, and ΔH
The local incidence angle (LIA) is defined as the angle that is formed by the negative of the SAR line-of-sight vector and the local surface normal (Krieger et al., 2010). LIA is known to influence InSAR coherence because of variation in decorrelation that is due to angular effects (Balzter et al., 2007.b). For this reason, we studied the relationships between TanDEM-X coherence, ΔH , and LIA. These values were extracted from the same systematic sample of 5 m pixels that was used in the previous general accuracy experiments (sampling interval of 100 m in X and Y). We report r^2 and P -values of the relationships between LIA and coherence, coherence and ΔH , and the relationship between LIA and ΔH .

2.5.6 Effects of forest structure on coherence and dominant ΔH
Forest structure is known to influence SAR backscattering and coherence (Izzawati et al., 2006; Balzter et al., 2007.b). To explore these effects in the case of the TanDEM-X products, we regressed four structural attributes, i.e., density (stems per ha), dominant height (DH), canopy surface height (CSH), and gap volume against InSAR coherence and dominant ΔH (Lidar DH – InSAR CSH) for the sampled pixels corresponding to the field plot locations. Density was obtained directly from the field measurements, while canopy heights were extracted from the lidar data. Gap volume was defined as the volume of empty space between a horizontal plane situated at a dominant height and the original (0.25 m) lidar canopy surface. We calculated r^2 and P -values for the relationships between the four forest structure parameters with coherence and dominant ΔH .

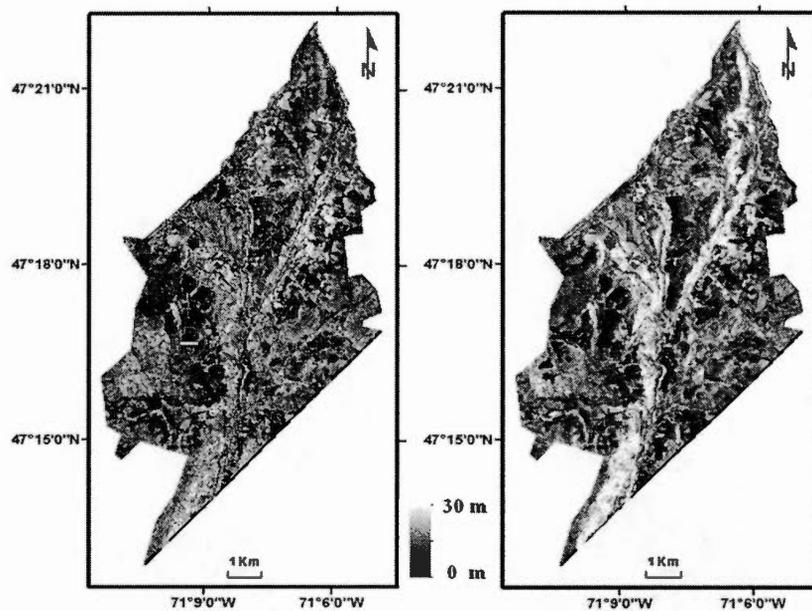
2.6 Results

The respective coherence histograms for the HH and VV polarizations were very similar and had the same average coherence of 0.68. Because the global TanDEM-X acquisitions are performed in HH polarization (Gruber et al., 2015), all experiments were conducted using only the HH interferogram.

Figure 2.3 presents the InSAR and lidar 5 m resolution CHM maps, together with difference and coherence maps. General patterns of canopy height are very similar to all height maps, but the lidar CHM was visually the sharpest. For example, narrow features such as forest roads are clearly visible on the lidar CHM, but these are sometimes almost invisible in the corresponding InSAR CHM. Some areas of the CHMs are markedly different, corresponding to clear-cuts that occurred between the respective acquisitions of lidar (2011) and InSAR (2013) data, to areas of very low coherence such as lakes, or to suspended power cables. Other problematic areas corresponded to layover or to phase unwrapping errors. All of these areas were manually masked out (see masks in Figure 2.3f). Apart from these features, the difference map showed that discrepancies between lidar and InSAR CHMs were generally small.

Figure 2.4a presents the respective profiles of the InSAR and lidar DSMs at 0.25 m and 5 m resolutions, together with the lidar DTM that was taken along a 470 m transect (Figure 2.3a). The transect shows InSAR CHM captures general trends while missing the apices of dominant conifer trees and small (< 10 m diameter) gaps. In the former case, lower heights, while in the latter case, higher heights result from gap filling. Elevations of InSAR over bare areas were very close to their corresponding lidar values. The TanDEM-X surface is much smoother overall than that produced by lidar. Figure 2.4b shows that coherence is lower in the presence of forests, but the generally inverse relationship between height and coherence is somewhat complex (for example,

see the local coherence peak that corresponds to a high canopy, and which is located at 240 m in Figure 2.4b).



(a)

(b)



(c)



(d)

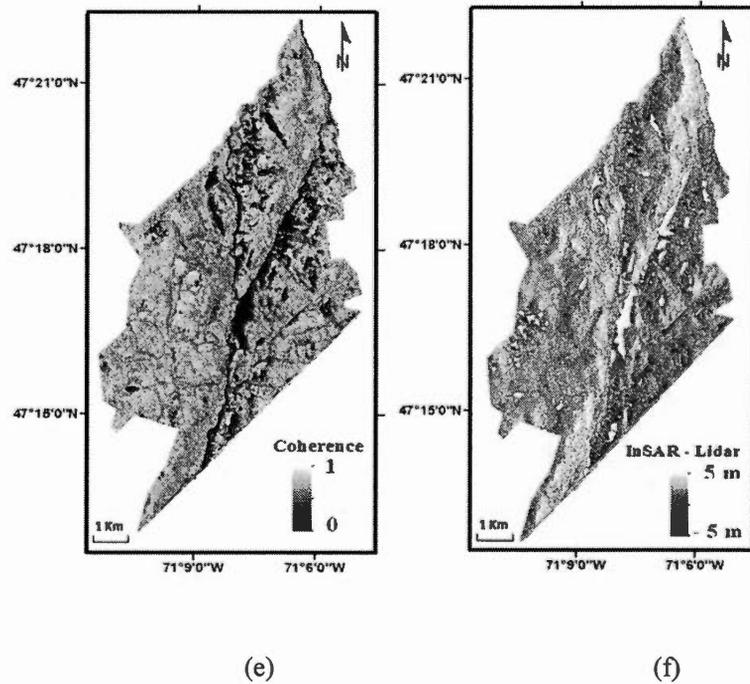


Figure 2.3 a) lidar CHM (unmasked), b) InSAR CHM (unmasked), c) and d) close-up view of the rectangular portion of CHMs shown in a and b, e) coherence map and f) map of the differences between the CHMs shown in a and b (with mask polygon outlines). The transect for which the profile that is shown in figure 2.4 was extracted appears in a within the red ellipse

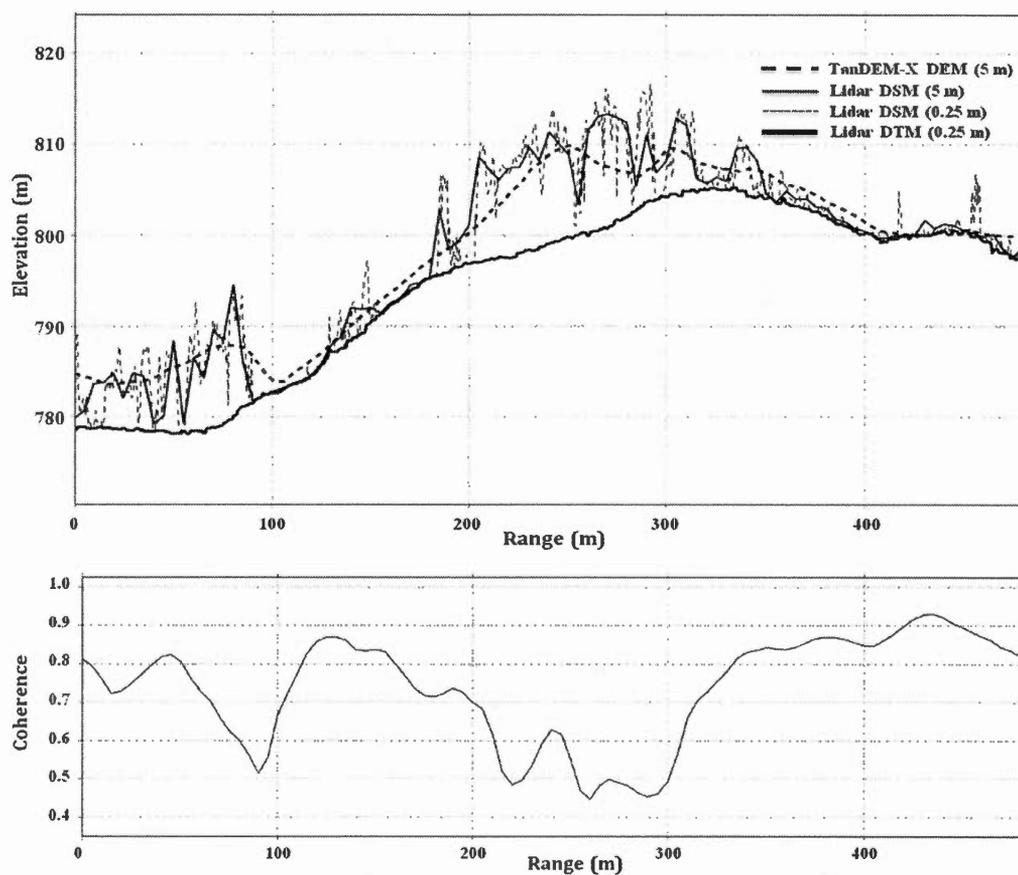


Figure 2.4 a) sample profiles of the InSAR DSM, the InSAR CHM, respectively at 0.25 m and 5 m, the lidar DTM, and b) coherence. The profiles were extracted along a transect that is shown in figure 2.3a.

To determine the error of lidar-measured individual tree heights relative to field-measured tree heights, we performed ordinary least-squares regression between lidar and field heights for each species (Table 2.2). The r^2 ranges from 0.92 for black spruce to 0.99 for trembling aspen, while the corresponding RMSE ranges from 1.31 m for balsam fir to 0.31 m for trembling aspen. Bias varies by one order of magnitude, being

much higher for the conifer species (-0.54 to -1.12 m) than for hardwood species (-0.42 in the case of paper birch, -0.10 for trembling aspen).

Table 2.2 Relationships between lidar and field-measured individual tree heights

| Tree type | <i>n</i> | Mean tree height | | Bias (m) | $H_{\text{Lidar}} = B_0 + B_1 H_{\text{Field}}$ | | r^2 | RMSE (m) |
|-----------------|----------|------------------|-----------|----------|---|-------|-------|----------|
| | | Lidar (m) | Field (m) | | B_0 | B_1 | | |
| Balsam fir | 326 | 13.78 | 14.89 | -1.12 | -1.05 | 0.99 | 0.93 | 1.31 |
| White spruce | 62 | 16.06 | 16.60 | -0.54 | -0.15 | 0.98 | 0.94 | 1.28 |
| Black spruce | 30 | 12.46 | 13.16 | -0.70 | -1.26 | 1.04 | 0.92 | 0.93 |
| Paper birch | 10 | 9.44 | 9.86 | -0.42 | 0.52 | 0.90 | 0.96 | 0.33 |
| Trembling aspen | 3 | 20.30 | 20.40 | -0.10 | -6.56 | 1.32 | 0.99 | 0.31 |
| All | 431 | 13.96 | 14.94 | -0.98 | -0.929 | 0.996 | 0.93 | 1.29 |

All the models are significant at $P < 0.001$

When all species were considered together, the resulting regression had a $r^2 = 0.93$ ($P < 0.001$), a bias of -0.98 m, and a RMSE of 1.29 m. Thus, errors were small. The slope of the regression (0.996) was close to the 1:1 line, but there is a significant offset of 0.929 m (Figure 2.5). Thus, we have used the regression equation in figure 2.5 for correcting lidar heights of individual trees that were used in the experiments using the lidar dataset as a reference for dominant heights.

In the first experiment, we directly compared the InSAR CSH to its lidar equivalent at cell sizes ranging from 5 m to 25 m, and at the stand level (Figure 2.6, Table 2.3). The

overall mean height of the canopy surface was rather low (5.38 m for lidar, 6.70 m for InSAR at the stand level), reflecting the fact that many gaps and some non-forested patches are present in the canopy surface model. The positive ΔH (around 1.4 m), indicates that the reference canopy surface was generally lower than the corresponding InSAR surface. However, the r^2 values of regressions between the respective CSHs increased with cell size, from 0.45 at 5 m resolution to 0.63 at 25 m, and reached 0.67 at the stand level. The greatest increase occurred between 5 m and 10 m resolutions. All regressions were highly significant ($P < 0.001$). The regression slope approached the 1:1 relationship as cell size increased, reaching a value of about 1.0 at 25 m resolution. RMSE between InSAR CSH and the predicted lidar CSH decreased from 2.77 m at 5 m resolution to 1.89 m at 25 m resolution and fell to 1.53 m at the stand level. The average difference between the two surfaces (ΔH) did not change substantially with resolution (1.32 to 1.48 m).

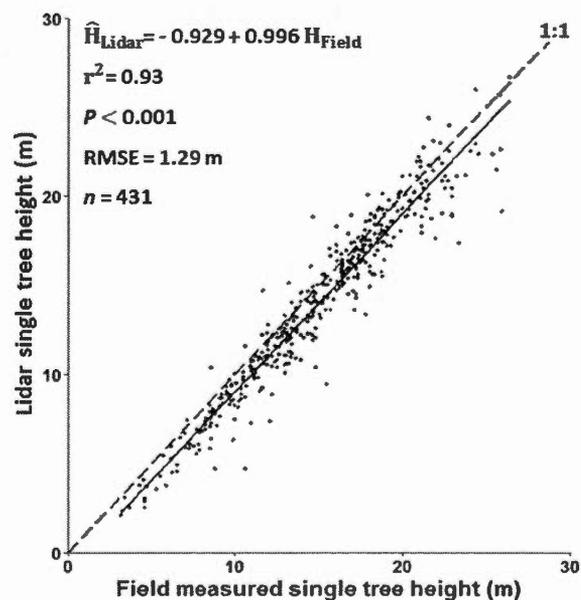
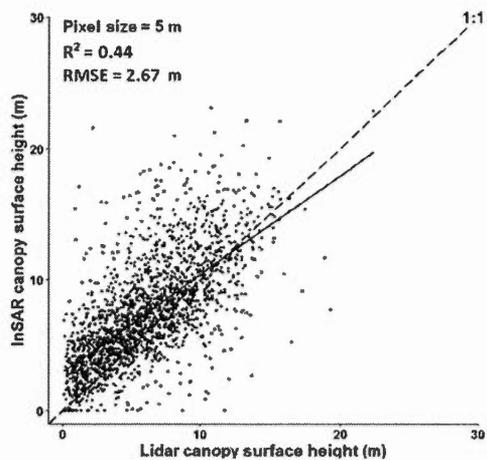
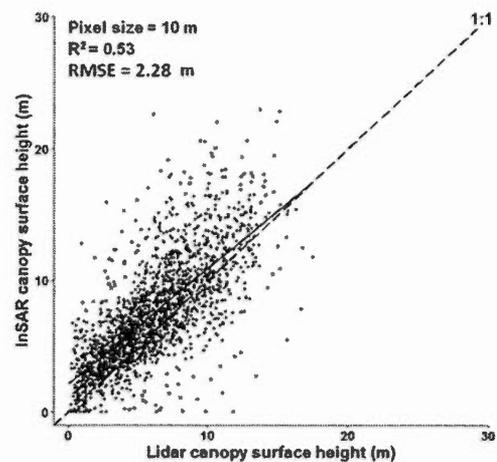


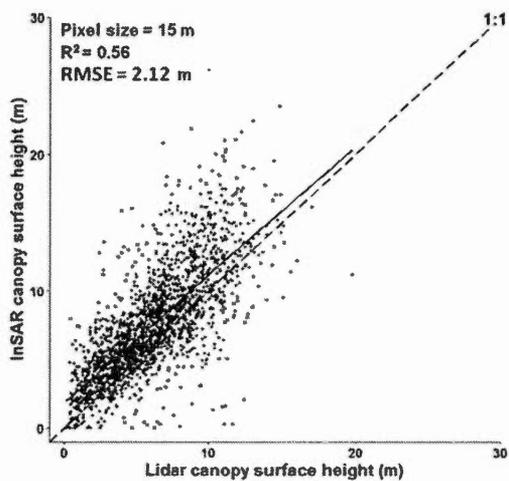
Figure 2.5 lidar vs. field-measured individual tree heights



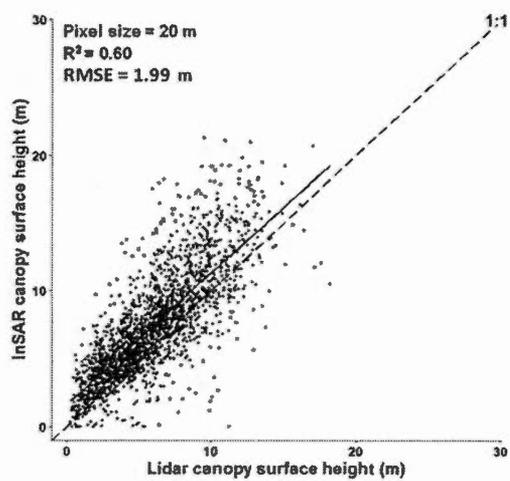
(a)



(b)



(c)



(d)

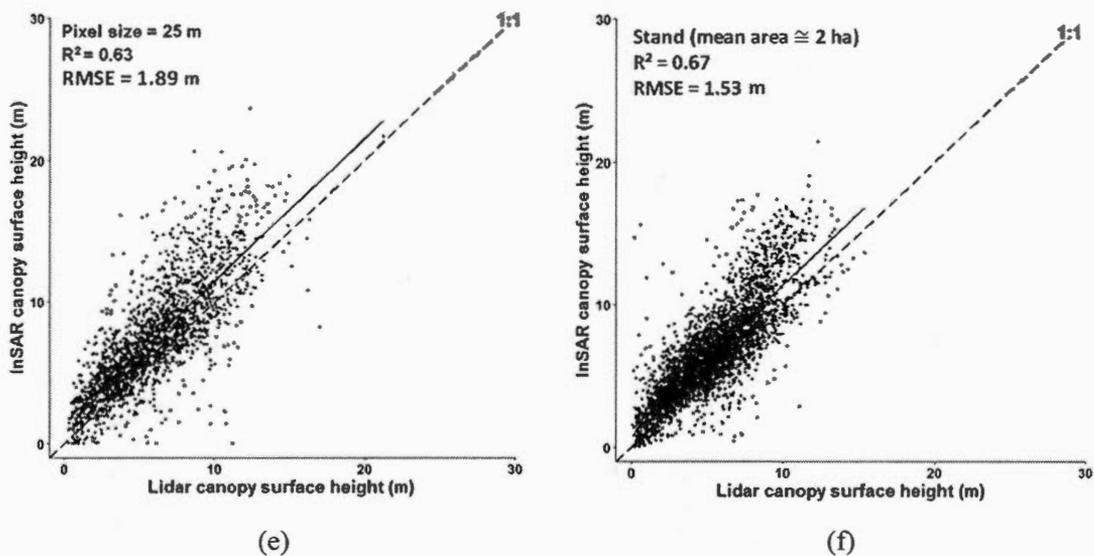


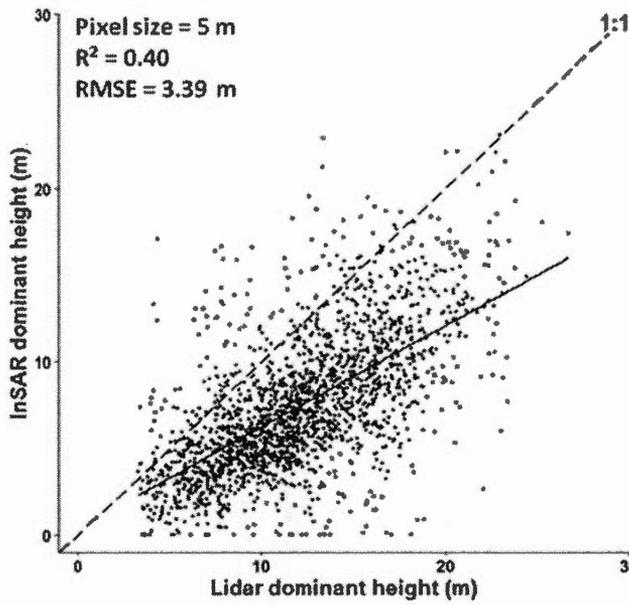
Figure 2.6 Relationships between InSAR and lidar canopy surface height (CSH) as a function of resolution (pixel size from 5 to 25 m), and at the stand level. Across all resolutions, there were 1933 observations and 3340 at the stand level. Table 2.3 provides detailed statistics for the regressions.

Table 2.3 InSAR CSH relationship with the lidar CSH

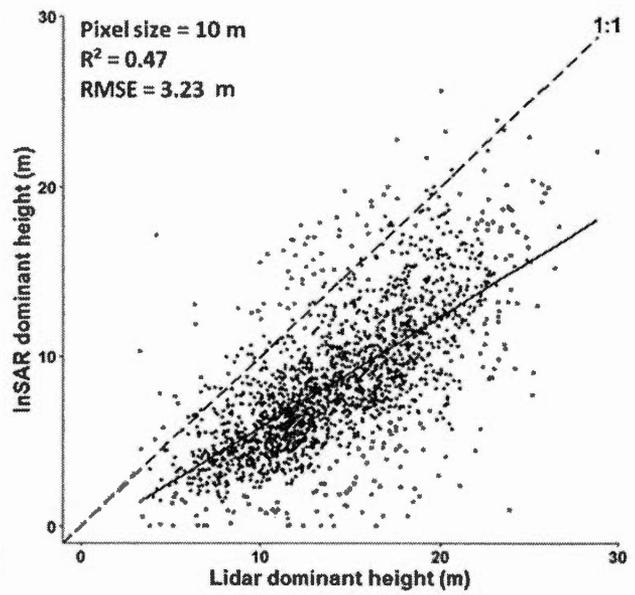
| Resolution (m) | Mean CSH | | ΔH (m) | $CSH_{InSAR} = B_0 + B_1 CSH_{Lidar}$ | | r^2 | RMSE (m) | n |
|-------------------|--------------|--------------|-------------------|---------------------------------------|-------|-------|-------------|------|
| | Lidar (m) | InSAR (m) | | B_0 | B_1 | | | |
| 5 | 6.08 | 7.49 | 1.41 | 2.94 | 0.75 | 0.44 | 2.67 | 1933 |
| 10 | 6.11 | 7.50 | 1.39 | 2.13 | 0.88 | 0.53 | 2.28 | 1933 |
| 15 | 6.04 | 7.48 | 1.44 | 1.83 | 0.93 | 0.56 | 2.12 | 1933 |
| 20 | 6.04 | 7.48 | 1.44 | 1.64 | 0.97 | 0.60 | 1.99 | 1933 |
| 25 | 6.00 | 7.48 | 1.48 | 1.46 | 1 | 0.63 | 1.89 | 1933 |
| Stand | 5.38 | 6.70 | 1.32 | 1.22 | 1.02 | 0.67 | 1.53 | 3340 |

All the models are significant at $P < 0.001$

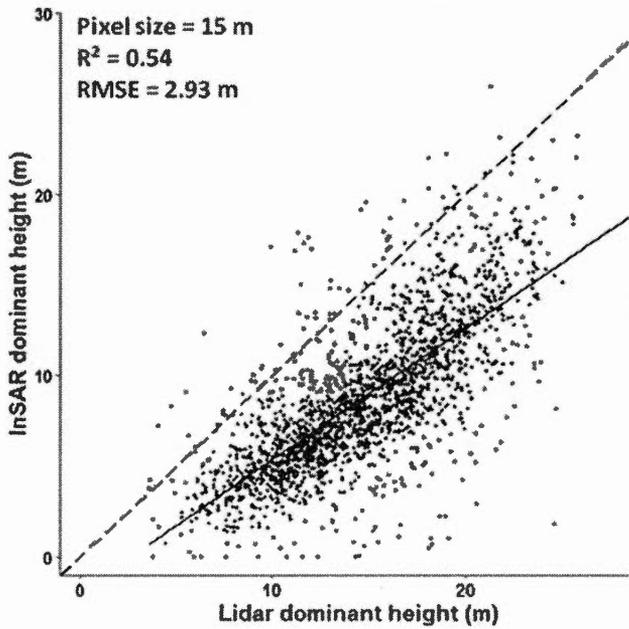
The second experiment compared maximum InSAR heights within a cell to the local DHs (based on H_{100}) that were predicted from the lidar data at various resolutions. InSAR local maxima are located well below the corresponding DHs (Figure 2.7, Table 2.4). For example, the average lidar DH was 13.4 m at the stand level, while the average InSAR DH was only of 7.7 m, which was a discrepancy that was much larger than that observed in surface-to-surface comparisons of the first experiment. RMSE between InSAR DH and the dominant height prediction decreased from 3.4 m at 5 m resolution to 2.7 m at 25 m resolution, and fell to 2.1 m at the stand level, thereby showing an inverse tendency compared to results of the first experiment. The r^2 varied from 0.40 to 0.63, which was slightly lower than those estimates from the surface-to-surface comparisons. All relationships were significant at $P < 0.001$.



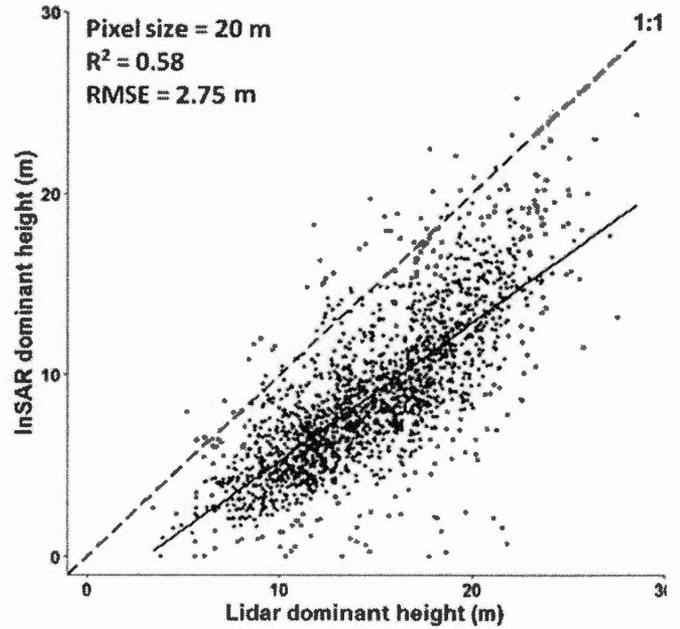
(a)



(b)



(c)



(d)

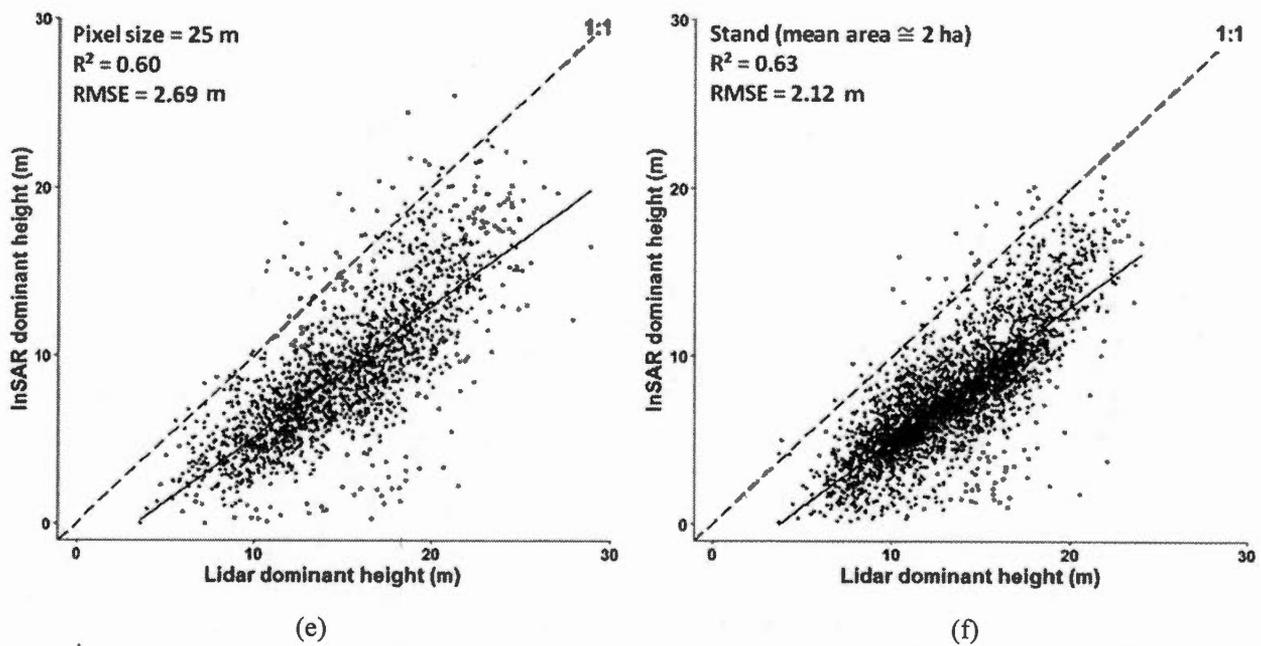
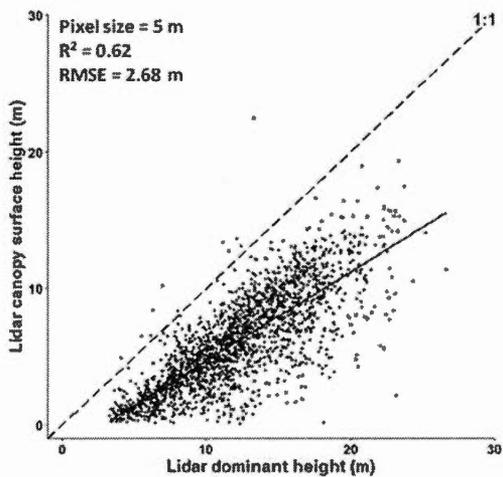
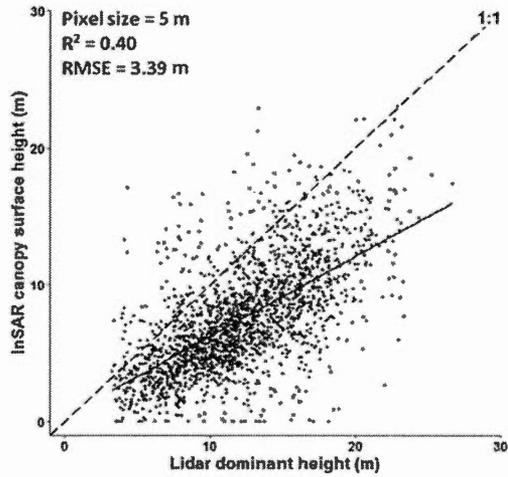


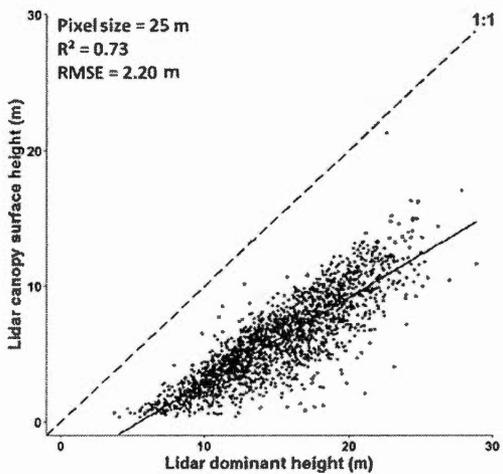
Figure 2.7 Relationships between InSAR DH and lidar DH as a function of resolution (5 m to 25 m), and at the stand level. RMSE is expressed in m. Table 2.4 provides details of each regression



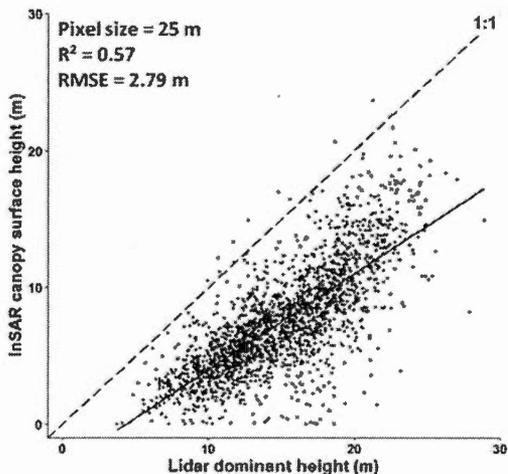
(a)



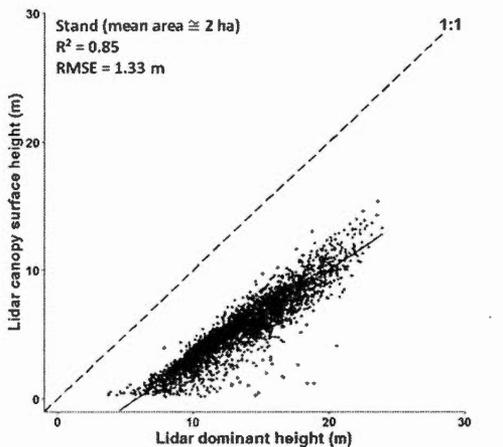
(b)



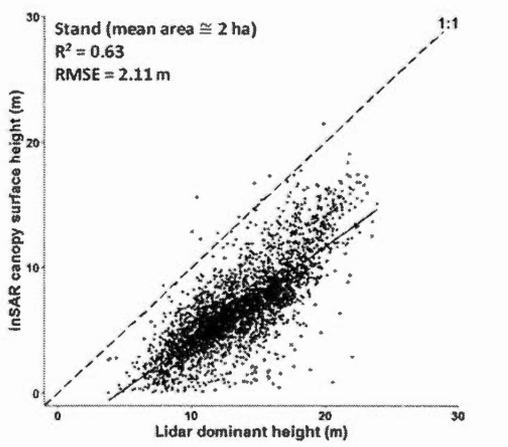
(c)



(d)



(e)



(f)

Figure 2.8 Relationships between InSAR CSH and lidar DH, compared to the relationships of lidar CSH with lidar DH as a function of resolution (pixel sizes 5 m and 25 m), and at the stand level. RMSE is expressed in m. Table 2.5 provides the details of each regression

Table 2.4 InSAR DH relationship with the lidar DH

| Resolution (m) | Mean DH | | ΔH (m) | $DH_{InSAR} = B_0 + B_1 DH_{Lidar}$ | | r^2 | RMSE (m) | n |
|-------------------|--------------|--------------|-------------------|-------------------------------------|-------|-------|-------------|------|
| | Lidar (m) | InSAR (m) | | B_0 | B_1 | | | |
| 5 | 12.12 | 7.49 | -4.63 | 0.43 | 0.58 | 0.40 | 3.39 | 1933 |
| 10 | 14.18 | 8.53 | -5.65 | -0.71 | 0.65 | 0.47 | 3.23 | 1933 |
| 15 | 14.67 | 8.75 | -5.92 | -1.84 | 0.72 | 0.54 | 2.93 | 1933 |
| 20 | 14.76 | 8.89 | -5.87 | -2.36 | 0.76 | 0.58 | 2.75 | 1933 |
| 25 | 14.98 | 9.03 | -5.95 | -2.56 | 0.77 | 0.60 | 2.69 | 1933 |
| Stand | 13.46 | 7.73 | -5.73 | -3.03 | 0.80 | 0.63 | 2.12 | 3340 |

All the models are significant at $P < 0.001$

The third experiment showed the level to which the average lidar DH was higher than average CSHs, whether given by InSAR or lidar. This difference is very apparent in figure 2.8 and Table 2.5. For example, the height of the lidar surface, on average, lies 9.0 m below the dominant height, while in the case of the InSAR is 7.5 m at a resolution of 25 m. At all resolutions, the r^2 of the lidar CSH versus lidar DH relationship was significantly higher than that of the corresponding InSAR regression. It should be noted that the corresponding r^2 of the lidar-based relationship was much less than 1.00 (i.e., 0.62 to 0.85), which underscores the effect that is imposed by canopy surface generalization, i.e., averaging heights at a coarser resolution. The RMSE of the InSAR

CSH versus lidar DH relationship was only 0.6 m to 0.8 m higher than that associated with the lidar CSH versus lidar DH relationship (Table 2.5).

Table 2.5 InSAR CSH and lidar CSH respective relationships to DH

| Resolution (m) | CSH-DH (m) | Equation | r^2 | RMSE (m) | n |
|-------------------|---------------|---|-------|-------------|------|
| 5 | -6.04 | $CSH_{Lidar} = -1.78 + 0.65 DH_{Lidar}$ | 0.62 | 2.68 | 1933 |
| | -4.63 | $CSH_{InSAR} = 0.43 + 0.58 DH_{Lidar}$ | 0.40 | 3.39 | 1933 |
| 25 | -8.98 | $CSH_{Lidar} = -3.38 + 0.62 DH_{Lidar}$ | 0.73 | 2.20 | 1933 |
| | -7.50 | $CSH_{InSAR} = -2.96 + 0.70 DH_{Lidar}$ | 0.57 | 2.79 | 1933 |
| Stand | -8.05 | $CSH_{Lidar} = -4.08 + 0.70 DH_{Lidar}$ | 0.85 | 1.33 | 3340 |
| | -6.74 | $CSH_{InSAR} = -3.43 + 0.75 DH_{Lidar}$ | 0.63 | 2.11 | 3340 |

All the models are significant at $P < 0.001$

Figure 2.9 and Table 2.6 present the relationships between LIA, coherence, and ΔH . These values were extracted from the same systematic sample of 5 m pixels as used in the previous experiments with a sampling interval of 100 m in X- and Y-dimensions. LIA values ranged from -14° to 60° , with the LIA of horizontal surfaces corresponding to SAR incidence angles of about 30° . LIA values lower than 30° are obtained for surfaces that were tilted toward the sensor, and vice versa. The relationship between LIA and coherence is very weak; maximum coherence generally occurs on horizontal

areas (with an LIA of about 30°, Figure 2.9). The relationship between LIA and ΔH was stronger ($r^2 = 0.31$, Figure 2.9b) than that between coherence and ΔH ($r^2 = 0.01$), which was not significant (Figure 2.9c). In general, the InSAR CSH was lower than lidar for pixels that were facing towards the sensor and higher for areas that were tilted away from the sensor. This is further evidenced by the light and dark patterns in figure 2.3f showing elevation differences between lidar and InSAR surfaces.

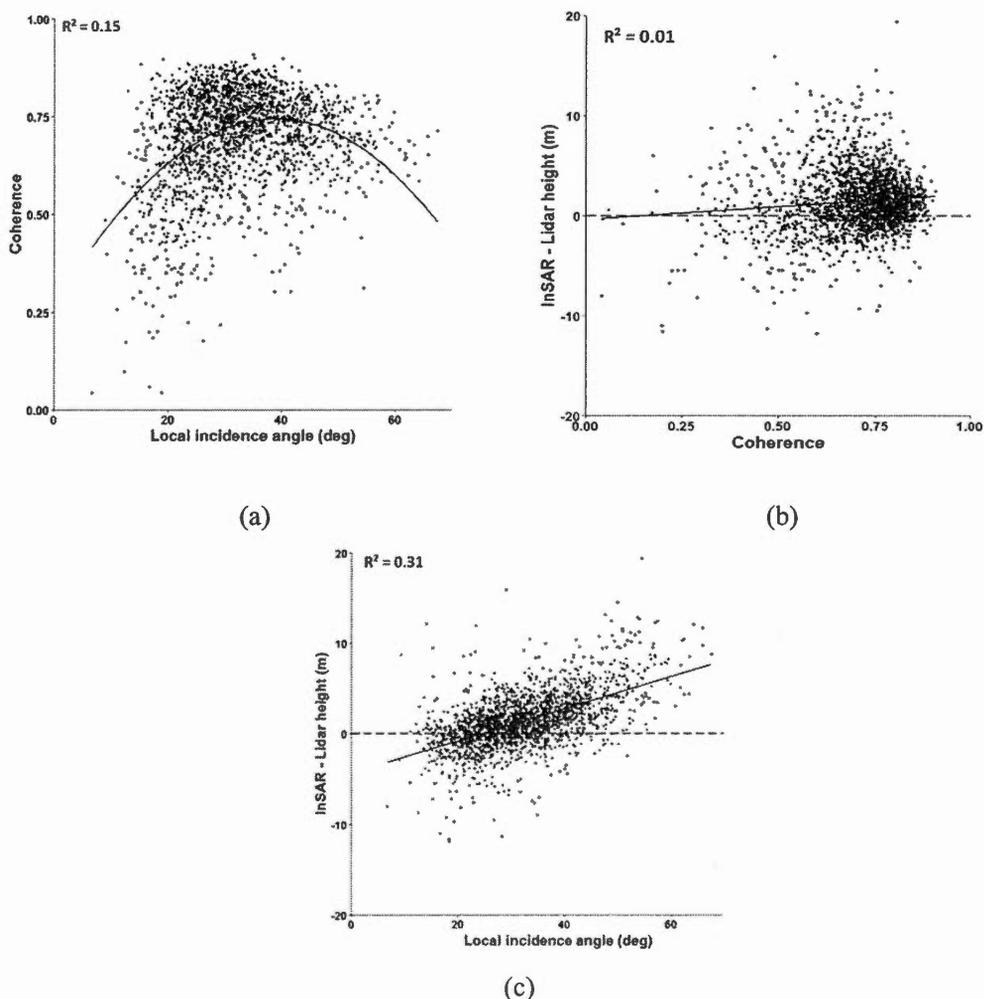
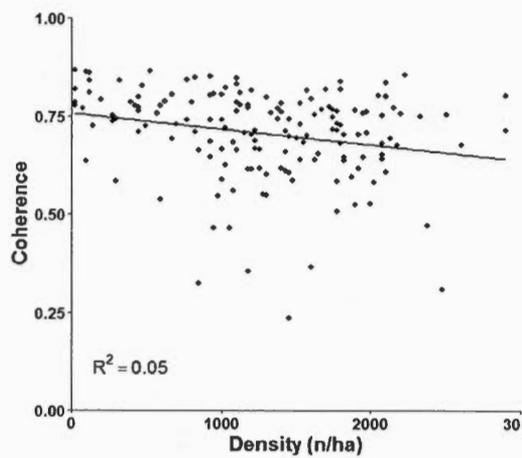


Figure 2.9 Relationship between a) coherence and local incidence angle, b) Δh and coherence, and c) Δh and local incidence angle (see Table 2.6 for additional details)

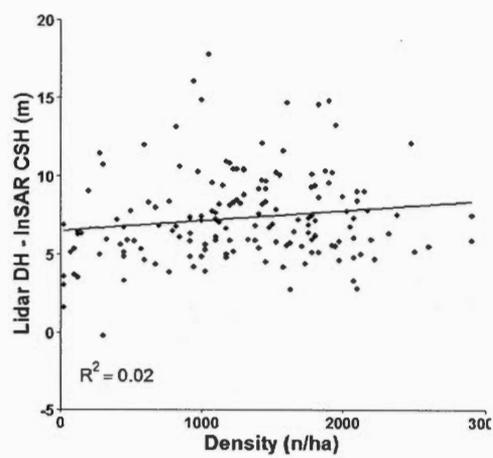
Table 2.6 Relationships between LIA, coherence, and ΔH

| Equation | r^2 | n |
|--|-------|------|
| Coherence = $2.63e-01 + 2.48e-02 \text{ LIA} - 3.20e-04 \text{ LIA}^2$ | 0.15 | 1933 |
| $\Delta H = -0.38 + 2.56 \text{ Coherence}$ | 0.01 | 1933 |
| $\Delta H = -4.3 + 0.18 \text{ LIA}$ | 0.31 | 1933 |
| All the models are significant at $P < 0.001$ | | |

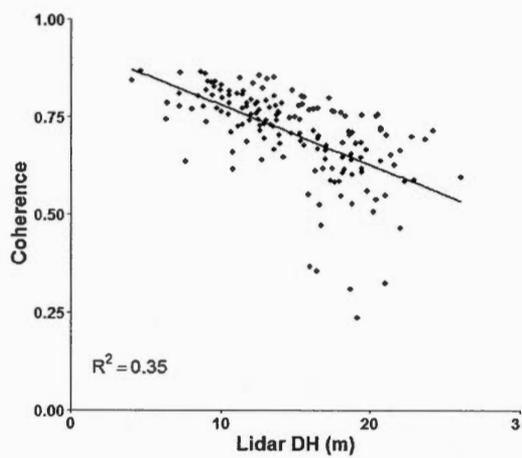
Last, we present relationships between forest parameters (density, DH, CSH, gap volume), and the coherence or dominant ΔH (Figure 2.10, Table 2.7). These relationships were estimated for 157 field plots. Density had a very weak effect on coherence or dominant ΔH ($r^2 = 0.05$ and 0.02 respectively, $P < 0.001$; Figure 2.10, Table 2.7). In contrast, dominant height had significant effect, with a $r^2 = 0.35$ for coherence, and $r^2 = 0.30$ for dominant ΔH . As for height increases, coherence decreases; underestimation of dominant height not only increases, but the variability of the estimates also increases (Figure 2.10 c, d). A r^2 of 0.24 between lidar CSH and coherence, and 0.06 between lidar CSH and dominant ΔH were observed (Figure 2.10 e, f). A relatively strong relationship between gap volume and dominant ΔH was found (Figure 2.10g), with a r^2 of 0.53, gap volume had much less influence (r^2 of 0.21) on coherence (Figure 2.10h).



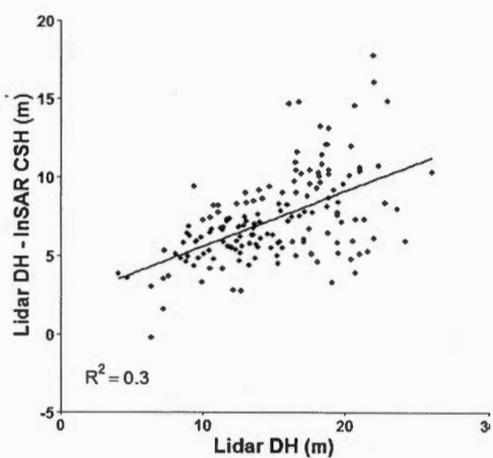
(a)



(b)



(c)



(d)

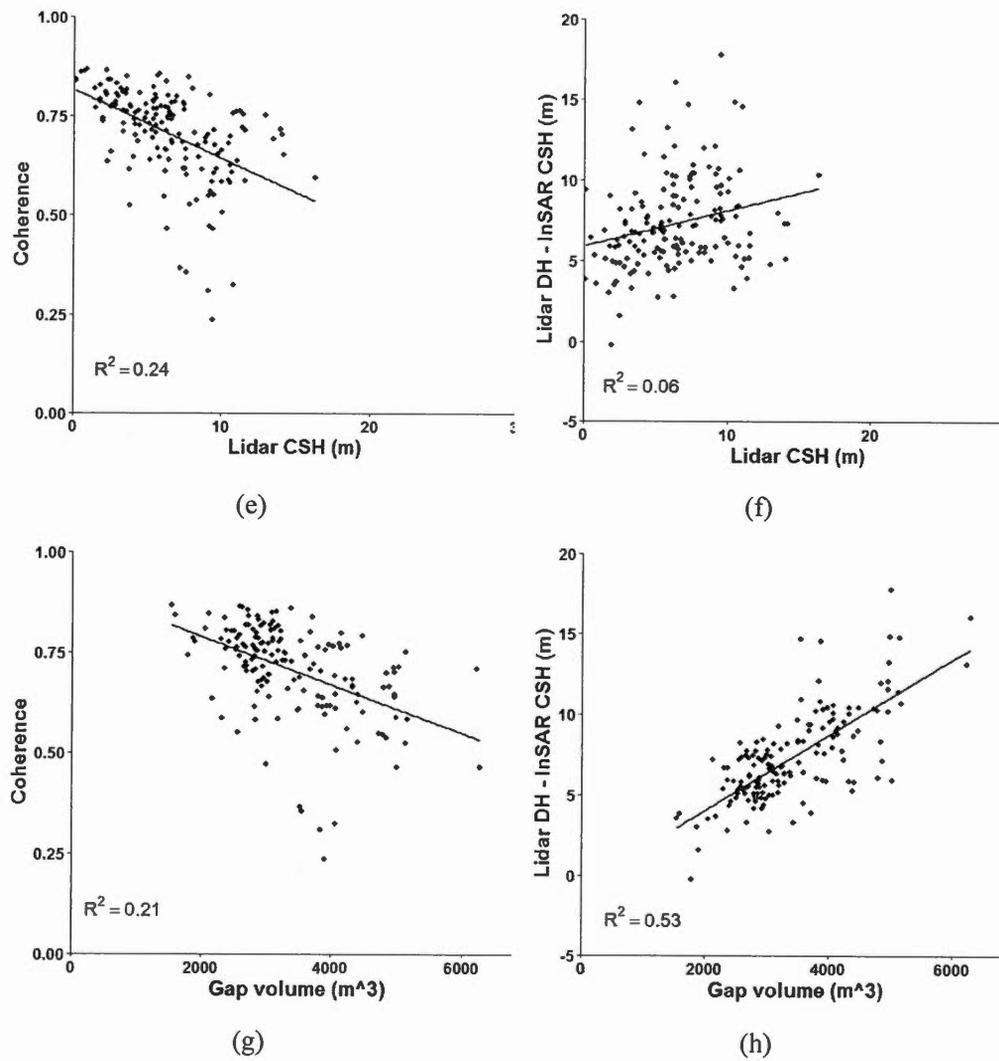


Figure 2.10 Relationships of stem density (a & b), dominant height (DH) (c & d), canopy surface height (CSH) (e & f), and gap volume (g & h) with coherence, and dominant ΔH

Table 2.7 Relationship between stem density, DH, CSH, gap volume, coherence, and dominant ΔH

| Equation | r^2 | n |
|--|-------|-----|
| Coherence = $7.57e-01 - 4.0e-05$ Density | 0.05 | 157 |
| Dominant ΔH = $6.48 + 6.4e-04$ Density | 0.02 | 157 |
| Coherence = $0.93 - 0.01$ Lidar DH | 0.35 | 157 |
| Dominant ΔH = $2.05 + 0.35$ Lidar DH | 0.30 | 157 |
| Coherence = $0.82 - 0.02$ Lidar CSH | 0.24 | 157 |
| Dominant ΔH = $5.9 + 0.22$ Lidar CSH | 0.06 | 157 |
| Coherence = $9.1e-01 - 6.0e-05$ Gap Volume | 0.21 | 157 |
| Dominant ΔH = $-0.65 + 0.002$ Gap Volume | 0.53 | 157 |
| All the models are significant at $P < 0.001$ | | |

2.7 Discussion

InSAR CHMs that were created by combining a TanDEM-X DSM and lidar DTM were compared with reference data, i.e., lidar CSH and lidar DH. Given that lidar was the source of the reference data, demonstrating its inherent quality was critically important. We accomplished this goal by comparing lidar height measurements of 431 trees with their corresponding field-measured heights. With respect to the most common species encountered in the area (326 balsam fir out of 431 trees), the lidar measurement was 1.12 m lower than the field value ($r^2 = 0.93$). This discrepancy can be attributed to the extreme narrowness of the crown that characterizes the apex (the extremity of the crown) of this species. Honer and Mitchell (Gilmore and Seymour 1997) modelled crown radius as a function of the vertical distance from the top of the tree. For co-dominant balsam fir, a point lying horizontally 0.19 m from the top of a balsam fir tree

(i.e., the average distance between a lidar first return and the tree apex; see the aforementioned *data* section) would be 50 cm lower than the tree height. The effect of tree shape on underestimation is further evidenced by the small values that were obtained for species with wide tree apices (e.g., -0.54 m, white spruce; -0.42 m, paper birch). The unexplained portion of the 1.12 m height lower for balsam fir (about 0.6 m) could have been caused by a time delay that was incurred during the triggering of the first return by the sensor electronics (Hopkinson 2007). In this specific case, the delay would be equivalent to 2 ns. Considering that the surface area of the plant material that is intercepting incident laser energy near the balsam fir apex is extremely small, it is highly likely that the quantity of energy that was returned to the sensor is quite small. This response is likely not as pronounced when the sides of trees are hit, or when less elongated tree species are involved. For this reason, and also considering the high sensitivity of the lidar sensor, we conclude that the first lidar return penetration in the canopy is very low and the lidar surface can be subsequently used as a reliable comparison basis for TanDEM-X DSMs. Because our estimates of dominant heights are bias-corrected, lidar estimates of dominant heights were very accurate.

In comparing InSAR and lidar CHMs, we noted that general spatial patterns of forest height were well represented by the TanDEM-X data. However, the InSAR CHM was smooth compared to the lidar version, i.e., the former has a markedly coarser spatial resolution. The TanDEM-X surface was very close to the 5 m resolution lidar surface forest in closed canopies. The ΔH between InSAR and lidar CHMs were small (between 1.32 m and 1.48 m, depending on the resolution), with the InSAR surface being slightly higher, on average. Sharp tree apices were smoothed out (both in the InSAR and 5 m lidar CHMs); however, small gaps were observed with lidar at 5 m resolution but appeared to be filled in the InSAR CHM. This is well illustrated by the profiles that are depicted in figure 2.4. At 5 m resolution, the slope of the relationship between lidar and InSAR (Figure 2.6a), also indicated that when the canopy was very low (e.g., in small gaps that are not resolved by InSAR), the InSAR CHM was higher

than the lidar CHM. Conversely, when it is high (i.e., tree tops), the InSAR CHM was lower than the lidar CHM. It should be noted that at larger cell sizes (e.g. 25 m, as in Figure 2.6e), the slope was very close to 1.0, and bias was very low. At this resolution, small gaps were also filled in the generalized lidar surface model. These observations make it clear that resolution effects on the representation of small troughs (gaps) and sharp protuberances (tree apices) dominate the relationship between InSAR and lidar canopy surfaces.

In the case of dominant height, the situation is quite different but leads us to the same conclusions as those reached in the canopy surface experiments. InSAR heights were consistently much lower than dominant heights (-4.63 m to -5.95 m). Again, this indicates that the tips of (conical) trees are not well represented in the InSAR CHM. When comparing these InSAR results with those of the deviations from the smoothed lidar surface relative to dominant height, the decrease in resolution explains most of the discrepancy between canopy surface height and dominant height. In this regard, the behavior of InSAR and the lidar surface relative to the dominant height was quite similar at the stand level. Dominant heights were higher by 8.05 m compared to the lidar surface stand average and by 6.74 m compared to the InSAR surface. These values are in the same range as those reported by other studies (Soja et al., 2014; Askne et al., 2013; Kugler et al., 2014; Solberg et al., 2013). Our findings were also consistent with other X-band InSAR studies, including those of (Neeff et al., 2005) with airborne data and (Solberg et al., 2010) with SRTM. These authors also indicated that InSAR and lidar models agreed when comparing their respective canopy surfaces. In contrast, X-band InSAR exhibited significant "penetration" with respect to dominant height.

Furthermore, factors such as local incidence angle, coherence, forest height, stem density, and gap volume have been studied to determine their effects on InSAR height retrieval. Our results show that incidence angle, forest height, and gap volume are the most important factors. InSAR canopy heights were lower for surfaces that were

oriented toward the sensor and higher for those that were facing away from the sensor. This response was consistent with results that were presented in (Solberg et al., 2010). This study used local incidence angle (LIA) and InSAR height as predictors of height, thereby indicating the importance of LIA. What is responsible for this effect is likely the tilting of the surface towards the sensor. The X-band signal is better able to propagate into gaps (fewer shadowed areas), which results in overall lower height. In contrast, more oblique pulses on surfaces that are oriented away from the sensor create less favorable propagation conditions. A higher canopy generally induces a coherence decrease, as has been shown in previous studies (Treuhaft et al., 2015; Cloude 2010; Krieger et al., 2005), and leads to greater errors (greater scattering) and more apparent penetration. We found the relationship between height and coherence to be non-linear. Although this relationship was not the main focus of our study, we hypothesize that coherence sometimes rises in very dense forests, and is lower in the presence of relatively small gaps, in such a way that a consistent inverse proportionality between height and coherence does not emerge. Furthermore, coherence might be partly influenced by SNR fluctuations caused by topography or low backscattering in addition to height-related volume decorrelation. However, the height-coherence relationship was calculated over forested plots located on the relatively flat ground so these external effects should not play a strong role. Again, propagation is likely attributable to the greater volume of empty spaces between the horizontal planes that represent the dominant height and the canopy surface. Incident microwaves, therefore, can travel in an unobstructed fashion for a greater distance between the level of the tallest trees and the general foliage surface.

2.8 Conclusions

InSAR CHM was generally similar to the corresponding lidar CHM, and this similarity increased as the spatial resolution of the lidar CHM was artificially decreased. RMSE for CSHs drops from 2.67 m at 5 m resolution to 1.53 m at the stand level, while

attempting to estimate DH, RMSE decreased from 3.4 m at 5 m resolution to 2.1 m at the stand level. Our results indicated that the TanDEM-X surface that was retrieved by the InSAR technique was similar to a reference lidar DSM and that penetration beneath the foliage surface was very limited. The amount of difference between the TanDEM-X canopy surface height and (lidar based-) dominant heights were found to be very similar to that of the smoothed lidar canopy surface and dominant height. This shows that “penetration” that has been reported in other studies was largely caused by effects of the lower spatial resolution of TanDEM-X relative to the reference lidar data used for dominant height, and the presence of canopy gaps. Apparent penetration beneath the level of dominant height results from unobstructed propagation of microwaves between tree apices, and possibly through smaller openings (e.g. gaps between branches), which are not resolved before they reach the primary foliage surface. In addition to the resolution effect, the side-looking geometry of SAR generally did not favor penetration within small gaps. In turn, this caused the InSAR surface height to be locally higher than the lidar height, particularly when the terrain slope faces away from the sensor. Thus, the local angle of incidence affects the apparent degree of penetration of the retrieved InSAR SPC. While it was affected by height, InSAR coherence by itself did not strongly influence InSAR height. However, these height differences increased significantly with the canopy height.

Our findings show that single polarization InSAR TanDEM-X data have a very great potential for mapping forest heights over wide areas, provided that a high-quality DTM is available. In areas with strong topographic relief, layover, shadows, and phase unwrapping problems preclude height estimation. This problem could be alleviated by using more than one interferometric pair with different view angles, for example, by combining data from ascending and descending orbits (Solberg et al., 2013) or using dual baselines (Lachaise et al., 2012). Obtaining a reliable DTM represents a greater challenge, as lidar DTMs are currently the only type that offers sufficient accuracy but cover only a small portion of the Earth’s surface. The future TanDEM-L mission could

possibly generate worldwide DTMs with sufficient accuracy because of greater canopy penetration in the L-band (Krieger et al., 2009). Moreover, lidar coverage in certain areas or the world is rapidly increasing, and state-wide coverage already exists in Europe, in some provinces of Canada, and in states of the USA. If it is taken for granted that a near-worldwide DTM will certainly exist in the near future, and because the topography of terrain does not change rapidly over time, a global CHM could be created retrospectively with previous global TanDEM-X acquisitions (circa 2013). This global CHM would subsequently enable precise monitoring of changes in forest height, together with likely changes in biomass and carbon stocks.

Many questions remain to be addressed, such as the influence of vegetation phenology on the accuracy of height estimates at the time of TanDEM-X acquisition. Furthermore, the influence of meteorological conditions just prior to the acquisition, such as intense rain-inducing changes in foliage dielectric properties or thick snow covering conifer trees, may complicate the extraction of forest heights from InSAR data. Nevertheless, X-band single-pass image acquisition from space opens up for the possibility of creating the first accurate and spatially continuous global canopy height mode

Foreword to the second research paper

In the first research paper (Chapter II), we have demonstrated that the elevation models extracted from TanDEM-X are clearly surface models, i.e. DSMs. A few months after the publication of this paper, it was announced that the first version of the WorldDEM dataset was completed. WorldDEM is the commercial name under which a worldwide mosaic of TanDEM-X DSMs is sold by Airbus Defence and Space. This product would enable the rapid creation of CHMs inasmuch as an accurate DTM can be obtained without the need to obtain the raw TanDEM-X images and go through the time-consuming and sometimes difficult interferometric processing. Our concern is that because this global elevation model is made from a patchwork of interferograms, it may vary in its accuracy, a situation that could complicate the creation of a global CHM, and biomass map. For this reason, we have studied, in the next paper (Chapter III), the effects of various acquisition conditions on the accuracy of the DSMs. These conditions relate to the configuration of the sensors themselves, such as the interferometric baseline, and the state of the forest (leaf-on vs. leaf-off deciduous trees)

CHAPTER III

EFFECTS OF TANDEM-X ACQUISITION PARAMETERS ON THE ACCURACY OF DIGITAL SURFACE MODELS OF A BOREAL FOREST CANOPY

This chapter has been accepted in *Canadian Journal of Remote Sensing* on 15 Dec 2016

Y. Sadeghi, B. St-Onge, B. Leblon, M. Simard, “Effects of TanDEM-X acquisition parameters on the accuracy of digital surface models of a boreal forest canopy,” *Canadian Journal of Remote Sensing*, 15 December 2016.

3.1 Résumé

L’exactitude des données d’altitude TanDEM-X sur la forêt boréale a été évaluée en utilisant cinq jeux de données TanDEM-X acquises sous différentes conditions géométriques, phénologiques et météorologiques. Cinq MHC InSAR ont été produits par la soustraction d’un MNT lidar sur le MNS TanDEM-X. Ces MHC InSAR ont ensuite été comparés à un MHC lidar (MNS lidar – MNT lidar); le biais se situe entre 0.77 et 1.56 m, le r^2 entre 0.38 et 0.68 et l’EMQ entre 2.1 et 3.67 m. Le pire résultat fut obtenu avec le jeu de données TanDEM-X avec la plus courte ligne de base, 21 m, comparée à une moyenne de 148 m pour les quatre autres jeux de données TanDEM-X. La hauteur de l’ambiguïté (HdA), qui résulte de la combinaison entre la ligne de

base et l'angle d'incidence, a un effet significatif sur les hauteurs calculées. Une valeur HdA de 40 m a été trouvée comme valeur optimale pour le calcul de hauteurs en milieu forestier. Une EMQ de 1.27 m a été calculée entre deux jeux de données TanDEM-X avec des paramètres d'acquisition identiques, ce qui démontre d'importantes fluctuations aléatoires. Pour les MNS TanDEM-X, une grande partie de cette erreur peut être attribuée aux erreurs introduites lors de l'opération du déroulement de phase. Les changements de température et de phénologie sur la forêt boréale ne semblent pas avoir d'effet matériel sur l'estimation des hauteurs comparé aux changements de ligne de base. Un MNS global représentant une surface stable et homogène sur la forêt boréale est possible en construisant une mosaïque TanDEM-X avec les HdA optimales.

3.2 Abstract

The accuracy of digital surface models (DSMs) derived from TanDEM-X interferograms of a dense and mostly evergreen boreal forest area was evaluated across five datasets acquired under various geometrical and phenological conditions. For each, an interferometric synthetic aperture radar (InSAR) canopy height model (CHM) was produced by subtracting a lidar digital terrain model from the TanDEM-X DSM. These InSAR CHMs were compared to a lidar CHM at a resolution of 25 m and led to biases from 0.77 m to 1.56 m, r^2 s from 0.68 to 0.38, and root-mean-square errors (RMSEs) from 2.06 m to 3.67 m. Two datasets acquired in similar conditions differed by 1.27 m (RMSE). Differences in the interferometric baseline had the strongest effect on the DSMs (RMSE of 3.27 m between short and long baseline DSMs). The height of ambiguity, therefore, had a significant effect on the resulting canopy height. The effect of phenological changes on canopy height estimations was lower (RMSE of 2.30 m between leaf-on and leaf-off DSMs) and not highly significant. These results indicate that, despite variations in the acquisition conditions, a continuous TanDEM-X mosaic acquired with proper baselines could produce a reliable estimate of canopy surface elevations of evergreen closed-canopy boreal forests.

3.3 Introduction

Forest canopy height is a key parameter often utilized as a predictor of forest and above-ground biomass (Saatchi et al., 2011.b; Eggleston et al., 2006; Houghton 2005; Lefsky et al., 2005). It can be measured by subtracting bare-earth terrain elevations from the outer canopy surface level to generate a canopy height model (CHM) (St-Onge et al., 2008.a,b; Balzter et al., 2007.a,b; Zagalikis et al., 2005; Treuhaft and Siqueira 2004; Lim et al., 2002). Then, forest biomass can be derived from the CHM using statistical relationships calibrated with field observations (Patenaude et al., 2002; Næsset 2002). A global CHM would provide critical information for accurate biomass and carbon stocks estimation which would in turn help in estimating the global carbon fluxes between the atmosphere and forests. Following this approach, the requirements for producing a global CHM are a worldwide and continuous digital terrain model (DTM) and a digital surface model (DSM). Both the DTM and DSM surfaces should optimally be acquired using satellite sensors, as it is impossible to acquire a complete global dataset using airborne platforms rapidly. Microwave interferometric remote sensing, because of its capacity to acquire images through clouds and generate three-dimensional data appears to be one of the means best adapted to the task. While obtaining a DTM in such a way is still out of reach, the new elevation model called WorldDEM Core provides a DSM surface for the world's forests (Schlund et al., 2016; Solberg et al., 2015.b; Persson et al., 2014.b). WorldDEM Core is a commercial product sold by Airbus Defence & Space (Airbus 2016), at CAD \$12/km² (as of November 2016). It was produced using the first twin formation of spaceborne synthetic aperture radar (SAR) sensors—TanDEM-X (TerraSAR-X add-on for Digital Elevation Measurements, a mission designed by the German Space Agency, DLR)—with a 12 m resolution and a reported relative height error of 2 m. For any given WorldDEM tile, several interferograms may have been combined to produce the final DSM, especially in high relief areas where layover effects create interferometric problems (Krieger et al., 2007). Researchers may have access to the original

interferograms used to create the WorldDEM data through a scientific agreement with DLR.

The elevation maps extracted from TanDEM-X interferograms correspond to the vertical location of the scattering phase center (SPC) (Praks et al., 2012). The precise SPC location relative to the canopy surface is affected by wavelength, forest structure and the viewing geometry of the sensor. In the case of TanDEM-X, the X-band ($\lambda=3.6$ cm) SPC is close to the upper surface of the forest canopy (topmost foliage). Therefore, the elevations computed from TanDEM-X images constitute a DSM (Sadeghi et al., 2016; Solberg et al., 2015.a; Sadeghi et al., 2014). Subtracting, a known DTM (e.g. from lidar) from a TanDEM-X DSM, a hybrid InSAR-lidar CHM, hereafter called “InSAR CHM” can be created. This leads to estimations of forest heights having accuracies from 0.8 m to 5.0 m (RMSE error) depending on forest type and acquisition conditions (Sadeghi et al., 2016; Schlund et al., 2016; Solberg et al., 2015.a).

By the end of 2016, the WorldDEM Core high-resolution worldwide DSM created from more than 500,000 TanDEM-X interferograms will have been finalized (Zink et al., 2014). All the individual interferograms used for this product were acquired in the HH polarization and in the bistatic strip-map mode for areas spanning 30 by 50 km with varying heights of ambiguity (HoA). HoA has defined as:

$$HoA = \frac{\lambda R \sin(\theta)}{B} \quad [3.1]$$

Where λ , R , θ and B respectively expresses the wavelength, the slant range from the sensor to the imaged object, the local incidence angle (i.e., the angle formed by the negative of the SAR line-of-sight vector and the local surface normal) and the effective baseline.

So far, two global TanDEM-X image acquisitions phases with varying ranges of HoA were used to create a global TanDEM-X DSM. The first had HoAs ranging from 40 to 55 m, and the second one HoAs ranging from 30 to 40 m (Gruber et al., 2015). In difficult areas, including Antarctica, regions with steep topographies and some desert areas, an additional acquisition was performed with a much higher HoA. In these cases, both the baseline and incidence angle were adjusted to improve the coherence and the signal-to-noise ratio over the forested or desert areas (Gruber et al., 2015). While the TanDEM-X image acquisitions were made using ascending and descending right-looking passes, some areas still have a shadow or layover problems. To minimize these deficiencies, further acquisitions were made after shifting the orbit of the TanDEM-X formation (Rizzoli et al., 2015).

Furthermore, due to the helix pattern of the twin satellite orbits, the effective baseline of the TanDEM-X formation changes with the latitude. Since the start of the TanDEM-X mission, across- and along-track baselines ranging between 100 m to 10 km, and 0 to 100 km, respectively, were used. From December 2010 to mid-2014, the helix formation was configured in such a way that effective baselines varied between 120 m and 500 m for obtaining accurate elevation measurements, except for some areas for which these baselines were out of this range (Zink et al., 2014). Moreover, the TanDEM-X incidence angles at the center of the scenes varied from 28° to 48° (Gruber et al., 2015). To overcome the layover and shadow problems in areas with slopes higher than 20%, further acquisitions were performed with steeper incidence angles.

As discussed above, the global TanDEM-X DSM is a patchwork of interferograms produced under different baselines and incidence angles configurations. Among other factors, it is known that the effective baseline influences the accuracy of elevations estimated from interferograms (Krieger et al., 2007). Larger baselines provide more accurate measurements of small height changes, whereas short baselines lead to less phase unwrapping problems, at the cost of poorer elevation accuracies. Furthermore,

because the acquisitions were made on a continuous basis over about four years, seasonal and weather variations also occurred. All of these variations may induce changes in the response of forest canopies to the incident microwave energy and affect the interferogram generation. Indeed, in most parts of the temperate or mixed boreal forest zone, deciduous trees lose their leaves during the cold season. These phenological variations may affect the SPC location in the forest canopy. Greater penetration in leafless trees should cause a downward bias in DSMs. In a study about the effect of phenology on the location of the SPC using five TanDEM-X datasets with small baselines (20–34 m) acquired from September (leaf-on) to November (leaf-off) in Finland, it was found that beam penetration into the canopy of deciduous forests was 4 m greater under leaf-off conditions than under leaf-on conditions (Demirpolat 2012). A bias of 2 m was also found for evergreen conifer trees in November, suggesting that the absence of leaves in deciduous trees was not the sole factor explaining the lower SPC heights. Furthermore, the temperature at the time of acquisition can change the response of forest covers to the incident microwave energy. Temperature changes significantly affect the dielectric constant, thus modifying the amount of backscattered energy, and possibly also the degree of microwave penetration in the canopy (Solberg et al., 2015.a). This phenomenon could, in turn, change the vertical location of the SPC. The fact that the TanDEM-X single-polarization interferograms used to produce the global DSM were generated from images acquired at different dates and different configurations, therefore, raises the question of data consistency. Specifically, if we attempt to map forest heights over large regions, accuracy variations, especially biases, may significantly affect the reliability of height estimates. This would hinder our capacity of using standard sampling and mapping procedures. The latter consist of building a height prediction model based on a remotely sensed variable, using strategies such as regression or k-NN methods (Yim et al., 2011) calibrated using field plots sparsely distributed throughout the region of interest. The resulting (single) predictive model is then applied to the entire region to generate a continuous and calibrated height map. In the case of regions covered by a mosaic of different TanDEM-X

interferograms, a single predictive model generated from a sparse set of well-distributed field plots would be most effective. However, if the mosaicked interferograms each have different acquisition parameters, using a generalized predictive model applied to the mosaic (Gruber et al., 2015) may lead to systematic under- or over-estimations depending on location in the mosaic.

The objective of this study was therefore to compare the accuracy of TanDEM-X DSMs obtained under different baseline, incidence angle, tree phenology and weather conditions over a mixed boreal forest area. This was investigated through four experiments, the first serving to verify that images acquired in almost identical conditions produce the same results. The other experiments respectively consisted in assessing the effects of variations in the baseline, incidence angle, and phenology. Weather data was also considered to control for the potential effect of precipitation or temperature on the experiments. The accuracy has been expressed regarding the amount of deviation and bias compared to reference values. For this purpose, we used five TanDEM-X interferometric pairs over a boreal forest region acquired in 2012 and 2013. A lidar DTM, DSM, and CHM were used as reference data to evaluate all TanDEM-X DSMs at a resolution of 25 m. Datasets were also directly compared to one another. Deviations from a general regression model calibrated with all the merged datasets were also analyzed for each dataset to assess the reliability of a general canopy height prediction model.

3.4 Materials

3.4.1 Study area and field data

The research was conducted at the Montmorency Forest, a 6 600 ha boreal research forest located in the south of the province of Quebec, Canada (47° 18' N, 71° 08' W, Figure 3.1). Elevations are between 600 and 1000 m, with steep slopes in some areas.

The average annual temperature is 0.2 °C, and fluctuates on average from a maximum of 15 °C in July, to a minimum of -15 °C in January. Balsam fir (*Abies balsamea* (L.) Miller) and paper birch (*Betula papyrifera* Marshall) are the main tree species found in the area. Most stands are composed of conifers only, but several mixed stands (i.e., having conifer and deciduous trees) can be found. Pure deciduous stands are, however, rare. The stand age ranges from 10 to 90 years, with several stands resulting from regeneration after harvesting.

Two hundred well distributed permanent plots are re-measured on a five-year cycle (i.e. 40 per year). Standard forestry related mensuration is performed, such as diameter at breast height and species of all trees larger than 5 cm, as well as the height of a few sample trees per plot. In addition, 431 height measurements of dominant trees were made in 2012 and 2013 using a Vertex III clinometer (Haglöf Sweden AB) to evaluate the accuracy of lidar single-tree height estimates specifically. The data from these plots indicate that the average stand density is 1300 stems/ha, and that most stands are closed, with a very high percentage of canopy cover. Moreover, the dominant tree height varies between 3.2 m and 26.4 m, with an average of 14.9 m. The area receives annual precipitations totaling 1400 mm, of which 34% is snow, mainly from December to April. Overall, this site is representative of closed canopy boreal forest conditions dominated by evergreen conifers.

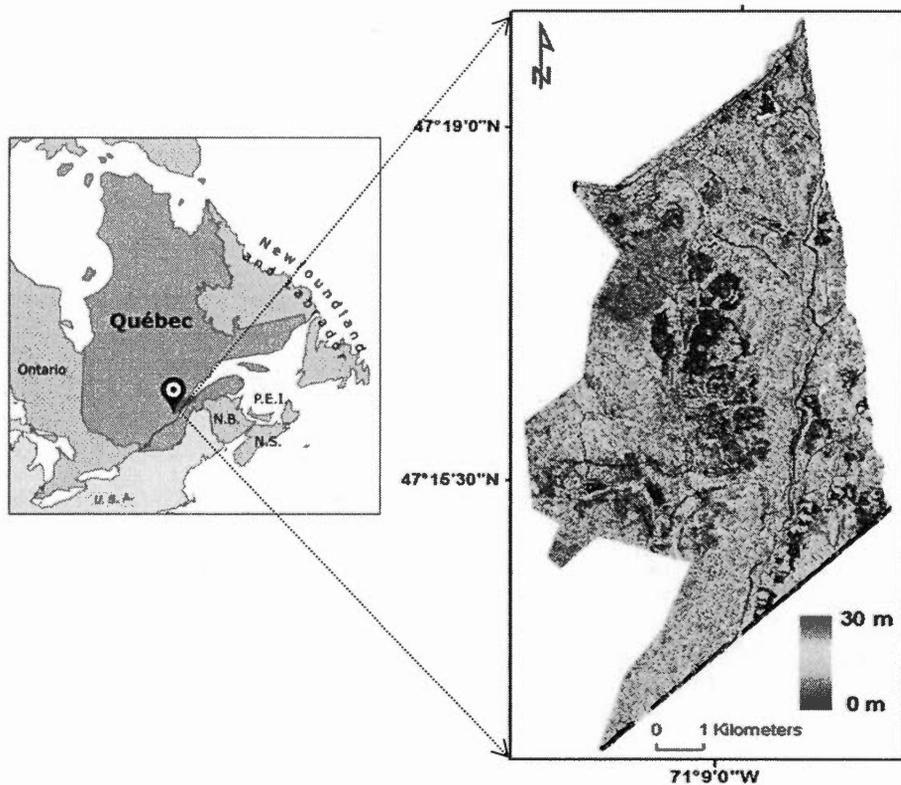


Figure 3.1 Location of the Montmorency Research Forest location in the south of the province of Quebec, Canada

3.4.2 TanDEM-X images

Five TanDEM-X interferometric pairs were acquired over the research area using a bistatic strip-map mode resulting in images measuring 50 km in the azimuthal direction and 30 km in the range direction. They were obtained using a single-look slant range complex (SSC) format. The analysis was conducted over a 4 000 ha area where all images overlapped. The TanDEM-X acquisitions were carried out between 15 November 2012 and 28 August 2013 on ascending paths with baselines ranging from 21 m to 192 m (See Table 3.1). For simplicity, we refer to these five acquisitions with the letters *a* to *e*. The *a* and *b* datasets were taken under leaf-off conditions at respective

temperatures of -6.0 and -1.5 °C, in the absence of snow cover. The HoA of the *a* and *b* datasets were 25.7 m and 38.1 m, respectively. Datasets *c*, *d*, and *e* were acquired with identical incidence angles, under leaf-on conditions and with similar temperatures (16.8 to 19.9° C). However, their HoAs varied considerably between 41.9 m and 218.4 m. The *a*, *d* and *e* datasets were acquired under dry conditions, while 0.3 mm and 2.3 mm of rain fell on the acquisition day before image capture in the case of datasets *b* and *c*.

Table 3.1 The characteristics of TanDEM-X acquisitions

| Acquisition ID | Date | Effective baseline (m) | Height of ambiguity (HoA) (m) | Incidence angle at the scene centre (°) | Temperature (°C) | Precipitation (mm) |
|----------------|------------|------------------------|-------------------------------|---|------------------|--------------------|
| a | 15-11-2012 | 182 | 25.7 | 30.6 | - 6.0 | 0 |
| b | 20-11-2012 | 192 | 38.1 | 43.3 | -1.5 | 0.3 |
| c | 04-7-2013 | 112 | 41.9 | 30.6 | 19.9 | 2.3 |
| d | 15-7-2013 | 108 | 43.6 | 30.6 | 19.2 | 0 |
| e | 28-8-2013 | 21 | 218.4 | 30.6 | 16.8 | 0 |

3.4.3 Lidar data

An Optech ALTM 3100 airborne laser scanner (Teledyne Optech Inc., Vaughan, ON, Canada) was used to collect the lidar data over the study site on 6 and 9 August 2011. The altitude of the sensor was 1000 m above ground level (AGL), which resulted in a median first-return density of 5 returns/m². The accuracy of tree height information extracted from this lidar dataset was verified using the field measurements of single

tree heights and found to be very accurate with a RMSE of 1.29 m (see Sadeghi et al. 2016 for details).

3.5 Methods

3.5.1 Creating the DTM, DSMs, and CHMs

A lidar DTM and a lidar DSM were created by interpolating the ground-classified returns and the first returns respectively. The initial interpolated rasters were created with an initial ground pixel size of 0.25 m though inverse distance weighting (IDW) interpolation of the returns. The lidar CHM was obtained by subtracting the DTM elevations from the DSM.

The TanDEM-X image pairs were used to produce interferograms that were flattened using a 30 m SRTM DEM. An adaptive filter was used to smoothen the flattened interferograms. Coherence and phase were calculated from these interferograms. Finally, phase unwrapping and conversion-to-canopy heights were performed using the minimum cost flow method (Costantini 1998). All InSAR processing steps were conducted using the ENVI SARscape 5.0 processor. The initial products, including coherence images and InSAR DSMs from the HH polarization, were created at an initial resolution of 5 m and georeferenced to a WGS84 datum in a UTM zone 19N projection. The resulting five InSAR DSMs were locally masked using a polygon layer including lakes, anachronic forest harvests (clear cuts having occurred during the time interval between the respective acquisitions of lidar and TanDEM-X data), radar layover areas, pylons and suspended power cable locations, areas with coherence values below 0.3, and areas with phase unwrapping problems caused by very steep slopes or the presence of edges with sharp coherence changes. Overall, 0.7% of the study area was left out from the analyses. The lidar models and the five masked TanDEM-X DSMs were aggregated at a resolution of 25 m for the subsequent analyses.

The TanDEM-X DSMs were converted into InSAR CHMs by subtracting the (25 m) lidar DTM elevations. This pixel size was selected because we considered it represented an adequate compromise between maintaining a good spatial resolution and reducing the noise in the InSAR surfaces. See Sadeghi et al. (2016) for a comparison of results from pixel sizes of 5 m to 25 m in the context of InSAR height accuracy for the same study area.

3.5.2 Accuracy assessment of the InSAR CHMs

The masked InSAR CHMs were evaluated by comparing them to the lidar CHM. For each TanDEM-X dataset, a linear regression between the lidar and InSAR CHMs was computed, and the corresponding bias (mean difference), r^2 and RMSE values were calculated. Because of the vast number of 25 x 25 m resolutions cells in the study area, a systematic sampling at 150 m intervals in the X and Y directions was done, bringing the total number of samples for regressions down to 1206 per dataset. In parallel, we have made a similar comparison for bare earth pixels to verify that potential elevation discrepancies in forested areas were not simply attributable to, say, vertical misregistration between the InSAR and lidar datasets. For this purpose, 608 pixels in bare areas were selected throughout the study region, in locations where InSAR coherence was high, and the slope was below 20°.

3.5.3 Consistency experiment

The first experiment aimed at verifying whether two TanDEM-X datasets acquired under the same conditions lead to similar CHMs. For this, we used datasets *c* and *d*. Both were acquired within an 11-day interval (on the 4th and 15th of July 2013, Table 3.1), with an identical incidence angle of 30.6°, and very similar effective baselines of 112 m and 108 m, respectively. The meteorological conditions were also comparable except for a small amount of rain that fell before the acquisition of *c*.

3.5.4 Baseline variation experiment

To study the influence of baseline variations on the CHMs, the accuracy of two InSAR CHMs having highly contrasted baselines but otherwise very similar acquisition conditions were compared. For this experiment, we used datasets *c* and *e*, which were acquired during leaf-on conditions on 4 July and 28 August 2013 (Table 3.1) with identical incidence angles (30.6°) and under similar meteorological conditions, although with an effective baseline of 112 m and 21 m, respectively. The baseline differences were considerable with regard to HoA (41.9 m vs. 218.4 m).

3.5.5 Incidence angle variation experiment

To assess the effect of incidence angles, the *a* and *b* InSAR datasets acquired with respective incidence angles of 30.6° and 43.3° were used. Both were acquired under leaf-off conditions, sub-zero temperatures with comparable baselines (182 m and 192 m) and related HoA's (25.7 m and 38.1 m) and very low levels of precipitation.

3.5.6 Phenology variation experiment

In a fourth experiment, the phenology variation was studied using datasets acquired under leaf-on and leaf-off conditions. Neither of the two leaf-off datasets could be paired with a leaf-on dataset having very similar baseline and incidence angle characteristics. The *a* (leaf-off), *d* (leaf-on), and *e* (leaf-on) datasets had identical incidence angles. Dataset *b* (leaf-off) differed from *d* and *e* with regard to both incidence angle and baseline, as a result of which it was not used to study the effect of phenology. The *a* and *d* datasets were the most consistent ones with yet a 74 m baseline difference. Our phenology experiment relies mostly on this pair. However, to assess the respective importance of phenology and baseline, we also compared the *a* and *e* datasets (despite a 161 m baseline difference).

For all experiments, we have reported the bias and RMSE, as well as the regression coefficients (intercept and slope) and the r^2 between the compared datasets. In addition,

we have performed a standard paired t-test between each compared datasets, the null hypothesis being that different acquisition conditions do not lead to discrepancies between the interferometric DSMs.

3.5.7 Analysis of the impact of acquisition conditions on canopy height predictions

An analysis was conducted to assess the effect of the various acquisition conditions on the error level of a general canopy height prediction model calibrated and applied to a heterogeneous set of interferograms. Our goal was to simulate a situation in which canopy surface height predictions would be sought for a large region covered by several mosaicked TanDEM-X interferograms having different acquisition characteristics but using a single set of calibration plots. This mimics the setting in which, for example, wide-area WorldDEM Core data would be employed for mapping forest canopy height. For this, we calibrated a single regression model using 1206 sample locations of the InSAR and lidar canopy surface heights, i.e. InSAR heights were extracted for the same locations in all five InSAR datasets. The general model was thus calibrated using 6030 observations (5×1206). It was then applied separately to the five InSAR CHMs. The bias and relative bias between the predictions of the general model and the reference lidar-based values were then computed separately for each of the TanDEM-X datasets.

3.6 Results

3.6.1 Coherence maps and CHMs

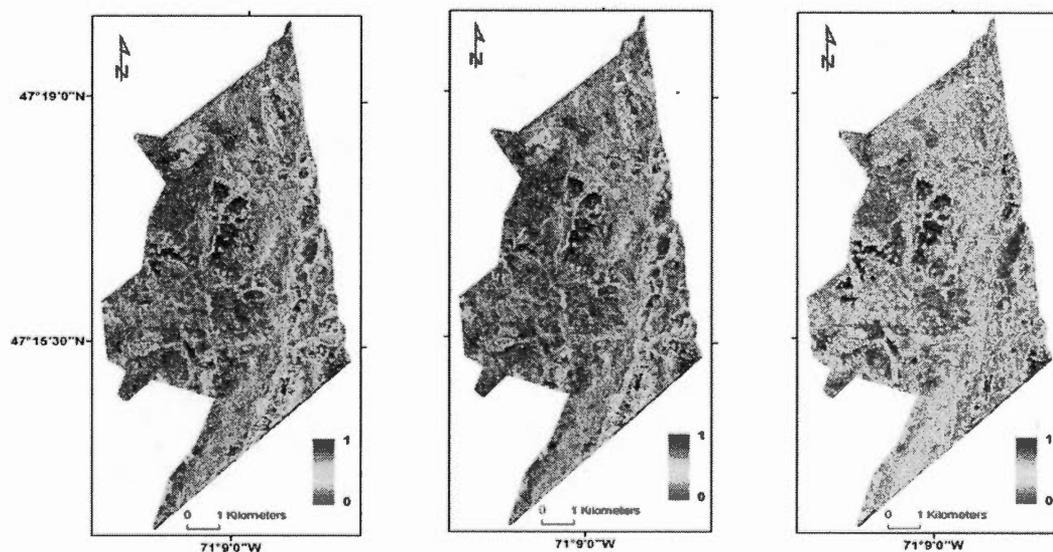
Figure 3.2 presents the TanDEM-X unmasked coherence maps for the five datasets (Figure 3.2(a) to 3.2(e) corresponding to datasets *a* to *e* described in Table 3.1). The TanDEM-X coherence is not affected by temporal decorrelations and only by the surface structure and acquisition conditions. High coherence values were seen in clear-cut and low vegetation areas. Low coherence occurred in high vegetation, lakes, and steep slope areas. Datasets *a* to *e* had respective average coherence values of 0.49, 0.59,

0.68, 0.69 and 0.82. The highest mean coherence value of 0.82 corresponded to dataset *e* having the lowest baseline value (21 m). Overall, the correlation between the mean coherence and baseline was -0.94, indicating a strong influence of the baseline on the coherence. The two summer datasets with nearly identical baselines (*c* and *d*) had essentially equal mean coherences. Among the leaf-off datasets (*a* and *b*), the coherence changed significantly as both the baseline (182 m vs. 192 m) and the incidence angle (30.6° vs. 43.3°) varied. Datasets with a low mean coherence exhibited more variation in the coherence values than those with a high mean coherence.

(a) 15 November 2012

(b) 20 November 2012

(c) 04 July 2013



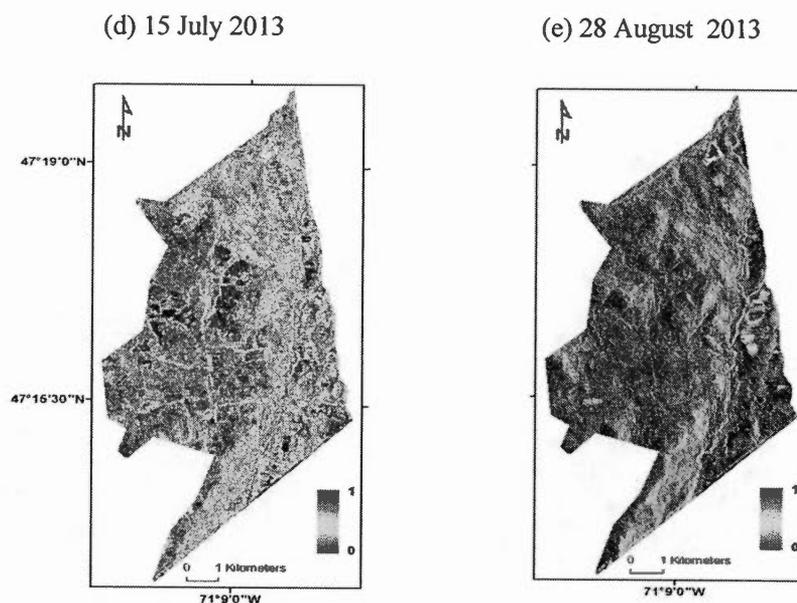


Figure 3.2 Coherence maps corresponding to the five TanDEM-X acquisitions (a to e)

Figure 3.2 compares the InSAR CHMs [Figure 3.2(a) to 3.2(e)] with the lidar CHM at a 25 m resolution [Figure 3.2(f)]. The general patterns of forest canopy heights are visually very similar. The lidar CHM map is sharper than the InSAR CHMs. The clear-cut areas and roads are obvious in all CHMs. The small baseline dataset (e) stands out from the other InSAR CHMs with regard to its high degree of noise. In bare areas (lidar CHM height = 0 m), the InSAR surface corresponded to the lidar DTM, indicating that the InSAR and lidar models were correctly coregistered vertically.

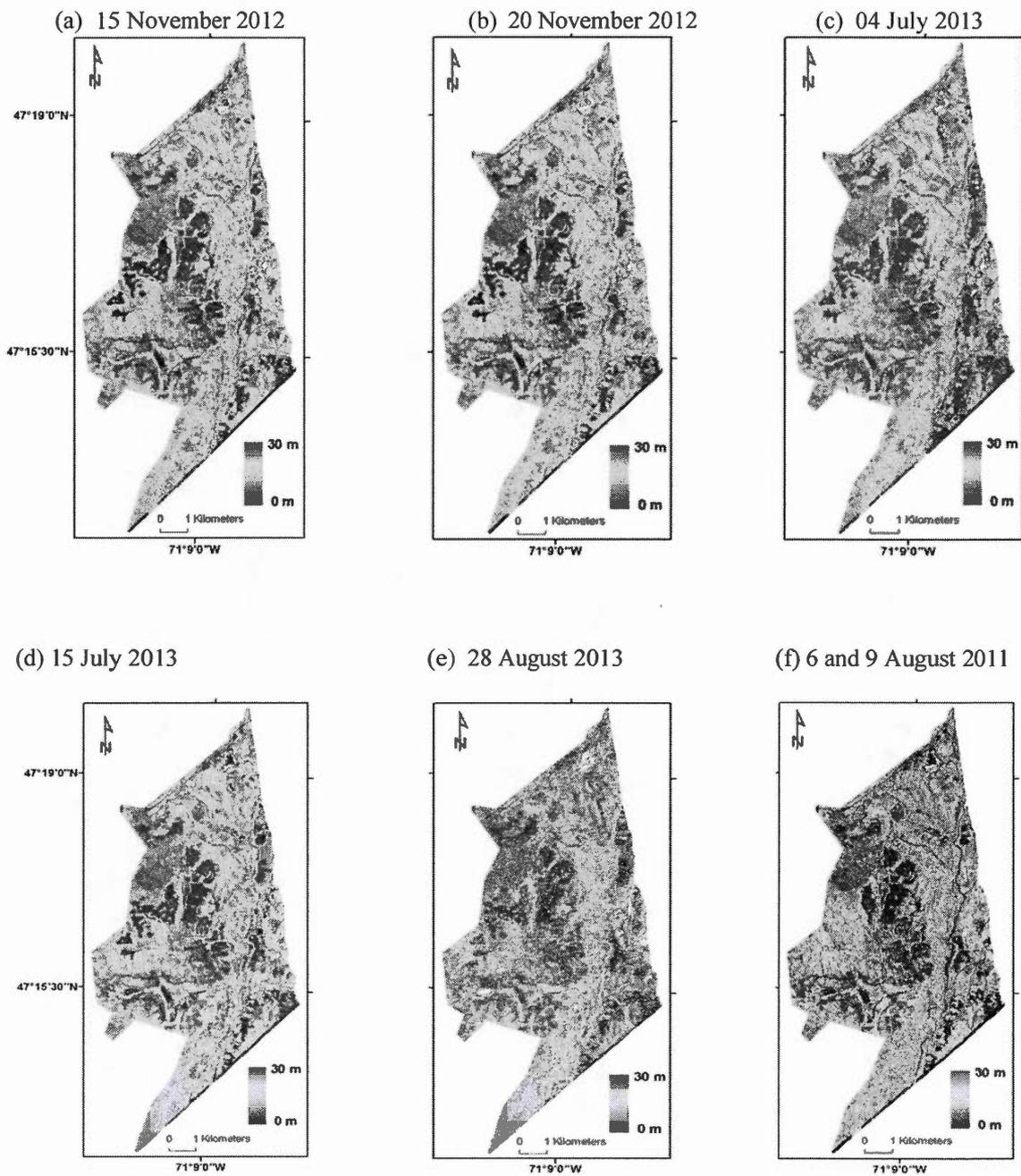
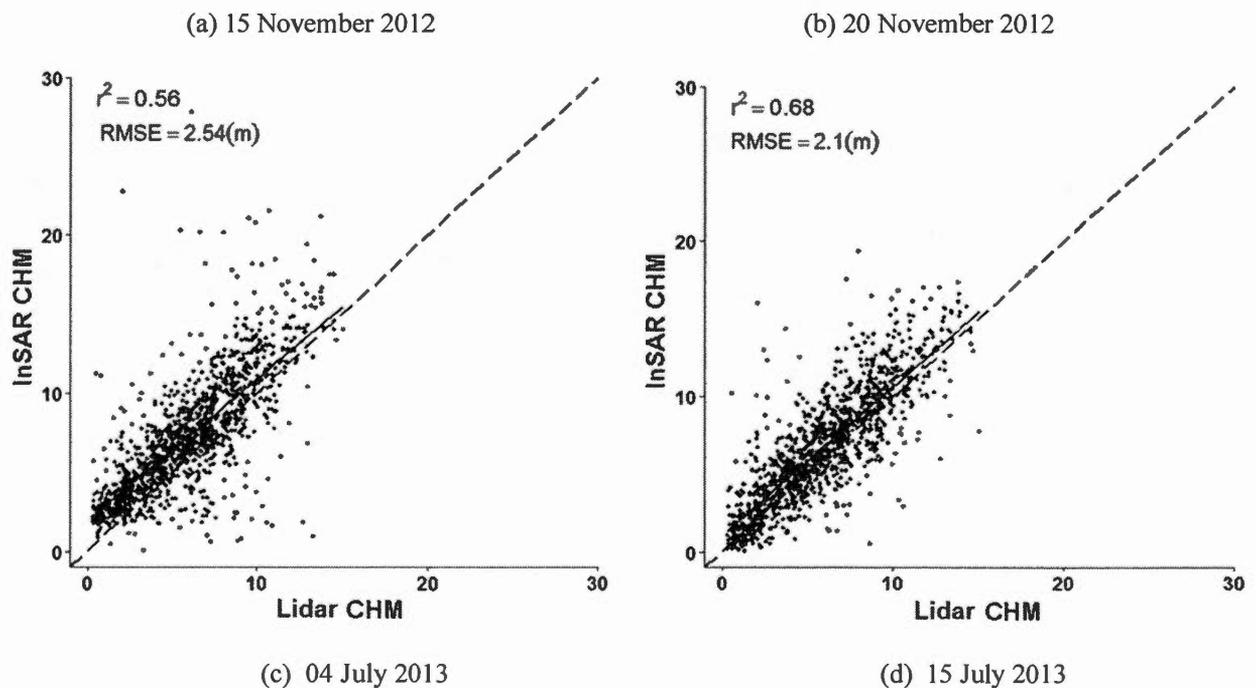


Figure 3.3 InSAR CHM maps (a to e according to the acquisitions in Table 3.1) and lidar CHM map.

InSAR TanDEM-X CHMs were quantitatively compared to the corresponding lidar CHM (Figure 3.4 and Table 3.2). The mean InSAR CHM heights varied from 6.72 m to 7.51 m and were in all cases greater than that of the lidar CHM (5.95 m), the difference varying between 0.77 m (dataset *b*) to 1.56 m (dataset *e*). The coefficients of determination (r^2) between InSAR and lidar CHMs were highly significant with p -values < 0.001 and varied between 0.38 and 0.68. The RMSEs ranged from 2.06 m and 3.67 m. With regard to all accuracy variables, the *e* dataset (short baseline) stood out as having the largest discrepancy with the lidar CHM. The variations in r^2 and RMSE did not follow closely coherence variations correlations of respectively -0.55 and 0.55; instead; they were somewhat more similar to baseline variations (correlations of respectively -0.70 and 0.74). Both the bias and RMSEs were lower in bare areas compared to forested ones. No significant correlation between the bias in forests and the bias in bare areas existed ($r = -0.19$, p -value = 0.75).



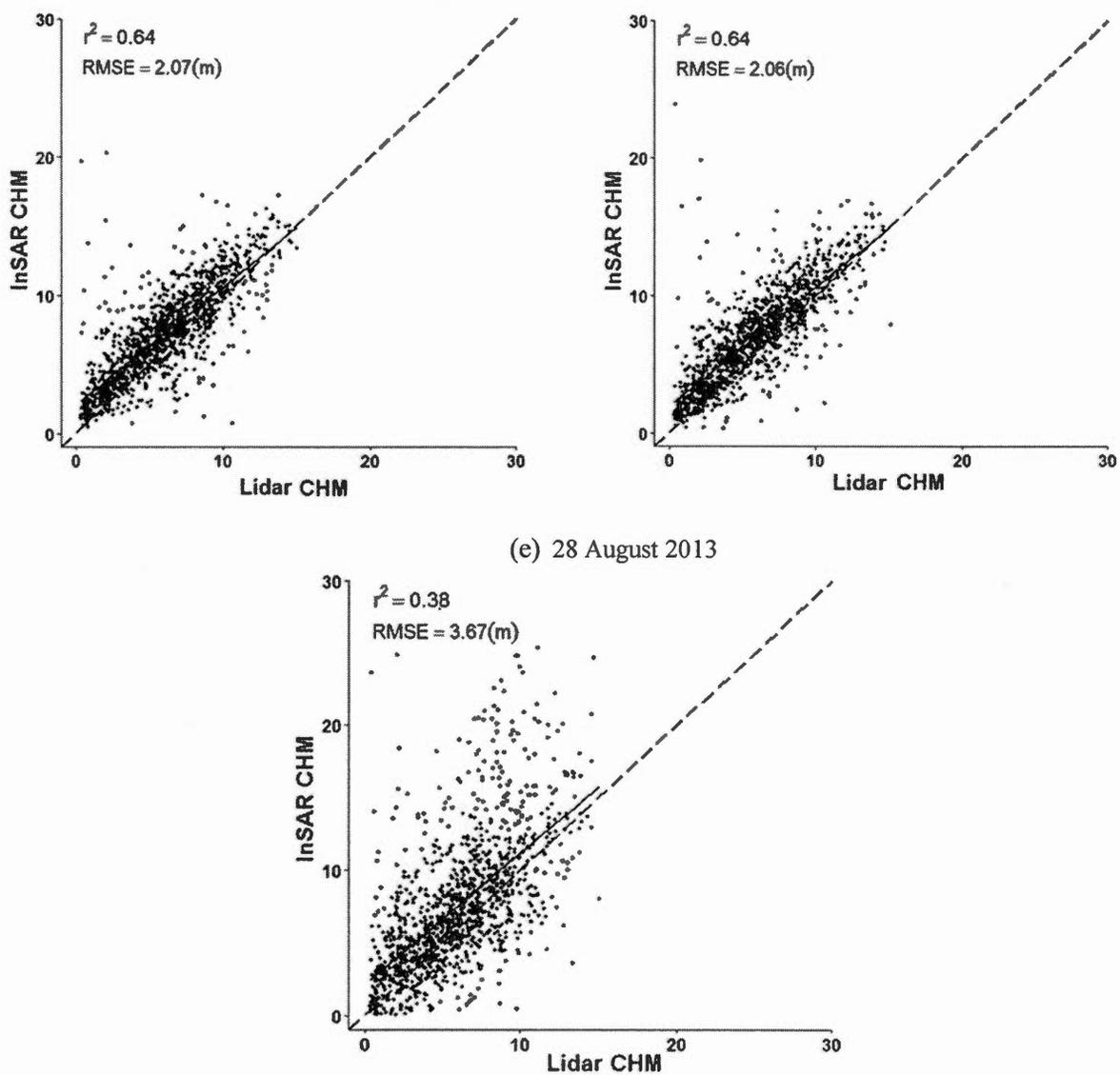


Figure 3.4 Relationships between InSAR CHMs and the lidar CHM for the five InSAR datasets [(a) to (e) of Table 3.1], $n = 1206$

Table 3.2 Statistics of the regression between InSAR CHM and lidar CHM

| Acquisition ID | Mean CHM value | | Forest area ^(*) | | | | Bare area ^(*) | | |
|----------------|----------------|-----------|----------------------------|---------------|-------------|----------|--------------------------|----------|----------|
| | Lidar (m) | InSAR (m) | $B_0^{(***)}$ | $B_1^{(***)}$ | $r^{2(**)}$ | Bias (m) | RMSE (m) | Bias (m) | RMSE (m) |
| a | 5.95 | 7.25 | 1.88 | 0.90 | 0.56 | 1.30 | 2.54 | 0.02 | 0.74 |
| b | 5.95 | 6.72 | 0.99 | 0.96 | 0.68 | 0.77 | 2.10 | 0.07 | 0.80 |
| c | 5.95 | 7.13 | 2.04 | 0.85 | 0.64 | 1.18 | 2.07 | 0.61 | 0.88 |
| d | 5.95 | 7.09 | 1.93 | 0.87 | 0.64 | 1.14 | 2.06 | -0.71 | 1.03 |
| e | 5.95 | 7.51 | 2.11 | 0.91 | 0.38 | 1.56 | 3.67 | -0.30 | 2.07 |

(*) Sample number for forest areas $n=1206$ and bare areas $n=608$

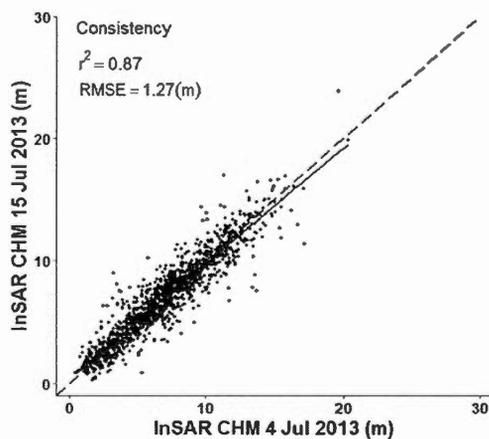
(**) All r^2 are significant at $p < 0.001$

(***) $CHM_{InSAR} = B_0 + B_1 CHM_{Lidar}$

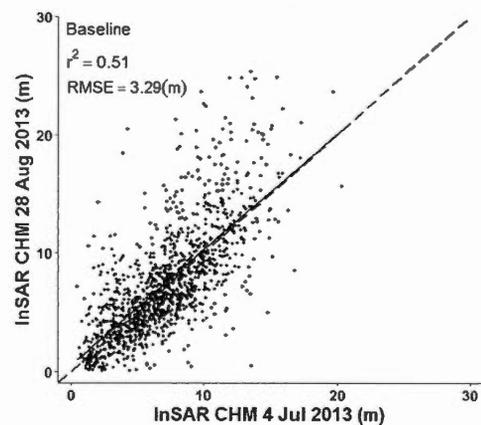
The results of the consistency, baseline, incidence angle and phenology variation experiments are presented in Figure 3.5 and Table 3.3. The two datasets acquired in the summer (*c* and *d*) with an 11-day interval and identical acquisition parameters were found to be highly consistent. The r^2 between the two datasets was 0.87, with a RMSE of 1.27 m. The slope of the relationship (0.94) between these datasets is close to, but not equal to, 1.0. The bias between the CHMs is negligible, at 0.04 m. Considering the high p -value (0.2329) of the paired t -test, we fail to reject the null hypothesis with a confidence level of 95%. A distinct difference in the baseline between two summers InSAR CHMs (*c* with 112 m, and *e* with 21 m) led to a poor relationship between these CHMs, with a r^2 of 0.51, a RMSE of 3.29 m and a bias of 0.38 m. The effect of the

incidence angle (30.6° and 43.3°) was tested with the fall imagery (November 2012). The impact of this parameter on the two CHMs (a and b) was lower than that of the baseline. The r^2 between the two datasets was 0.69, with a RMSE of 2.26 m. However, the bias between these two datasets, namely 0.53 m, was the largest observed in all comparisons. The lowest average height was obtained for the largest incidence angle (43.3° from vertical). In the case of both the baseline and incidence angle results, there was only a 1/10 000 chance that the observed discrepancies would be observed considering the null hypothesis. We, therefore, should clearly reject it. Finally, the phenology variation experiment was conducted in two parts. The first part used datasets a and d , having a moderate baseline difference (74 m). It showed a r^2 of 0.65, a RMSE of 2.3 m and a bias of 0.16 m. The second, with a much higher baseline difference (161 m), led to the greatest discrepancy between any two datasets ($r^2 = 0.43$, RMSE = 3.61 m, bias = 0.26 m). Overall, the best correspondence between any two InSAR CHMs was obtained when all parameters (phenology, baseline, and incidence angle) were nearly constant. In both phenology experiments, the p values of the paired t -test are less decisive (0.1 - 0.2).

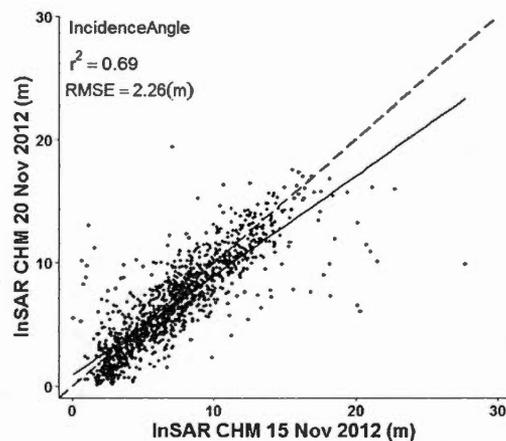
(a) Consistency



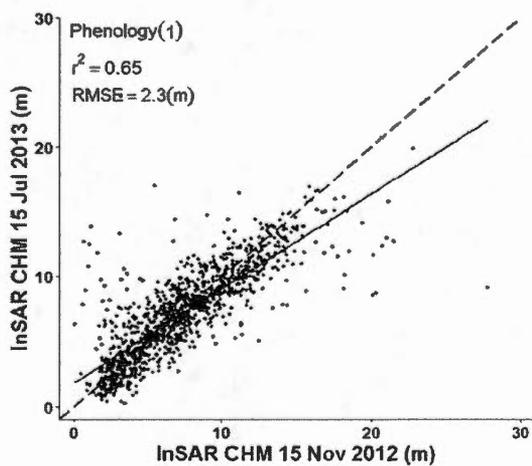
(b) Baseline



(c) Incidence Angle



(d) phenology (1)



(e) phenology (2)

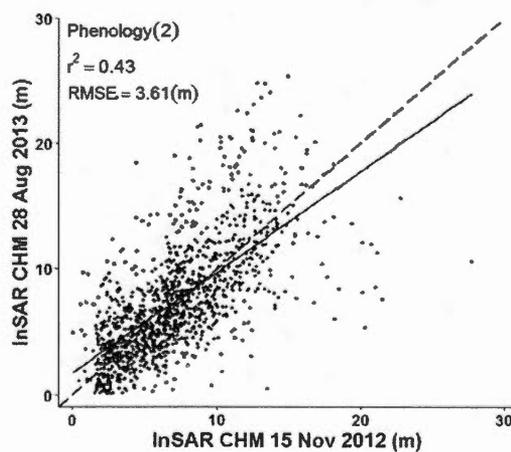


Figure 3.5 Relationships between the InSAR CHMs as a function of different experiences to study the effect of each parameter on the InSAR CHM. Consistency (a), baseline (b), incidence angle (c), and phenology [(d) and (e)], $n=1206$

Table 3.3 Statistics of the comparison experiments

| Experiment | X | Y | Mean CHM | | Bias (m) | Y = B ₀ + B ₁ X | | r ² (*) | RMSE (m) | Paired <i>t</i> -test (<i>P</i> -value) | n |
|----------------|---|---|----------|------|-------------|---------------------------------------|----------------|--------------------|-------------|---|------|
| | | | X | Y | | B ₀ | B ₁ | | | | |
| Consistency | c | d | 7.13 | 7.09 | 0.04 | 0.39 | 0.94 | 0.87 | 1.27 | 0.2329 | 1206 |
| Baseline | c | e | 7.13 | 7.51 | 0.38 | 0.55 | 0.98 | 0.51 | 3.29 | 0.0001 | 1206 |
| IncidenceAngle | a | b | 7.25 | 6.72 | 0.53 | 0.87 | 0.81 | 0.69 | 2.26 | 0.0000 | 1206 |
| Phenology (1) | a | d | 7.25 | 7.09 | 0.16 | 1.82 | 0.73 | 0.65 | 2.30 | 0.0208 | 1206 |
| Phenology (2) | a | e | 7.25 | 7.51 | 0.26 | 1.70 | 0.80 | 0.43 | 3.61 | 0.0113 | 1206 |

(*) All models are significant at $p < 0.001$

3.6.2 General prediction models

For analyzing the impact of acquisition conditions on canopy height predictions based on a general model, all InSAR observations from various datasets (*a–e* for the model I and *a–d* for model II) were pooled together and then regressed against the corresponding lidar heights. This led to an overall r^2 of 0.56, and a RMSE of 2.14 m for the model I (Table 3.4). Comparing these results with those of Table 3.2, the RMSE of general model (Model I) was lower than that of models *a* and *e* in Table 3.2. However, the intercept (1.54 m) and slope (0.62) of this model were far from the ideal theoretical values (0 m and 1). Because the *e* dataset was shown to yield a poor CHM compared to the corresponding lidar CHM due to its very short baseline, we computed an alternative general model (model II) that excluded this particular dataset. The overall r^2 and RMSE changed respectively to 0.64 and 1.92 m.

Table 3.4 General model to predict lidar CHM

| Model | Mean CHM | | Bias (m) | CHM _{Lidar} = B ₀ + B ₁ CHM _{InSAR} | | r ² (*) | RMSE (m) | n |
|-------|--------------|--------------|-------------|---|----------------|--------------------|-------------|------|
| | Lidar (m) | InSAR (m) | | B ₀ | B ₁ | | | |
| I | 5.95 | 7.14 | 1.19 | 1.54 | 0.62 | 0.56 | 2.14 | 6030 |
| II | 5.95 | 7.05 | 1.10 | 1 | 0.71 | 0.64 | 1.92 | 4824 |

(*) All models are significant at $p < 0.001$

The general models of Table 3.4 were applied to each InSAR CHM to predict (reference) lidar CHM heights that were considered as being the reference (Table 3.5). The average of the predicted lidar CHM heights was calculated for each InSAR dataset and compared to the mean lidar CHM. We thus obtained bias values of 0.08, 0.24, 0.05, 0.02 and 0.24 m for dataset *a*, *b*, *c*, *d*, and *e* respectively using the first version of the general model (model I). When using the other general model based only on datasets *a* to *d*, (Model II), the biases changed to 0.15, 0.20, 0.09 and 0.06 m respectively. The *p*-values within the paired t-test table of Table 3.6 is the highest for both the model I and II between dataset *c* and *d* (0.2329), while they are the smallest between datasets *b*, and (*d*, *e*) with the values of $< 2.2 \text{ e-}16$.

Table 3.5 Statistics of the comparison between the predicted and reference CHM using general models (*)

| Model | Dataset | Lidar predicted CHM (m)(**) | Bias (m) | Relative Bias (%) |
|-------|---------|-----------------------------|----------|-------------------|
| I | a | 6.03 | + 0.08 | 1.3 |
| | b | 5.71 | - 0.24 | 4.0 |
| | c | 5.90 | - 0.05 | 0.8 |
| | d | 5.93 | - 0.02 | 0.3 |
| | e | 6.19 | + 0.24 | 4.0 |
| II | a | 6.10 | + 0.15 | 2.5 |
| | b | 5.75 | - 0.20 | 3.4 |
| | c | 6.04 | + 0.09 | 1.5 |
| | d | 6.01 | + 0.06 | 1.0 |

(*) All the models are significant at $p < 0.001$, with $n = 1206$ observations for each model.

(**) Mean lidar CHM height = 5.95 m

3.7 Discussion

Five TanDEM-X DSMs generated under different geometric and phenological conditions were converted to CHMs using a lidar DTM and then compared to a reference lidar-derived CHM model. It was found that the general spatial patterns of canopy heights were well represented in all TanDEM-X DSMs. All InSAR CHMs were overestimated compared to the lidar CHM, with a bias ranging from 0.77 m to 1.56 m. As demonstrated in a previous study based on similar data (Sadeghi et al., 2016), such discrepancy originates from small forest gaps being "filled" in the InSAR CHM, while being detected by lidar. As expected, the interferogram created with a very short baseline, (21 m), compared to an average of 148 m for the four other ones, led to the worst result. A short baseline generally has the highest coherence with reduced volume decorrelation. However, as the uncertainty in the measurement of the canopy surface

height increases with HoA (Martone et al., 2012; Rizzoli et al., 2014), the short-baseline interferograms should not be used to create DSMs over forested regions. Similar conclusions were reached in (Solberg et al., 2015.a) who used TanDEM-X interferograms over boreal forests in Norway. Among the long baselines in the present study (108 m - 192 m), the effect of baseline on the DSM error was not significant.

The other geometric factor varying between the acquisitions was the incidence angle at the center of the scene. It had very different values among the two largest baselines and led to low CHM errors when it is shallow (43.3° , dataset *b*). In contrast, the combination of the long baseline (182 m, dataset *a*) and moderate incidence angle (30.6° , dataset *a*), which resulted in a very low HoA (25.7 m), created large errors (RMSE of 2.54). Ideally, the HoA should be around 40 m for the type of forest present in our study area (datasets *b-d*) (Solberg et al., 2015.a). This corresponds to suggested values between 20 m to 50 m made for boreal forests in Norway (Solberg et al., 2015.a). Other studies also showed HoA should remain higher than the height of the tallest trees in the case of a flat region and be increased in areas with large terrain slopes (Kugler et al., 2014 and 2015). Therefore, a configuration using optimal baseline and incidence angle is needed to achieve an appropriate HoA and accurate surface elevation, following the relations between baseline, incidence angle, and HoA.

Besides the effect of geometrical acquisition parameters on coherence and related InSAR CHM errors, the phenology could be a potential factor affecting both variables. Datasets *a* and *b*, corresponding to winter acquisitions showed the lowest mean coherence of 0.49 and 0.59, but these datasets also had the longest baseline, making it difficult to conclude on the effect of the phenology. However, datasets *b*, *c*, and *d* had very similar HoA, with almost identical InSAR CHM RMSEs. The winter CHM (dataset *b*) giving the same RMSE as the summer ones (datasets *c* and *d*) seems to indicate that the phenology may not be as important as the other factors. The bias for dataset *b* was lower (0.77 m) than that of the two comparable summer datasets (1.18

(dataset *c*), and 1.14 m (dataset *d*). This could be caused by a significant penetration in the leafless parts of the canopy, as reported in (Demirpolat 2012) for boreal forests. Indeed, we would expect less penetration in gaps with large incidence angle giving a significant positive bias, but the reverse is observed. One potential reason to explain the small impact of phenology may reside in the much more abundant evergreen conifer species with respect to deciduous tree species.

Differences in the InSAR CHMs should be expected for the cases where geometrical or phenological conditions are different during image acquisition, but some differences were also observed between datasets *c* and *d* that were acquired in nearly identical conditions. Although very high, the r^2 of 0.87 and RMSE of 1.27 m between *c* and *d* (Figure 3.5a) showed an appreciable departure from the theoretically expected values (respectively 1.0 and 0.0 m), but bias (0.04 m) was very low, and the null hypothesis (no difference between datasets) could not be rejected. A visual examination of the image of InSAR CHM difference (not shown) reveals essential random fluctuations without clear patterns in the discrepancies. The differences were great on edges (e.g. forest/non-forest edges) but were not due to misregistration as no shift could be detected. Explaining the discrepancies between datasets *c* and *d* is not trivial. The only variation in acquisition conditions was related to moisture. Precipitations totaling 2.3 mm of rain occurred on July 4th, just before the acquisition that occurred around midnight while dry conditions prevailed on July 15th. This would change the dielectric constant of the canopy, which leads to more volume decorrelation, and a decrease in coherence and signal to noise ratio as shown in (Solberg et al., 2015.a) over Norway boreal forests. This could affect the interferometric solution, resulting in a somewhat degraded CHM for dataset *c*. Table 3.2, however, shows that the accuracy of datasets *c* and *d* are nearly identical hinting to other underlying causes. We hypothesize that the phase unwrapping itself might have induced the observed variation. We have used the minimum cost flow algorithm, which uses different paths in each run. Thus, small changes in the initial conditions might produce significant changes in the solution. This

hypothesis is supported to some degree by the fact that larger discrepancies were observed in areas where rapid phase changes are to be expected (forest edges).

The fact that differences in the CHMs exist even when acquisition conditions are the same means that part of the observed discrepancies when HoA or phenology differs between datasets is due to random fluctuations caused, among other factors, by potential phase unwrapping problems. Therefore, in the case of the incidence angle difference (datasets *a* and *b*) for example, the RMSE (2.26 m) should be interpreted as being less than twice that of the *c* and *d* datasets for the consistency experiment (1.27 m).

When mapping height, or height-derived attributes such as forest biomass, from a mosaic of TanDEM-X images covering large areas (say several hundreds of thousands of km²), and from a sparse set of field plots (eventually less than one plot per TanDEM-X scene for example), variations in the RMSE of the prediction may be less important as variations of the bias. The analysis of the impact of the acquisition conditions on the height predictions revealed significant biases, whether we included the short baseline dataset or not, but these biases were smaller in the latter case. In the worst case, for the second general model (based on datasets *a-d*), the bias reached 0.20 m, i.e. 3.4% of the average height, with limited potential consequences on the regional mapping of boreal canopies. Even the inclusion of the short baseline dataset (*e*), which slightly increases the observed bias to +/- 4.0%, demonstrate the potential of a mosaicked TanDEM-X dataset to estimate canopy height at large scales. The findings of this paper were made on a closed canopy boreal forest dominated by evergreen conifers. These forests populate large parts of North America, Scandinavia, and parts of Russia. Results could differ in the case of open canopy forests, due to the presence for example, or exposed ground covered by snow in winter. They would also most probably be different in the case of deciduous conifers, such as the Siberian larch forests.

3.8 Conclusions

Though this study, we were able to demonstrate that CHMs derived from a TanDEM-X DSM and a lidar DTM vary slightly due to changes in the acquisition conditions. First, when two interferometric pairs are acquired in nearly identical conditions over a short interval, the resulting CHMs are very similar, though not identical, thus demonstrating the relative consistency of solutions in invariable conditions (RMSE of 1.27 m). Secondly, large differences in baseline clearly affect the DSM estimation (RMSE of 3.29 m). Combined baseline and incidence angle create differences in the height of ambiguity (HoA). When HoA is not optimal (i.e., too small or too high), uncertainty in the CHMs was observed. Moreover, the conditions, such as phenology or weather, seemed to have a certain impact, but isolating the effect of each of these variables from the HoA effect, has proved to be difficult to achieve in this study. However, CHMs differences remain at such a low levels that it should allow mosaicking DSMs, and pooling calibration field plots together to predict the canopy height over large boreal forest areas, with very low local biases (not exceeding 3.4% of the canopy height in this study). This opens up the possibility of mapping a large proportion of closed-canopy forests dominated by evergreen conifers using this approach.

Foreword to the third research paper

Having studied the effect of the acquisition conditions on TanDEM-X DSMs (Chapter III), and having learned that, except for cases where the interferograms were created with very short baselines, the worldwide DSM mosaic can be used for creating CHMs, at least for closed canopy boreal forest, we turn our attention to the next chapter (Chapter IV) towards finding a spaceborne alternative to ALS for obtaining a DTM. For this, we have explored the correction of SRTM DEMs to bring them down to a quasi-DTM level. We then use this DTM, in conjunction with a TanDEM-X DSM, to create an approximate CHM used as one predictor for forest biomass. The set of biomass predictive variables was also comprised of the TanDEM-X coherence images, and vegetation indices derived from Landsat 8 images. The accuracy of the biomass map generated with these methods is compared to a highly accurate biomass map generated using ALS and calibrated using dense field plots.

CHAPTER IV

MAPPING BOREAL FOREST BIOMASS FROM A SRTM AND TANDEM-X BASED CANOPY HEIGHT MODEL AND LANDSAT SPECTRAL INDICES

This chapter was submitted to International Journal of Applied Earth Observation and Geoinformation as:

Y. Sadeghi, B. St-Onge, B. Leblon, J.F. Prieur, M. Simard, "Mapping boreal forest biomass from an SRTM and TanDEM-X based canopy height model and Landsat spectral indices,"

4.1 Résumé

Un modèle pour le calcul de la biomasse (Mg ha^{-1}) en forêt boréale à partir d'un MHC obtenu via les capteurs spatioportés à RSO (TanDEM-X et SRTM) a été élaboré. Le MNA du SRTM brut est corrigé pour les effets topographiques ainsi que la densité de la couverture végétale. Ce quasi-MNT du SRTM est ensuite utilisé avec un MNS de TanDEM-X pour produire un MHC RSO ('MNS de TanDEM-X' - 'quasi-MNT du SRTM'). La courbure ($r^2=0.29$) et le GNDVI ($r^2=0.18$) sont les variables qui sont le plus liées à l'erreur du SRTM (MNA du SRTM - MNT du lidar) et sont utilisées pour corriger le MNA du SRTM. Le MHC RSO est ainsi comparé au MHC du lidar. Une EMQ de 2.45 m, un biais de 0.07m et un r^2 de 0.43 ont été obtenus au niveau du peuplement forestier. Le modèle de calcul de la biomasse utilise le MHC RSO, la cohérence, le NDVI, l'humidité et la brillance avec un classificateur de type Random Forest. En comparaison avec le lidar, une EMQ de 25 Mg ha^{-1} , un biais de 0.19 ton/ha

et un r^2 de 0.64 ont été observés. Les variables spectrales proviennent d'imagerie Landsat 8 en été alors que la cohérence provient de TanDEM-X.

4.2 Abstract

We propose a method for mapping the biomass (Mg ha^{-1}) of boreal forests based principally on a canopy height model generated using interferometric synthetic aperture radar (InSAR), namely from the Shuttle Radar Topographic Mission (SRTM), and TanDEM-X mission, as well as Landsat images. The initial SRTM digital elevation model (DEM) was corrected by modelling the respective effects of landform and land cover on its errors and then subtracted from a TanDEM-X DSM to produce a SAR canopy height model (CHM). Among landform factors, terrain curvature had the largest effect on SRTM elevation errors, with a r^2 of 0.29. The NDVI was the best predictor of residual SRTM land cover error, with a r^2 of 0.18. The final SAR CHM (TanDEM-X DSM minus corrected SRTM) had a 2.45 m RMSE, with 0.07 m bias, compared to a lidar CHM. A biomass prediction model was developed based on a combination of the SAR CHM, TanDEM-X coherence, Landsat 8 NDVI, and other vegetation indices. The best results were obtained using random forest regression, at the stand level, yielding a RMSE of 26 Mg ha^{-1} (34% of average biomass), with a r^2 of 0.62. This method has the potential for creating spatially continuous biomass maps over entire biomes, using only spaceborne sensors, and requiring only low-intensity calibration.

4.3 Introduction

Despite intensive research efforts devoted to understand the role of vegetation in the global carbon (C) cycle, the spatial distribution of the above-ground biomass density of forests (expressed in Mg ha^{-1} , and hereafter simply termed “biomass”) remains uncertain (Neigh et al., 2013; Hall et al., 2011; Gonzalez et al., 2010). Average biomass varies greatly between biomes, being estimated for example at 390, 270, and 83 Mg ha^{-1}

¹ respectively for tropical, temperate, and boreal forests (Houghton et al., 2012; Houghton 2010; Houghton et al., 2009). Documenting its changing spatial distribution within biomes is a challenging task. Several researchers have attempted to produce biomass maps over entire biomes at spatial resolutions varying from 500 m to 1000 m (Saatchi et al., 2011.b; Baccini et al., 2012; Lefsky 2010). For such large extents, only remote sensing from orbital platforms can provide the necessary data in a timely manner and proper resolution (Lindberg et al., 2012; Hall et al., 2011; Le Toan et al., 2011). There is, however, no scientific consensus on an optimal approach, and results from studies employing similar methods sometimes strongly disagree (Ometto et al., 2014; Houghton et al., 2001).

Three-dimensional (3D) remote sensing techniques for mapping forest biomass, as opposed to image intensity- or backscatter-based strategies, should in theory constitute an effective approach because of the strong relation between forest height and biomass (or wood volume) in closed forests (Molto et al., 2014; Feldpausch et al., 2012; Dubayah et al., 2010; Kellner et al., 2009; Pflugmacher et al., 2008). For their implementation, the development of a large-scale modeling capacity of forest canopy height from space is critical. Airborne laser scanning (ALS, often simply referred as lidar) has shown its efficacy for mapping forest biomass at very high resolution (e.g. 20 by 20 m cells) with great accuracy (relative root mean square error [RMSE]) sometimes as low as 15% (Hyypä et al., 2012; Benoit et al., 2008; Næsset 2002). To achieve such result, predictive models relying on ALS-derived canopy height (canopy surface elevations minus terrain elevations) are calibrated using field measurements of biomass in ground plots (Theodor Ene et al., 2017; Næsset 2002; Næsset and Gobakken 2008). Spatially continuous ALS coverages are not available for entire biomes however. As there are no currently funded projects for spaceborne laser scanning pertaining to forestry, these coverages could not be rapidly updated even if they existed. Therefore, alternative methods based on satellite data are being sought after. Spaceborne multispectral images of reflected energy, or radar backscatter images, have

been tested for estimating biomass based on statistical relationships with ground plots (Lu 2006; Goetz et al., 2009; Næsset et al., 2016). One of these attempts led to the first large-scale biomass map of African tropical forests in 2008 using MODIS (Moderate Resolution Imaging Spectroradiometer) images. The RMSE of this map was estimated to be of 50.5 Mg ha⁻¹ (Baccini et al., 2008). However, the relationship between image intensity (reflected or backscattered energy) and biomass are non-linear over the range of possible biomass levels. For example, long wavelength SAR backscatter (P-band) saturates at biomass levels of approximately 300 Mg ha⁻¹, and short wavelength (C and X-band) at around 100 Mg ha⁻¹ (Saatchi et al. 2011.a, Ahmed et al. 2014, Imhoff 1995.b). Similar relationships occur when using optical imagery (Landsat, MODIS, etc.). Therefore, models relying solely on intensity are not invertible at higher biomass levels. Spectral information has however proved useful when combined to ALS-derived forest heights (Baccini et al. 2008, Saatchi et al. 2011.b). The r^2 and RMSE of biomass predictions were for example improved respectively from 0.57 to 0.73, and 66.0 % to 14.9 % based on only Landsat 8 spectral information (Karlson et al. 2015), and a combination of forest height from TanDEM-X and spectral information of EO1-Hyperion (Kattenborn et al. 2015). Therefore, large-scale forest height maps combined with widely available spaceborne imagery could provide the necessary data for biomass mapping over entire biomes. Although large-scale spatially continuous lidar coverages are still out of reach, sparse spaceborne lidar data and multispectral imagery combination were attempted. Saatchi et al. 2011.b used data from the orbital GLAS instrument (Geoscience Laser Altimeter System, onboard the ICESat satellite, decommissioned in 2010) combined with MODIS imagery and SRTM (Shuttle Radar Topographic Mission) data to create a biomass map over tropical forests with a 100 ha pixel size. The overall uncertainty of 30% was attributed mainly to forest height errors. A similar approach was employed independently by Simard et al. (2011), and Lefsky (2010), leading to forest biomass maps showing large discrepancies. Since these studies were published, significant progress was made in radar interferometry sensors and methods, opening up new possibilities for producing continuous maps of forest heights

globally, regardless of cloud cover. InSAR (Interferometric Synthetic Aperture Radar) uses the phase differences between two SAR images acquired either simultaneously (single-pass) or in two passes (repeat-pass) to estimate the elevation of terrestrial objects, or the height above ground of forests (Liu et al. 2008, Thirion-Lefevre and Colin-Koeniguer 2007, Balzter et al. 2007.b, Izzawati et al. 2006).

Interferometric coherence, a parameter normally used to evaluate the accuracy of the phase of repeat-pass interferograms, has also been employed to estimate forest height (Kugler et al., 2014; Askne et al., 2013; Praks et al., 2012; Hajnsek et al., 2009). The prediction of forest height from coherence is based on the inversion of the RVoG (Random Volume over Ground) model. A single-pass system with full polarimetry and long wavelength (L and P-band) is needed for applying this approach in closed forests. However, due to lack of such of a single system, using two interferometric systems with a different wavelength is the alternative approach. For a given forest structure, the SPC (Scattering Phase Centre) is closer to the canopy surface at short wavelengths (e.g., in X-band) and deeper within the canopy at longer wavelengths (e.g., in C-band). In three studies (Balzter et al. 2007.a-b; Neeff et al. 2005), a combination of short and long wavelengths was used to produce forest height by subtracting the InSAR height extracted with L or P bands (ground signal) from the InSAR height based on X-band (surface signal). Spaceborne imaging interferometers in the L or P bands do not yet exist, and the first such potential sensor in L-band, TanDEM-L (DLR) is still in early funding stages (Irena Hajnsek, *personal communication*; TanDEM-X Science Coordination, Deutsches Zentrum für Luft- und Raumfahrt – DLR). The TanDEM-X mission, the only current single-pass InSAR system, has recently completed a series of acquisitions leading to the generation of a global high-resolution DSM (12 m resolution). The accuracy of forest height extracted by combining TanDEM-X DSMs with ALS DTMs was estimated to be around 0.8 m to 5.0 m (RMSE error) depending on forest type and acquisition conditions (Sadeghi et al., 2014 and 2016; Schlund et al.,

2015 and 2016; Solberg et al., 2013 and 2015.a) and in areas where topography did not cause greater interferometric errors.

The problem of obtaining a global DTM still remains. Currently, the only available height surface under the forest canopy and relatively close to the terrain surface is the SRTM DEM, produced using a single-pass C-band imaging interferometer. However, in densely vegetated areas the penetration of C-band SAR signal through the canopy is often only partial, i.e. the SAR SPC is located above the ground, but below the vegetation surface (Bourgine and Baghdadi 2005; Simard et al., 2006; Kelldorfer et al., 2004). The amplitude of this over-estimation depends, among other factors, on the height and density of vegetation.

Several studies investigated the SRTM error in vegetated areas, using reference height data GPS (Rodriguez et al., 2006), ALS (Su et al., 2015), LVIS (Laser Vegetation Imaging Sensor, Hofton et al., 2006), or GLAS (Carabajal and Harding 2006; Bhang et al., 2008), with errors ranging from 6.2 m to 22.4 m depending on the vegetated areas structure. It follows that the SRTM DEM needs to be corrected for the overestimation caused by vegetation before being considered as a DTM (Farr et al., 2007). Several methods for performing such a correction have been proposed, but most assume a uniform vegetation-induced error, which is not the case (Coe et al., 2008; Paiva et al., 2013). More recent approaches by Su et al. in 2014 and 2015 modeled the vegetation error as a function of land cover, which resulted in an improved DEM with an overall bias of less than 1 m compared to a corresponding lidar DTM. Moreover, Shortridge and Messina (2011) also found a relation between local terrain slope and SRTM error. Other authors suggested that terrain curvature, or local incidence angle of SRTM pulses, may also be correlated with SRTM errors (Wilson et al., 2007; Coe et al., 2008; Paiva et al., 2013). Correction methods for landform were for example proposed by Rodriguez et al (2006), Castel and Oettli (2008). Whether for correcting the effect of vegetation or landform, ancillary data, such as land cover maps or data on topography

needs to be acquired for implementing a correction method. In the case of topography, an apparent problem with circular logic arises, as topographical data is needed to correct a DEM. This problem could theoretically be circumvented if a DSM could be considered sufficiently correlated to the underlying DTM for extracting general landform metrics. In this paper, we use a TanDEM-X DSM to represent the top of the canopy elevations, and SRTM elevations corrected for vegetation and landform effects as a pseudo-DTM. A CHM is computed based on the difference between these two layers. The CHM is then used as a predictor of boreal forest biomass, along with other biomass predictors. Our goal is to show that the ancillary information needed to calculate the vegetation density and topographical corrections can be obtained from Earth observation satellites, respectively Landsat 8 and TanDEM-X itself. Lidar data is first used to calibrate this method, and then later in assessing the accuracy of the results, but is not an integral part of the biomass mapping strategy. The overarching goal of this study is to design a new method for mapping forest biomass for entire biomes using only spaceborne sensors, provided local elevation calibration data is available.

4.4 Study area and data

4.4.1 Study area

The Montmorency forest, a research site North of Quebec City, Canada (centered at 47°18' N - 71°08' W), was selected for this study (Figure 4.1). It covers 6 600 ha of boreal forest composed mostly of balsam fir (*Abies balsamea* (L.) Miller), and paper birch (*Betula papyrifera* Marshall). The forest inventory map of this site divides it in more than 3 000 stands, most of which are a pure conifer, with some mixed stands (conifer/deciduous). Pure deciduous stands are quite rare in this area. The topography is hilly, with elevations that vary between 600 m to 1000 m, and slopes that can locally

reach 53°. The study area is covered by snow from December to April, and the average annual temperature is 0.3° (Bélanger et al. 2001).



Figure 4.1 Location of the study site

4.4.2 Reference data

Two hundred well distributed permanent circular plots having a radius of 11.28 m fall within the study area. These are re-measured every five years, with 20% of plots visited annually. The plot database containing all dendrometric values was acquired in winter of 2016, thus covering all inventory years up to 2015. Within each plot, the DBH (diameter at breast height) and species of all trees with a DBH ≥ 9 cm are measured, as well as the height of three randomly selected dominant trees. Table 4.1 presents general statistics on the main structural parameters of these 200 plots. The plot geolocations were recently measured with an SX-Blue GNSS receiver having an estimated error not exceeding 1 m. The forest inventory reference data was also comprised of the stand map which was used for the sole purpose of averaging certain

predictions standwise (see Methods section). The photo-interpreted stands of this map had an individual area of 0.5 ha and 4 ha.

Table 4.1 General statistics of the main structural attributes of the 200 field plots of the Montmorency Forest

| | Tree DBH (cm) | Tree height (m) | Plot density (stems/ha) | Plot basal Area (m ² /ha) |
|---------|---------------------|-----------------------|-------------------------------|--|
| Minimum | 2.0 | 1.9 | 25 | 0.1 |
| Maximum | 52.1 | 32.3 | 13000 | 15.6 |
| Mean | 13.5 | 11.6 | 116 | 0.5 |
| SD | 5.5 | 3.9 | 506 | 0.7 |

Airborne laser scanner data used for calibration and validation was collected in August of 2011 using an ALTM 3100 system from Optech Inc. The resulting density of the first returns was approximately 5 points/m². Ground returns were identified by classifying the point cloud using TerraScan (Terrasolid, Helsinki). The field height of more than 400 trees was compared to the corresponding lidar heights, yielding a r^2 of 0.93 and a RMSE of 1.23 m (see Sadeghi et al. 2016 for further details). This indicated that both the lidar DTM and DSM were accurate.

4.4.3 Satellite data used to generate predictive variables

A 30 m resolution SRTM DEM (version SRTMGL1) created by the National Geospatial-Intelligence Agency (NGA) and Jet Propulsion Laboratory (JPL, NASA) was acquired for the study area. This elevation data was derived from single pass SAR interferograms acquired in February 2000 in C-band at a wavelength of 5.6 cm (Farr et al. 2007). The SRTMGL1 version is corrected for all modelable errors and is considered as the definitive version of the SRTM DEMs (NASA_SRTM_V3, USGS).

A TanDEM-X interferometric image pair created in a strip-map mode with a baseline of 107.54 m and a height of ambiguity (HoA) of 43.56 m was acquired for the purpose of this study in August of 2013. The raw single image resolution was 1.2 m and 6.6 m respectively in range and azimuth direction. The TanDEM-X images in slant range complex (SSC) format were processed using ENVI SARscape 5.0 functions to extract height information (interferogram generation, flattening, and phase unwrapping). We have initially extracted the elevations at 5 m resolution, which we then aggregated to 20 m for the biomass prediction experiments. The resulting TanDEM-X height surface was demonstrated to be located at the top of the canopy, meaning it is analogous to a DSM (Sadeghi et al. 2016, Solberg et al. 2015.a).

Landsat images were acquired to describe the land cover at two different moments in time: 1) close to the acquisition of SRTM data (February 2000), and 2) close to the acquisition of TanDEM-X data (winter and summer of 2013, see Table 4.2 for exact dates). The Landsat scenes were obtained in orthorectified geometry, and in their surface reflectance version (Vermote et al., 2016). The Landsat 7 ETM+ image from March of 2000 serves the purpose of documenting the state of the land cover, with a focus on vegetation, at the time of SRTM acquisition in order to model its effects on the elevation error. The Landsat 8 OLI images acquired in 2013 were chosen in order to be concomitant with the TanDEM-X interferogram. The winter image was used to map areas of very low or absent vegetation based on the evidence of ground snow (see Methods section). The summer image was used to compute vegetation indices as additional predictors of biomass. All images were resampled to a 20 m pixel size from their original resolution. Table 4.2 summarizes the characteristics of all the remote sensing data used in this study.

Table 4.2 Summary characteristics of the remote sensing data

| Dataset | Provider | Date | Native Resolution (m) |
|------------------------|--|-------------------------|-----------------------|
| Optech ALTM 3100 laser | Optech Inc., Vaughan, ON, Canada | 2011, 6 and 9 August | 0.25 |
| SRTM | NASA/USGS | 2000, 11-22 February | 30 |
| TanDEM-X | DLR | 2013, 15 July | 5 |
| Landsat 7 | NASA/USGS | 2000, 10 March | 30 |
| Landsat 8 | NASA/USGS | 2013, 15 January | 30 |
| Landsat 8 | NASA/USGS | 2013, 12 July | 30 |

4.5 Methods

4.5.1 Calculation of field plot biomass

The total aboveground biomass of trees in each plot was predicted using the species-specific equations found in Lambert et al. (2005). A standard error of 4.5% for estimating biomass for each single tree using these equations has been reported. These equations use DBH and height (H) to predict biomass. Since only the DBH was available for all trees, height also had to be predicted. Using the trees for which both DBH and H were measured in the 200 field plots, a species-specific model was calibrated for predicting H from DBH. The DBH and H were first adjusted to their 2013 value (TanDEM-X Acquisition year), by interpolation (for those plots measured before and after 2013) or extrapolation of previous growth (for the plots measured for the last time before 2013). The measured DBH and a predicted H were input into the

equations mentioned above to predict tree-wise total above-ground biomass. These values were summed for each plot, and biomass density (Mg ha^{-1}) was obtained by simply dividing this total by 400 m^2 .

4.5.2 Creation of lidar biomass maps

Because our aim was to evaluate the accuracy of InSAR-based biomass predictions over the entire study area, we needed a spatially continuous reference map of biomass density having both high resolution and high accuracy. We prepared this map using airborne lidar data and following the classical area-based approach (e.g. Næsset 2002; Lim and Treitz 2004). We used the USDA Forest Service's (USFS) FUSION software to extract lidar points for each plot and to calculate plotwise metrics (height percentiles, etc. - see the FUSION documentation for further details). Parsimonious subsets of metrics that were highly correlated with the reference biomass data, but that were not too inter-correlated ($r < 0.8$) were retained. Multiple linear regression models based on these metrics and calibrated using the field plot data were then applied to the points in every $20 \text{ m} \times 20 \text{ m}$ cell of the study area to create a reference map of biomass. In addition, a map of average biomass per stand was created by averaging the 20 m resolution predictions standwise.

4.5.3 SRTM DEM correction

Prior studies have suggested, and in some case demonstrated, that the error in the SRTM data, defined as the deviation from a "true" representation of the ground elevation (generally provided by airborne lidar) is caused by geometrical factors as well as by the vegetation (Guth 2006). Therefore, this error can be corrected to a certain degree. Our aim was to produce a 20 m approximate DTM by resampling and correcting the original SRTMGL1 elevations. Based on Prieur (2016), we contend that the difference between the (resampled) SRTM elevations and a lidar DTM used as reference are caused principally by 1) terrain curvature, and 2) the presence of treed vegetation bearing leaves in February 2000. The curvature effect stems from the

difference in resolution between the 20 m lidar DTM and the true resolution (size of the smallest resolvable 3D element) of the SRTM elevations. This resolution has been estimated to be around 45 m (Guth 2006), despite the SRTM products being delivered at 30 m resolution. It is to be expected that the average elevation within say, a 45 m x 45 m SRTM patch, will be higher than the 20 m lidar elevation at the center of this patch in the case of convex terrain, and lower in the case of concave terrain. Resampling SRTM pixels to 20 m on rectilinear (curveless) terrain should not lead to discrepancies relative to the lidar DTM. In other words, the elevation of crests is expected to be underestimated, while that of valleys should be overestimated. Although strongly convex or concave terrain covers only a fraction of typical hilly landscapes, this effect must be taken into account in the correction process. We also postulate that the curvature of a DSM closely follows that of the underlying terrain as the topographical variations have much bigger amplitude than the forest canopy height variations. Therefore, if the relationship between curvature and error in the proposed resampling process can be first established using calibration data, curvature could be measured on the TanDEM-X DSM and then used to correct the resampled SRTM DEM elevations. Furthermore, this effect and the correction factors should be fairly constant between different landscapes, lessening the need for intensive calibration, i.e., the model developed based on this study site should be applicable to other sites without the need for more reference data. In this study, we used lidar data to calibrate the relationship between SRTM error and curvature. For this, curvature (C) was calculated by fitting a six parameter ($a-f$) polynomial surface (Eq. 4.2, Hurst et al. 2012) onto local neighborhoods of the TanDEM-X DSM [$z(x, y)$]:

$$z = ax^2 + by^2 + cxy + dx + ey + f \quad [4.1]$$

$$C = 2a + 2b \quad [4.2]$$

where a-f are coefficients, and x and y are the planimetric position coordinates. Because other studies have suggested that slope or LIA can affect the SRTM errors (e.g. Shortridge and Messina 2011), we have calculated slope (S) based on Eq. 4.3:

$$S = \sqrt{d^2 + e^2} \quad [4.3]$$

The local incidence angle data for the SRTM mission is not available, to the best of our knowledge. The local incidence angle was therefore estimated from local terrain slope and aspect (A , this parameter is calculated based on Spatial Analyst Tools in ArcGIS 10) using Eq. 4.4 (Castel and Oetli 2008).

$$\text{LIA} = (\cos S \cos \theta - \sin S \sin \theta \cos(A - \varphi)) / \cos \theta \quad [4.4]$$

where θ is the SRTM average beam incidence angle (54.5°), and φ is the SRTM azimuth (147°).

A moving window was centered on each TanDEM-X pixel, and curvature and slope were computed after fitting the polynomial (Eq. 4.1) by least squares to the TanDEM-X DSM using the `scipy.linalg.leastsq` function of the Python programming language (*scipy*). Window sizes from 15x15 m to 145x145 m were tested and the size providing the highest correlation between landform attributes and SRTM error was kept for building an SRTM error prediction model. A simple linear regression was used for the relation between landform attributes (curvature, slope, and local incidence angle) and SRTM error (SRTM DEM – Lidar DTM), and the RMSE and R^2 were calculated. This yielded raster layers containing the predicted amount of error caused by landform. The SRTM elevations were then adjusted by subtracting this error layer. Figure 4.2 summarizes the final process for curvature (see Results section for explanations on the exclusion of the other landform parameters).

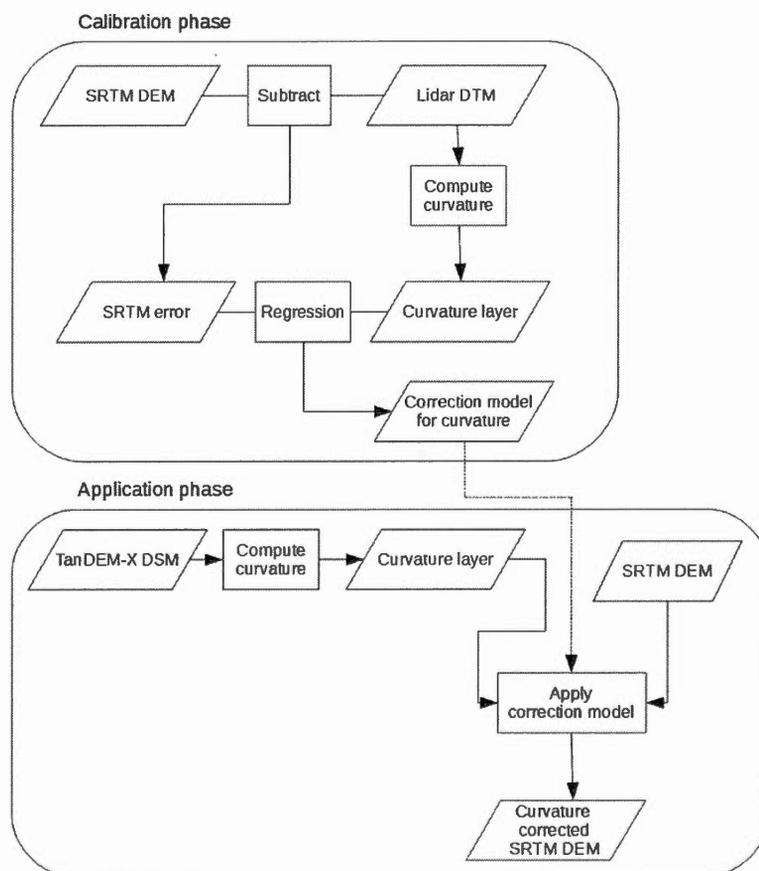


Figure 4.2 SRTM correction process for landform

As previously mentioned, land cover is also an important source of elevation errors in SRTM DEMs (Baugh et al., 2013). Radiometric information from Landsat images was used to determine the degree to which vegetation causes overestimation of the ground elevations. The Normalized Difference Vegetation Index (NDVI, Tucker 1977), the Green Normalized Difference Vegetation Index (GNDVI, Gitelson et al., 1996), and the Green Vegetation Index or Tasseled Cap (GVI, Kauth and Thomas 1976), were computed respectively in winter and summer seasons of y. 2000. The Normalized Difference Snow Index (NDSI) in winter season (during the presence of ground snow)

was calculated, also from Landsat calibrated surface reflectance values. It is, on the one hand, hypothesized that the various vegetation indices would be correlated to a certain degree to the amount of vegetation, and therefore to the amount of overestimation caused by the forest layer acting as a screen for incident microwave pulses (Simard et al., 2011; Su et al., 2015). On the other hand, because deciduous trees are leafless in winter, we surmised that the degree to which ground snow can be seen through the canopy could be related to the "transparency" of the forest layer to SAR pulses, and therefore, to the amount of elevation overestimation. Following that theory, a pure deciduous stand, leafless in winter, should allow for SRTM pulses to travel through it to a great extent, and the radiometry of the corresponding Landsat pixel should be greatly influenced by visible ground snow, i.e. be relatively bright compared to pixels of mixed or coniferous stands. The same logic applies to the various densities of (evergreen) coniferous trees. These variations should be represented to some extent in the y. 2000 NDSI values. All spectral indices were regressed against the residual SRTM error (the error left after correction for curvature effects). The best predictor was then used to correct for the overestimation caused by vegetation, yielding the final corrected DEM, assumed to be an approximate DTM.

4.5.4 Creation and validation of the SAR CHM

A SAR CHM was created by first subtracting the final corrected SRTM DTM from the TanDEM-X DSM. Then, areas with a high NDSI value in the winter of 2013 (year of the TanDEM-X Acquisition) were considered to be devoid of vegetation and used to adjust the CHM to zero height under the hypothesis that a continuous and bright snow cover reveals the absence of a forest layer at the time the TanDEM-X data was acquired. The accuracy of this SAR CHM was evaluated at a resolution of 20 m, and at stand level, by computing the coefficient of determination between the lidar and SAR CHMs, as well as the bias (average height difference) and RMSE values. At 20 m resolution, a sample of 3310 cells was used to assess the accuracy, a number equal to that of stands.

4.5.5 Biomass prediction from remote sensing variables

Biomass prediction was performed at the two above resolution levels using the SAR CHM, SAR coherence, and vegetation indices such as NDVI, GVI, GNDVI, Enhanced Vegetation Index (EVI, Jiang et al., 2008), Soil Adjusted Vegetation Index (SAVI, Huete 1988), Difference Vegetation Index (DVI, Richardson and Wiegand (1977)), Green-Red Vegetation Index (GRVI, Motohka et al., 2010), Leaf Area Index (LAI, Duchemin et al., 2006), wetness, greenness, and brightness (Franklin et al., 2002) indices extracted from the Landsat image.

Two types of predictions models were developed, respectively using linear regression and random forest regression (both using R statistical software package (R Core Team 2005)). The random forest modelling used the default value for *n*tree of 500 and *m*try of 3. The best biomass predictors were selected for linear and random forest regressions based on the backward elimination method (Powell et al., 2010). This method starts with all predictors and then progressively eliminates the weaker variables. The best model was applied to the selected predictors, and the r^2 , RMSE, and %RMSE values between the predicted and the observed biomasses (based on ALS) were calculated. Predictions were only made in areas where the TanDEM-X DSM was deemed free of gross errors caused by radar layover or very low interferometric coherence. Areas corresponding to anachronic forest harvests (clear cuts having occurred during the time interval between the respective acquisitions of lidar and TanDEM-X data) have also been removed. In total, 99% of the study area could be used.

4.6 Results

Table 4.3 presents the biomass density statistics of 191 field plots of the used study area, as estimated using Lambert et al. (2015) prediction models.

Table 4.3 Plot level biomass statistics

| | Minimum | Maximum | Mean | SD |
|------------------|---------|---------|-------|-------|
| Biomass (ton/ha) | 0.81 | 306.06 | 87.29 | 54.67 |

The best FUSION lidar metrics for predicting biomass density were *Elev.mean* (the point cloud mean elevation value) and *Elev.IQ* (the point cloud interquartile range elevation value), yielding a r^2 of 0.90, a RMSE 17.9 Mg ha⁻¹, which translates to 20% RMSE (Figure 4.3). The resulting model was applied to the entire lidar CHM to create a reference biomass map at a resolution of 20 m, and at the stand level.

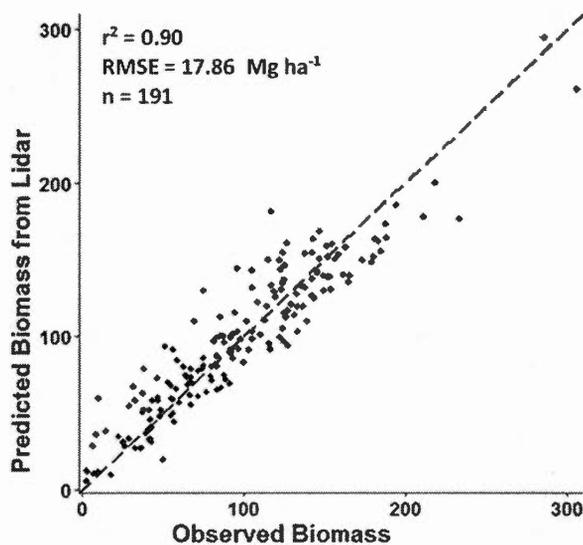


Figure 4.3 Lidar-predicted vs. Observed biomass for 191 field plots

The resulting height range of the SAR DSM (476-980 m) was comparable to that of the reference lidar DSM, i.e. 468-979 m (Figure 4.4c).

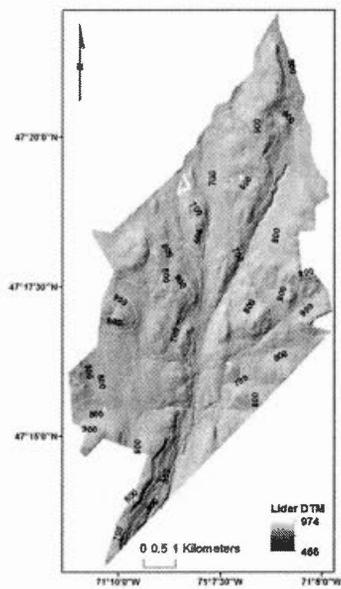
The extraction of the landform variables involved the calculation of local curvature, slope, and LIA using an optimal window size, which was found to be 75x75 m based on its capacity to predict the amount of SRTM height error. In the best case, a r^2 of 0.29 was obtained between curvature and the height error. This relationship behaved as expected: negative curves (convex) were associated with negative height error (SRTM DEM – lidar DTM), and vice versa. The slope and LIA variables, however, presented very weak relationships with $r^2 = 0.050$, and $r^2 = 0.003$, respectively, so they were dropped from the SRTM DEM correction process. After correcting for the curvature effect using the optimal model, the corrected DEM showed a 12.9% RMSE improvement compared to the initial SRTMGL1 (30 m resolution).

In the second step, we sought a relationship between the remaining SRTM error and land cover. A simple linear regression model was adjusted between SRTM error and land cover parameters such as NDVI, GNDVI, and DNSI, using the R package. Coefficients of determination of the vegetation indices were 0.24 for the NDVI, 0.30 for the NDSI and 0.25 for the GNDVI in the winter season, but much lower in the summer season, with r^2 s of 0.05 and 0.06 for GNDVI and NDVI, respectively. After applying the effect of land cover parameters (NDSI index) to the curvature corrected SRTM DEM, an improvement of 29% RMSE was achieved compared to SRTMGL1. Table 4.4 presents the characteristics of enhanced SRTM DEM compared to the initial SRTM DEM.

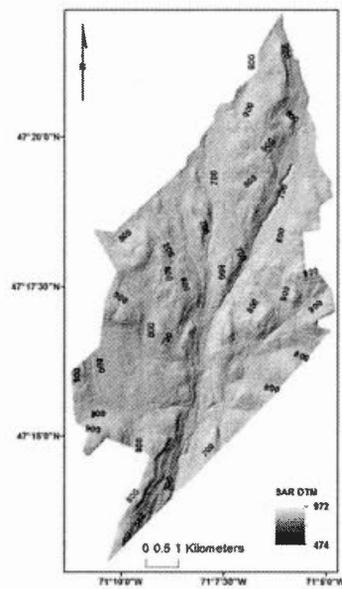
Table 4.4 The statistical summary of the SRTM DEMs relative to the lidar DTM

| | RMSE (m) | RMSE Improvement (%) |
|---|-------------|-------------------------|
| SRTMGL1 resampled at 20 m | 3.1 | |
| After curvature correction | 2.7 | 12.9 |
| After curvature and land cover corrections | 2.2 | 29.0 |

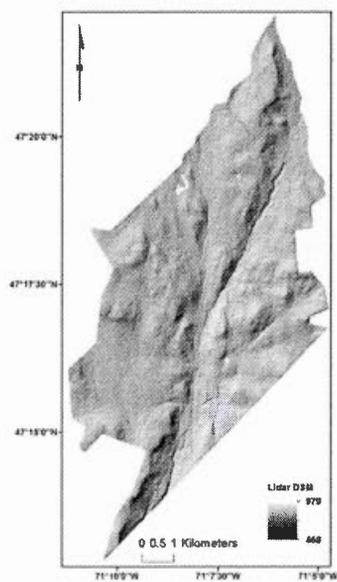
The corrected SRTM DEM, or "SAR DTM" (Figure 4.4b) with a height range of 474-972 m at 20 m resolution, is globally comparable to the reference lidar DTM (Figure 4.4a), which has a height range of 466-974 m. Once the SAR CHM (TanDEM-X DSM – corrected SRTM DEM) was calculated, the mask for non-forested areas derived by thresholding the NDSI values was applied to it to bring the heights under the mask to 0.0 m. The resulting SAR CHM is presented in figure 4.4g. This SAR CHM is generally comparable to the lidar CHM (Figure 4.4e). The patterns in the both CHMs are resembling, but some discrepancies can be seen.



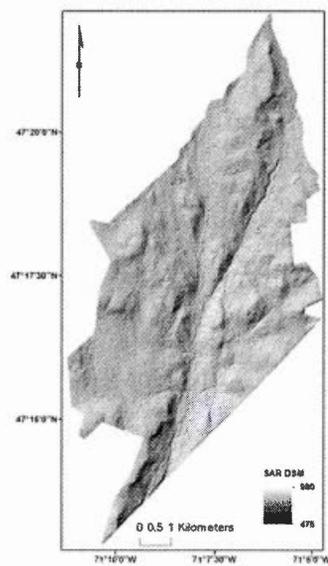
(a)



(b)



(c)



(d)

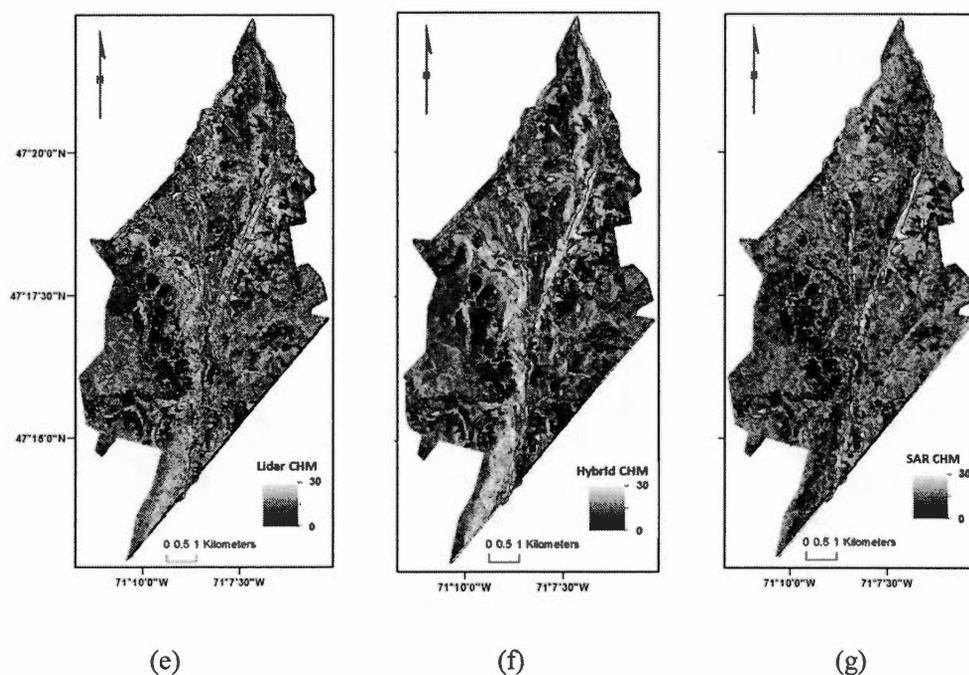


Figure 4.4 The height maps for lidar and SAR (20 m resolution). A) lidar DTM, b) SAR DTM (corrected SRTM DEM), c) lidar DSM, d) TanDEM-X DSM, e) lidar CHM, f) Hybrid CHM (SAR DSM – lidar DTM), and g) SAR CHM (SAR DSM – corrected SRTM). Red outlines in e, f, and g represent areas masked due to gross InSAR errors or anachronic clear cuts.

The relation between the lidar CHM and SAR CHM at 20 m resolution and stand level are presented in Table 4.5 and Figure 4.5. At the cell level, the correspondence is rather low, with a r^2 of 0.18, and a RMSE of 4.11 m. After the data was aggregated at the stand level, the relationship strengthened significantly, reaching a r^2 of 0.43, and a RMSE of 2.45 m. The height bias at stand level was very low, at 0.07 m. Moreover, the scatterplots of figure 4.5 indicate that the relationships are linear, without clear outliers. In figure 4.5b, the SAR CHM with values 0 against of variable height values for lidar is because the stands with a prediction < 0 were given a 0 value.

Table 4.5 Statistical relationships between the lidar CHM and the SAR CHM

| Resolution | Mean CHM | | Bias (m) | SAR CHM = $B_0 + B_1$ Lidar CHM | | r^2 | RMSE (m) | n |
|------------|--------------|------------|-------------|---------------------------------|-------|-------|-------------|------|
| | Lidar (m) | SAR (m) | | B_0 | B_1 | | | |
| 20 m | 6.52 | 8.15 | 1.63 | 5.61 | 0.39 | 0.18 | 4.11 | 3310 |
| Stand | 5.38 | 5.45 | 0.07 | 1.29 | 0.77 | 0.43 | 2.45 | 3310 |

Both models are significant at $P < 0.001$

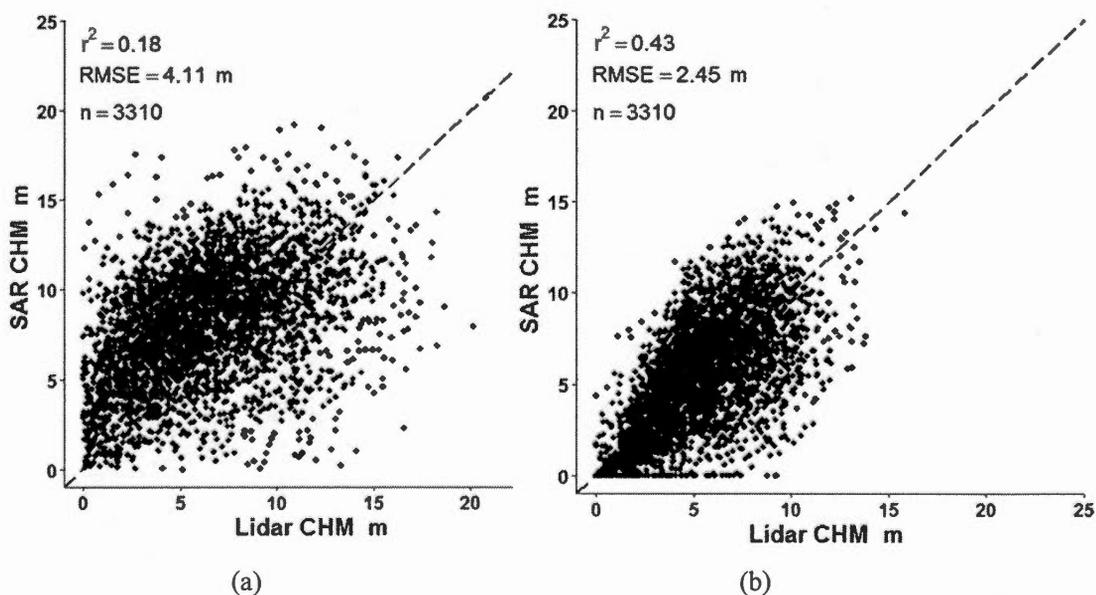


Figure 4.5 Scatterplots of the lidar CHM and the SAR at 20 m (a) and stand (b) resolutions.

The next stage involved combining the SAR CHM, SAR coherence, Landsat bands for summer and winter seasons (b2-b7), Landsat band ratios, Landsat indices such as NDVI, GNDVI, GVI, EVI, NDSI, SAVI, GVI, GRVI, RVI, LAI, as well as the wetness, brightness, and greenness indices, to predict biomass. Table 4.6 and Figure

4.6 present the relationships between the lidar-based predicted biomass and the values predicted using SAR and Landsat-based variables. At stand level, this relationship reached a relatively high r^2 of 0.62 using random forest regression (but only 0.44 using a linear regression). The relative RMSEs remain rather high, at 34% in the best case. The best predictor variables at both resolutions were NDVI and SAR CHM, both for the Model I linear regression and random forest. In the latter case, the SAR CHM was either the first or second variable, depending on the type of importance assessment performed. However, the overestimation at low biomass level and underestimation at high biomass level can be due to the effect of resolution and gap effect on the SAR CHM predictor. Figure 4.7 shows the mean decrease GINI and the mean decrease accuracy of the predictor variables of the random forest model. These two provided a measure of the importance of each input variables in the RF regression. The mean decrease accuracy of a variable is calculated by normalizing the difference between the out-of-bag (OOB) accuracy of the observations. This process is repeated for all variables. The GINI index is calculated by summarizing all decreases for each variable and normalizing by the *n*tree value. The higher values of %IncMSE or IncNodePurity correspond to the more significant variables.

Table 4.6 Statistical relationships between SAR-Landsat and lidar predictions of biomass using linear regression and random forest (RF)

| Resolution | r^2 | RMSE (Mg ha ⁻¹) | RMSE % | <i>n</i> |
|----------------|-------|--------------------------------|-----------|----------|
| 20 m (linear) | 0.27 | 50 | 54 | 3310 |
| Stand (linear) | 0.44 | 31 | 40 | 3310 |
| 20 m (RF) | 0.36 | 42 | 46 | 3310 |
| Stand (RF) | 0.62 | 26 | 34 | 3310 |

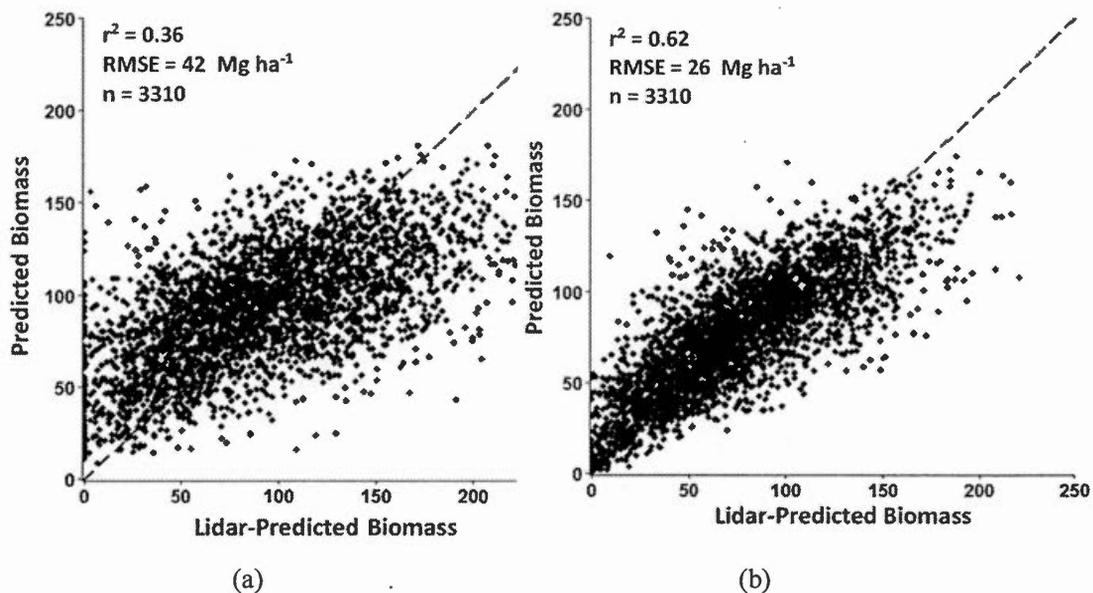


Figure 4.6 The relationship between lidar-predicted and SAR-Landsat biomass at a) 20 m resolution and b) stand level

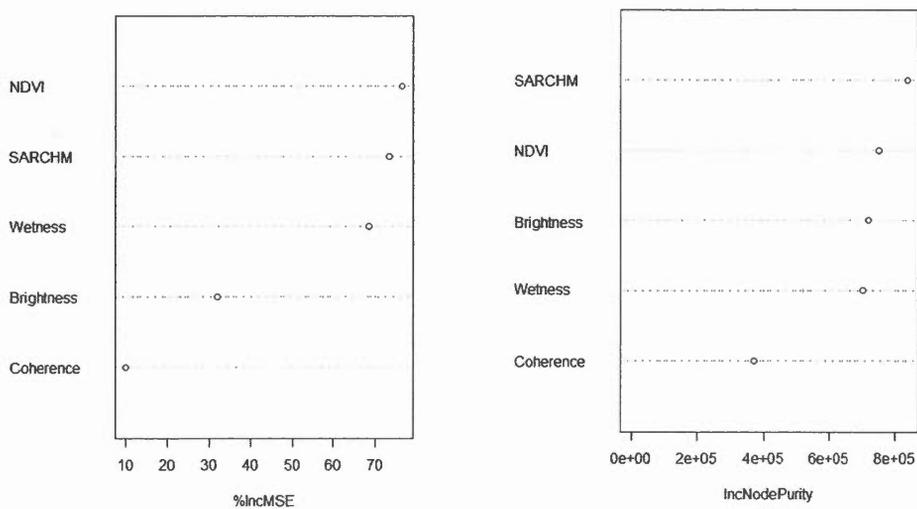


Figure 4.7 Random forest regression, mean decrease accuracy (%IncMSE) and mean decrease Gini (IncNodePurity)

4.7 Discussion and conclusion

A global DSM and DTM were created using InSAR spaceborne sensors (TanDEM-X and SRTM), were combined over the study area to produce an InSAR canopy height model. This model was then used as a predictive variable for mapping forest biomass. Correcting the SRTM DEM for landform and land cover reduced the SRTM RMSE by 29 %, from 3.1 to 2.2 m, therefore improving the accuracy of the CHM. Among three landform-related parameters (local curvature, slope, and incidence angle) only the first was useful for predicting the SRTM error. This is consistent with the hypothesis that resolution effects cause an underestimation of elevations on convex terrain, and vice-versa, but that rectilinear sloping terrain does not need correction (Prieur 2016). The r^2 between curvature and SRTM error (0.29) may seem low, but this must be carefully interpreted. Curvature effects only occur on strongly curved terrain, which covers only a small proportion of the study region. Elsewhere, land cover effects dominate. The overall relationship between curvature and error is thus dampened by the varying overestimation caused by vegetation. However, using a variable window approach to identify the optimal resolution for assessing the effects of landform proved very useful. In many other studies, the default 3x3 pixel local operators available in image processing software were used, leading to a situation where the effect of landform went undetected because it occurs at a different scale. As for the effect of LIA, it may be that our LIA approximation was too crude for finding a relationship between incidence angles and SRTM error. Data from NASA on SRTM incidence angles are unfortunately unavailable, thus preventing a more precise estimation of LIA.

Moreover, the bias caused by the vegetation screen that partially blocks the propagation of the SRTM microwave pulses, could be partially modelled with Landsat images acquired at approximately the same time as the SRTM interferograms. The relationships were somewhat weak, with best r^2 not exceeding 0.3. However, they help to lower the SRTM DEM RMSE error from 2.7 m to 2.2 m. This latter number represents approximately 10% of maximum tree height in the study region. The final

canopy height model remains much coarser and less accurate than those created using airborne lidar but is to the best of our knowledge, the first time a spatially continuous CHM is created based on the difference between a DSM and a DTM acquired entirely using spaceborne sensors.

This approximate CHM proved to be the first or second most useful variable for predicting forest biomass, preceded only, when it was second, by the NDVI derived from Landsat images acquired during the same period as the TanDEM-X image. In this study area, biomass levels top at approximately 300 Mg ha^{-1} , i.e. a level equivalent to the backscatter saturation threshold of long wavelength radar biomass studies, but well above that of short wavelength SAR (100 Mg ha^{-1}). Our data shows a linear relationship between reference and predicted forest height over the full $0 - 300 \text{ Mg ha}^{-1}$ range. These results are encouraging as they suggest that using a CHM among biomass predictors, instead of just the amount of reflected or backscattered energy, might help to overcome saturation effects. The method should, however, be tested on taller forests, such as those of the Pacific Maritime Ecozone of Canada (Wulder et al., 2008) as STRM pulse penetration through these higher canopies may prevent an adequate reconstruction of the DTM, despite landform and land cover corrections. Furthermore, penetration in tall and dense tropical forest is likely much more difficult, and the use of the NDSI snow index in the creation of the CHM cannot be envisaged. The relationship between forest height and biomass is itself non-linear in many situations, as trees continue to grow in width (DBH) even after reaching their top height (Mugasha et al., 2013). Also, there are limitations in using vegetation indices as predictors of either the effectiveness of the vegetation screen in 2000 or the biomass at the time of the TanDEM-X acquisition. As the ground progressively becomes covered by denser vegetation, the NDVI, among other vegetation indices, is expected to increase. However, past a certain density, it will saturate (Wang et al., 2005). It has however here proved useful when combined to the SAR CHM.

Resolution played a big role in the accuracy of the results. The 20 m resolution is useful for calibrating the ALS-based prediction model, because this cell size is similar to that of typical forest inventory field plots in Canada (400 m²), and other countries (Gobakken and Næsset 2008). However, biomass predictions at a resolution of 20 m were noisy and could be markedly improved by aggregating them at the stand level. The cancellation of positive and negative errors at this coarser level, decreasing scatter, is the likely cause of this improvement. We consider that the stand level is a fine enough resolution for mapping biomass or changes therein. Alternatively, larger square cells (e.g. 100x100 m) could be used, if stand polygons are not available.

In the best case (random forest regression applied at stand level), biomass predictions had a RMSE of 26 Mg ha⁻¹, which represents 34% of the average biomass in the study region. This is similar, for example, to the value reported by Saatchi et al. (2011.a) for tropical forests where average biomass is at least four times higher. In the Tanzanian forest, biomass was modeled using Landsat 8 NDVI, with a RMSE of 44 Mg ha⁻¹, i.e. 49% (Gizachew et al., 2016). Moreover, by using the InSAR coherence of TanDEM-X as the sole predictor to estimate tropical forest biomass, Treuhaf et al. (2015) obtained a RMSE varying between 52 and 62 Mg ha⁻¹ (29% to 35%). It is difficult to predict if the %RMSE associated to our method would drop for forests having a much larger biomass because random errors would remain at their current level, or if these errors would also increase. In any case, we must keep in mind that the accuracies that we have been reported are calculated relative to lidar predictions of biomass, which themselves are already in error by 17.9 Mg ha⁻¹, compared to field plots (at 20 m resolution).

The method we have here described relies on spaceborne remote sensing data but requires some amount of calibration data, notably from field plot databases, and lidar. The curvature correction model for the SRTM data should in principle be valid for the entire SRTM DEMs, thus requiring just a single, definitive, and low-intensity

calibration phase. Land cover correction, however, is likely determined by forest types because of different penetration rates of C-band pulses depending on vegetational characteristics. For this reason, at least one calibration per ecozone is likely needed. This calibration is however required only once, since the SRTM data can be considered as a snapshot, having been acquired over a very short period. The relationship between the SAR CHM and biomass should also be quite stable within a given ecozone. We have demonstrated in Sadeghi et al. (2016.b), as also showed by Solberg et al. (2015.a), that TanDEM-X DSMs, at least in the boreal zone, are not markedly affected by acquisition conditions. The relationship between the top of canopy elevations and the TanDEM-X elevation should, therefore, be quite stable throughout the WorldDEM dataset for given forest types, and could likely be calibrated once and for all, per ecozones to account for specific forest canopy conditions (such as density).

Improvements to the method we have here proposed should first target the correction of the SRTM DEM, or more broadly, should aim at finding a better source of terrain elevation data. While options exist for bettering the land cover correction, such as analyzing a 1984-2000 time series of Landsat images to predict better the amount of vegetation acting as a screen to SRTM pulses in y. 2000, future InSAR sensors could also be considered. The most promising is the TanDEM-L mission, currently being designed by DLR (Moreira et al., 2015). The use of the L-band would provide a much-improved penetration through forest canopies, albeit still not complete. The extraction of a DTM using that sensor could be carried out using a similar procedure, i.e. through landform and land cover correction, but would produce a much-improved bare earth elevation model as TanDEM-L will have a higher resolution than SRTM, and a better raw accuracy. In any case, we submit that using the difference between a DSM and DTM acquired from space over a latitude range that comprises most forested land is an approach that holds a unique potential for producing global, wall-to-wall maps of forest biomass, without serious saturation effects.

CHAPTER V

SYNTHESIS, CONCLUSIONS AND FUTURE RESEARCH

This research thesis was set out to explore the capability of current spaceborne remote sensing to produce high-resolution and wall-to-wall forest height and biomass maps over large areas, i.e. entire biomes or even for the entire world's forests. Such forest biomass maps on a global scale are essential to quantify carbon stocks density, its distribution, and its dynamics in an era where dramatic ecological changes are expected to occur due to climate change and the destruction of habitats (Schimel et al., 2015). For this, we have chosen a path consisting of creating a canopy height model using solely satellite-based sensors and use forest height as the main predictor of biomass for an area located in the Canadian boreal forest. This leads us to demonstrate that:

- TanDEM-X elevation models represent the average elevation of the canopy surface within each resolution cell;
- these DSMs are quite robust to changes in phenology or minor changes in the interferometric baseline and the local incidence angle in closed-canopy boreal forests, opening the possibility of using the WorldDEM dataset for mapping forests;
- an approximate canopy height model (CHM) can be generated from space, using existing data from satellite missions, namely SRTM and TanDEM-X;
- this CHM represents a significant variable for predicting and mapping forest biomass.

We first discuss the scientific contribution of this thesis, and then present generalization possibilities, as well as limitations. Finally, we address future research options and provide concluding remarks.

5.1 Scientific contribution

There has been debate over the last several years about the degree of "penetration" of SAR pulse "within" forest canopies (Kugler et al., 2014; Praks et al., 2012; Treuhaft and Siqueira 2000). This question is crucial for two types of uses in InSAR and PolInSAR: a) creating elevation models, and b) inverting models to retrieve canopy height, for example, the RVoG model. In theory, the short wavelengths of the X-band can only interact with the topmost layer of tree material (Mougin et al., 1993), but several authors have contended that a ground signal or at least a deep penetration is nevertheless achieved in that band (Kugler et al., 2014 and 2015; Praks et al., 2012). Throughout the scientific literature, we have often found lack of detail, or imprecision about how forest height was defined or measured, or lack of adequate data for providing a reliable reference of forest height by which the degree of penetration was evaluated. We found out, or guessed in the case of some studies, that forest height was most often taken as being the average height of dominant trees at stand level (e.g., Kugler et al., 2014; Balzter et al., 2007.b), approximated through samples of tree heights within stands. In the case of interferometric elevation models, the height above ground of the scattering phase centre was reported as being much lower than the "forest height," creating the impression that the SAR pulses do travel considerably through vegetation layers. This inexact conception is in our opinion more damageable for studies involving model inversion where interferometric coherence is a key element (Kugler et al., 2015; Krieger et al., 2005), or in the case of SAR tomography (Mariotti d'Alessandro et al., 2013). Such approaches rely on the deep penetration of SAR pulses within vegetation canopies such that the vertical extension (forest height) of these canopies can be estimated. We do not contest the fact the PolInSAR approaches in L- or P-band has the

potential for such height retrieval, but put in serious doubt the possibility of using X-band in such way. In the first paper of this thesis (Chapter II), we have precisely defined and measured a) canopy surface height, and b) average dominant height, at several resolutions. This enabled us to clearly demonstrate, using in one instance ALS data alone, that canopy surface height is much lower than the dominant height (by 5.6 m in our ALS data). The magnitude of this apparent "penetration" was very similar to that reported by other InSAR studies. We, therefore, hope that our first thesis paper will help guide future research and disambiguate the above notion.

The clear conclusion coming out of our study is that the TanDEM-X SPC, whether in HH or VV mode, represents the average height of the canopy surface. One might argue that assessing dominant height is more important than canopy surface height. We, however, contend that in the case of quantifying biomass, the reverse is true. Recalling that forest biomass can be conceived as the product of forest volume and tree material density ($m^3 \cdot g m^{-3} = g$), it is easy to show that the average height (in m) of the forest surface within a resolution cell (area measured in m^2) represents volume ($m \cdot m^2 = m^3$). Assuming a rather constant forest material density for tree species of a given ecoregion with one forest species, the average height of a canopy surface as measured by X-band InSAR becomes, at least theoretically, a good volume, and therefore, a good biomass predictor. This was demonstrated numerous times in ALS studies where the average height of laser returns on the canopy was identified as one of the best predictors of timber volume (Næsset 2002; Lefsky et al., 2002). This was also the case in our third thesis paper (Chapter IV) where the *Elev.mean* FUSION ALS predictor came out first for predicting biomass. We therefore contend that this thesis has clearly demonstrated the usefulness (and accuracy) of the TanDEM-X DSMs for mapping forest biomass of closed-canopy boreal forests, provided a good DTM exists.

For applying such an approach over, say to the entire circumpolar boreal biome - putting aside for now the DTM issue - we need a wall-to-wall DSM. We have found that within a single interferogram of a mountainous forested region, part of the DSM is unusable due to radar layover (Rizzoli et al., 2015) or ambiguity in the phase unwrapping process (Rizzoli et al., 2014). Before we go further, we would like to point out that, to the best of our knowledge, the study region used for this thesis has a much stronger topography than that found in most other studies, with $>30^\circ$ slopes being common, and the largest part of the study region being on the non-horizontal terrain. This could have weakened certain relationships between remote sensing data and field or ALS reference data, compared to other studies. It, however, may provide a "worst case" scenario, as topographical conditions in the boreal zone are rarely more difficult than in the Forêt Montmorency. In any case, the areas devoid of usable data need to be filled in if a truly spatially continuous DSM is to be created. This has been accomplished by DLR and Airbus Space and Defence when creating the WorldDEM dataset. Data acquisitions from ascending and descending TanDEM-X orbits have been combined in such a way that unusable data in one passage was replaced by that coming from another one. This means that the only current worldwide DSM is made of a complicated assemblage of several pieces, even over small areas in the case of mountainous regions. Each of these pieces may have been acquired in different conditions. This is why we have assessed the effect of these conditions on the accuracy of these DSMs. As reported in the second paper of this thesis (Chapter III), acquisition conditions, at least within the scope of those we were able to study; do not sharply affect the RMSE or the bias of the TanDEM-X DSM. This represents an empirical scientific contribution that has a great impact on the usability of the WorldDEM dataset as a global DSM for forest studies. The selected forest site in this study is located over strong sloped topography where the DSM accuracy of the WorldDEM product is less variable compared to the results from this study. Among other scientific or practical considerations, it means that a small set of field calibration plots can be used to evaluate the available allometric equations to predict height, timber volume, and more

importantly biomass for the specific area. This calibration approach is quite different from what is used in reflectance-biomass or backscatter-biomass models, which have to be recalibrated for each image as the amount of returned energy may depend on several factors having to do with the state of the forest or the acquisition geometry. We have found the SPC location to be much more stable in the face of changes in these conditions than reflectance (prone to bidirectional effects) or backscatter (sensitive to dielectric changes caused by wetness or temperature).

Our thesis also made an important contribution in showing that a DTM created using a satellite mission, SRTM, can be used to create an approximate CHM useful in biomass mapping. The SRTM correction method is not ours, having been developed by St-Onge and Prieur during the research work of Prieur's masters (Prieur 2016), in collaboration with First Resource Management Group, a company based in Ontario, Canada. The method has however been adapted to the characteristic of our doctoral work dataset and used to predict and map forest biomass. It is, to the best of our knowledge, the first time such a CHM is created and used for that purpose. Numerous studies that have used a DSM made from satellite data resorted to an ALS DTM to create a CHM (see for example Solberg et al. 2015.a for a TanDEM-X example or St-Onge et al. 2008.a for a stereo Ikonos example). The satellite-based CHM (TanDEM-X minus corrected SRTM) came out as the best, or second best predictor in the random forest regression model, but a close second. It must be recognized that the amplitude of the errors remaining in the corrected SRTM elevations is important relative to the small canopies of the study region. This creates a rather large relative uncertainty in the CHM, especially as the errors present in the TanDEM-X DSM are also non-negligible. In any case, the proposed approach also requires cloud-free optical imagery, such as Landsat, to be applicable. Cloud-free Landsat mosaics of circa 2000 (concomitant to SRTM acquisition) do exist or can be assembled using data collected over a few years up to 2000 (say, 1996-2000). The bigger challenge is to find cloud-free images closely concomitant to the acquisition of TanDEM-X DSM for generating vegetation indices.

Notwithstanding those operational issues, the fact that a spaceborne CHM was proven useful for forest biomass assessment is a first. In other studies, the SRTM data was used to normalize GLAS vertical profiles to height profiles, which were then used to represent the height and biomass of forest regions, but no corrected SRTM elevation data was produced before and used as presented in our third thesis paper (Chapter IV). We contend that the DTM component of the approach can be still improved (see section 5.3).

As we proceeded with the research reported in our second paper, another researcher in Norway was following a parallel path, unbeknownst to us, until his paper was published, some months before ours, but after our poster presentation in the PolInSAR Workshop Proceedings (January 2015, Frascati). Our efforts were conducted independently, but corroborate Solberg's results and conclusions (being independent, our respective research projects can be said to be mutually corroborative). Solberg's studies being the closest to ours, we took some time here to compare our results to his (Solberg et al., 2015.a). He investigated the stability of eight TanDEM-X InSAR datasets over the boreal forest in Norway. A TanDEM-X dataset with a height of ambiguity of 20-50 m was found to be optimal. In our study, based on five datasets, a similar result with an optimal height ambiguity identified as being around 40 m. In his study, the winter dataset had more significant penetration into forest canopy compared to ours. This could be attributed to the different forest structure and density of our respective forest study sites.

Other researchers attempting to retrieve forest height from TanDEM-X data have followed different paths, most notably Kugler et al. (2014). They have used an approach based on applying a RVoG inversion on single or dual polarization datasets to extract forest height from the coherence differences between polarizations. The best results, with a r^2 0.86, were achieved for the summer acquisition of the boreal site. Kugler showed the SPC locates around 8 m under the top of dominant height for a

managed boreal forest with an average height of 18 m. However, in our study, the corresponding penetration (4.6 m - 7.5 m) could be simply explained by the difference between a dominant tree and average surface height (Sadeghi et al., 2016). Moreover, some information about the DTM is needed in an inversion approach for dual polarization in order to solve the phase unwrapping problem. This information was extracted from an (uncorrected) SRTM DEM in the case of Kugler et al (2014). They have reported that this inversion approach based on SRTM information will not be applicable for 20% of the area due to existing of difficult topography areas with a very high slope while in the case of study by Sadeghi et al (2016) in a very sloping area; only less than 1% of study site was masked.

It should be noted that our results and those of other similar recent studies based on TanDEM-X surpass in accuracy those that could be obtained from reflectance- or backscatter based approach. The Landsat Tasseled Cap (TC) was used to estimate forest biomass in boreal forests over hilly topography in the north of British Columbia, Canada, and the best performance from combination of simple and complex metrics from all TC components corresponded to a r^2 of 0.62, and a relative RMSE of 49% (Frazier et al., 2014). For a mixed-conifer region in eastern Oregon USA with an elevation range of 500 m to 2700 m, forest biomass was estimated using Landsat-based disturbance and recovery (DR) metrics (Pflugmacher et al., 2014), and an improved biomass prediction with a RMSE of 27%, compared to models based on only single-date reflectance, with a RMSE 35%, was achieved. P-band backscatter was used to estimate a sloping boreal forest's biomass (Soja et al., 2013), and a RMSE of 40–59 Mg ha^{-1} (22–33% of the mean biomass) was obtained. In any case, the calibration of the biomass-height relationship remains easier, less time-consuming, and more generalizable, compared to biomass prediction models using return intensities from single images.

5.2 Generalization and limitations

Our entire thesis relies on observations made over a 66 km² study region. It used five different interferometric pairs (10 TanDEM-X images), high-density lidar data, 200 field plots, and more than 400 individual tree height measurements. Despite the interest of building a scientific study on more study sites, obtaining and processing similar data on other sites would have represented a much greater effort, exceeding a reasonable amount of work and financial load for a doctoral thesis. Although limiting a Ph.D. project to a single area is fairly common, it does entail some limitations, particularly when trying to generalize the conclusions to other regions.

The Forest Montmorency is mainly composed of dense forests populated by balsam fir stands, at various growth stages from very young to mature, laid over a rugged topography. A first observation about limiting the generalization of our results is based on the role of forest gaps that was identified in our first thesis paper. We found that at the highest resolution, for example, small gaps between trees tend to be filled in the TanDEM-X DSM, compared to the ALS DSM. Moreover, the elevation value of a given resolution cell represents the average elevation of the canopy surface. Some of the world's forests have open canopies, with very sparsely populated stands. Open subarctic forests, for example, are constituted of thin black spruces growing far apart. Moreover, in open woody savannahs, acacia trees in Africa, for example, or oak trees in American savannahs, are also sometimes very sparse. The Cerrado vegetational formations of Brazil also do not form closed-canopy forests. We could question the capacity of TanDEM-X DSMs to properly represent the elevation of vegetation canopies in these areas. Would the SPC still represent the average height of the canopy surface? Could it be used for spatially quantifying biomass? Very little is known about the interaction of SAR pulses with these types of canopies.

In other circumstances, other complications may arise. For example, in very tall and dense tropical forest, it is reasonable to expect that canopy surface would be well modeled from TanDEM-X interferograms. However, we can question if a corrected SRTM DTM is feasible in these situations. These forests being very tall, it is likely that the SRTM pulses traveled only through the upper part of the canopy and were extinguished well above the ground. Furthermore, the relationship between the NDVI, or another spectral index, and the amount of overestimation in the SRTM elevations due to land cover may not hold.

Moreover, the selected study area is populated mostly by evergreen conifers, such as balsam fir and white spruce, and shows a low number of deciduous trees. When these occur, they are most often interspersed in a coniferous matrix, and never occur as pure stands. Although our findings concerning the effect of phenology (leaf-on vs. leaf-off) are not completely at odds to those that have been found in other studies (e.g. Solberg et al., 2015.a), a more thorough investigation would be needed to elucidate this question definitively. It is possible that deciduous trees with dense crowns bear a sufficient amount of twigs to backscatter SAR pulses in their topmost part. Similar behaviour was observed with ALS in which the height of leaf-less tree could still be recovered from dense point clouds. Due to the lack of similar studies in X-band InSAR, we can only hypothesize that the behaviour would be somewhat similar for short wavelength microwave pulses. Finally, our study site has very strong slopes, and phase unwrapping will not be accurate for very steep areas. However, these locations have been masked from our process (less than 1%). By using multi-baseline interferograms and developing the unwrapping phase algorithms, this limitation should be solved. This topic is however out from the scope of this PhD thesis. Moreover, TanDEM-X global coverage is based mostly on single baseline acquisitions, making access multi-baselines data difficult over most regions.

5.3 Future research directions

The considerations about limitations in the previous section point to certain research directions that need to be taken to expand the generalization possibilities. In this section, we would like to add other promising lines of research. It seems clear that the greatest challenge remains the acquisition of a high resolution and accurate DTM under dense vegetation canopies. We consider being possible to improve the SRTM DEMs further. First, local incidence data, or at least orbit positions allowing the precise computation of LIA, could increase our capacity to model SRTM errors. The combination of topography and forest canopies is quite likely to have an impact on the location of the SRTM SPC. Slope facing the incoming pulses, compared to those sloping away could change the degree of overestimation of ground elevation due to complex interactions with the canopy (Farr et al., 2007). At the time of this writing, there was still no expressed intention by NASA for releasing incidence angle or orbital data. Another possible path would be to increase the accuracy of the land cover correction by studying times series of Landsat images, from 1972 to 2000. Any clear cuts having occurred during that period could be traced in time to predict the amount of vegetation, and hence, better estimate the efficacy of the vegetation screen stopping the SRTM pulses above ground. Any forest present in 1972 and still observable in 2000 could be considered as a mature forest (fully grown trees) and its effect on SRTM elevation estimated. This would lead to a better DTM, and therefore a better CHM. This time series could be pursued until the time of TanDEM-X acquisition, and used, in conjunction with the improved CHM, to better predict and map forest biomass. Already, time series of Landsat images have been used for such purpose (Ota et al., 2014). This, however, comes again with the caveat of finding sufficient cloud-free imagery.

It should be considered that the upcoming spaceborne missions that focus on forest biomass have important design limitations. The BIOMASS mission (Le Toan et al.,

2011) will carry SAR sensor operating in P-band. However, it will not have single-pass InSAR capabilities, precluding the creation of coherent interferograms over forested regions. Moreover, the operation of P-band, due to its possible interference with military telecommunication, is prohibited over certain territories, including the United States. Moreover, the GEDI lidar sensor will operate briefly on the international space station (ISS) before it is decommissioned near 2020. However, GEDI will only sample the world's forests, during cloud-free periods, and only within the orbit of the ISS, i.e. 50° N to 50° S, thus missing large parts of the boreal forests (Qi and Dubayah 2016; Stysley et al., 2015).

In our opinion, the greatest potential stays with the TanDEM-L mission, for which the methods we have developed, particularly in our third thesis paper, could apply. The penetration within vegetation of L-band SAR pulses is much greater than in the case of C-band SRTM. Several studies based on airborne acquisitions have demonstrated that L-band InSAR has the capacity to gather signal from levels that are close to the ground level, and in some cases (sparser or shorter forests), from the terrain itself (Lucas et al., 2010; Praks et al., 2012; Ahmed et al., 2014). Considering the expertise of DLR developed during the SRTM mission (experimental X-band InSAR), and through the TanDEM-X highly sophisticated mission, the success of the TanDEM-L mission appears to be very probable. With its 12 m resolution, correction for curvature may not be necessary. However, land cover effects could still be present, albeit much less pronounced than those affecting the SRTM elevations. Calibration of this correction could likely be helped through the use of GEDI lidar data. Corrections based on spectral indices or time series of Landsat images could also be applied in the TanDEM-L case and could help produce a relatively accurate DTM, for the entire Earth. This, in turn, would considerably improve CHMs, including using the TanDEM-X data used in the first phases of acquisition (2010-2015).

This brings us to consider the potential offered by time series of global CHMs for tracking the changes in time and space of the biomass of the world's forests. The first

edition of the WorldDEM dataset, the first reliable global DSM, constitutes a baseline for studying future changes. Although the accurate estimation (say, with a RMSE < 10%) is still not attained, multi-temporal DSMs can at least be already envisaged as tools for tracking biomass losses in the case of forest clear-cut or, more importantly deforestation. Once a block of forest total is totally removed, a DSM difference before and after the clear cut will locally provide both a DSM and a DTM, since on the bare terrain left after forest disappearance, the DSM and DTM coincide. This allows a precise estimation of the height of the removed canopy, and likely, a very good assessment of the loss of forest biomass. Furthermore, once a good DTM is produced worldwide, whether by TanDEM-L or other future means, the current version of TanDEM-X elevation data, having been ascertained as a DSM through this thesis, will be leveraged into retrospective accurate CHMs of the entire world's forests. Not only will the generation of future CHM be made possible, but at the same time, we will inherit a time series of global CHMs, from the 2010-2015 period on.

What also remains, aside from the consideration of species on the height-biomass relationship, is modeling of the non-linear relationship between height and biomass. After trees have reached their maximum height, a threshold set by their genetics, site richness, and other fixed factors, they usually continue to grow in diameter, and therefore see their biomass increase without any further height gains (Köhl et al., 2017; Weiskittel et al., 2011). This "DBH signal" is much more difficult to sense remotely than height changes. There is evidence that P-band backscatter does carry information on the size of tree stems (Neeff et al., 2005; Rignot et al., 1995; Neumann et al., 2012). It is unclear, however, if the signal of the marginal DBH gains after height growth has peaked can be detected in such manner. Perhaps the best approach would be to track the height changes through times series of CHM and infer the additional DBH growth after stands have ceased gaining in height.

5.4 Concluding remarks

The orientation of this doctoral works stems from a review of existing, and near future methods for mapping forest biomass worldwide circa 2011-12. This analysis has led us to choose an approach in which forest biomass is estimated principally based on canopy height and to define the optimal remote sensing approach in terms of spatial and temporal resolutions, and accuracy. This brought us to study the characteristics of the TanDEM-X interferograms, and to seek means for producing an approximate DTM while avoiding the need for intensive calibration or airborne data, and as much as possible, spaceborne optical data. At this point, we feel confident to conclude that:

- TanDEM-X data, and data from similar missions, i.e. TanDEM-L, have the necessary spatial and temporal resolutions while offering the necessary coverage of the world's forests;
- TanDEM-X DSMs have the necessary accuracy and robustness to acquisition conditions to characterize the elevation of the surface of closed-canopy forest, at least in the boreal case;
- approximate DTMs derived from the SRTM data can be used to some extent for creating global CHMs, but are still error-laden, and there is no guarantee that they can be generated under dense tropical forests;
- there exists a high potential for creating a time series of CHM using TanDEM-X and TanDEM-L;
- these successive CHMs offer the possibility of tracking the biomass of the world's forests from the 2010-2015 and on, provided that a solution to the non-linear nature of the height-biomass relation be found.
- the biomass map was produced at a relatively high resolution of 20 m which shows the great potential of current remote sensing systems for producing detailed maps. However, the lower RMSE of biomass estimation at stand level,

compared to 20 m pixel size, does not refute the proposition that the optimal resolution should take field plot size into account.

These new (2010-) remote sensing missions, data, and approaches have a very high importance, and pertinence, relative to the assessment of the biomass of the world's forests. A much better knowledge of the amounts of forest carbon currently stored, or trapped and released in the future, will greatly help us understand the drivers of climate changes, and the effects of increased temperatures, modified precipitation regimes, and CO₂ fertilization as factors of vegetation growth, in the global Earth ecosystem. Let us hope that this will lead us to ameliorate our tending of the planet.

APPENDIX A

MAPPING FOREST CANOPY HEIGHT USING TANDEM-X DSM

AND AIRBORNE LIDAR DTM

This chapter has been published as:

Y. Sadeghi, B. St-Onge, B. Leblon, M. Simard, and K. Papathanassiou, "Mapping Forest Canopy height Using tandem DSM and Airborne Lidar DTM," Proc. IEEE Geoscience and Remote Sensing Symposium, IGARSS 2014, Quebec, Canada, July 2014.

A.1 Abstract

This study assesses the potential of single-pass TanDEM-X interferometric SAR (InSAR) data to map forest canopy height when a corresponding accurate DTM is available. In the proposed method, the forest CHM is extracted by subtracting an airborne lidar DTM from the TanDEM-X DSM. We showed that the TanDEM-X coherence is influenced by local incidence angles and tree basal area. These factors should be taken into account when estimating forest canopy height using TanDEM-X combined to lidar data.

Index Terms- Canopy height model (CHM), TanDEM-X, TanDEM-X-Lidar CHM, InSAR, SAR

A.2 Introduction

High-resolution maps of forest canopy height are needed to improve estimates of tree aboveground biomass and carbon stocks. In the recent years, airborne lidar methods have improved to a point where certain forest attributes can be measured with an accuracy equivalent to that of field measurements (Wulder et al. 2012). The cost per unit area of land of lidar data, however, remains too high to consider using it for periodic surveys over large areas. Interferometric SAR (InSAR) acquisition from a spaceborne platform, although less accurate, can provide 3D data on forests with little limitations on spatial extent or revisit rate. The TanDEM-X mission, a twin SAR satellite mission that operates as a spaceborne single-pass SAR interferometer at a frequency of 9.6 GHz, offers new InSAR possibilities, as it is the first long-term spaceborne single pass InSAR formation. In this study, we quantified the accuracy of forest heights estimated from TanDEM-X interferogram. This data has here been used only to derive a DSM. This DSM was then converted to canopy height by subtracting a lidar DTM to produce a hybrid CHM. This hybrid CHM was then compared to a lidar-only CHM. Despite the absence of temporal decorrelation in single pass InSAR data, the InSAR coherence over forest canopy may still be influenced by other factors such as forest species composition (i.e. coniferous versus deciduous trees), or tree basal area (Perko et al. 2011), thus affecting height accuracy. Basal area is used here as a surrogate for local aboveground tree biomass. This accuracy may also depend on local incidence angles (Perko et al. 2011). In this study, we specifically analyzed the influence of these factors on the TanDEM-X coherence and hybrid CHM error.

A.3 Study Site and Data

The test site for this study is the Montmorency forest, a 64 km² research forest located approximately 70 km north of Quebec City, Canada (47° 18' N, 71° 08' W). Located in the balsam fir (*Abies balsamea*) – white birch (*Betula papyrifera*) climatic domain, its altitude varies between 600 and 1000 m with locally strong slopes. Both temperate and

boreal tree species can be found, in pure and mixed stands. Airborne lidar data were collected on 6 & 9 August 2011, using an *Optech ALTM 3100* laser scanner having a wavelength of 1064 nm and a pulse repetition frequency of 100 kHz leading to an average first return density of 7 hits/m². A dual-polarization (HH/VV) high-resolution strip-map mode dataset was acquired on 15 July 2013 in bistatic mode by TanDEM-X. Only the HH images have been used here. The images were acquired with a range and azimuth resolution of 1.2 m and 6.6 m, respectively, and obtained in the slant range (CoSSC) format for interferometric processing.

A.4 Methodology

A.4.1 Lidar CHM

The lidar DSM and DTM were produced by interpolation, respectively of the first returns, and of points classified as ground. They were first created at a 0.25 m resolution and then resampled to 5 m for comparison with the TanDEM-X images. The lidar DTM was subtracted from the lidar DSM to produce the lidar CHM (Figure A.1).

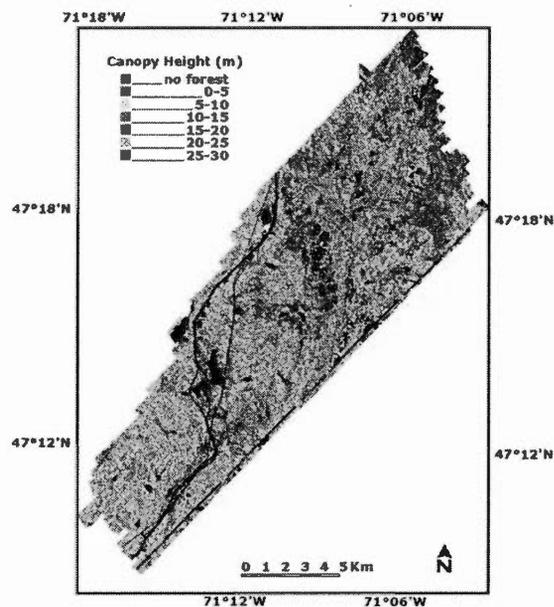


Figure A.1 Lidar CHM

A.4.2 TanDEM-X-lidar CHM

Interferometric processing and TanDEM-X DSM generation were done using *ENVI SARscape 5.0 Processor*. The image pair with HH polarization was coregistered with sub-pixel accuracy. Both co-registered SAR images were combined to form an interferogram that was processed as described in (Richards 2007). The interferogram was subjected to flattening using the 90 m SRTM DEM. The coherence image was derived from the interferogram filtered using a 3×3 adaptive spatial filter. The resulting image was then subjected to phase unwrapping using the minimum cost flow algorithm (Richards 2007). The resulting unwrapped phase values were converted into height values with a range-Doppler approach (Richards 2007). The final product was georeferenced into a WGS-84 datum and UTM 19N projection using the SRTM DEM coordinate system and the SAR sensor orbit parameters. The phase offset and phase

ramp errors originating from possible orbit inaccuracies were removed by fitting the following model to 20 Ground Control Points (GCPs). The final InSAR DSM had a 5 m ground pixel size. It was reduced to heights by subtracting the lidar DTM from it, thus yielding the TanDEM-X-lidar CHM (Figure A.2).

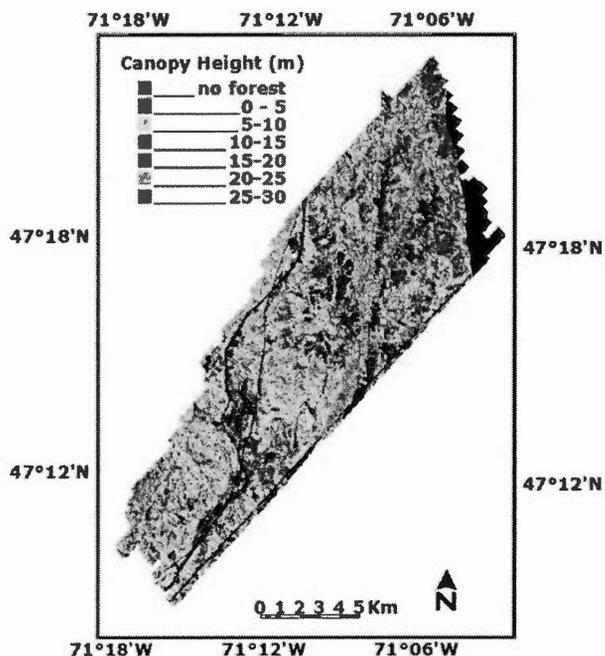


Figure A.2 TanDEM-X-lidar CHM (TanDEM-X DSM minus lidar DTM)

A.5 Results

The hybrid TanDEM-X-lidar CHM model was compared to the lidar-only CHM over 167 forest sample plots for which field measurements were available. A RMSE of 1.9 m and a r^2 of 0.75 were observed (Figure A.3). The TanDEM-X incidence angles vary only slightly (29.5° to 31.5°), but the local incidence angle relative to terrain slope varies from -14.5° to 77.5° . The negative values correspond to layover areas. As seen in figure A.4, the interferometric coherence is too low for extreme local topography, which causes layover, shadowing, or foreshortening.

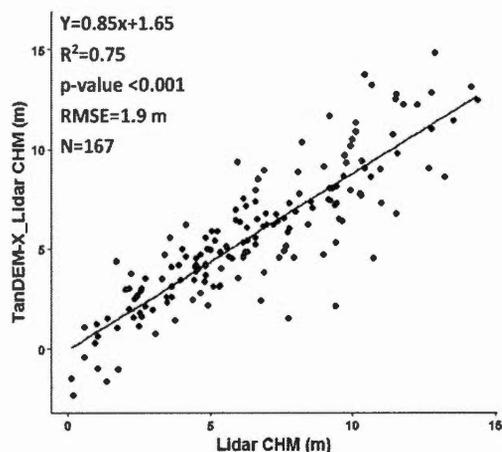


Figure A.3 TanDEM-X-lidar CHM vs. Lidar CHM

This low coherence is a limitation for the phase unwrapping, and therefore for forest height estimation. For the positive coherence values only, we found a quadratic relationship with the local incidence angle with an overall r^2 of 0.78 (Figure A.4).

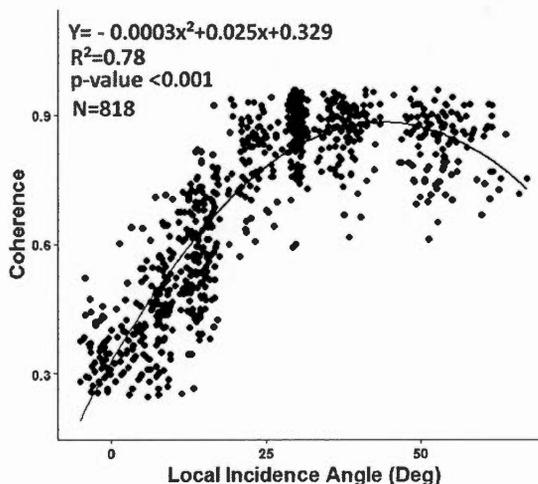


Figure A.4 Relationship between TanDEM-X InSAR coherence and local incidence angles

The various forest scattering mechanisms and their relative contributions depending on the incidence angles are known to influence coherence (Sarbandi and Lin 2000). In forest areas, according to the Michigan Microwave Canopy Scattering Model (MIMICS) (Ulaby et al., 1990), microwave scattering of a forest canopy is caused by

three main mechanisms, 1) volume scattering by the canopy itself, 2) surface scattering by the underlying ground surface, and 3) multiple interactions (that include double-bounce scattering) involving both the canopy volume and the ground surface. The volume scattering by the canopy is produced from three regions a) the crown, b) the trunk, and c) the underlying ground. In the InSAR DSM, the forest height corresponds to the extracted scattering phase centre heights from TanDEM-X DSM. These should be determined by volume scattering in the upper canopy because of low X-band penetration in dense forests. The high coherence over surface scatterers (roads, grassland or bare fields) is degraded only by SNR effects, while the coherence over forest areas is affected by volume decorrelation. The latter depends on the forest vertical structure (height, basal area, and tree type), dielectric constant, imaging geometry and image-related properties (polarization, wavelength) (Papathanassiou and Cloude 2001). We found a linear negative relationship between the interferometric coherence and the lidar CHM heights (Figure A.5), or the basal area reported for field plots (Figure A.6).

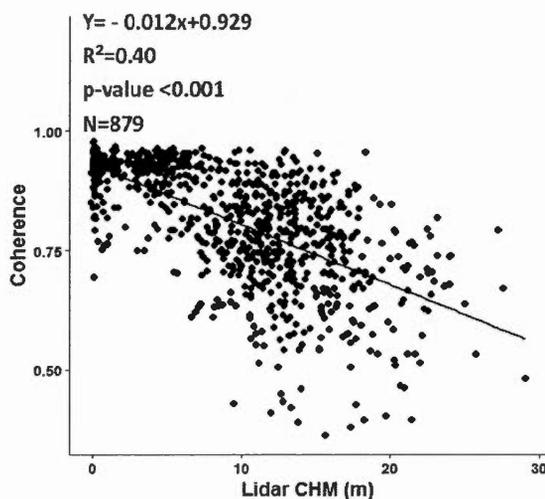


Figure A.5 Relationship between TanDEM-X InSAR coherence and the lidar CHM heights. The relationship is stronger with the lidar CHM than with the basal area. Both parameters are related directly to the volume decorrelation. High values of forest height

and tree basal area are associated to low coherence due to high volume decorrelation by such canopies. However, high coherence values for other canopies having a high basal area were also observed, possibly because the forest canopy can sometimes be an impenetrable volume to the X-band radar beams. In this case, scattering occurs only on the top of the canopy without volume decorrelation and loss of coherence. With respect to forest composition, we were not yet able to test this hypothesis by the lack of plots having a high percentage of deciduous in the field dataset. However, literature results suggest that X-band beams should have a high canopy penetration in coniferous forests compared to deciduous forests. Therefore the underestimation of forest height in coniferous forests should be greater than in deciduous forests (Demirpolat 2012).

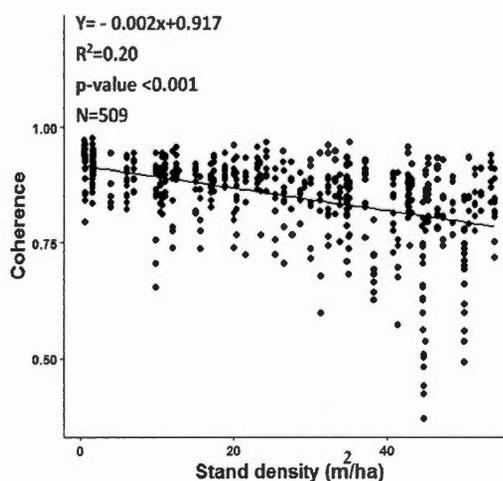


Figure A.6 Relationship between InSAR coherence and forest basal area

A.6 Conclusions

Our results show that TanDEM-X interferometric SAR coherences can be used to estimate forest canopy height, when combined with an accurate DTM, such as Lidar DTM. The TanDEM-X-Lidar CHM approach has the potential to allow rapid updating of forest canopy height maps over large areas with a high accuracy. However, several factors such as tree basal area, local incidence angles and forest composition should be taken into account.

REFERENCES

- Abdalati, W., Zwally, H., Bindschadler, R., Csatho, B., Farrell, S., Fricker, H., Harding, D., Kwok, R., Lefsky, M., Markus, T., Marshak, A., Neumann, T., Palm, S., Schutz, B., Smith, B., Spinhirne, J., and Webb, C. (2010). The ICESat-2 laser altimetry mission, *Proc. IEEE*, 98, 735–751.
- Ahmed, R., P. Siqueira, and S. Hensley, (2014). Analyzing the uncertainty of biomass estimates from L-band radar backscatter over the Harvard and Howland forests. *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3568-3586.
- Airbus Defence and Space, (2016). WorldDEMTM, <http://www.intelligence.airbusds.com/worlddem/>.
- Anderson, R.S. and Bolstad, P. (2013). Estimating aboveground biomass and average annual wood biomass increment with airborne leaf-on and leaf-off lidar in Great Lakes forest types. *North. J. Appl. For.* 30(1): 16–22.
- Askne, J. I. H., J. E. S. Fransson, M. Santoro, M. J. Soja and L.M.H. Ulander, (2013). Model-based biomass estimation of a hemi-boreal forest from multitemporal TanDEM-X acquisitions. *Remote Sens.*, vol. 5, no. 11, pp.5574-5597.
- Asner PG, Rudel KT, Aide TM, De Fries R, Emerson R. (2009). A contemporary assessment of change in humid tropical forests. *Conservation Biology* 23: 1386-1395.
- Avitabile V, Herold M, Heuvelink G, Lewis SL, Phillips OL, Asner GP, et al. (2016). An integrated pan-tropical biomass maps using multiple reference datasets. *Global Change Biology*, 22: 1406–1420.
- Baccini, A., Laporte, N., Goetz, S. J., Sun, M., & Don, H., (2008). A first map of Tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters*, 3, 1–9.
- Baccini, A., S. J. Goetz, W. S. Walker, N. T. Laporte, M. Sun, D. Sulla-Menashe, J. Hackler, P. S. A. Beck, R. Dubayah, M. A. Friedl, S. Samanta and R. A. Houghton, (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nat. Clim. Chang. Lett.*, vol. 2, pp.182-185.
- Baldasso, M., L. Birigazzi, et al. (2012). Tutorial for tree allometric equation database development. Rome, Italy: 27.

- Balzter, H., A. Luckman, L. Skinner, C. Rowland and T. Dawson, (2007.b). 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.*, vol. 28, no. 6, pp.1173-1197.
- Balzter, H., C. S. Rowland, P. Saich, (2007.a). Forest canopy height and carbon estimation at monks wood national nature reserve, UK, using dual-wavelength SAR interferometry. *Remote Sens. Environ.*, vol. 108, no. 3, pp.224-239.
- Baugh, Calum A., Bates, Paul D., Schumann, Guy, and Trigg, Mark A., (2013). SRTM vegetation removal and hydrodynamic modeling accuracy. *Water resources research*, vol. 49, 5276–5289.
- Bélanger, L., (2001). La forêt mosaïque comme stratégie de conservation de la biodiversité de la sapinière boréale de l'Est: l'expérience de la forêt Montmorency. *Le Naturaliste Canadien*, 125(3), 18-25.
- Ben-Arie, J. R., G. J. Hay, R. P. Powers, G. Castilla, B. St-Onge, (2009). Development of a pit filling algorithm for lidar canopy height models. *Computers & Geosciences*, Vol. 35, no. 9, pp. 1940-1949.
- Bergen, K., S. Goetz, R. Dubayah, G. Henebry, C. Hunsaker, G. Imhoff, (2009). Remote sensing of vegetation 3D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research*, 114 (G00E06).
- Bhang, K. J., and Schwartz, F., (2008). Limitations in the hydrologic applications of C-Band SRTM DEMs in low-relief settings, *IEEE Trans. Geosci. Remote Sens. Lett.*, 5(3), 497–501.
- Bourgine, B., Baghdadi, N., (2005). Assessment of C-band SRTM DEM in a dense equatorial forest zone. *Comptes Rendus Geoscience*, Volume 337, Issue 14, Pages 1225–1234.
- Brunt, K., Neumann, T., Amundson, J., Kavanaugh, J., Moussavi, M., Walsh, K., Cook, W., and Markus, T., (2016). MABEL photon-counting laser altimetry data in Alaska for ICESat-2 simulations and development, *The Cryosphere Discuss.*, 1–31.
- Bustamante, M. M. C., Roitman, I., Aide, T. M., Alencar, A., Anderson, L. O., Aragao, L., et al. (2016). Toward an integrated monitoring framework to assess the effects of tropical forest degradation and recovery on carbon stocks and biodiversity. *Global Change Biology*, 22, 92–109.

- Canadell, J. G., C. Le Que´re´, M. R. Raupach, C. B. Field, E. T. Buitenhuis, P. Ciais, T. J. Conway, N. P. Gillett, R. A. Houghton, and G. Marland, (2007). Contributions to accelerating atmospheric CO₂ growth from economic activity, carbon intensity, and efficiency of natural sinks, *Proc.Natl. Acad. Sci. U. S. A.*, 104, 18,866 – 18,870.
- Carabajal, C. C., and Harding, D. J., (2006). SRTM C-band and ICESat laser altimetry elevation comparisons as a function of tree cover and relief, *Photogramm. Eng. Remote Sens.*, 72(3), 287–298.
- Castel, T. and P. Oetli. (2008). Sensitivity of the C-band SRTM DEM vertical accuracy to terrain characteristics and spatial resolution. Dans *Headway in Spatial Data Handling: 13th International Symposium on Spatial Data Handling*, sous la dir. de A. Ruas et C. Gold, p. 163-176.
- Chave, J., C. Andalo, et al. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* 145, 87-99.
- Cloude, S. R., (2010). Polarisation applications in remote sensing. *London, U.K.: Oxford Univ. Press.*
- Coe, M. T., M. H. Costa, and E. A. Howard, (2008). Simulating the surface waters of the Amazon River basin: Impacts of new river geomorphic and flow parameterizations, *Hydrol. Processes*, 22(14), 2542–2553.
- Costantini, M., (1998). A novel phase unwrapping method based on network programming. *IEEE Trans. Geosci. Remote Sens.*, vol. 36. no. 3, pp.813-821.
- Dalponle, M., Martinez, C., Rodeghiero, M., Gianelle, D., (2011). The role of ground reference data collection in the prediction of stem volume with ALS data in mountain areas. *ISPRS J. Photogramm. Remote Sens.* 66 (6), 787–797.
- De Zan, F., G. Krieger, and P. López-Dekker, (2013). On some spectral properties of TanDEM-X interferograms over forested areas. *IEEE Letter. Geosci. Remote Sens.*, vol. 10, no. 1, pp.71-75.
- Demirpolat, C. (2012). X-band interferometric radar for mapping temporal variability in forest. *Master thesis, Aalto University, Department of Radio Science and Engineering.*
- DLR-TanDEM-L, (2016). http://www.dlr.de/hr/en/desktopdefault.aspx/tabid-8113/14171_read-35837.

- Doneus, M., C. Briese, N. Studnicka, (2010). Analysis of full-waveform ALS data by simultaneously acquired TLS data: towards an advanced DTM generation in wooded areas. In: *W. Wagner und B. Székely (Hg.): 100 Years ISPRS, Advancing remote sensing science. ISPRS technical commission VII symposium*. Vienna, 05.-07.07.2010 (The International archives of photogrammetry, remote sensing and spatial information sciences, Vol. XXXVIII, Part 7B), pp. 193–198.
- Dubayah, R. O., S. L. Sheldon, D. B. Clark, M. A. Hofton, J. B. Blair, G. C. Hurtt, and R.L. Chazdon, (2010). Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. *J. Geophys. Res.*, vol. 115, no. G00E09, pp. G00E09-1-17.
- Duchemin, B., Hadriab, R., Errakib, S., Bouleta, G., Maisongrandea, P., Chehbounia, A., Escadafala, R., Ezzaharb, J., Hoedjesa, J.C.B., Kharroud, M.H., Khabbab, S., Mougenota, B., Olioso, A., Rodriguezf, J.C., Simonneauxa, V. (2006). Monitoring wheat phenology and irrigation in Central Morocco: On the use of relationships between evapotranspiration, crops coefficients, leaf area index and remotely-sensed vegetation indices. *Agric. Water Manage*, 79, 1-27.
- Eggleston, H. S., Buendia, L., Miwa, K., Ngara, T., and Tanabe Eds., K. IGES, Japan, (2006). IPCC guidelines for national greenhouse gas inventories. Prepared by the National Greenhouse Gas Inventories Programme.
- Farr, T., P. 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, (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, volume, 45.
- Feldpausch, T. R., Liyd, J., Lewis, S. L., Brienen, R. J. W., Gloor, M., Monteagudo, A. Mendoza, L. B. G. Lopez-Gonzalez, K. Abu Salim, K. Affum-Baffoe, M. Alexiades, S. Almeida, I. Amaral, L. E. O. Andrade, A. Araujo Murakami, E. J. M. M. Arets, L. Arroyo, G. A. Aymard C, *et al.*, (2012). Tree height integrated into pantropical forest biomass estimates. *Biogeosciences*, vol. 9, pp. 3381–3403.
- Fisher, J.I., Hurtt, G.C., Thomas, R. & Chambers, J.Q. (2008). Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecol. Lett.*, 11, 1–10.
- Franklin, Steven E., Michael B. Lavigne, Michael A. Wulder, and Thomas M. McCaffrey. (2002). Large- area forest structure change detection: an example. *Canadian Journal of Remote Sensing* 28 (4):588-592.

- Frazier, R. J., Coops, N. C., Wulder, M. A., & Kennedy, R. (2014). Characterization of aboveground biomass in an unmanaged boreal forest using Landsat temporal segmentation metrics. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92, 137–146.
- Garestier, F., P. Dubois-Fernandez, X. Dupuis, P. Paillou, and I. Hajnsek, (2006). PolInSAR analysis of X-band data over vegetated and urban areas. *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 2, pp.356-364.
- Garestier, F., P.C. Dubois-Fernandez, K. P. Papathanassiou, (2008). Pine forest height inversion using single-pass X-band PolInSAR data. *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp.59-68.
- GEDI-NASA, (2015), <http://science.nasa.gov/missions/gedi/>
- Geomatics Montreal, (2016). http://publications.gc.ca/collections/collection_2015/rncan-nrcan/M103-1-4-2014-eng.pdf.
- Gibbs, H. K., Brown, S., Niles, J. O., & Foley, J. A. (2007). Monitoring and estimating tropical forest carbon stocks: Making REDD a reality. *Environmental Research Letters*, 2, 045023.
- Gilmore, D. W., R. S. Seymour, (1997). Crown architecture of *Abies balsamea* from four canopy positions. *Tree Physiology*, no. 17, pp. 71-80.
- Gitelson, A.A., Y.J. Kaufman, and M.N. Merzlyak, (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.*, 58:289-298.
- Gizachew, B. Solberg, S., Næsset, E., Gobakken, T., Ole Martin Bollandsås, Johannes Breidenbach, Eliakimu Zahabu, and Ernest William Mauya, (2016). Mapping and estimating the total living biomass and carbon in low-biomass woodlands using Landsat 8 CDR data. *Carbon Balance Manag.* 11(1): 13.
- Gobakken, T., & Næsset, E. (2008). Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, 5, 1095–1109.
- Goetz S., Baccini A., Laporte N., Johns T., Walker W., Kellndorfer J., Houghton R., Sun M., (2009). Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. *Carbon Bal Manag*, 4: 2. 10.1186/1750-0680-4-2.

- Gonzalez, P., G. P. Asner, J. J. Battles, M. A. Lefsky, K. M. Waring, M. Palace, (2010). Forest carbon densities and uncertainties from lidar, QuickBird, and field measurements in California. *Remote Sens. Environ.*, no. 114, pp.1561–1575.
- Graham, L. C. (1974). Synthetic interferometer radar for topographic mapping. *Proceedings of the IEEE*, 62(6), pp. 763-768.
- Gruber, A., Wessel, B., Martone, M., and Roth, A. (2015). The TanDEM-X DEM mosaicking: fusion of multiple acquisitions using InSAR quality parameters. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 9, No. 3, pp.1047-1057.
- Guth, P. L. (2006). Geomorphometry from SRTM: Comparison to NED. *Photogrammetric engineering and remote sensing*, vol. 72, no 3, p. 269-277.
- Hajnsek, I., Kugler, F., Lee, S. -K., & Papathanassiou, K.P. (2009). Tropical-forest-parameter estimation by means of Pol-InSAR: The INDREX-II campaign. *IEEE Transactions on Geoscience and Remote Sensing*, 47(2), 481–493.
- Hall, F. G., K. Bergen, J. B. Blair, R. Dubayah, R. Houghton, G. Hurtt, J. Kellndorfer, M. Lefsky, J. Ranson, S. Saatchi, H. H. Shugart, and D. Wickland, (2011). Characterizing 3D vegetation structure from space: mission requirements. *Remote Sens. Environ.*, vol. 115, no. 11, pp. 2753– 2775.
- Hansen, M. C., Defries, R. S., Townshend, J. R. G., Carroll, M., Dimiceli, C., & Sohlberg, R. A. (2003). Global percent tree cover at a spatial resolution of 500 meters: First results of the modis vegetation continuous fields algorithm. *Earth Interactions*, 7, 1–15.
- Henry, M., A. Bombelli, et al. (2013). GlobAllomeTree: International platform for tree allometric equations to support volume, biomass, and carbon assessment. *iForest - Biogeosciences and Forestry* e1-e5.
- Hensley, S., E. Chapin, A. Freedman, C. Le, S. Madsen, T. Michel, E. Rodriguez, P. Siqueira, and K. Wheeler, (2001.a). First P-band results using the GeoSAR mapping system. In Proc. Int. Geosci. Remote Sens. Symp., Sydney, Australia, pp. 126–128.
- Hensley, S., R. Munjy and P. A. Rosen, (2001.b). Interferometric synthetic aperture radar. Digital Elevation Model Technologies and Applications: The DEM User's Manual, Chapter 6, *ASPRS*.

- Herzfeld, U. C., B. W. McDonald, B. F. Wallin, T. A. Neumann, T. Markus, A. Brenner, and C. Field, (2013). Algorithm for detection of ground and canopy cover in micropulse photon-counting Lidar altimeter data in preparation for the ICESat-2 mission. *To appear in IEEE Transactions on Geoscience and Remote Sensing*.
- Hilbert, C., and C. Schmillius. (2012). Influence of surface topography on ICESat/GLAS forest height estimation and waveform shape. *Remote Sensing* 4 (12): 2210–2235.
- Hofton, M. A., Dubayah, R., Blair, J. B., & Rabine, D., (2006). Validation of SRTM Elevations over vegetated and non-vegetated terrain using medium footprint Lidar. *Photogrammetric Engineering and Remote Sensing*, 72, 279–285.
- Hopkinson, C., (2007). The influence of flying altitude, beam divergence, and pulse repetition frequency on laser pulse return intensity and canopy frequency distribution. *Can. J. Remote Sensing*, vol. 33, no. 4, pp. 312-324.
- Houghton, R. A. (2005). Aboveground forest biomass and the global carbon balance. *Global Change Biology*, vol. 11, pp. 945-958.
- Houghton, R. A. (2010). How well do we know the flux of CO₂ from land-use change?. *Tellus B*, vol. 62, pp. 337–351.
- Houghton, R. A., D. Butman, A. G. Bunn, O. N. Krankina, P. Schlesinger, T. A. Stone, (2007). Mapping Russian forest biomass with data from satellites and forest inventories. *Environ. Res. Lett.*, vol. 2, pp. 045032 (7pp).
- Houghton, R. A., F. Hall and S. J. Goetz, (2009). Importance of biomass in the global carbon cycle. *Journal Geophys. Res.*, vol. 114, pp. 1-13.
- Houghton, R. A., J. I. House, J. Pongratz, G. R. van der Werf, R. S. DeFries, M. C. Hansen, C. Le Quere, and N. Ramankutty, (2012). Carbon emissions from land use and land-cover change. *Biogeosciences*, vol. 9, pp. 5125–5142.
- Houghton, R. A., K. T. Lawrence, J. L. Hackler, and S. Brown, (2001). The spatial distribution of forest biomass in the Brazilian Amazon: a comparison of estimates. *Global Change Biology*, vol. 7, pp. 731-746.
- Huete, A.R., (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, vol. 25, issue 3, pp. 259-309.
- Hurst, M. D., S. M. Mudd, R. Walcott, M. Attal et K. Yoo. (2012). Using hilltop curvature to derive the spatial distribution of erosion rates. *Journal of Geophysical Research*, vol. 117, no F2, p. F02017.

- Hurt, G. C., J. Fisk, R. Q. Thomas, R. Dubayah, P. R. Moorcroft, and H. H. Shugart, (2010). Linking models and data on vegetation structure. *J. Geophys. Res.*, 115 (G00E10).
- Hyypä, J., Yu, X., Hyypä, H., Vastaranta, M., Holopainen, M., Kukko, A., et al., (2012). Advances in forest inventory using airborne laser scanning. *Remote Sensing*, 4(5), 1190–1207.
- Imhoff, M. L. (1995.a). A theoretical analysis of the effect of forest structure on SAR backscatter and the remote sensing of biomass. *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 341–352.
- Imhoff, M. L. (1995.b). Radar backscatter and biomass saturation: Ramifications for global biomass inventory. *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 511–518.
- IPCC, (2003). Definitions and methodological options to inventory emissions from direct human-induced degradation of forests and devegetation of other vegetation types, prepared by the National Greenhouse Gas Inventories Programme. Timo Karjalainen (Finland) and Gary Richards (Australia), Tomas Hernandez (Mexico), Samuel Kainja (Malawi), Gerry Lawson (Uk), Shirong Liu (China), and Steve Prisley (USA) ed.: Iges, Japan.
- Izzawati, E. D. Wallington, and I. H. Woodhouse, (2006). Forest height retrieval from commercial X-band SAR products. *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 4, pp. 863–870.
- Jiang, Z., Huete, A.R., Didan, K., Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.*, 112, 3833–3845.
- Karila, K., Vastaranta, M., Karjalainen, M. and Kaasalainen, S. (2015). Tandem-X interferometry in the prediction of forest inventory attributes in managed boreal forests. *Remote Sensing of Environment*, Volume 159, Pages 259–268.
- Karlson, M., Ostwald, M., Reese, H., Sanou, J., Tankoano, B., Mattsson, E., (2015). Mapping tree canopy cover and aboveground biomass in sudano-sahelian woodlands using Landsat 8 and random forest. *Remote Sens.* 7(8), 10017-10041.
- Kattenborn, T., Maack J., Faßnacht F., Enßle F., Ermert J., Koch B., (2015). Mapping forest biomass from space-fusion of hyperspectral EO1-hyperion data and Tandem-X and WorldView-2 canopy height models. *International Journal of Applied Earth Observation and Geoinformation*, 35: 359-367.

- Kauth, R.J., and G.S. Thomas, (1976). The tasseled cap - a graphic description of the spectral temporal development of agricultural crops as seen by Landsat. *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Perdue University, West Lafayette, Indiana*, 41-51.
- Kellndorfer, J., Walker, W., Pierce, L., Dobson, C., Fites, J. A., Hunsaker, C., Vona, J., Clutter, M., (2004). Vegetation height estimation from Shuttle Radar Topography Mission and National Elevation Datasets. *Remote Sensing of Environment*, Volume 93, Issue 3, Pages 339–358.
- Kellner, J.R., D.B. Clark, S.P. Hubbell, (2009). Pervasive canopy dynamics produce short-term stability in a tropical rain forest landscape. *Ecology Letters*, 12 (2), pp. 155–164.
- Köhl, M., Neupane, P. R., Lotfiomran, N., (2017). The impact of tree age on biomass growth and carbon accumulation capacity: A retrospective analysis using tree ring data of three tropical tree species grown in natural forests of Suriname. *PLoS ONE* 12(8): e0181187.
- Krieger, G., et al., (2009). The Tandem-L Mission Proposal: Monitoring Earth's dynamics with high resolution SAR interferometry. *IEEE Radar Conference*, Pasadena, CA, pp.1-6.
- Krieger, G., I. Hajnsek, K. P. Papathanassiou, M. Younis, A. Moreira, (2010). Interferometric synthetic aperture radar (SAR) missions employing formation flying. *Proc. IEEE.*, vol. 98, no. 5, pp.816-843.
- Krieger, G., K. P. Papathanassiou, S. R. Cloude, (2005). Spaceborne polarimetric SAR interferometry: performance analysis and mission concepts. *EURASIP Journal of Applied Signal Processing*, Vol. 20, pp 3272-3292.
- Krieger, G., Moreira, A., Fiedler, H., Hajnsek, I., Werner, M., Younis, M., Zink, M. (2007). TanDEM-X: A satellite formation for high-resolution SAR interferometry. *IEEE Trans. Geosci. Remote Sens.*, 45, 3317–3341.
- Kugler, F., D. Schulze, I. Hajnsek, H. Pretzsch, and K. P. Papathanassiou, (2014). TanDEM-X pol-InSAR performance for forest height estimation. *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp.6404–6422.
- Kugler, F., S-K. Lee, I. Hajnsek, and K. P. Papathanassiou, (2015). Forest height estimation by means of Pol-InSAR data inversion: the role of the vertical wavenumber. *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5294–6422.

- Lachaise, M., T. Fritz, U. Blass, R. Bamler, and M. Eineder, (2012). Phase unwrapping correction with dual-baseline data for the TanDEM-X mission. In *Proc. Int. Geosci. Remote Sens. Symp. (IGARSS)*, Munich, Germany, pp. 5566–5569.
- Lambert, M.-C., Ung, and Raulier, F., (2005). Canadian national tree aboveground biomass models. *Can. J. For. Res.* 35:1996–2018.
- Lavalle, M., M. Simard, and S. Hensely, (2012). A temporal decorrelation model for polarimetric radar interferometers. *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2880–2888.
- Le Toan, T., S. Quegan, M.W.J. Davidson, H. Balzter, P. Paillou, K. Papathanassiou, S. Plummer, F. Rocca, S. Saatchi, H. Shugart, L. Ulander, (2011). The BIOMASS mission: mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sens. Environ.*, no. 115, pp.2850-2860.
- Lee, S.-K., F. Kugler, K. Papathanassiou, and I. Hajnsek, (2013). Quantification of temporal decorrelation effects at L-band for polarimetric SAR interferometry applications. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 6, no. 3, pp. 1351–1367.
- Lee, S.K., Fatoyinbo, T.E. (2015). TanDEM-X Pol-InSAR inversion for mangrove canopy height estimation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 8, 3608–3618.
- Lefsky, M. A. (2010). A global forest canopy height map from the moderate resolution imaging spectroradiometer and the geoscience laser altimeter system. *Geophysical Research Letters*, 37, L15401.
- Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Espirito-Santo, F. D., Hunter, M. O., and Jr, R. O. (2005). Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, Vol. 32, L22S02.
- Lefsky, M. A., W. B. Cohen, D. J. Harding, G. G. Parker, S. A. Acker, and S. T. Gower. (2002). Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology and Biogeography*, 11 (5): 393–399.
- Lim, K. S, Treitz, P. M., (2004). Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scand J For Res* 19:558–570.
- Lim, K., Treitz, P., Wulder, M., St-Onge, B. and Flood, M. (2002). LiDAR remote sensing of forest structure. *Progress in Physical Geography*, 27, pp. 88–106.

- Lindberg, E., K. Olofsson, J. Holmgren, H. Olsson, (2012). Estimation of 3D vegetation structure from waveform and discrete return airborne laser scanning data. *Remote Sens. Environ.*, vol. 118, pp.151-161.
- Liu, D., Du, Y., Sun, G., Yan, W., Wu, B., (2008). Analysis of InSAR sensitivity to forest structure based on radar scattering model. *Prog Electromagnet Res.* 84:149–171.
- Liu, D., G. Sun, Z. Guo, K. J. Ranson, Y. Du, (2009). Three-dimensional coherent radar backscatter model and simulations of scattering phase center of forest canopies. *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 1, pp.349–357.
- Lu, D., (2006). The potential and challenge of remote sensing-based biomass estimation. *Int J Rem Sens*, 27: 1297–1328.
- Lu, D., Qi Chen, Guangxing Wang, Lijuan Liu, Guiying Li & Emilio Moran, (2016). A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth*, Volume 9, 2016 - Issue 1, p: 63-105.
- Lucas, R. M., J. D. Armston, R. Fairfax, R. Fensham, A. Accad, J. Carreiras, J. Kelley, P. Bunting, D. Clewley, S. Bray, D. Metcalfe, M. Dwyer, M. Bowen, T. Eyre, and M. Laidlaw, (2010). An evaluation of the ALOS PALSAR L-band backscatter—Above ground biomass relationship Queensland, Australia: Impacts of surface moisture condition and vegetation structure. *IEEE J. Sel. Topics Earth Obs. Remote Sens.*, vol.3, no. 4, pp. 576–593.
- Maltamo, M., Bollandsas, O.M., Næsset, E., Gobakken, T., Packalén, P., (2011). Different plot selection strategies for field training data in ALS-assisted forest inventory. *Forestry* 84 (1), 23–31.
- Mariotti d’Alessandro, M., S. Tebaldini, and F. Rocca, (2013). Phenomenology of ground scattering in a tropical forest through polarimetric synthetic aperture radar tomography. *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 8, pp. 4430–4437.
- Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., Jasinski, M., Kwok, R., Magruder, L., Lubin, D., Luthcke, S., Morison, J., Nelson, R., Neuenschwander, A., Palm, S., Popescu, S., Shum, C., Schutz, B., Smith, B., Yang, Y., and Zwally, J.: (2016). The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation, *Remote Sens. Environ.*, in review.

- Martone, M., Brautigam, B., Rizzoli, P., Gonzalez, C., Bachmann, M., Krieger, G. (2012). Coherence evaluation of TanDEM-X interferometric data. *ISPRS Journal of Photogrammetry and Remote Sensing.*, Vol. 73, pp. 21–29.
- Mette, T., K. P. Papathanassiou, and I. Hajnsek, (2004). Biomass estimation from polarimetric SAR interferometry over heterogeneous forest terrain. *In Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Anchorage, AK, USA, pp. 511–514.
- Molto, Q., B. Hérault, J.-J. Boreux, M. Daullet, A. Rousteau, and V. Rossi, (2014). Predicting tree heights for biomass estimates in tropical forests – a test from French Guiana. *Biogeosciences*, no. 11, pp.3121–3130.
- Moreira A., Gerhard Krieger, Irena Hajnsek, Konstantinos Papathanassiou, Marwan Younis, Paco Lopez-Dekker, Sigurd Huber, Michelangelo Villano, Matteo Pardini, Michael Eineder, Francesco De Zan, and Alessandro Parizzi, (2015). Tandem-L: A highly innovative bistatic SAR mission for global observation of dynamic processes on the Earth's surface. *IEEE Geoscience and Remote Sensing Magazine*, Volume: 3, Issue: 2, P:8-23.
- Motohka T, Nasahara KN, Oguma H, et al. (2010). Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sens* 2:2369–87.
- Mougin, E., A. Lopes, M. A. Karam, A. K. Fung, (1993). Effect of tree structure on X-band microwave signature of conifers. *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 3, pp.655-667.
- Moussavi, M. S., W. Abdalati, T. Scambos, and A. Neuenschwander, (2014). Applicability of an automatic surface detection approach to micropulse photon-counting lidar altimetry data: implications for canopy height retrieval from future ICESat-2 data. *Int. J. Remote Sens.*, vol. 35, no. 13, pp.5263–5279.
- Mugasha, WA, Bollandas OM, Eid T. (2013). Relationships between diameter and height of trees for natural tropical forest in Tanzania. *Southern Forest*, 75, 221–237.
- Murooka, J., T. Kobayashi, T. Imai, K. Suzuki, D. Sakaizawa, S. Yamakawa, R. Sato, H. Sawada, K. Asai, (2013). Overview of Japan's spaceborne vegetation lidar mission. *SPIE Proc., Lidar Tech., Techn., and Measur. Atmo. Remote Sens. IX*, 88940B.
- Muukkonen, P., Heiskanen, J. (2005). Estimating biomass for boreal forests using ASTER satellite data combined with standwise forest inventory data. *Remote Sensing of Environment*, 99, 434–447.

- Næsset E., Hans Ole Ørka, Svein Solberg, Ole Martin Bollandsås, Endre Hofstad Hansen, Ernest Mauya, Eliakimu Zahabu, Rogers Malimbwi, Nurdin Chamuya, Håkan Olsson, Terje Gobakken, (2016). Mapping and estimating forest area and aboveground biomass in miombo woodlands in Tanzania using data from airborne laser scanning, TanDEM-X, RapidEye, and global forest maps: A comparison of estimated precision. *Remote Sensing of Environment*, Volume 175, Pages 282–300.
- Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80, pp. 88–99.
- Næsset, E. and Gobakken, T., (2008). Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment*, Volume 112, Issue 6, Pages 3079–3090.
- Neeff, T., L. Vieira Dutra, J. Roberto dos Santos, C. da Costa Freitas, L. Spinelli Araujo, (2005). Tropical forest measurement by interferometric height modeling and P-band radar backscatter. *For. Sci.*, vol. 6, no. 51, pp.585-594.
- Neigh, C.S.R., R. F. Nelson, K. J. Ranson, H. A. Margolis, P. M. Montesano, G. Sun, V. Kharuk, E. Næsset, M. A. Wulder, H. Andersen, (2013). Taking stock of circumboreal forest carbon with ground measurements, airborne and spaceborne lidar. *Remote Sens. Environ.*, no. 137, pp.274-287.
- Neumann, M., S. S. Saatchi, L. M. H. Ulander, J. E. S. Fransson, (2012). 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.
- Nichol, JE and Sarker MLR. (2011). Improved biomass estimation using the texture parameters of two high-resolution optical sensors. *IEEE Transaction on Geosciences and Remote Sensing*. 49: 930–948.
- Olander, L.P., H.K. Gibbs, M. Steininger, J.J. Swenson, and B.C. Murray,(2008). Reference scenarios for deforestation and forest degradation in support of REDD: A review of data and methods. *Environmental Research Letters* 3(2): 025011.
- Ometto, J. P., A. P. Aguiar, T. Assis, (2014). Amazon forest biomass density maps: tackling the uncertainty in carbon emission estimates. *Climatic Change*, vol. 124, pp. 545–560.
- Ota, T., Ahmed, O. S., Franklin, S.E., Wulder, M. A., et al. (2014). Estimation of airborne lidar-derived tropical forest canopy height using Landsat time series in Cambodia. *Remote Sensing*, 6, 10750–10772.

- Paiva, R. C. D., W. Collischonn, and D. C. Buarque, (2013). Validation of a full hydrodynamic model for large-scale hydrologic modelling in the Amazon, *Hydrol. Processes*, 27(3), 333–346.
- Papathanassiou, K.P., S.R. Cloude, (2001). Single Baseline Polarimetric SAR Interferometry. *IEEE Transactions Geoscience and Remote Sensing*, 39(11), pp. 2352- 2363.
- Patenaude, G. L., Hill, R. A., Milne, R., Rowland, C. S., Dawson, T. P. (2002). Forest carbon accounting using Airborne Laser Scanning remote sensing and modelling approaches. *Proc. of ForestSAT 2002*, 5-9 August Edinburgh, UK: Forest Research.
- Perko, R., H. Raggam, J. Deutscher, K. Gutjahr, M. Schardt, (2011). Forest assessment using high resolution SAR data in X-band. *Remote Sensing*, 3(4), pp. 792– 815.
- Persson, H., (2014.a). Estimation of forest parameters using 3D satellite data stereogrammetry, radargrammetry and interferometry. *Ph.D. dissertation*, Dept. For. Res. Manag., Swedish University of Agricultural Sciences, Umeå.
- Persson, H., Fransson, JES. (2014.b). Forest variable estimation using radargrammetric processing of TerraSAR-X images in boreal forests. *Remote Sens.* 6:2084–107.
- Pflugmacher, D., W. B. Cohen, R. E. Kennedy, and Z. Yang. (2014). Using Landsat-derived disturbance and recovery history and Lidar to map forest biomass dynamics. *Remote Sensing of Environment*. 151: 124–137.
- Pflugmacher, D., W. Cohen, R. Kennedy, and M. Lefsky, (2008). Regional applicability of forest height and aboveground biomass models for the geoscience laser altimeter system. *For. Sci.*, vol. 6, no. 54, pp. 647-657.
- Powell, S. L., Cohen, W. B., Healey, S. P., Kennedy, R. E., Moisen, G. G., Pierce, K. B., et al. (2010). Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sensing of Environment*, 114, 1053–1068.
- Praks, J., O. Antropov, and M. T. Hallikainen, (2012). 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.
- Prieur, (2016). Amélioration de la résolution et de l'exactitude de modèles numériques d'altitude srtm pour leur utilisation dans le calcul de modèles de hauteur de couverts forestiers. *Master thesis in Geography*, UQAM.

- Qi, Wenlu, Ralph O. Dubayah, (2016). Combining Tandem-X InSAR and simulated GEDI lidar observations for forest structure mapping. *Remote Sensing of Environment* 187, 253–266.
- R Development Core Team. R: A Language and environment for statistical computing, reference index version 2.2.1; R foundations for statistical computing: Vienna, Austria, (2005). Available online: <http://www.r-project.org/>
- Raggam, H., K.Gutjahr, R. Perko, and M. Schardt, (2009). Assessment of the stereo-radargrammetric mapping potential of TerraSAR-X multibeam spotlight data. *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 0196-2892.
- Richards, M., (2007). A Beginner's Guide to Interferometric SAR Concepts and Signal Processing. *IEEE Aerospace and Electronic*, 22(9), pp. 5-29.
- Richardson, A.J., and C.L. Wiegand, (1977). Distinguishing vegetation from soil background information, *Photogrammetric Engineering and Remote Sensing*, 43(12):1541-1552.
- Rignot, E., R. Zimmerman, and J. V. Zyl, (1995). Space-borne applications of p-band imaging radars for measuring forest biomass. *IEEE Trans. Geosci. Remote Sensing*, vol. 33, p. 1162.
- Rizzoli, P., Martone, M., Brautigam, B. (2014). Global interferometric coherence maps from TanDEM-X quicklook data. *IEEE Letter Geoscience and Remote Sensing*, Vol. 11, No. 11, pp.1861-1865.
- Rizzoli, P., Martone, M., Brautigam, B. (2015). Global mosaics of the relative height error from TanDEM-X quick looks. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 12, No. 9, pp.1928-1932.
- Rodriguez, E., Morris, C. S., and Belz, J. E., (2006), A global assessment of the SRTM performance, *Photogramm. Eng. Remote Sens.*, 72(3), 249–260.
- Rodríguez-Veiga, P. Saatchi, S., Tansey, K., Balzter, H., (2016). Magnitude, spatial distribution and uncertainty of forest biomass stocks in Mexico. *Remote Sensing of Environment*, Volume 183, Pages 265–281.
- Saatchi, S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., and Morel, A. (2011.b). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences*, Vo.108, No. 24, pp.9899–9904.

- Saatchi, S., M. Marlier, R. L. Chazdon, D. B. Clark, A. E. Russell, (2011.a). Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. *Remote Sens. Environ.*, vol. 115, pp. 2836–2849.
- Sadeghi b, Y., St-Onge, B., Leblon, B., and Simard, M., (*in press*). Effects of TanDEM-X acquisition parameters on the accuracy of digital surface models of a boreal forest canopy. *Canadian Journal of Remote Sensing*, (It was accepted on 15 December 2016).
- Sadeghi, Y., St-Onge, B., Leblon, B., and Simard, M. (2016). Canopy height model (CHM) derived from a TanDEM-X InSAR DSM and an airborne lidar DTM in boreal forest. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 9, No. 1, pp. 381-397.
- Sadeghi, Y., St-Onge, B., Leblon, B., Simard, M., and Papathanassiou, K. (2014). Mapping forest canopy height using TanDEM-X DSM and airborne Lidar DTM. *In Proc. Int. Geoscience and Remote Sensing Symposium (IGARSS'2014)*, pp. 76-79.
- Sarbandi, K., and Y. Lin, (2000). Simulation of Interferometric SAR Response for Characterizing the Scattering Phase Centre Statistics of Forest Canopies. *IEEE Transactions on Geoscience and Remote Sensing*, 38, pp. 115–125.
- Schimel, D., Pavlick, R., Fisher, J. B., Asner, G. P., Saatchi, S., et al. (2015). Observing terrestrial ecosystems and the carbon cycle from space. *Global Change Biology*, Volume 21, Issue 5, pages 1762–1776.
- Schlund, M., Poncet, F., Kuntz, S., Boehm, H. V., Hoekman, D. H., Schmullius, C. (2016). TanDEM-X elevation model data for canopy height and aboveground biomass retrieval in a tropical peat swamp forest. *International Journal of Remote Sensing*, Volume 37, Issue 21.
- Schlund, M., von Poncet, F., Kuntz, S., Schmullius, C., & Hoekman, D.H., (2015). TanDEM-X data for aboveground biomass retrieval in a tropical peat swamp forest. *Remote Sensing of Environment*, 158, 255–266.
- Schutz, B. E., H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio, (2005). Overview of the ICESat mission. *Geophys. Res. Lett.*, vol. 32, L21S01.
- Shortridge, A. and J. Messina. (2011). Spatial structure and landscape associations of SRTM error. *Remote Sensing of Environment*, vol. 115, no 6, p. 1576-1587.

- Simard, M., K. Q. Zhang, V. H. Rivera-Monroy, M. S. Ross, P. L. Ruiz, E. Castaneda-Moya, R. R. Twilley, and E. Rodriguez, (2006). Mapping height and biomass of mangrove forests in Everglades National Park with SRTM elevation data. *Photogram. Eng. Remote Sens.*, vol. 72, no. 3, pp. 299–311.
- Simard, M., N. Pinto, J. Fisher, and A. Baccini, (2011). Mapping forest canopy height globally with spaceborne lidar. *Geophys. Res.*, vol. 116, G04021.
- Simard, M., S. Hensley, M. Laval, R. Dubayah, N. Pinto, M. Hofton, (2012). An empirical assessment of temporal decorrelation using the uninhabited aerial vehicle synthetic aperture radar over forested landscapes. *Remote Sens.*, no. 4, pp.975-986.
- Simard, M., V. H. Rivera-Monroy, J. E. Mancera-Pineda, E. Castaneda-Moya, and R. R. Twilley, (2008). A systematic method for 3D mapping of mangrove forests based on shuttle radar topography mission elevation data, ICESat/GLAS waveforms and field data: application to Ciénaga Grande de Santa Marta, Colombia. *Remote Sens. Environ.*, vol. 112, pp.1231-1244.
- Soja, M. J., G. Sandberg, and L. M. H. Ulander. (2013). Regression-based retrieval of boreal forest biomass in sloping terrain using P-band SAR backscatter intensity data. *IEEE Transactions on Geoscience and Remote Sensing*, 51(5):2646–2665.
- Soja, M. J., H. Persson, and L. M. H. Ulander, (2014). Estimation of forest height and canopy density from a single InSAR correlation coefficient. *IEEE Let. Geosci. Remote Sens.*, vol. 12, no. 3, pp.646-650.
- Solberg, S., D. J. Weydahl, and R. Astrup, (2015.a). Temporal stability of X-Band single-pass InSAR heights in a spruce forest: effects of acquisition properties and season. *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp.1607-1614.
- Solberg, S., E. Næsset, T. Gobakken, and O. Bollandsås. (2014). Forest biomass change estimated from height change in interferometric SAR height models. *Carbon Balance and Management* 9: 5. doi:10.1186/s13021-014-0005-2.
- Solberg, S., Gizachew, B., Næsset, E., Gobakken, T., Bollandsås, O.M., Mauya, E.W., Zahabu, E, Malimbwi, R., and Olsson, H. (2015.b). Monitoring forest carbon in a Tanzanian woodland using interferometric SAR: a novel methodology for REDD+. *Carbon Balance and Management*. 10:14. DOI 10.1186/s13021-015-0023-8.
- Solberg, S., R. Astrup, J. Breidenbach, B. Nilsen, and D. Weydahl, (2013). Monitoring spruce volume and biomass with InSAR data from TanDEM-X. *Remote Sens. Environ.*, vol. 139, pp. 60–67.

- Solberg, S., R. Astrup, O. M. Bollandsas, E. Næsset, and D. J. Weydahl, (2010). Deriving forest monitoring variables from X-band InSAR SRTM height. *Can. J. Remote Sens.*, vol. 36, no. 1, pp. 68–79.
- St-Onge, B., (2008). Methods for improving the quality of a true orthomosaic of vexcel ultracam images created using a lidar digital surface model. *In Proc. Silvilaser*, pp. 555-562.
- St-Onge, B., Hu, Y., Vega, C. (2008.b). Mapping the height and above-ground biomass of a mixed forest using lidar and stereo Ikonos images. *International Journal of Remote Sensing*, Vol. 29, No. 5, pp.1277-1294.
- St-Onge, B., Vega, C., Fournier, R. A., and Hu, Y. (2008.a). Mapping canopy height using a combination of digital stereo-photogrammetry and lidar. *International Journal of Remote Sensing*, Vol. 29, No. 11, pp. 3343-3364.
- Stysley, P.R., Coyle, D.B., Kay, R.B., Frederickson, R., Poullos, D., Cory, K., Clarke, G., (2015). Long term performance of the High Output Maximum Efficiency Resonator (HOMER) laser for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar. *Opt. Laser Technol.* 68, 67–72.
- Su, Y. and Guo, Q., (2014). A practical method for SRTM DEM correction over vegetated mountain areas. *ISPRS J. Photogramm. Remote Sens.* 87, 216–228.
- Su, Y., Q. Guo, Q. Ma, W. Li, (2015). SRTM DEM correction in vegetated mountain areas through the integration of spaceborne LiDAR, airborne LiDAR, and optical imagery. *Remote Sensing* 7(9): 11202-11225.
- Theodor Ene, L., Næsset, E., Terje Gobakken, Ole Martin Bollandsås, Ernest William Mauya, Eliakimu Zahabu, (2017). Large-scale estimation of change in aboveground biomass in miombo woodlands using airborne laser scanning and national forest inventory data. *Remote Sensing of Environment*, Volume 188, Pages 106–117.
- Thirion-Lefevre, L., E. Colin-Koeniguer, (2007). Investigating attenuation, scattering phase center, and total height using simulated interferometric SAR images of forested areas. *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3172–3179.
- Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., et al. (2014). Carbon stock and density of northern boreal and temperate forests. *Global Ecology and Biogeography*, 23, 297–310.
- Toutin, T., I. Zakharov, C. Schmitt, (2010). Fusion of Radarsat-2 polarimetric images for improved stereoradargrammetric DEM. *Int. J. Image Fusion*, vol.1, no.1,

pp.67-82.

- Treuhaft, R. N., P. R. Siqueira, (2000). The vertical structure of vegetated land surfaces from interferometric and polarimetric radar. *Radio Sci.*, vol. 35, no. 1, pp.141-177.
- Treuhaft, R. N., S. N. Madsen, M. Moghaddam, and J. J. Van Zyl, (1996). Vegetation characteristics and surface topography from interferometric radar. *Radio Sci.*, vol. 31, no. 6, pp. 1449–1485.
- Treuhaft, R., F. Gonçalves, J. Roberto dos Santos, M. Keller, M. Palace Palace, S. Madsen, *et al.* (2015). Tropical-forest biomass estimation at X-band from the spaceborne TanDEM-X interferometer. *IEEE Geoscience and Remote Sensing Letters*, 12 (2), pp. 239–243.
- Treuhaft, R.N. and Siqueira, P.R. (2004). The calculated performance of forest structure and biomass estimates from interferometric radar. *Waves in Random Media*, 14, pp. S345–S358.
- Tucker, C.J., (1977). Asymptotic nature of grass canopy spectral reflectance. *Applied Optics* 16, 1151–1156.
- Ulaby, F.T., K. Sarbandi, K. McDonald, M. Whitt, M. C. Dobson, (1990). Michigan Microwave Canopy Scattering Model. *International Journal of Remote Sensing*, 11(7), pp.1223–1253.
- Vanderwel, M.C., Coomes, D.A., Purves, D.W., (2013a). Quantifying variation in forest disturbance, and its effects on aboveground biomass dynamics, across the eastern United States. *Glob. Change Biol.* 19, 1504–1517.
- Vanderwel, M.C., Lyutsarev, V.S., Purves, D.W., (2013b). Climate-related variation in mortality and recruitment determine regional forest-type distributions. *Glob. Ecol. Biogeogr.* 22, 1192–1203.
- Vastaranta, M., M.Holopainen, M. Karjalainen, V. Kankare, J. Hyypä, S. Kaasalainen, (2014). TerraSAR-X stereo radargrammetry and airborne scanning lidar height metrics in imputation of forest aboveground biomass and stem volume. *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1197-1204.
- Vermote, E., C. Justice, M. Claverie, B. Franch, (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, Volume 185, Pages 46–56.
- Walker, W. S., J. M. Kelndorfer, L. E. Pierce, (2007). Quality assessment of SRTM C- and X-band interferometric data: implications for the retrieval of vegetation canopy height. *Remote Sens. Environ.*, vol. 106, no. 4, pp.428-448.

- Wang, Q, Adiku S, Tenhunen J, Granier A., (2005). On the relationship of NDVI with leaf area index in a deciduous forest site. *Remote Sens Environ* 94:244 -255.
- Wang, Y., F. W. Davis, J. M. Melack, E. S. Kasischke, and N. L. Christensen, Jr., (1995). The effects of changes in forest biomass on radar backscatter from tree canopies. *Int. J. Remote Sens.*, vol. 16, no. 3, pp. 503–513.
- Weiskittel, A. R., Hann, D. W., Kershaw, J. A., Vanclay, J. k., (2011). Forest growth and yield modeling. Chichester, West Sussex, UK: John Wiley&Sons.
- White, Joanne C., John T.T.R. Arnett, Michael A. Wulder, Piotr Tompalski, and Nicholas C. Coops, (2015). Evaluating the impact of leaf-on and leaf-off airborne laser scanning data on the estimation of forest inventory attributes with the area-based approach. *Can. J. For. Res.* 45: 1498–1513.
- White, M. A., Brunsell, N., and Schwartz, M. D. (2003). Vegetation phenology in global change studies, in Phenology: An integrative environmental science, edited by Schwartz, M. D., *Kluwer Academic Publishers*, New York, NY, 453–466.
- Wilson, M., P. Bates, D. Alsdorf, B. Forsberg, M. Horritt, J. Melack, F. Frappart, and J. Famiglietti, (2007). Modeling large-scale inundation of Amazonian seasonally flooded wetlands, *Geophys. Res. Lett.*, 34, L15404.
- Wulder, M. A., J. C. White, R. F. Nelson., E. Næsset, H. Ole Ørka, N. C. Coops, T. Hilker, C. W. Bater, T. Gobakken, (2012). Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, 121, pp. 196–209.
- Wulder, M.A.; White, J.C.; Cranny, M.M.; Hall, R.J.; Luther, J.E.; Beaudoin, A.; Goodenough, D.G.; Dechka, J.A., (2008). Monitoring Canada's forests. Part 1: Completion of the EOSD land cover project. *Canadian Journal of Remote Sensing* 34(6): 549-562.
- Yim, J. S. Kim, Y. H. Kim, S. H. Jeong, J. H., and Shin, M. Y. (2011). Comparison of the k -nearest neighbor technique with geographical calibration for estimating forest growing stock volume This article is one of a selection of papers from Extending Forest Inventory and Monitoring over Space and Time. *Can. J. For. Res.*, vol. 41, no. 1, pp. 73–82.
- Zagalikis, G., Cameron, A.D. and Miller, D.R. (2005). The application of digital photogrammetry and image analysis techniques to derive tree and stand characteristics. *Canadian Journal of Forest Research*, 35, pp. 1224–1237.
- Zink, M., Bachmann, M., Brautigam, B., Fritz, T., Hajnsek, I., Moreira, A., Wessel, B., and Krieger, G. (2014). TanDEM-X: The new global DEM takes shape. *IEEE Geoscience and Remote Sensing Magazine*, Vol. 2, No. 2, pp.8-23.