ABSTRACT

Title of Dissertation: FUSING GEDI LIDAR AND TANDEM-X INSAR OBSERVATIONS FOR IMPROVED FOREST STRUCTURE AND BIOMASS MAPPING

Wenlu Qi, Doctor of Philosophy, 2018

Dissertation directed by: Professor Ralph O. Dubayah
Department of Geographical Sciences

The upcoming NASA's Global Ecosystem Dynamics Investigation (GEDI) mission presents an unprecedented opportunity to advance current global biomass estimates. However, gaps are expected between GEDI's ground tracks, requiring the development of fusion-based methodologies to contiguously map forest biomass at satisfactory resolutions and accuracies. This dissertation is built on the complementary advantages of observations from GEDI and DLR's TerraSAR-X/TanDEM-X (TDX)) Interferometric Synthetic Aperture Radar (InSAR) mission. To meet the goal of mapping forest structure and biomass contiguously and accurately, three types of fusion strategies have been investigated.

First, a simulated GEDI-derived digital terrain model (DTM) was utilized to improve height estimation from TDX. Forest heights were initially derived from TDX coherence alone as a baseline using the widely used Random Volume over Ground (RVoG) scattering model. Here, assumptions about RVoG parameters – extinction coefficient (σ) and ground-to-volume amplitude ratio (μ) – were made. Using an

external DTM derived from simulated GEDI lidar data, RVoG model was used to calculate spatially varied σ values and derived forest heights with better accuracy. TDX forest height estimation was further improved with the aid of simulated GEDI-derived DTM and canopy heights. The additional use of simulated GEDI canopy heights as RVoG input not just refined σ but also enabled the estimation of μ. Based on these parameters, forest heights were improved across three different forest types; biases were reduced from 1.7–3.8 m using only simulated GEDI DTMs to -0.9–1.1 m by using both simulated GEDI DTMs and canopy heights. Finally, wall-to-wall TDX heights were used to improve biomass estimates from simulated GEDI data over three contrasting forest types. When using simulated GEDI sampled observations alone, uncertainties were estimated statistically to be 9.0–19.9% at 1 km. These were improved to 5.2–11.7% at the same resolution by upscaling simulated GEDI footprint biomass with TDX heights. The GEDI/TDX data fusion also enabled the generation of biomass maps at a fine spatial resolution of 100 m, with uncertainties estimated to be 6.0–14.0%.

Through the exploration of these fusion strategies, it has been demonstrated that a fusion-based mapping method could realize the generation of forest biomass products from GEDI with unprecedented resolutions and accuracies, while taking advantage of global seamless observations from TDX.

FUSING GEDI LIDAR AND TANDEM-X INSAR OBSERVATIONS FOR IMPROVED FOREST STRUCTURE AND BIOMASS MAPPING

By

Wenlu Qi

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:
Professor Ralph Dubayah (Chair)
Professor Christopher Justice
Research Assistant Professor John Armston
Research Scientist Dr. Temilola Fatoyinbo
Professor Joseph Sullivan (Dean's Representative)

Acknowledgements

I would like to express my deepest gratitude to my adviser Dr. Ralph Dubayah, who has been giving me invaluable guidance and tremendous help in this research. I had many ups and downs and made lots of mistakes during the past five years, but Dr. Dubayah has always been patient and never stopped encouraging me to learn from my mistakes and to grow as a serious scientist. He provided me lots of research opportunities I would never be able to have from anywhere else. Thanks to Dr. Dubayah, I also got the chance to collaborate with Dr. Konstantinos Papathanassiou and Dr. Matteo Pardini from German Aerospace Agency, from whom I learn the most advanced knowledge in forestry application of InSAR remote sensing and receive lots of help in TanDEM-X InSAR data processing and forest parameter modeling.

Special thanks go to my committee members - Dr. Chris Justice for his valuable advice in my research direction, Dr. Temilola Fatoyinbo and Dr. Guoqing Sun for their constructive suggestions in the radar/lidar fusion methods, Dr. John Armston for the big help in forest biomass estimation, and Dr. Joseph Sullivan for his insights on the ecological aspect of this study. I would also like to thank the research team of Dr. Chengquan Huang for providing the VCT disturbance products from Landsat. My gratitude also go to Dr. George Hurtt, Dr. Laixiang Sun, Dr. Shunlin Liang, Dr. Rachel Berndtson, Liz Smith and Marc Lennon for their supports in the department.

A big thank you to Dr. Seungkuk Lee, Dr. Hao Tang, Dr. Wenli Huang, Dr. Katelyn Dolan, Dr. Steven Hancock, Dr. John Armston, Suzanna Marselis and Dr. Svetlana Saarela for your generous help in data processing, paper comments, and results discussions. Thank you to all current and former members in VCL lab for making this chilly lab full of love, passion, energy and warmth. I am so lucky to know and be friends with you guys.

Thanks to the NASA Earth Space and Science Fellowship and the NASA GEDI mission for the support.

Table of Contents

Acknowledgements ii
Table of Contents iii
List of Tables vi
List of Figures vii
Chapter 1: .... Forest Structure and Biomass Estimation Using Fused Spaceborne Lidar and InSAR Data 1
1.1 Motivation 1
1.2 Background 3
1.2.1 Significance of forest biomass estimation to global carbon cycle study .. 3
1.2.2 Significance of forest biomass estimation to terrestrial ecosystem and biodiversity studies 6
1.2.3 Relevance of forest structure to biomass estimation and limits of field data 7
1.2.4 Lidar observations of forest structure and biomass 9
1.2.5 SAR observations of forest structure and biomass 10
1.2.6 InSAR observations of forest structure and biomass 12
1.3 Research Objectives 14
1.4 Dissertation Outline 15
Chapter 2: Improved TanDEM-X InSAR Height Estimates Using GEDI-derived DTM 17
2.1 Introduction 17
2.2 Study Area and Dataset 23
2.2.1 Study area 23
2.2.2 Lidar data 24
2.2.3 TDX data 25
2.3 Methods 27
2.3.1 Overall Experimental Design 27
2.3.2 Lidar data processing 31
2.3.3 TDX data processing 35
2.3.4 Overall processing strategies 38
2.4 Results 40
2.4.1 Simulating TDX volume decorrelation from lidar forest height 40

2.4.2 Canopy structure retrievals without an ancillary DTM to correct ground phase 45
2.4.3 Canopy structure retrievals using LVIS and GEDI DTM 46
2.5 Discussion 54
2.6 Conclusion 59
2.7 Acknowledgements 60
Chapter 3: ..... Improved Forest Height Estimation by Fusion of Simulated GEDI Lidar Data and TanDEM-X InSAR Data 61
3.1 Introduction 61
3.2 Test Sites and Data 65
3.2.1 Test sites 65
3.2.2 Datasets 68
3.3 Single-polarization RVoG Inversion and Combination with GEDI Data .. 72
3.3.1 RVoG model and height inversion from single-pol InSAR data 72
3.3.2 Combining RVoG single-pol inversion with GEDI data 75
3.4 Results 80
3.4.1 Case A – Only TDX 80
3.4.2 Case B – Using simulated GEDI DTM 82
3.4.3 Cases C1 and C2 – Using simulated GEDI DTM and canopy heights ... 84
3.5 Discussion 89
3.6 Conclusions 94
3.7 Acknowledgements 94
Chapter 4: ..... Forest Biomass Estimation over Three Distinct Forest Types Using Tandem-X InSAR Data and GEDI Lidar Data 96
4.1 Introduction 96
4.2 Study Sites and Datasets 100
4.2.1 Study sites and field data 100
4.2.2 Lidar data 104
4.2.3 TDX data 105
4.3 Methods 107
4.3.1 Establishment of relationships between field biomass and lidar data ... 108
4.3.2 Establishment of relationships between GEDI biomass and TDX heights 109
4.3.3 Uncertainty analysis for the biomass products at mapping scales of 100 m and 1 km 112

4.4 Results 116
4.4.1 Biomass estimation 116
4.4.2 Uncertainty estimation 122
4.5 Discussion 126
4.6 Conclusions 128
4.7 Acknowledgements 129
Chapter 5: ..... Conclusions and Future Research 130
5.1 Summary of the Research 130
5.1.1 Evaluating the effectiveness of GEDI elevation data to improve forest height estimations from TDX 131
5.1.2 Improving the parameterization of TDX inversion model to enhance forest height estimation using GEDI canopy height and elevation data 132
5.1.3 Assessing the efficacy of TDX and GEDI data fusion to improve estimations of forest biomass 134
5.2 Future Research 136
Bibliography 141

List of Tables

Table 2-1. Summary of acquisition parameters for both TDX images. 26

Table 2-2. Relationships of TDX scattering phase height and RVoG-derived forest height with LVIS relative heights. 53

Table 3-1. Summary of TDX acquisitions over the study areas. 71

Table 3-2. RVoG parameter summary for different cases performed in this study. ... 76

Table 3-3. Validation results of RVoG heights from all cases at 90 m resolution. 83

Table 4-1. Summary of acquisition time for the used field, lidar, and InSAR data. 101

Table 4-2. Summary of scales of field data used for each forest type. 103

Table 4-3. Summary of TDX acquisitions over the study areas. 106

Table 4-4. Relationship between field biomass and lidar RH metrics for each site. 117

List of Figures

Figure 1-1. (Data: CDIAC/NOAA-ESRL/GCP) Perturbation of the global carbon cycle caused by anthropogenic activities, averaged globally for the decade 2007-2016 (Gt CO_2/yr) (Quéré et al. 2018). How terrestrial processes are absorbing the "missing carbon" and how long they can continue is one of the critical and challenging questions for understanding future climate change (Hall et al. 2011). 5

Figure 1-2. Framework of the primary three chapters in this dissertation. 16

Figure 2-1. Hubbard Brook Experimental Forest (HBEF) (a) Study Site. Blue box shows the area of the used TDX data (for both 2011 and 2012 acquisitions); green box represents the area of the cropped TDX images shown in (b) (c) and (d); the boundary of HBEF is delineated in red. (b) TDX magnitude in dB (black to white: 10 – 40 dB). (c) TDX interferometric coherence (magnitude) (black to white: 0 - 1). (d) The associated TDX DEM (green to light yellow: 100 m – 1200 m) from which the steepness of the area can be observed. To avoid confusion, InSAR DEM refers to the scattering phase elevation whereas the DTM in this manuscript refers to the bare ground elevation. Their difference is the so-called scattering phase height. Figures (b), (c) and (d) were all generated using TDX 2011 acquisition. 26

Figure 2-2. (a) Multi-beam lidar coverage of GEDI. GEDI's three lasers, including one full power laser and two coverage lasers, generate 14 ground tracks (Dubayah et al. 2014). (b) Different tracks (colored differently) at the latitude of HBEF in nominal one-year period. The irregular spacing between tracks is because some are obtained during different passes (that are from a different set of 14 tracks). This spacing occurs because the ISS is in a non-repeating orbit. 32

Figure 2-3. Simulated GEDI tracks over HBEF (watershed shown in green) in nominal 1 operation year. Blue solid lines: full power laser tracks; red dashed lines: coverage laser tracks. 33

Figure 2-4. Simulated GEDI elevation and canopy top height (RH100) coverages over HBEF. 34

Figure 2-5. RVoG parameterization for inversion of forest height from single-polarized InSAR data. 37

Figure 2-6. Variation of simulated volume decorrelation (γ_{vol}) corresponding to an increase of forest height at extinction coefficient σ values of 0.001, 0.01, 0.02, 0.05, 0.1 and 0.2 dB/m at the incidence angle (θ_0) of ~36° without considering the effect of terrain slopes. 38

Figure 2-7. A framework for estimating height variables from TDX. Three different strategies were performed in the study: (1) Derive RVoG forest height without using a DTM to correct ground phase; (2) Derive RVoG forest height using ground phase

simulated from LVIS/GEDI DTMs; (3) Derive scattering phase height using simulated ground phase from LVIS/GEDI DTMs. 39

Figure 2-8. Panels (a) and (b) are the observed volume coherence for the two TDX acquisitions; Panels (c) and (d) are the corresponding simulated volume coherence using LVIS canopy top height and assuming a fixed extinction coefficient. Two transect lines shown in (a) are for detailed comparisons between the observation and simulation in Figure 2-9. 42

Figure 2-9. Comparisons between the observed and simulated volume decorrelations at 90 m resolution over the two HBEF transect lines shown in Figure 2-8 (a). Figures (a) and (b) correspond to the upper transect line; figures (c) and (d) correspond to the lower transect line. 43

Figure 2-10. Observed vs. simulated volume coherences at 90 m resolution. P-value < 0.0001 for both r^2. 44

Figure 2-11. Forest height from RVoG model without using an external DTM to correct the ground phase vs. LVIS RH100 at 90 m resolution. 46

Figure 2-12. 2011 and 2012 scattering phase height derived by using different DTMs to correct the topographic phase. (e) is the reference canopy top height from LVIS RH100. 47

Figure 2-13. Scattering phase height of TDX compared against RH100, RH75, RH50 and RH25 from LVIS at 90 m resolution over the two transect lines in Figure 2-8 (a). Figures (a) and (b) correspond to the upper transect line; figures (c) and (d) correspond to the lower transect line. Migration of scattering phase height away from RH50 to RH75 or RH25 can be observed in the black dotted boxes. 48

Figure 2-14. 2011 and 2012 scattering phase height using different DTMs vs. LVIS RH100 at 90 m resolution. 50

Figure 2-15. Forest height derived from the RVoG model respectively using LVIS and GEDI DTMs. Reference canopy top height from LVIS RH100 is given in (e)... 51

Figure 2-16. Forest height derived from the RVoG model using different DTMs to estimate the ground phase vs. LVIS RH100 at 90 m resolution. 52

Figure 3-1. Simulated two-year GEDI tracks over HBEF, TEF and LSBS test sites respectively based on cloud free and 50% cloud cover conditions. 67

Figure 3-2. Simulated GEDI observations of elevation and canopy height over nominal two-year period based on cloud-free and 50%-cloud cover conditions..... 70

Figure 3-3. The basis of the RVoG model. Forest structure in (a) is modeled using the two-layer scattering model in (b) with ground elevation z_0 and volume height (h_V). Scatterers are randomly distributed and oriented inside the forest volume (Cloude and

Papathanassiou 2003). F(z), radar reflectivity of forest scatterers at different height z, decays as a function of extinction coefficient (σ) as shown in (c). The term φ_0 denotes the ground phase ($e^{i\varphi_0} = e^{i\kappa_z z_0}$) and μ is the ground-to-volume amplitude ratio (Cloude and Papathanassiou 2003). 72

Figure 3-4. Main procedures for the four different fusion approaches – cases A, B, C1 and C2. 75

Figure 3-5. An increase of h_V up to HoA corresponds to the decrease of $|\tilde{\gamma}_V|$ for a fixed σ. For the same $|\tilde{\gamma}_V|$ value, a higher σ' value derived a larger forest height h_V'. 78

Figure 3-6. (Case A) (a) Forest heights derived using fixed extinction (σ) values of 0.3 dB/m (for leaf-off, 0.2 dB/m was used for leaf-on condition), 0.2 dB/m, and 0.3 dB/m respectively for HBEF, TEF, and LSBS. (b) Comparisons of the derived heights and reference lidar heights at 90 m resolution. 80

Figure 3-7. (Case B) (a) Forest heights derived from complex TDX coherence using simulated GEDI DTM, based on cloud-free and 50%-cloud cover conditions. (b) Comparisons of the derived heights and reference lidar heights at 90 m resolution. . 82

Figure 3-8. (Case C1) (a) Forest heights derived from complex TDX coherence using DTM and canopy height derived from simulated GEDI observations, respectively based on cloud-free and 50%-cloud conditions, to constrain σ. (b) Comparisons of the derived heights and reference lidar heights at 90 m resolution. 85

Figure 3-9. (Case C2) (a) Forest heights derived from complex TDX coherence using DTM and canopy height derived from simulated GEDI observations, respectively based on cloud-free and 50%-cloud cover conditions, to constrain σ and μ. (b) Comparisons of the derived heights and reference lidar heights at 90 m resolution. . 87

Figure 3-10. Different model performance corresponding to the four different cases over HBEF, TEF and LSBS under both cloud-free and 50%-cloud cover conditions. 90

Figure 4-1. TDX coherence maps overlaid by in situ plots (red) and simulated GEDI tracks (blue) for the nominal 2-year period after cloud decimation based on 50% cloud cover over (a) Hubbard Brook, (b) Teakettle and (c) La Selva test sites..... 101

Figure 4-2. Methods designed to estimate biomass and uncertainty maps respectively using wall-to-wall (reference) and GEDI-track lidar data, and fused GEDI/TDX data. 108

Figure 4-3. Area selection for establishing regression models for Grid cell #1 (each grid has a size of 1 km × 1 km). (a) When Grid cell #1 contained GEDI tracks, co-located GEDI biomass and TDX height (derived with the aid of GEDI observations) in Grid #1 were used to train the regression model. (b) When Grid cell #1 did not

contain GEDI observations, co-located GEDI biomass and TDX height (derived using TDX coherence alone) in an enlarged area around Grid #1 (2 km × 2 km) were used to develop the regression model. (c) When the enlarged area in (b) did not contain enough GEDI observations, the area was enlarged to 3 km × 3 km for establishing the GEDI biomass–TDX height model. 111

Figure 4-4. Lidar-predicted biomass vs. in situ biomass at the subplot level (25 m over HBEF and LSBS, or 33 m over TEF). 117

Figure 4-5. Biomass maps predicted from the wall-to-wall- (left) and GEDI- (middle) lidar data, and TDX heights (right) respectively at 25 m, 100 m and 1 km over (a) HBEF, (b) TEF and (c) LSBS. 120

Figure 4-6. GEDI biomass vs. co-located TDX biomass at 25 m respectively over (a) HBEF, (b) TEF, and (c) LSBS. 120

Figure 4-7. GEDI vs. Reference (wall-to-wall-lidar) biomass (top row), and TDX vs. Reference biomass (bottom row) at 1 km. 122

Figure 4-8. Uncertainties (Mg/ha) for biomass mean estimates at 100 m and 1 km scales respectively using the wall-to-wall- (left) and GEDI- (middle) lidar data, and TDX heights (right) over (a) HBEF, (b) TEF, and (c) LSBS. 125

Figure 4-9. Uncertainty (%) of wall-to-wall- and GEDI-lidar biomass, and GEDI/TDX-fusion biomass at 100 m (top row) and 1 km (bottom row). Uncertainty (%) was calculated as uncertainty (Mg/ha) divided by mean biomass (Mg/ha) of each grid cell. For consistency, uncertainties at 100 m were only compared over grid cells with valid GEDI estimates. 125

Chapter 1: Forest Structure and Biomass Estimation Using Fused Spaceborne Lidar and InSAR Data

1.1 Motivation

Accurate estimation of global forest structure and above ground biomass are critical to advancing our understanding of terrestrial carbon dynamics and their implications for climate, habitat and biodiversity (Goetz and Dubayah 2011). Numerous studies have demonstrated the particular strength and efficacy of the lidar remote sensing technique in estimating biomass via accurate measurements of forest structure. However, given the limited and inconsistent distribution of worldwide lidar data, it is still a challenging issue to provide global forest biomass maps at the spatial resolutions (100 m–1 km) and accuracies (≤50 Mg/ha, or better than 20%, whichever is lower) (Houghton et al. 2009) needed to reduce the large uncertainties in the global carbon budget and projections of future climate from surface-atmosphere carbon exchange analysis (Hall et al. 2011).

Much more complete and consistent lidar coverage will become available with the launch of Global Ecosystem Dynamics Investigation (GEDI) – the National Aeronautics and Space Administration (NASA)'s spaceborne lidar mission. The billions of lidar observations that will be collected by GEDI will revolutionize our understanding and quantification of forest vertical structure and biomass globally between 51.5° north and south latitude (Stysley et al. 2015). Although gaps will

remain between the adjacent 25-m footprints of GEDI, in both the along- and across-track directions, observations from this mission alone will allow the creation of moderate-resolution global biomass maps at approximately 1 km using statistical methods. These GEDI-based products, if provided with appropriate characterization of uncertainties and sufficient accuracy, will present superior forest structure and biomass information compared to current global estimates. Nevertheless, there is still a pressing need to integrate GEDI observations with ancillary data from other satellite missions, since GEDI cannot directly provide a wall-to-wall map with individual footprints, to provide biomass estimates at finer resolutions and accuracies and minimize uncertainties in the terrestrial carbon budget (Pardini et al. 2016; Qi and Dubayah 2016).

Driven by this need, this dissertation explores the efficacy of fusing simulated GEDI lidar data and data from the German Aerospace Center (DLR)'s TerraSAR-X (2007 launch)/TanDEM-X (2010 launch) (a combined mission abbreviated simply as TDX) Interferometric Synthetic Aperture Radar (InSAR) mission to improve forest structure and biomass mapping. TDX single-pol InSAR data have a resolution of approximately 3 m and contiguous global coverage, and have been successfully applied to provide forest structure and biomass maps at regional scales, using external bare ground topography or other *a priori* forest parameters to constrain the inversion model (Askne et al. 2017; Kugler et al. 2014). Yet accuracies of the estimated forest attributes are subject to the availability and accuracy of such prerequisite information.

While there are multiple data sources available, GEDI would certainly provide one of the most accurate estimates of ground topography and forest vertical structure at its footprint level globally. Therefore, it forms the fundamental basis for a complementary observation of forest structure between GEDI lidar and TDX InSAR — GEDI data are more accurate but spatially limited, whereas TDX InSAR data are spatially contiguous but require certain regional calibrations. There is a timely opportunity for exploiting the two advanced datasets in a synergistic way to map forest structure and biomass at a finer spatial resolution and higher accuracy than either mission can achieve alone.

1.2 Background

1.2.1 Significance of forest biomass estimation to global carbon cycle study

Carbon dioxide has been increasingly emitted into the atmosphere through human activities for the past decades, contributing to global warming that potentially leads to environmental changes such as rising sea levels, melting Arctic sea ice, increasing ocean temperature, severe weather and warmer winters (IPCC 2013). Fossil fuel use, forest deforestation and land-use change have been identified to be the largest contribution of carbon dioxide emission (Quéré et al. 2018), and each of these sources has different levels of uncertainties (Figure 1-1), leading to an imbalanced global carbon cycle known as “missing carbon”(Houghton et al. 2012). Significant efforts have been made to address this issue and to reduce the uncertainties in the

global carbon cycle. For example, a recent study indicates an accelerated growth in forests that may have a major effect on the global carbon and that the world's trees may have sequestered enough carbon during the period from 1990 to 2007 to account for the entire "missing sink" (Pan et al. 2011). However, uncertainties in these estimates are large, particularly in tropical forests, because of the sparse sampling of field data at local scales and the widely-varied quality (Popkin 2015). Reducing the high uncertainty levels in the quantification of terrestrial carbon budget will significantly help to better understand global carbon cycle and its implication for climate change (Hall et al. 2011).

As one of the major carbon sinks and sources in the global ecosystem, forest stores ~45% of terrestrial carbon and contributes to ~50% of terrestrial net primary production (Canadell et al. 2007; Houghton and Goetz 2008). Carbon stocks in forests are usually stored in the form of biomass (including both living and dead aboveground biomass (AGB, which includes stems, stumps, branches, bark, seeds and foliage) and belowground biomass (BGB, which includes live roots)). One way to reduce uncertainties in the terrestrial carbon budget is to improve the measures of amount and distribution of global forest AGB (hereafter just "biomass") (Malhi et al. 1999). Spatially explicit measurements of biomass also support the definition of policies in the context of climate change mitigation strategies which are part of Reducing Emissions from Deforestation and forest Degradation in developing countries (REDD+) program under the United Nations Framework Convention on

Climate Change (UNFCCC) (Langner et al. 2014; Naesset et al. 2013). Therefore, large-scale forest biomass monitoring has been a key mission and measurement objective for NASA (Dubayah et al. 2014; Joshi et al. 2017; Rosen et al. 2006).

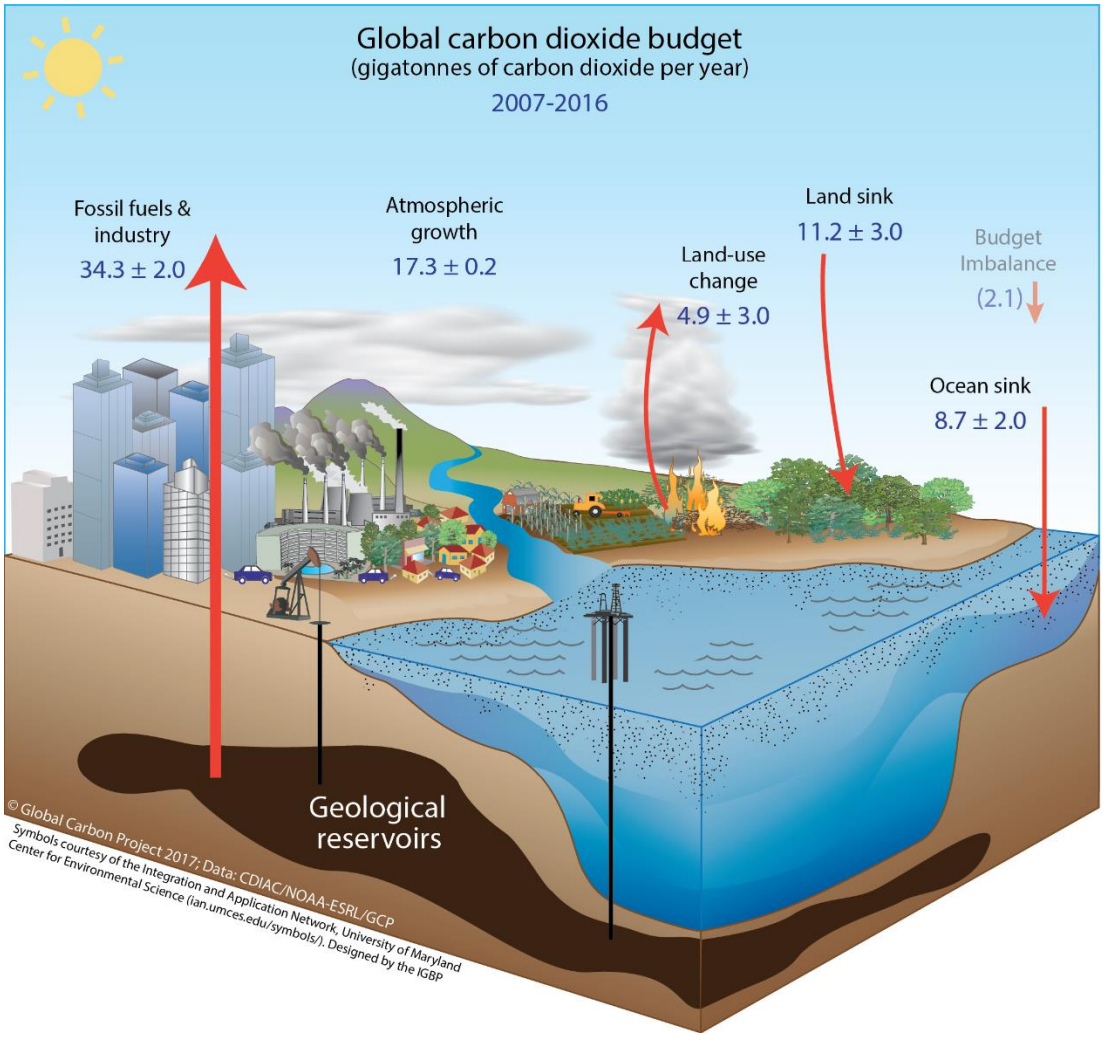


Figure 1-1. (Data: CDIAC/NOAA-ESRL/GCP) Perturbation of the global carbon cycle caused by anthropogenic activities, averaged globally for the decade 2007-2016 (Gt CO_2/yr) (Quéré et al. 2018). How terrestrial processes are absorbing the "missing

carbon" and how long they can continue is one of the critical and challenging questions for understanding future climate change (Hall et al. 2011).

1.2.2 Significance of forest biomass estimation to terrestrial ecosystem and biodiversity studies

Recent studies show that important terrestrial ecosystem and biodiversity are under increasing pressure from human-induced land-use change and global climate change (Bergen et al. 2009; Millennium Ecosystem Assessment 2005). In the past decades, biodiversity has been lost at an alarming rate, with a 1,000-fold increase in species extinction in the last century alone and a continuous loss of about 50,000 species annually (Millennium Ecosystem Assessment 2005). The losses greatly reduce Earth's biological heterogeneity and ecological complexity, and ultimately jeopardize the function and stability of the terrestrial ecosystem and the sustainable development of human society. Therefore, there is an urgent need to monitor the biodiversity of wild species and to protect important habitat which are often carbon sinks (Whitehead 2011).

Biodiversity is often distributed in a way that some areas with biodiversity far richer than others. Identification and mapping of these areas are imperative in biodiversity and conservation studies. Forest biomass monitoring system, particularly those using remote sensing data that give important biophysical variables and biomass information, are highly suitable for recognizing these areas given their high spatial and temporal mapping capabilities. These data can be used to support the

mapping of certain individual species, making predictions based on their habitat requirements and exploring patterns by fusing with field observations (Nagendra 2001). Some key biophysical variables are also important parameters in terrestrial ecosystem and help to improve our understanding of global carbon, water and energy fluxes between biosphere and atmosphere (Mu et al. 2007; Myneni et al. 2002; Zhao and Zhou 2005).

1.2.3 Relevance of forest structure to biomass estimation and limits of field data

Biomass is a developing outcome of many complex processes that act at different temporal and spatial scales. It is determined by the balance between inputs, i.e. growth, and outputs, i.e. mortality, which are influenced by many limiting factors such as water supply, temperature and nutrient, etc. An environment that favors high growth rates will thus tend to correlate with high biomass and vice versa. These factors affect carbon stocks principally by determining stand structure, such as basal area (Álvarez-Dávila et al. 2017).

Direct measurements of biomass are typically conducted through the destructive sampling method, which requires harvesting sample trees that are then dried and weighed (Nunery and Keeton 2010). This method is expensive and time- and labor-intensive. To avoid this process, empirical relationships have been developed to estimate biomass indirectly through field measurements of tree heights, diameter at breast height (dbh), basal area, crown volume, and other parameters that are highly correlated with biomass (Jenkins et al. 2003; Li et al. 2015; Yang et al.

2013). These relationships have been used to accurately estimate biomass through properly-conducted field inventories of forest structural properties.

Many national-scale field inventories have been carried out in the past few decades (Blackard et al. 2008; Pollard et al. 2006; Wilson et al. 2013). These field data have been successfully combined with land use and land cover change maps to track changes in biomass for carbon stock and flux analysis. Nevertheless, field measurements of biomass are inefficient; they are not practical ways to produce spatial biomass maps over a landscape (Blackard et al. 2008; Treuhaft et al. 2009). Therefore, forest maps based on remote sensing data are often combined with field-measured biomass to provide large-scale estimates of forest biomass.

Various remote sensing methods have been developed to complement field measurements and produce spatial biomass maps efficiently (Benson et al. 2010; Chen et al. 2009; Hyde et al. 2006). While biomass maps have been generated using passive optical data, these maps tend to have large uncertainties at scales less than a few kilometers due to the limited sensitivity of sensors to forest vertical structure, particularly after canopy closure (Adam et al. 2010; Anderson et al. 2008; Fassnacht et al. 2014). On the other hand, active remote sensing data, in particular those acquired by Light Detection and Ranging (lidar) and Radio Detection and Ranging (radar) sensors, are more sensitive to forest structural attributes and have been widely explored for biomass estimation (Hall et al. 2011).

1.2.4 Lidar observations of forest structure and biomass

The efficacy of lidar remote sensing in biomass estimation has been widely acknowledged due to its capability to provide accurate three-dimensional (3D) estimates of forest structure that are highly correlated with forest biomass (Drake et al. 2002; Goetz and Dubayah 2011; Lefsky et al. 2002). Lidar measures ground elevation and canopy structure by recording the time that elapses between the emission of a short-duration laser pulse (usually in the infrared of the electromagnetic spectrum) and the arrival of the reflection from ground and canopy to the receiver. The size of a footprint (the area illuminated by the laser beam) varies in different lidar systems and can be classified as small- (< 2 m), medium- (10–30 m) and large- (> 50 m) footprint. The way lidar instruments record the returned signal may also differ, either in discrete returns or fully-digitized waveforms. Lidar remote sensing of forests is currently dominated by the deployment of airborne platforms, including commercial small-footprint sensors, both discrete return and waveform, as well as research sensors, such as the medium-footprint waveform Land, Vegetation and Ice Sensor (LVIS) (Anderson et al. 2006; Dubayah et al. 2010; Huang et al. 2013).

Most airborne lidar data are samples at the stand level or densely cover small areas due to the high costs of flight time, the limit of scanning to near nadir to prevent ranging errors, and gaps caused by the pitch and roll of the aircraft. To obtain stable and consistent estimates of forest structure and biomass, there is an increasing interest in exploring data acquired by spaceborne lidar missions (Frolking et al. 2009;

Hayashi et al. 2015a; Hayashi et al. 2015b). The Geoscience Laser Altimeter System (GLAS) aboard NASA's Ice, Cloud, and land Elevation satellite (ICESat) has been used successfully to obtain various forest structural properties, such as canopy height, LAI and canopy cover, as well as biomass (Lefsky et al. 2005). However, since this mission was optimized for ice sheets, GLAS provided low sampling density over mid-latitude and tropical forests (Dolan et al. 2011; Dolan et al. 2009; Fatoyinbo and Simard 2013; Hall et al. 2011). The lack of comprehensive ecosystem structure observations has been a major obstacle for reducing uncertainties in the terrestrial carbon budget and for understanding the role of forests as carbon sinks and sources.

GEDI has been scheduled for launch by NASA in 2018 to solve this problem. During its nominal two-year mission, GEDI will provide unprecedented measurements of forest height and structure at a medium-footprint size (around 25 m). Although gaps will remain between GEDI's ground-tracks, this mission alone will enable the generation of forest structure and biomass products at resolutions of 1 km. Higher resolution products can also be created by combining accurate GEDI lidar measurements with data from other spaceborne missions that provide more complete coverage but require calibration and validation from lidar (Qi and Dubayah 2016).

1.2.5 SAR observations of forest structure and biomass

Synthetic Aperture Radar (SAR) is a method that synthesizes a very long radar aperture by integrating the magnitude and phase of the echoes returned from a feature over the entire time it is viewed, thus providing high-resolution images of the

feature. The amplitude of the returned radar signal for a given ground resolution cell, the so-called backscatter coefficient, is determined by the signal wavelength, the characteristics of the target, and the acquisition geometry (Askne and Santoro 2005; Brown and Sarabandi 2001; Cloude and Papathanassiou 2003). In a forest setting, the SAR wavelength determines which components of the forest the SAR signal will interact with most. For example, shorter microwave wavelengths (e.g. ~3 cm for the X-band) interact mainly with the surface of the forest canopy and are backscattered by foliage and small branches, whereas longer wavelengths (e.g. ~23 cm for the L-band or ~74 cm for the P-band) often penetrate into the canopy, with reflections coming from large branches, stems, and the surface of the terrain (Englhart et al. 2011; Neumann et al. 2012; Praks et al. 2007). SAR backscatter from forests also depends on a variety of forest characteristics including structure, canopy roughness, and water content.

SAR systems often polarize the radar signal in different ways (often horizontally or vertically) to yield valuable information regarding the vegetation type and structure, by independently recording the reflection of like-polarized energy (e.g. vertical send–vertical receive (VV) or horizontal send–horizontal receive (HH)) and cross-polarized energy (e.g. vertical send–horizontal receive (VH) or horizontal send–vertical receive (HV)) (Carver 1988; Jensen et al. 2000). Some previous studies have used P- to C-band and L- to C-band ratios at HV polarizations (PHV/CHV and LHV/CHV) to successfully predict biomass in boreal forests (Ranson et al. 1995;

Ranson and Sun 2000; Ranson and Sun 1994). Other studies have also found it possible to accurately estimate biomass from SAR backscatters in plantations or very simple types of forest. However, in general, SAR backscatters tend to produce biomass maps with large uncertainties unless they are integrated to large scales, and are often not responsive to biomass levels above ~100 Mg/ha, depending on the signal wavelength (Englhart et al. 2011; Mitchard et al. 2012; Montesano et al. 2013; Saatchi et al. 2011; Tsui et al. 2013). This is mainly because the two-dimensional (2D) format of SAR images only allows for an indirect estimation of the 3D structural attributes of forests.

1.2.6 InSAR observations of forest structure and biomass

InSAR and Polarimetric InSAR (PolInSAR) technologies have been developed recently to enable direct measurement of forest vertical structure (Garestier et al. 2008; Hajnsek et al. 2009; Kugler et al. 2014; Praks et al. 2007). The key observation in InSAR/PolInSAR measurements is the complex interferometric coherence that includes the interferometric correlation coefficient (magnitude) and the interferometric phase (interferogram). When the two SAR images forming the coherence are acquired at different times, changes in the observed forests (for example from wind and soil moisture) cause temporal decorrelation and reduce the quality of coherence for the forest structure estimations (Brolly et al. 2016; Chowdhury et al. 2014). In 2000, the Shuttle Radar Topography Mission (SRTM) collected the first simultaneous spaceborne InSAR data that produced global

scattering phase vertical location (the elevation at which the scattering is focused) (Kellndorfer et al. 2004). Using the SRTM elevation data, studies have derived scattering phase heights and biomass by subtracting from existing topographic data (Fatoyinbo and Simard 2013; Files 2012; Kellndorfer et al. 2010; Kellndorfer et al. 2004; Kenyi et al. 2009).

DLR recently launched a dual platform mission, TDX (still operational), as the first long-term, simultaneous, spaceborne InSAR mission that could potentially provide a means for retrieving data on forest structure at a high spatial resolution (Askne et al. 2013; Kugler et al. 2014; Soja and Ulander 2013; Solberg et al. 2013). Previous studies have estimated scattering phase height from TDX interferograms using external lidar Digital Terrain Models (DTMs) (Kugler et al. 2014; Soja and Ulander 2013). Forest parameters, including height, have also been derived from (Pol)InSAR TDX data, based on the widely used Random Volume over Ground (RVoG) scattering model (Askne et al. 2013; Kugler et al. 2014; Lee and Fatoyinbo 2015; Solberg et al. 2013). However, using single-polarized TDX data in this model posed an underdetermined problem, requiring input from ancillary DTMs or other forest parameters that were usually provided by airborne lidar data (Kugler et al. 2014; Soja and Ulander 2013). Dual-polarized acquisitions do not require *a priori* conditions, but are only available in some experimental areas and generally provide limited height diversity in the X-band, especially over high-cover tropical forests (Cloude et al. 2013; Kugler et al. 2014).

The upcoming GEDI mission will provide consistent worldwide topographic and forest structural data, thus providing the *a priori* information required by single-polarized TDX InSAR data to drive the SAR scattering models for wall-to-wall forest height mapping over large areas (Coyle et al. 2015; Pardini et al. 2016; Stysley et al. 2015). There is a potential to provide experimental height and biomass products at unprecedented resolutions and accuracies by combining these GEDI and TDX observations.

1.3 Research Objectives

This research investigates approaches to fuse GEDI lidar and TDX InSAR data to improve forest height and biomass mapping by addressing the following objectives.

1) Investigate the relationships between lidar canopy height and TDX interferometric coherence and examine the improvement on TDX height retrieval using GEDI-derived DTM.
2) Evaluate the effectiveness of using GEDI canopy height, additional to the GEDI-derived DTM, to improve forest height estimation from TDX.
3) Explore the feasibility of improving GEDI-based biomass estimation by using wall-to-wall TDX InSAR heights.

Since data from the GEDI mission are not available yet, this research simulated GEDI lidar observations based on Land, Vegetation, and Ice Sensor (LVIS)

and airborne laser scanning (ALS) data to investigate GEDI/TDX fusion over three different types of forest.

1.4 Dissertation Outline

This dissertation is subdivided into five chapters. The three primary chapters are structured in the format of journal articles and are self-contained yet closely related to each other (Figure 1-2). Chapters 2 and 3 focus mainly on improving TDX height estimation using GEDI-derived variables, whereas Chapter 4 primarily investigates the utility of GEDI-aided TDX height variables on GEDI-based biomass estimates.

Chapter 2 establishes the relationship between canopy height and TDX coherence observations using the physically-based SAR scattering model – RVoG. The established relationship is subsequently applied to derive canopy height from TDX coherence and then improved with the use of simulated GEDI-derived topography. This chapter also examines the inversion performance under distinct forest structure conditions, using TDX data acquired in leaf-on and leaf-off seasons. Chapter 3 further improves the parameterization of TDX RVoG inversion using simulated GEDI DTM and canopy heights. The utility of each GEDI-constrained RVoG parameter in height estimation is assessed over diverse forest types. In addition, this chapter examines the impact of cloud cover and foliage conditions on the fusion results. Chapter 4 examines the capability of fused GEDI/TDX data to estimate biomass at resolutions of both 100 m and 1 km, compared to when this is

done using GEDI data alone, and characterizes uncertainties associated with these products. Chapter 5 summarizes this dissertation, and discusses the implications of its major findings and their potential linkage to current and future spaceborne missions.

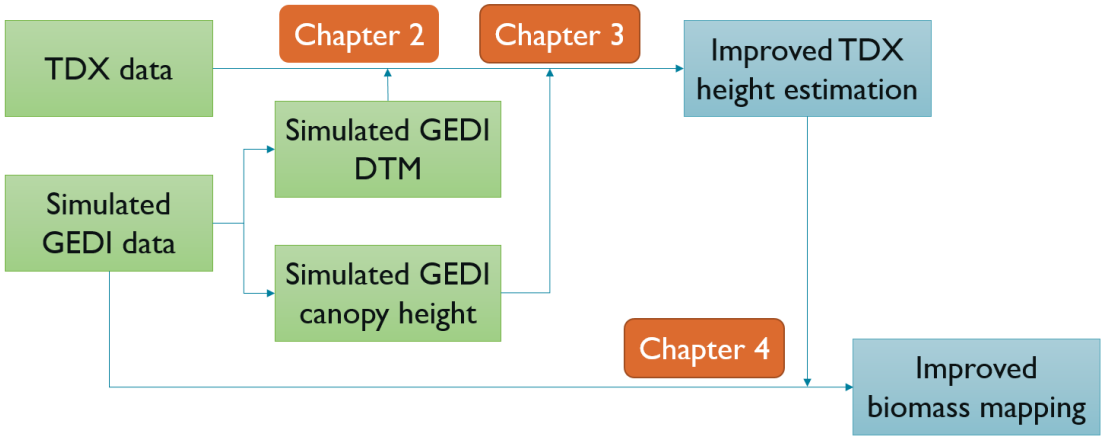


Figure 1-2. Framework of the primary three chapters in this dissertation.

Chapter 2: Improved TanDEM-X InSAR Height Estimates Using GEDI-derived DTM

2.1 Introduction

Forests play an important role in the terrestrial carbon cycle through their functioning as potential sources and sinks of carbon. Consequently, improved quantification of their carbon content in the form of above-ground biomass (AGB) at high spatial resolution and accuracy has been identified as a priority to reduce uncertainties on the net impact of deforestation and regrowth on atmospheric CO_2 concentrations (CEOS 2014; Drake et al. 2003; Lu and Liu 2014). Biomass is usually estimated from various elements of forest structural properties, particularly tree diameter and height, using allometric equations (Clark and Kellner 2012; Duncanson et al. 2015a; Duncanson et al. 2015b; Keller et al. 2001; Zhao et al. 2012). Though traditional inventory methods provide accurate biomass information at field scales, remote sensing technologies have demonstrated great advantages in estimating forest biomass at larger scales by measurement of biomass-related structural properties, such as canopy height and other vertical structure indices (Goetz and Dubayah 2011; Lu and Liu 2014).

Lidar has emerged as the one of the most effective remote sensing techniques for observing canopy vertical structure (Goetz and Dubayah 2011). Lidar remote sensing is currently dominated by deployment via airborne platforms. These include

commercial, small-footprint sensors, both discrete return and waveform, as well as research sensors, such as the medium-footprint waveform Land Vegetation and Ice Sensor (LVIS) (Anderson et al. 2006; Dubayah et al. 2010; Huang et al. 2013). Spaceborne deployment of lidar enables global coverage and potentially provides a means for obtaining stable and consistent estimates of biomass. The Geoscience Laser Altimeter System (GLAS) onboard ICESat has been used successfully to obtain forest structure measurements, but because it was optimized for ice sheets, had low sampling density over mid-latitude and tropical biomes (Dolan et al. 2011; Dolan et al. 2009; Fatoyinbo and Simard 2013; Hall et al. 2011). Obtaining a much more comprehensive set of ecosystem structure observations is thus a priority for advancing our understanding of the role of forests and reducing uncertainties in the global carbon cycle.

A new mission designed specifically to provide needed ecosystem structure measurements is now under development in the form of NASA's Global Ecosystem Dynamics Investigation (GEDI). Scheduled for deployment on the International Space Station (ISS) in 2018, GEDI will consist of three laser transmitters to produce billions of full waveform observations per year (Stysley et al. 2015). The GEDI baseline mission concept is comprised of a swath of 14 ground tracks of 25m diameter footprints separated by 60 m along-track and 500 m across-track. The measurements will produce a fine mesh of transect observations between 50° north and south latitude providing the most complete set of measurements of canopy

structure and bare earth (under canopy) topography yet achieved. However, GEDI is not a continuously imaging instrument, and gaps among its ground tracks and between adjacent swaths will remain. These gaps between swaths decrease as a function of latitude, mission length, and at orbital cross-over points. In the worst case, some persistently cloudy areas may never be observed.

One approach to providing continuous (that is wall-to-wall) mapping of structure and biomass is through the fusion of lidar data with data acquired from other spaceborne sensors. For example, large-scale height and biomass maps have been derived by fusing lidar observations from GLAS, with other spatially contiguous multispectral and SAR backscatter datasets, such as Landsat, MODIS and PALSAR (Mitchard et al. 2012; Saatchi et al. 2011). However, fusion based on multispectral or backscatter data tend to produce maps with large uncertainties at scales less than a few kilometers. This is primarily because neither passive optical nor backscatter data are sensitive to the full range of vertical structure and biomass. Leaf area index saturates early in successional state and SAR backscatter is often not responsive to biomass levels above about 100 Mg/ha, depending on the signal wavelength (Englhart et al. 2011; Mitchard et al. 2012; Montesano et al. 2013; Saatchi et al. 2011; Tsui et al. 2013).

Interferometric SAR (InSAR) and Polarimetric InSAR (Pol-InSAR) recently have been explored for providing better estimates of vertical structure (Garestier et al. 2008; Hajnsek et al. 2009; Kugler et al. 2014; Praks et al. 2007). InSAR methods are

based on obtaining at least two images from different vantage points at different times or simultaneous acquisitions from which interferograms may be constructed (Cloude and Papathanassiou 1998; Hajnsek et al. 2009; Papathanassiou and Cloude 2001). When images are acquired at different times, changes in the observed forests, for example from wind, soil moisture, etc., cause temporal decorrelation between the images and reduce the quality of the interferograms for forest structure estimation. The Shuttle Radar Topography Mission (SRTM) in 2000 collected the first spaceborne simultaneous InSAR data of global scattering phase vertical location, the elevation at which the scattering is focused (Kellndorfer et al. 2004). One common approach to estimating canopy height is to use InSAR scattering phase elevation (relative to the geoid) and subtract the elevation of the underlying topography -- provided it is known from some external source -- to produce what is called the scattering phase height (Fatoyinbo and Simard 2013; Kellndorfer et al. 2004; Kenyi et al. 2009). For example, heights and biomass have been derived based on this approach using the SRTM data over the United States (Files 2012; Kellndorfer et al. 2010), exploiting the existing topographic data available from the U.S. Geological Survey. The German Aerospace Center, DLR, recently launched a dual platform InSAR mission, TerraSAR-X (2007 launch) and TanDEM-X (2010 launch) (with the joint mission abbreviated as TDX). These instruments are capable of viewing the same portion of the Earth simultaneously at various spacecraft separation distances (baselines) to produce interferometric images. TDX is the first long-term SAR

interferometry mission in space and potentially provides a means for retrieving forest structural information at high spatial resolution (Askne et al. 2013; Kugler et al. 2014; Soja and Ulander 2013; Solberg et al. 2013).

Previous studies have estimated scattering phase height using TDX interferograms over forests where lidar Digital Terrain Models (DTMs) were available (Kugler et al. 2014; Soja and Ulander 2013). The feasibility of using TDX data to derive forest height has also been demonstrated by a widely used SAR scattering model – the Random Volume over Ground (RVoG) model (Askne et al. 2013; Kugler et al. 2014; Lee and Fatoyinbo 2015; Solberg et al. 2013). However, its efficacy is still unclear under varying landscape and terrain conditions, for example, high levels of canopy cover, forests composed of different tree species, and terrains with a wide range of slopes. Also, TDX acquires data in different seasons during which forests, especially deciduous forests, have varying environmental conditions. It is thus of interest to examine whether TDX retrievals of forest height are consistent as conditions vary seasonally (Abdullahi et al. 2016; Kugler et al. 2014; Olesk et al. 2015). Additionally, past studies deriving forest height estimates from single-polarized TDX data have relied heavily on high-resolution ancillary DTM datasets from airborne lidar measurements (Kugler et al. 2014; Soja and Ulander 2013). When suitable pre-existing topographic data do not exist, they require dual-polarized acquisitions which are only available in selected locations and which generally

provide limited height diversity at X-band, especially over forests with taller trees and higher densities (Cloude et al. 2013; Kugler et al. 2014).

One alternate approach for fusion is to examine the potential of using GEDI observations of topography with TDX for forest structure recovery. GEDI will provide billions of high quality but spatially limited samples over its global sampling grid, which when used to create a DTM may significantly improve TDX estimates of canopy height. These height estimates, if sufficiently accurate, could then be used as the basis for some type of blended height product from GEDI that combines laser-derived canopy structure with that derived from TDX, the latter being less accurate but spatially contiguous. The goal of this study is to take the first step towards such fusion by examining the efficacy of GEDI-derived topographic data for canopy height estimation from TDX.

This study investigates the performance of lidar/InSAR fusion using GEDI data simulated from airborne lidar data and TDX satellite data over a mountainous temperate forest – Hubbard Brook Experimental Forest (HBEF). HBEF provides a challenging experimental location given its mixture of broadleaf deciduous and conifer forests, moderate canopy heights and biomass, high canopy cover, rugged terrain, and large seasonal variation of environmental conditions. The investigation addresses the following questions regarding the estimation of canopy height and vertical structure from TDX: (1) Are TDX coherences responsive to height variations of HBEF and to what degree? (2) What accuracies are achievable using TDX with no

external DTM to correct the ground phase? (3) What accuracies are achievable using TDX when a high-resolution, spatially continuous DTM derived from airborne lidar is available? (4) What accuracies are achievable using TDX when a coarser-resolution DTM derived from simulated GEDI footprints is available? This study also examines the consistency of RVoG model height estimates under different environmental and canopy conditions: leaf-on vs. leaf-off and fall vs. winter.

The remainder of this study is organized as follows. 2.2 describes the overall experimental design to address the aforementioned questions. In 2.3, we briefly describe our study area, Hubbard Brook Experimental Forest, and the data sets used. We then outline the models, methods and processing approaches used for analyses of both TDX and GEDI/LVIS in 2.4. Our results with the full and GEDI-like DTM for deriving scattering phase heights and RVoG heights are presented in 2.5, where we also examine the relationship of these heights to common lidar relative height metrics. Finally, we discuss our results in more detail, focusing on both the potential and limitations of the methods presented for fusion of TDX and GEDI in 2.6.

2.2 Study Area and Dataset

2.2.1 Study area

Hubbard Brook Experimental Forest (HBEF) (43°56’ N, 71°45’ W) is located in the White Mountain National Forest in New Hampshire. It covers an area of 3,317 hectares and is characterized by a bowl-shaped valley with elevations ranging from

150-1050 m. HBEF is a temperate forest consisting of 80-90% deciduous northern hardwoods and 10-20% spruce-fir. Forest height in this study site varies between 5 m and 48 m, with a mean of 24 m. Average aboveground biomass of HBEF in 2001 is about 216 Mg/ha (Siccama et al. 2007). The remote sensing datasets used in this study include LVIS data, single-polarization acquisitions from TDX and auxiliary disturbance data products derived from Landsat.

2.2.2 Lidar data

LVIS is an airborne, medium-resolution (selectable 5- to 50-meter diameter) full-waveform footprint laser altimeter system developed by NASA Goddard Space Flight Center (GSFC) (Blair et al. 1999). By digitally recording the returned full waveform, LVIS provides accurate vertical structure of both the canopy-top and underlying topography within the illuminated area. LVIS data were collected in August 2009 at a footprint diameter of between 20 and 25 m, and used to derive ground elevation and height metrics at 25%, 50%, 75% and 100% (RH25, RH50, RH75 and RH100) of the cumulative waveform energy relative to the ground. LVIS ground elevations were used to create a 30 m resolution, wall-to-wall DTM for correcting the topographic phase component in the TDX interferogram. LVIS sub-canopy ground elevations are generally accurate to within tens of centimeters. The RH100 metric is assumed to be the canopy top (generally accurate to about 1 m) and the map is used for validation of the TDX-derived canopy height. The other RH

metric maps are used to help understand the physical significance of derived scattering phase heights (e.g. which RH metric is most similar to the phase heights).

2.2.3 TDX data

Both TDX acquisitions used in the study (Table 2-1; see an example in Figure 2-1) were collected in bistatic mode, where TerraSAR-X was transmitting the signal and both satellites were simultaneously receiving the returned signal. The images were acquired in Stripmap imaging mode with ground range resolution of 2.99 m and azimuth resolution of 3.30 m, and a scene size of ~35 km in width and ~56 km in length. The data were acquired at single-polarization (HH) at ~36° look angle from ascending orbits respectively on 10/21/2011and 01/28/2012. We used the data provided as Co-registered Single look Slant range Complex (CoSSC) format for the interferometric processing.

Between 10/14/2011 and 10/21/2011, the precipitation over our study site occurred only as rain or melted snow (with 42.4 mm in total within that week, 2.5 mm one day before the acquisition, and no rain during the acquisition date); whereas between 01/21/2012 and 01/28/2012, it was mostly snow, ice pellets or hail (with 109.7 mm in total within that week, 71.6 mm one day before the acquisition date and 5.1 mm during the acquisition date) due to the very low temperature (<-9°C in average) (NOAA 2011-2012). Therefore, we expect a higher frozen-state water content for the 2012 acquisition whereas a possibly higher fluid-state water content

for the 2011 acquisition. A frozen-state vegetation leads to a decreased vegetation dielectric constant (Anderson et al. 2006).

Table 2-1. Summary of acquisition parameters for both TDX images.

Date	10/21/2011	01/28/2012
θ **[º]**	36.2	36.1
Effective Baseline [m]	121.42	85.37
Height of Ambiguity [m]	-47.43	-68.12
κ_Z **[rad/m]**	-0.132	-0.092
Pol.	HH	HH
Mode	Ascending, Stripmap, Bistatic	Ascending, Stripmap, Bistatic

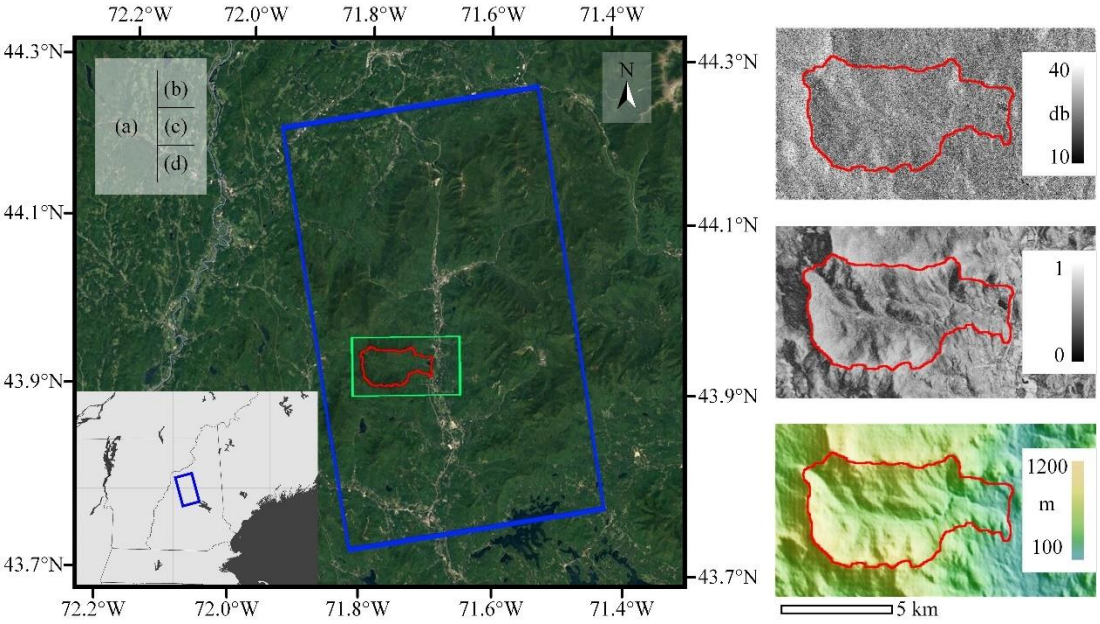


Figure 2-1. Hubbard Brook Experimental Forest (HBEF) (a) Study Site. Blue box shows the area of the used TDX data (for both 2011 and 2012 acquisitions); green box represents the area of the cropped TDX images shown in (b) (c) and (d); the

boundary of HBEF is delineated in red. (b) TDX magnitude in dB (black to white: 10 – 40 dB). (c) TDX interferometric coherence (magnitude) (black to white: 0 - 1). (d) The associated TDX DEM (green to light yellow: 100 m – 1200 m) from which the steepness of the area can be observed. To avoid confusion, InSAR DEM refers to the scattering phase elevation whereas the DTM in this manuscript refers to the bare ground elevation. Their difference is the so-called scattering phase height. Figures (b), (c) and (d) were all generated using TDX 2011 acquisition.

2.3 Methods

2.3.1 Overall Experimental Design

The logic of our experimental approach first dictates that we examine whether coherences respond to some aspect of height variations in the watershed. If they do not, or do so only weakly, it would be unlikely that height may be derived from TDX in our setting with an acceptable error. This is accomplished through testing the correlation between the observed TDX coherence and the reference canopy height. However, InSAR coherence is not only influenced by canopy height, but also by acquisition geometry, including the incidence angle, the spatial baseline and the local terrain slope, among others, which vary by acquisition and by pixel in a particular image (Hajnsek et al. 2009; Kugler et al. 2015; Kugler et al. 2014; Lee and Fatoyinbo 2015). For example, a lower TDX coherence for a certain acquisition is not necessarily related to a higher canopy height; it might come from a same or even

lower canopy height over a positively-slope terrain surface (i.e. a slope facing towards the radar line of sight) (Kugler et al. 2015). Therefore, we need a controlled experimental design that enables examining the response of coherence to canopy height variation under the same imaging geometry. In this study, we use an inverted RVoG model to generate synthetic coherence data under the controlled geometries. To be more specific, for a certain acquisition, we simulate an interferometric coherence image using airborne lidar canopy height in the inverted RVoG model and compare the simulation with the observed TDX coherence image acquired at that geometry. In this way, the responsiveness of TDX coherence to canopy height is examined independent of varying imaging geometrical parameters.

In addition, real TDX data are affected by the environmental conditions of forests (Olesk et al. 2015). In this study, we choose two TDX images acquired respectively during leaf-on and leaf-off seasons (see Section 2.2.3). The precipitation and temperature between the two acquisitions were also different. Since we do not incorporate these environmental factors in our inverted RVoG model, correlation between the observed coherence under actual conditions, and the simulated coherence, which is not affected by these conditions, enables us to evaluate the robustness of the sensitivity of TDX data to forest height over the study site under varying environments.

Secondly, given that both TDX acquisitions are responsive from above, we next want to test if these TDX coherences alone (without using external DTM

information) are sufficient to derive canopy heights at acceptable accuracies. When there is no external DTM, the ground phase is unknown and assumptions on certain parameters, in particular a fixed extinction coefficient value (for the entire study area), must be made in the RVoG model to establish a well-posed inversion problem (Hajnsek et al. 2009). The extinction coefficient represents the attenuation rate of microwave inside the forest volume. It is a function of the density of scatters in forest as well as their dielectric constant and also reflects the environmental conditions of forest (Cloude and Papathanassiou 2003; Hajnsek et al. 2009; Lee and Fatoyinbo 2015). Therefore, we would expect that the RVoG model would not perform as well with a fixed extinction, as opposed to a spatially and temporally varying extinction that are used when an external DTM is available to the model. Also, others have found better model performance with leaf-off vs. leaf-on conditions when there is no DTM (Olesk et al. 2015).

Third, the best possible derivations from TDX should occur when a high-resolution, spatially continuous DTM is available, say from airborne lidar, and thus serve as a kind of baseline from which retrievals using coarser DTMs from missions such as GEDI may be evaluated. Such a DTM allows the ground phase component in the TDX interferogram to be properly compensated in the RVoG inversion and allows the model to determine the spatially varying extinction coefficient (Hajnsek et al. 2009; Kugler et al. 2015; Kugler et al. 2014). The derived extinction coefficient image varies acquisition by acquisition depending on actual forest conditions at the

acquisition time, leading to (in an ideal inversion) a robust derivation of canopy height independent of environmental conditions. In contrast to the case with no DTM, better performance is expected during leaf-on conditions (Abdullahi et al. 2016; Kugler et al. 2014).

Related to our third question, if an external DTM is available, we can avoid the RVoG model entirely and estimate canopy height by subtracting the scattering phase height from the DTM elevation for each pixel. This method has been applied widely (Kellndorfer et al. 2004; Kenyi et al. 2009), and is methodologically and conceptually far simpler than the RVoG model. We include this analysis mainly because of its previous use, but also to assess whether improvements seen with the RVoG method justify the increased modeling complexity.

Our final question is the central thrust of our work, which is to explore how DTM data created from spaceborne lidar may help improve canopy retrievals using TDX. Wall-to-wall lidar coverage is rare, especially over tropical forests, and therefore quantifying the impact on accuracy from using elevation data derived from GEDI is of considerable interest. We apply the same procedures as those using the airborne lidar DTM to estimate scattering phase height and RVoG-height but with a coarser DTM. This DTM is created by simulating the orbital coverage patterns we expect from the ISS during a nominal one-year period and subsampling the continuous LVIS data appropriately.

2.3.2 Lidar data processing

We used LVIS waveform data to find the ground elevation and relative height metrics RH25, RH50, RH75 and RH100. The nominal spacing of LVIS footprints was ~20 m both along and cross track, but the density of footprints varied across the landscape. LVIS elevation and RH metric maps were created at 30 m resolution so that there were at least one LVIS footprint in every grid cell of each map and that the cell size was comparable to that of TDX coherence images, which were estimated using a 9 × 9 window, to facilitate subsequent fusion.

GEDI data were simulated from the LVIS full-waveform data. The LVIS data are similar in terms of radiometric characteristics, sensitivity, and footprint size to GEDI footprints. In addition, the expected elevation and height accuracies from GEDI are similar to those achievable from LVIS. GEDI is expected to obtain canopy heights to about 1 m (in leaf-on conditions) and sub-canopy elevations (again leaf-on) to about 70 cm resolution. GEDI geolocation accuracy, however, is expected to be within 7 - 10 m (1 sigma), in contrast to LVIS which has a sub-meter geolocation accuracy.

GEDI consists of 14 parallel beams of data (see Figure *2-2*) with approximately 500 m between tracks, and will acquire data day and night in leaf-on and leaf-off conditions. Three lasers are used to create 14 tracks through a combination of beam dithering and beam splitting. Two coverage lasers are split into 3 beams each, with 1/3rd the power of the strong beam (which is not split). These 7

beams are dithered across track every other line to produce 14 ground tracks of data. Note that other configurations of the mission are possible. For example, an alternate but equivalent baseline mission has a two-year mission length with 5 beams and 10 tracks while a threshold mission may consists of two lasers, 4 beams and 8 tracks. Additionally the spacing along track can be selectively chosen along orbital segments to provide up to n contiguous footprints, so that the dithering is done in groups of between 1 to n footprints. The coverage beams are designed to penetrate in conditions of up to 95% canopy cover and the strong beam up to 98% cover during daylight conditions. At night, the beams penetrate even higher canopy cover because of the absence of solar background noise. Canopy cover at HBEF is almost always less than 95% so there was no need to simulate the strong and coverage beams using LVIS (all were assumed to be strong enough to penetrate the canopy).

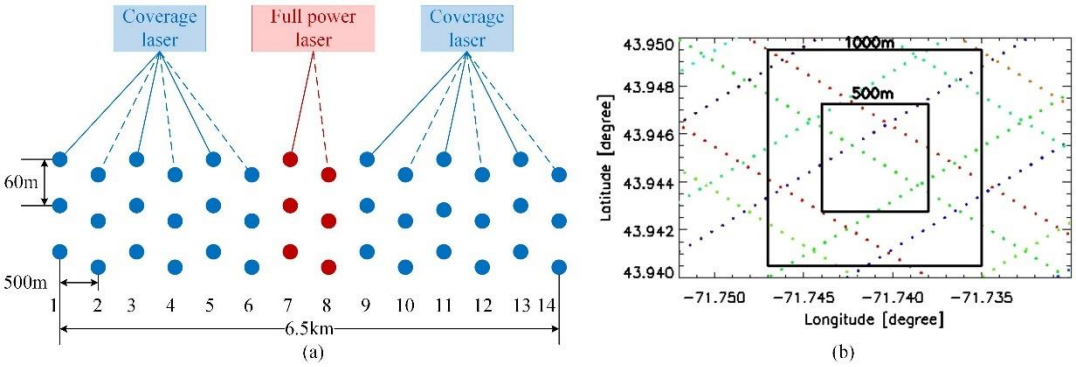


Figure 2-2. (a) Multi-beam lidar coverage of GEDI. GEDI's three lasers, including one full power laser and two coverage lasers, generate 14 ground tracks (Dubayah et al. 2014). (b) Different tracks (colored differently) at the latitude of HBEF in nominal one-year period. The irregular spacing between tracks is because some are obtained

during different passes (that are from a different set of 14 tracks). This spacing occurs because the ISS is in a non-repeating orbit.

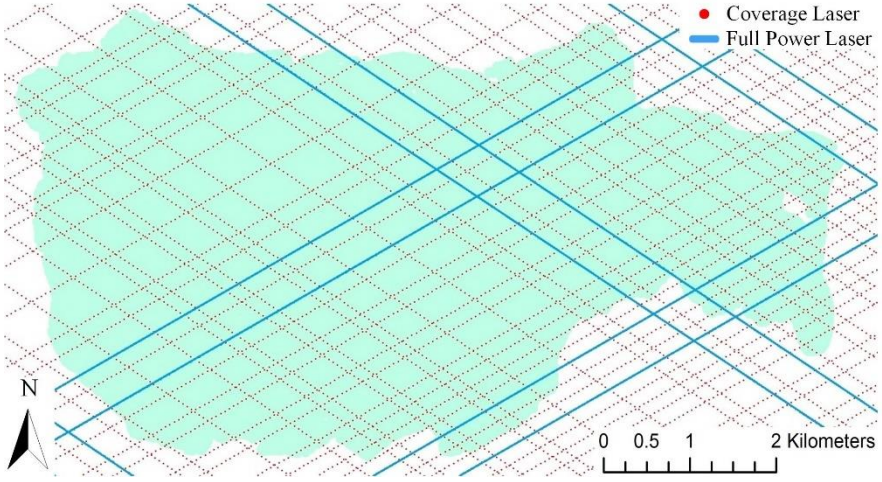


Figure 2-3. Simulated GEDI tracks over HBEF (watershed shown in green) in nominal 1 operation year. Blue solid lines: full power laser tracks; red dashed lines: coverage laser tracks.

ISS orbital simulations were performed to give ground track distributions for the 14 beams over nominal one-year period (Figure 2-2). The GEDI across track ground swath width is approximately 6.5 km (that is the distance from the first beam to the 14th beam) and the inclination of the beams relative to north is determined by the inclination of the ISS orbit. The ISS only operates between ~50° north and south latitude, and is in an approximate 4 day revisit cycle. However, the orbit drifts so that subsequent ISS ground tracks can become clustered or spaced farther apart (that is, it is not in a regularly repeating orbit with a fixed overpass time). The GEDI parallel track pattern can be laid down either in ascending or descending mode leading to the

crossing pattern shown in Figure 2-3. All of these effects were realistically modeled in the full-orbit simulations. For the latitude of HBEF, the most likely number of times the 14 beam pattern would cross HBEF in one year period was about 4 times total in ascending and descending mode, providing the pattern shown in Figure 2-3. This pattern was then used as a template for extraction of LVIS footprints over HBEF.

A grid of simulated GEDI elevations and canopy RH metrics over HBEF were created respectively using this track pattern (Figure 2-4). The scattered GEDI ground elevation data were kriged using the spherical semivariogram model to provide a continuous 30 m-DTM for removing the topographic phase in the TDX interferogram (Oliver and Webster 1990). An RMSE of ~5.3 m was observed for the simulated GEDI DTM against the LVIS DTM at 30 m.

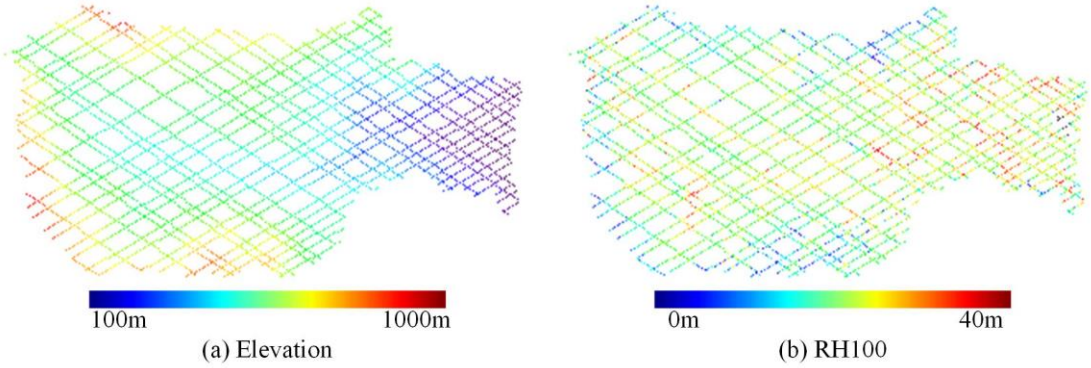


Figure 2-4. Simulated GEDI elevation and canopy top height (RH100) coverages over HBEF.

2.3.3 TDX data processing

Because dual-polarization TDX data are only available over a limited number of sites, our study focuses on forest parameter estimation using single-polarization data, which are available globally from TDX (Kugler et al. 2014).

A. Interferometric coherence and RVoG model

The most commonly used observable from InSAR for forest structure estimation is the complex interferometric coherence, which contains the coherence magnitude, i.e. the correlation coefficient, and the interferogram. If pre-existing topographic data are available, the TDX interferogram is often used to provide volume scattering phase by subtracting the ground phase simulated from the topographic data to derive the so-called scattering phase height (Fatoyinbo and Simard 2013; Lee and Fatoyinbo 2015; Soja and Ulander 2013). The interferometric coherence, once appropriately corrected for non-volumetric decorrelations, i.e. the decorrelation components that are not related to forest vertical structure, leads to the so-called volume coherence (Cloude and Papathanassiou 2003; Lee et al. 2013). The volume coherence can be established as a function of InSAR acquisition geometry, including the incidence angle θ_0 and effective vertical wavenumber κ_z, and parameterized forest structural properties via a suitable forest scattering model, discussed in the following paragraph of this section (Cloude and Papathanassiou 2003; Kugler et al. 2015). κ_z can be expressed as a function of the effective spatial baseline, the slant range distance and the local incidence angle. It is an important

InSAR parameter that translates/scales between forest height and the interferometric coherence, and is the key for InSAR data to provide consistent forest height measurements under varying acquisition geometries (Kugler et al. 2015).

The Random Volume over Ground (RVoG) model is a widely used forest scattering model for InSAR/Pol-InSAR data to estimate forest parameters using complex coherence. The efficacy of the model has been explored under various forest types and conditions (Askne et al. 2013; Garestier et al. 2008; Kugler et al. 2014). In general, the model simplifies forest environment as two layers -- a volume layer and a ground layer, and considers the forest canopy as a layer of thickness h_V containing a volume with randomly oriented particles and scattering amplitude per unit volume m_V and the ground as a scatterer with scattering amplitude m_G. The ratio of m_G and m_V accounting for the attenuation through the volume is called the effective ground-to-volume amplitude ratio *m* (Cloude and Papathanassiou 2003). The vertical volume scattering distribution is often represented as an exponential function of the attenuation rate (also named as the extinction coefficient σ), which is a function of the density of scatterers in the volume and their dielectric constant (Cloude and Papathanassiou 2003). Once appropriate assumption is made about *m*, and an additional assumption about σ or with topography known *a priori*, forest height h_V is derivable from the RVoG model using single-polarized InSAR data (Hajnsek et al. 2009; Kugler et al. 2014; Lee and Fatoyinbo 2015).

B. Model parameterization for single-polarized InSAR inversion

Since we use single polarization data (Figure *2-5*), if the ground topography is already known from an external source, volume decorrelation and volume scattering phase are calculated to derive forest height h_V and extinction coefficient σ assuming a zero ground-to-volume amplitude ratio *m* for HBEF (Kugler et al. 2014; Lee and Fatoyinbo 2015). In this case, the extinction coefficient σ, instead of being set with a fixed value, is automatically determined by the TDX data depending on the attenuation rate of the microwave signal which is affected by the environmental conditions during the acquisition time.

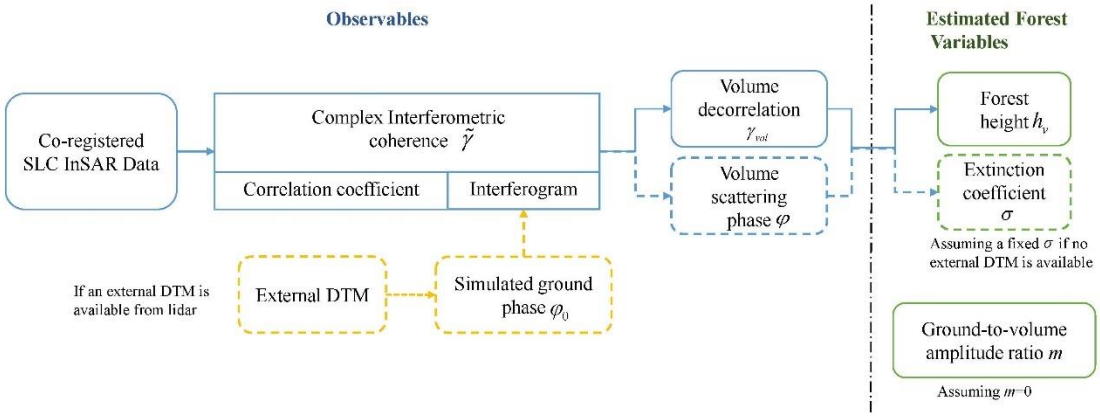


Figure 2-5. RVoG parameterization for inversion of forest height from single-polarized InSAR data.

However, if topography is not known, only volume decorrelation will be used to derive forest height h_V assuming a constant extinction coefficient σ and a zero ground-to-volume amplitude ratio *m* for HBEF (Hajnsek et al. 2009). In some boreal forests where tree density is very low, ground scattering m_G cannot be neglected (so

m cannot be assumed 0), and assumptions must be made about the extinction and/or the ground-to-volume ratio even if topography is known (Caicoya et al. 2016).

2.3.4 Overall processing strategies

As previously discussed, volume coherence can be established as a function of TDX acquisition geometry, forest height h_V and extinction coefficient σ based on the RVoG model. Figure *2-6* shows how the absolute value of volume coherence (volume decorrelation) varies with the increase of forest height assuming a series of extinction coefficient values for TDX data.

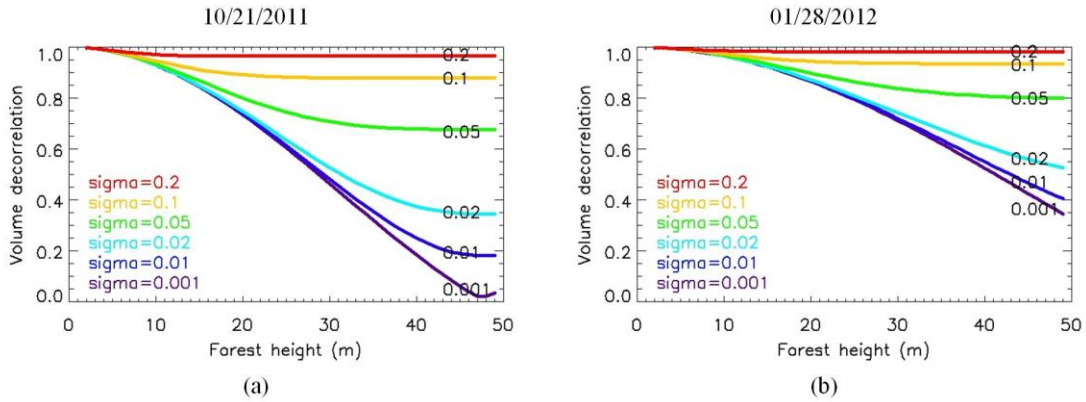


Figure 2-6. Variation of simulated volume decorrelation (γ_{vol}) corresponding to an increase of forest height at extinction coefficient σ values of 0.001, 0.01, 0.02, 0.05, 0.1 and 0.2 dB/m at the incidence angle (θ_0) of ~36° without considering the effect of terrain slopes.

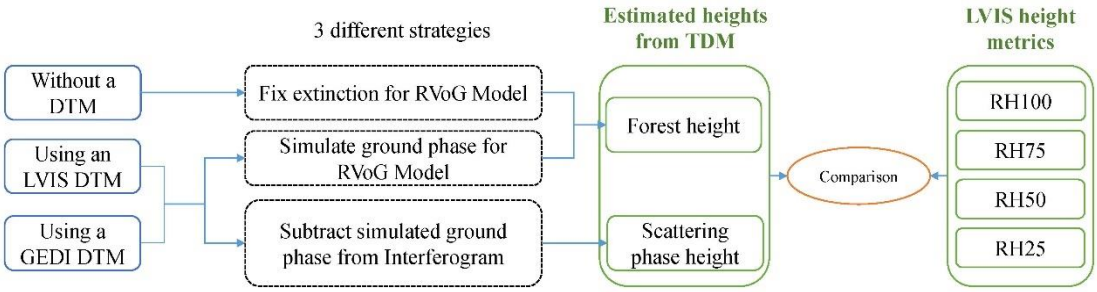


Figure 2-7. A framework for estimating height variables from TDX. Three different strategies were performed in the study: (1) Derive RVoG forest height without using a DTM to correct ground phase; (2) Derive RVoG forest height using ground phase simulated from LVIS/GEDI DTMs; (3) Derive scattering phase height using simulated ground phase from LVIS/GEDI DTMs.

These plots demonstrate that at certain acquisition geometries, the volume decorrelation is expected to decrease with the increase of forest height and saturate at the height of ambiguity given an extinction coefficient value. Therefore, it is necessary to test first if, in a real case, our TDX volume decorrelation is responsive to changes in canopy height by comparison with anticipated decorrelation giving lidar forest height. If strong agreement can be found between the observed and the simulated data, the TDX observation is expected to be closely related to lidar measurements of height enabling fusion between GEDI and TDX.

After testing correlation between the observed and simulated volume decorrelations, two procedures were performed to apply the single-pol inversion to the TDX data for forest height estimation (Figure 2-7). If a lidar DTM was not available for ground phase correction, forest height was retrieved from volume

decorrelation by fixing the extinction coefficient in the RVoG model, whereas if a DTM (from LVIS or simulated GEDI data) was available, forest height was estimated from the complex volume coherence by simulating the ground phase from the DTM and applying it to the model. Using either the LVIS or interpolated GEDI DTM, scattering phase height was also estimated. All the TDX height maps were resampled at 30 m resolution, and were compared against LVIS RH100 (for validation) as well as against other RH metrics (for understanding their relationships) at 90 m resolution (by averaging the 30-m maps using a 3 × 3 window). Auxiliary disturbance data products derived from Landsat were used to exclude forest disturbance that occurred after 2009 (the date of the LVIS acquisition) for more accurate validation of TDX height results.

2.4 Results

2.4.1 Simulating TDX volume decorrelation from lidar forest height

Following the methodology in 2.3.4, we first simulated TDX volume coherence (see Figure 2-8) using lidar forest height (RH100) based on the relationship between volume coherence and volume height in the RVoG model (Cloude and Papathanassiou 2003). To compare variation of the observed volume coherence with that of lidar simulation in more detail, we extracted two co-located transects from the simulated and observed volume coherence maps and plotted their absolute values (volume decorrelation) along each transect (Figure 2-9). Both variations of the

simulated and observed volume decorrelations are mainly driven by changes of forest height, with a lower decorrelation value corresponding to taller trees and vice versa. Their variations are also affected by the acquisition geometry which varies pixel by pixel for each acquisition. For example, a positive terrain slope (terrain surface facing towards the radar line of sight) generally leads to a larger κ_z and thus a lower decorrelation value for the same forest height, and vice versa (Hajnsek et al. 2009; Kugler et al. 2015). However, since our simulation is performed under the same imaging geometry as its associated observation, the variation of imaging geometry is not an issue for a reasonable comparison between the two to determine the height responsiveness of TDX data. The observed volume decorrelation is also slightly influenced by the environment; for example, less leaves and lower forest water content may generate relatively lower values of volume decorrelation under the same imaging geometry, and vice versa.

The calculation of the complex interferometric coherence from TDX acquisitions is affected by the number of looks, i.e. the number of pixels averaged in a certain window assumed as a homogeneous region (Seymour and Cumming 1994). In this study, a window size of 9×9 was used so that the coherence estimation bias (Cloude and Papathanassiou 2003) is assumed negligible and that its resolution approximates the resolution of LVIS simulation. A constant extinction of 0.02 dB/m was used so that the simulation in general has the closest relationship to the observation.

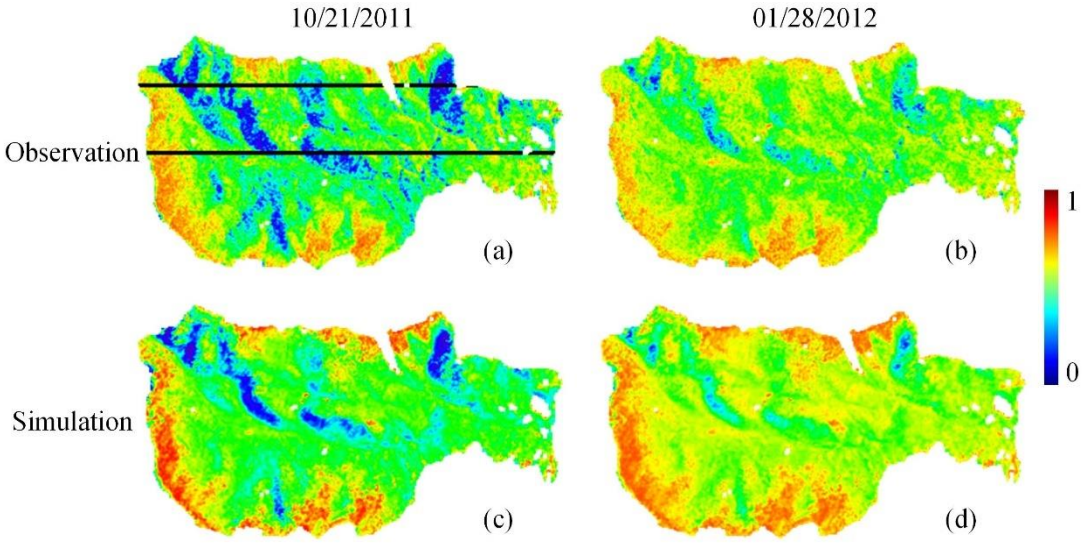


Figure 2-8. Panels (a) and (b) are the observed volume coherence for the two TDX acquisitions; Panels (c) and (d) are the corresponding simulated volume coherence using LVIS canopy top height and assuming a fixed extinction coefficient. Two transect lines shown in (a) are for detailed comparisons between the observation and simulation in Figure 2-9.

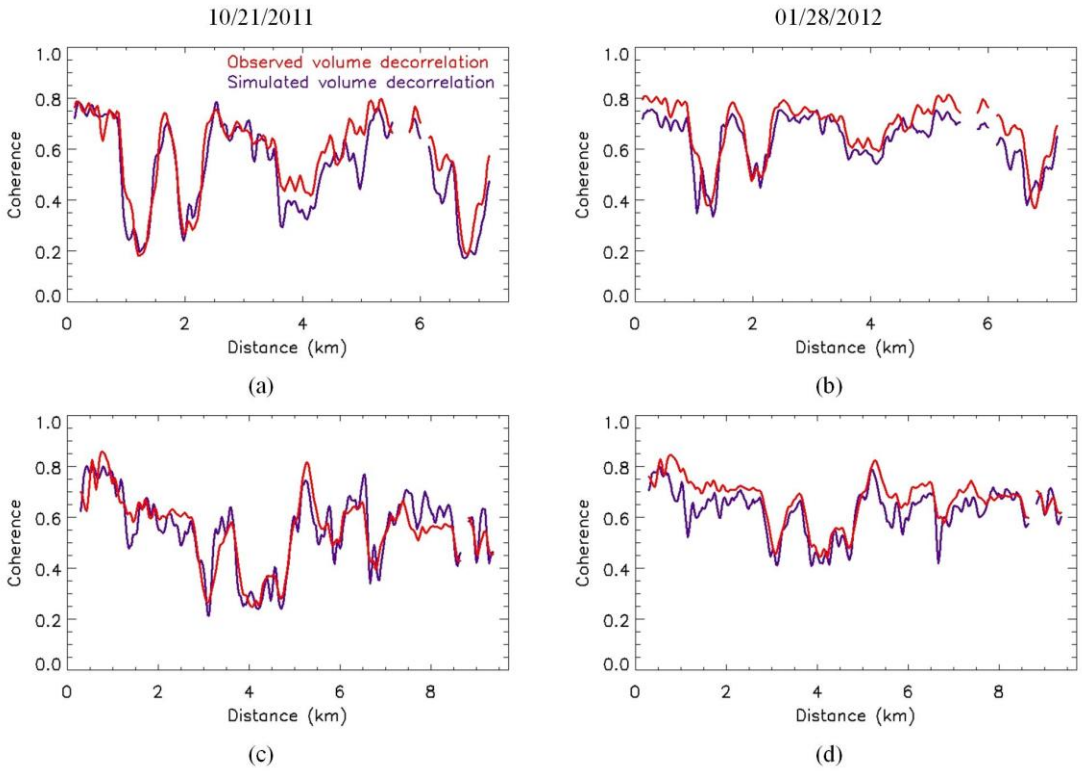


Figure 2-9. Comparisons between the observed and simulated volume decorrelations at 90 m resolution over the two HBEF transect lines shown in Figure 2-8 (a). Figures (a) and (b) correspond to the upper transect line; figures (c) and (d) correspond to the lower transect line.

Both our observed volume decorrelations show similar variability as the associated simulated ones (see Figure 2-8 and Figure 2-9), indicating that TDX should observe similar variations of forest height as those from lidar data. At a spatial resolution of 90 m, the observed and simulated volume decorrelations are highly correlated (Figure 2-10) with an average r^2 of about 0.74 (p-value < 0.0001). The high correlations for both acquisitions suggest that simulated volume decorrelations that

are driven by forest height (in the inverted RVoG model) show the same pattern of volume coherences observed from TDX under the two environmental conditions. We therefore conclude that TDX data is responsive to forest height for the two acquisitions.

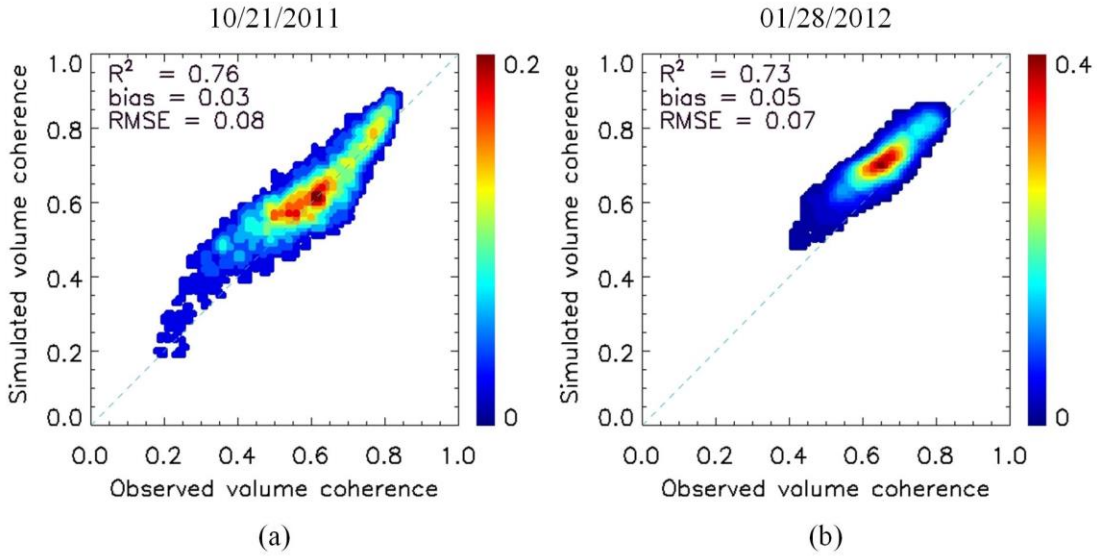


Figure 2-10. Observed vs. simulated volume coherences at 90 m resolution. P-value < 0.0001 for both r^2.

However, observed volume decorrelations are in general biased slightly low for both acquisitions. This may be because the non-volumetric decorrelation components, which are the spectral decorrelation and SNR (signal-to-noise ratio) decorrelation in this study, are undercompensated, leaving a residual decorrelation contribution (Kugler et al. 2015). The differences might also come from inappropriate assumption of mean extinction (0.02 dB/m) and inaccurate terrain slope estimation

when simulating the volume coherence, and real forest height changes between the TDX and LVIS observations.

2.4.2 Canopy structure retrievals without an ancillary DTM to correct ground phase

Using canopy height measurements from GEDI, the closest relationship between simulated and observed coherence was found using a mean extinction of close to 0.02 dB/m, similar as that in 2.4.1. Using this extinction value, forest height was derived based on the relationships between forest height and volume coherence in the RVoG model. The derived forest height from both acquisitions were compared against LVIS RH100 (Figure 2-11) as well as against other RH metrics (see Table 2-2). A low-resolution slope image was required to correct for the effective vertical wavenumber κ_z (Hajnsek et al. 2009; Kugler et al. 2014). We used the TDX DEM to estimate the slope images after applying a low-pass filtering method to the DEM (Kugler et al. 2015) since an external DTM was not available for this case. Coherence values lower than 0.3 were excluded from the inversion (same in the following sections) since low-coherence areas usually have an expectation of inferior height estimation performance (Kugler et al. 2015).

An r^2 of about 0.55 and 0.59 (p-value < 0.0001 for both) was respectively found for the 2011 and 2012 acquisitions at 90 m resolution (Figure 2-11). Model estimates of height were biased high in both acquisitions. This is likely because of low-biased observed volume decorrelations, as shown in 2.4.1. Our finding of marginally better performance from leaf-off TDX acquisition is consistent with

results found by Olesk et al. (2015) and may be the result of greater variability in the real extinction coefficient when leaves are present, relative to the fixed value used.

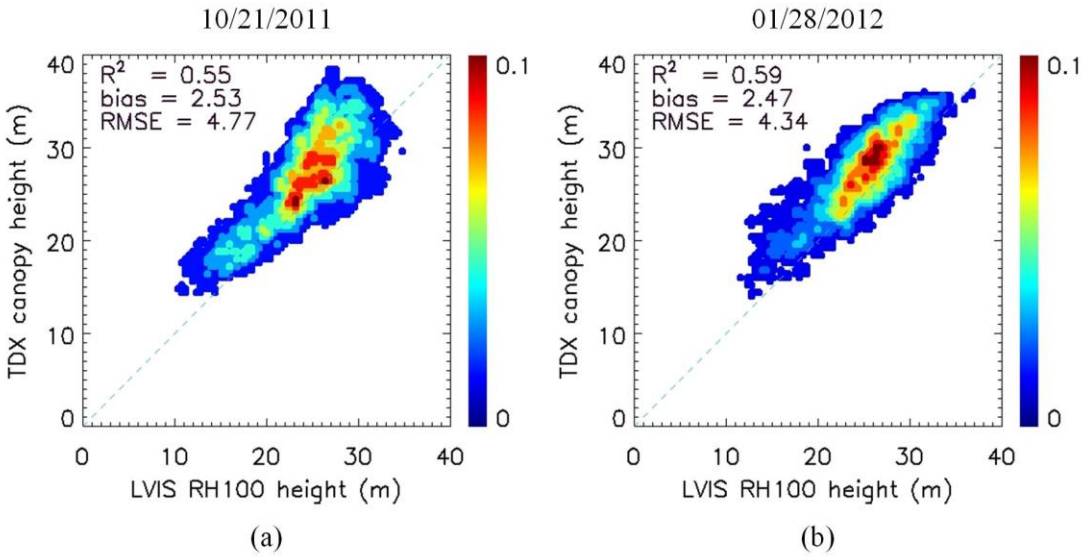


Figure 2-11. Forest height from RVoG model without using an external DTM to correct the ground phase vs. LVIS RH100 at 90 m resolution.

2.4.3 Canopy structure retrievals using LVIS and GEDI DTM

We next present our results for canopy retrievals when two types of external DTMs are available: a wall-to-wall, reference data set from LVIS and one made from GEDI transects of elevation.

A. Scattering phase height

The scattering phase height (SPH) (Figure 2-12) was estimated first by using the LVIS DTM and then using the DTM interpolated from simulated GEDI tracks to correct the ground phase from the TDX interferogram. SPH depends not just on forest

height but also on canopy vertical density and total canopy cover as well as the dielectric properties of the forest layer. Therefore, our SPH results vary spatially over HBEF for each acquisition as well as seasonally between acquisitions as the forest had very different phenology and forest water content during the acquisition dates. In most cases, the SPH (see Figure 2-13) is located quite close to RH50, the height of median energy in the lidar waveform. However, it sometimes migrates towards the canopy top and is located closer to RH75, or towards the ground and is located closer to RH25 probably due to the different effects of canopy cover, forest density and water content on the penetration capability X-band radar signals.

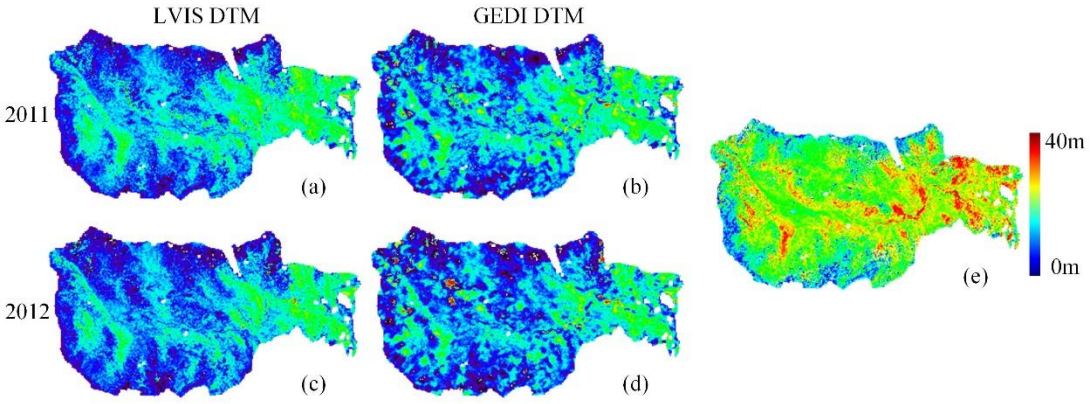


Figure 2-12. 2011 and 2012 scattering phase height derived by using different DTMs to correct the topographic phase. (e) is the reference canopy top height from LVIS RH100.

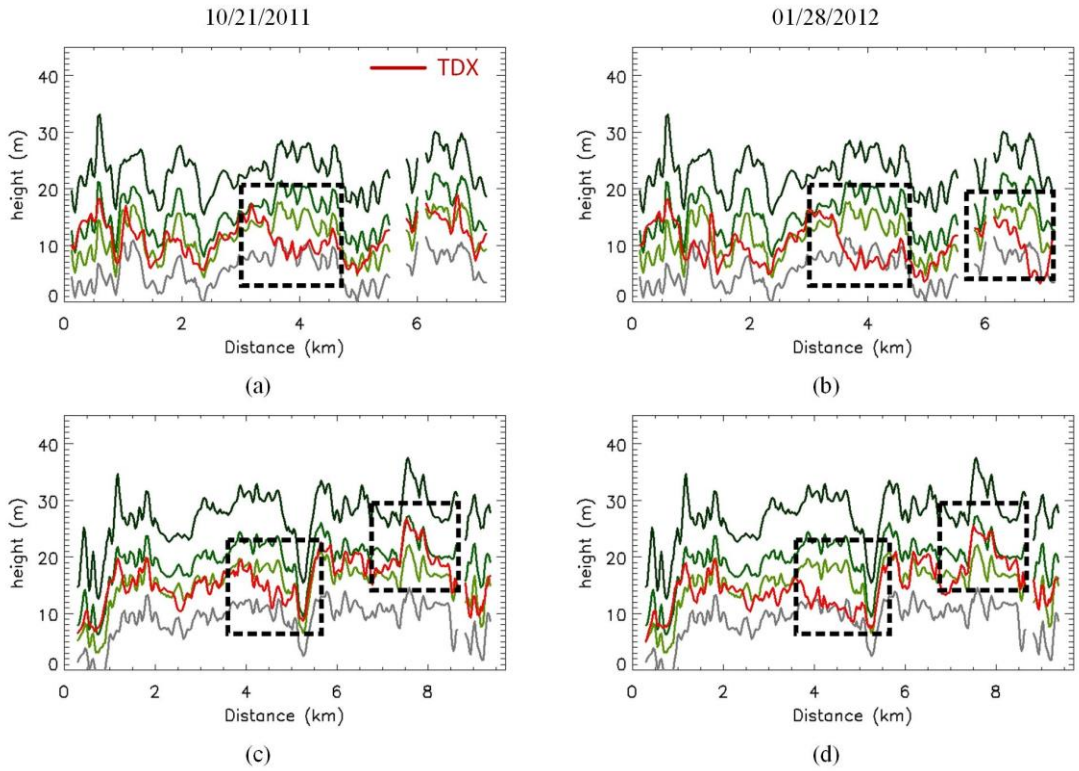


Figure 2-13. Scattering phase height of TDX compared against RH100, RH75, RH50 and RH25 from LVIS at 90 m resolution over the two transect lines in Figure 2-8 (a). Figures (a) and (b) correspond to the upper transect line; figures (c) and (d) correspond to the lower transect line. Migration of scattering phase height away from RH50 to RH75 or RH25 can be observed in the black dotted boxes.

We observed an r^2 of 0.54-0.71 (p-value < 0.0001 for both) and RMSE of 10.85-12.19 m at 90 m resolution for SPH and LVIS RH100 using the LVIS DTM (Figure 2-14 and Table 2-2). The 10/21/2011 acquisition in general has a higher scattering phase center and higher correlation to the reference canopy height than the 01/28/2012 acquisition. This is mainly due to the higher canopy cover and density in

HBEF during the leaf-on season. The possibly higher fluid-state forest water content for the 2011 acquisition may contribute towards its higher scattering phase center. Recent studies have demonstrated similar patterns of seasonal changes in scattering phase height over temperate broadleaf forests, with the largest influence contributed by forest phenology variation (Abdullahi et al. 2016; Kugler et al. 2014; Praks et al. 2012).

When using the GEDI DTM, an r^2 of 0.38 and RMSE of 11.33 m was found for the 2011 acquisition and an r^2 of 0.28 and RMSE of 12.59 m was found for the 2012 acquisition (p-value < 0.0001 for both r^2) (see Figure 2-14 and Table 2-2). The correlation of scattering phase height to canopy height decreased as DTM accuracy decreased and as leaf-off condition prevailed. These results suggest that SPH may be related to canopy height in leaf-on condition and for where an accurate DTM is available. However, the relationships are empirical and must be calibrated using actual canopy height data (so that a regression/calibration equation may be created).

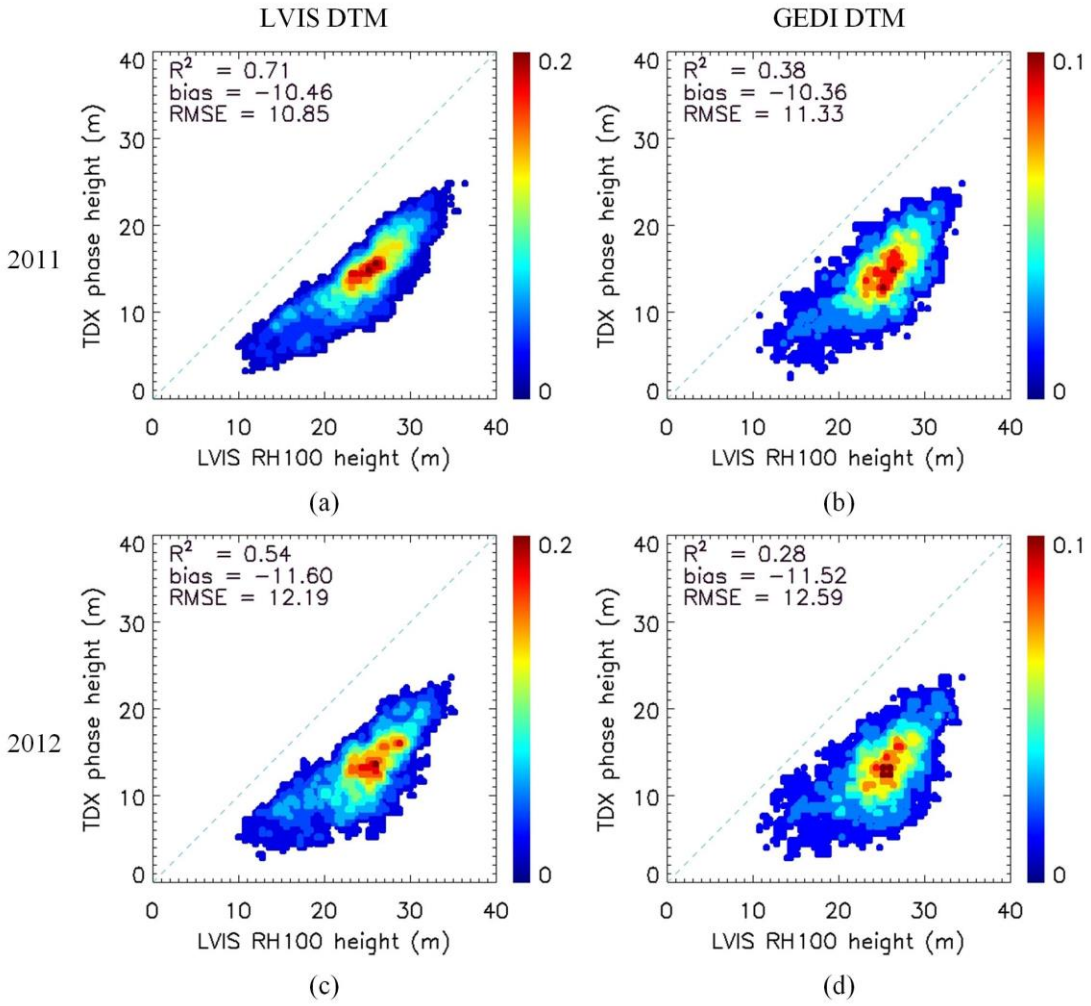


Figure 2-14. 2011 and 2012 scattering phase height using different DTMs vs. LVIS RH100 at 90 m resolution.

B. Forest height from RVoG model

Single-pol RVoG inversion using lidar DTM was applied in the study to derive forest height in 2011 and 2012 (see Figure 2-15).

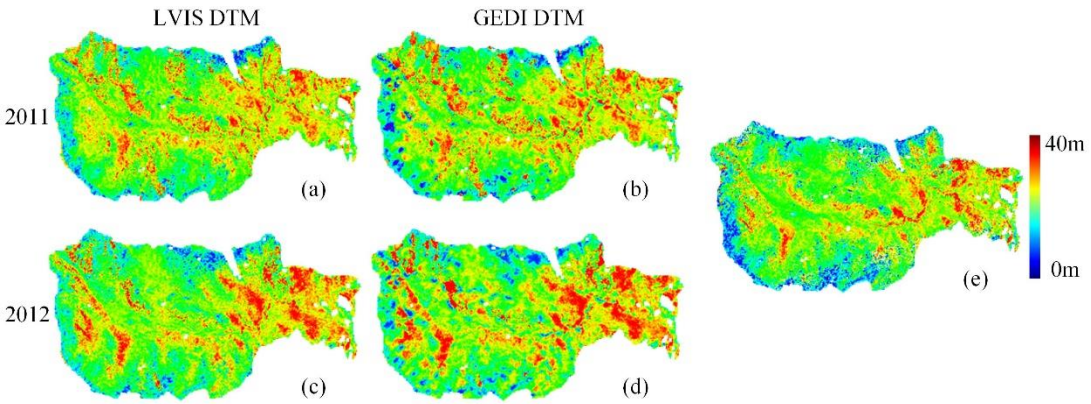


Figure 2-15. Forest height derived from the RVoG model respectively using LVIS and GEDI DTMs. Reference canopy top height from LVIS RH100 is given in (e).

We found an r^2 of 0.71 (p-value < 0.0001) and RMSE of 3.61 m from the 2011 acquisition and an r^2 of 0.66 (p-value < 0.0001) and RMSE of 3.53 m from the 2012 acquisition between TDX-derived forest height and LVIS RH100 using the LVIS DTM to correct the topographic phase. Using the coarser GEDI DTM, an r^2 of 0.48 (p-value < 0.0001) and RMSE of 4.59 m was observed from the 2011 acquisition, and an r^2 of 0.39 (p-value < 0.0001) and RMSE of 5.36 m was found from the 2012 acquisition (Figure 2-16 and Table 2-2). The performance of inversion was degraded as DTM accuracy decreased and as the leaf-off condition prevailed, but less so than observed on the scattering phase height. We found better results using leaf-on data, consistent with other studies using the RVoG model (Abdullahi et al. 2016; Kugler et al. 2014). This improvement may be the result of leaf-on data better matching assumptions of exponential distribution of scatters. Another reason cited in

the above studies is leaf-on better suits the assumption of no ground scattering in the RVoG model.

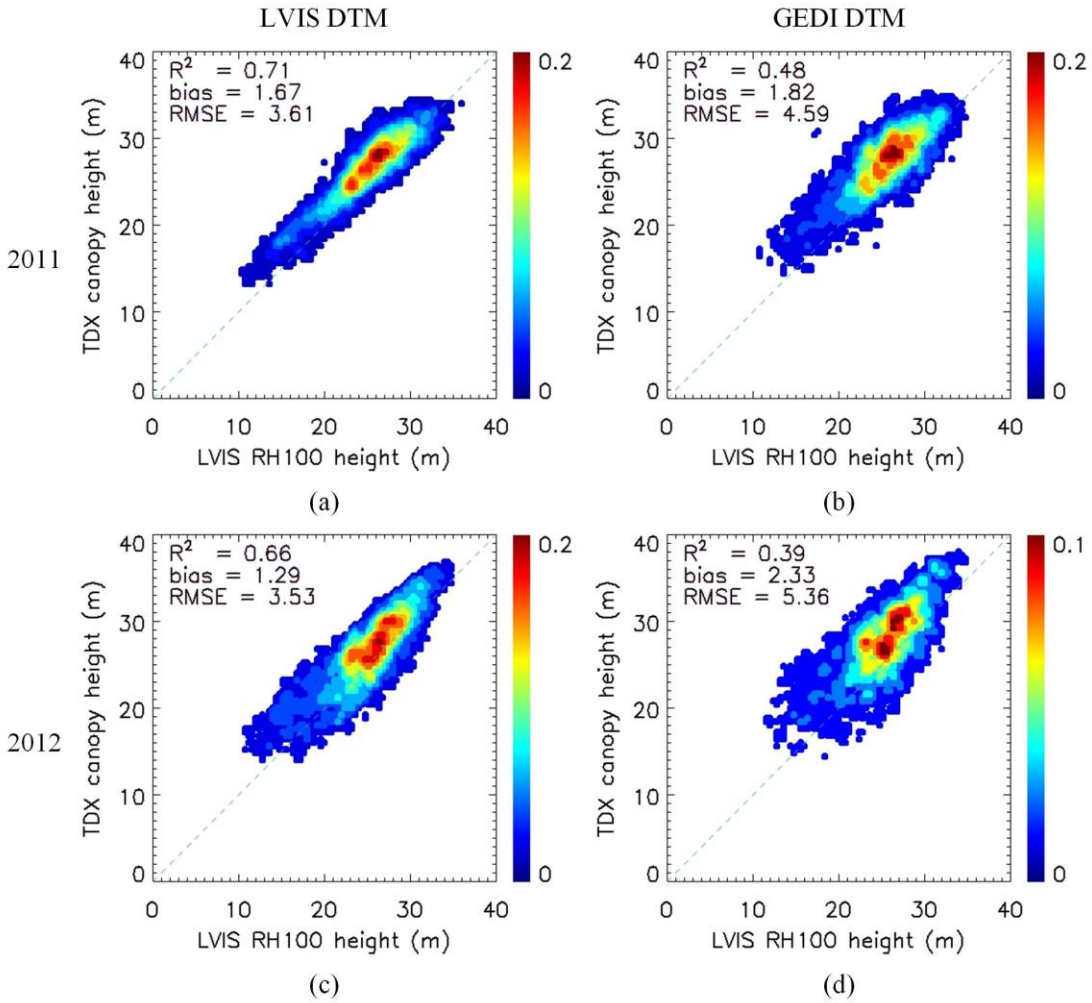


Figure 2-16. Forest height derived from the RVoG model using different DTMs to estimate the ground phase vs. LVIS RH100 at 90 m resolution.

Table 2-2. Relationships of TDX scattering phase height and RVoG-derived forest height with LVIS relative heights.

Forest height from TDX vs. LVIS quartile heights (90m)		**RH100**		**RH75**		**RH50**		**RH25**	
		2011	**2012**	**2011**	**2012**	**2011**	**2012**	**2011**	**2012**
RVoG Forest Height by fixing the extinction	r^2	0.55	0.59	0.61	0.56	0.60	0.53	0.54	0.47
	bias	2.53	2.47	9.96	9.90	13.93	13.88	19.79	19.73
	RMSE	4.77	4.34	10.62	10.54	14.41	14.35	20.17	20.07
Scattering Phase Height using LVIS DTM	r^2	0.71	0.54	0.71	0.51	0.70	0.49	0.68	0.47
	bias	-10.46	-11.60	-3.02	-4.16	0.95	-0.19	6.80	5.66
	RMSE	10.85	12.19	4.12	5.60	2.91	3.73	7.30	6.65
Scattering Phase Height using GEDI DTM	r^2	0.38	0.28	0.39	0.27	0.38	0.26	0.37	0.26
	bias	-10.36	-11.52	-2.92	-4.08	1.05	-0.11	6.90	5.75
	RMSE	11.33	12.59	5.3	6.45	4.47	4.92	8.03	7.37
RVoG Forest Height using LVIS DTM	r^2	0.71	0.66	0.73	0.62	0.72	0.59	0.67	0.55
	bias	1.67	1.29	9.10	8.72	13.07	12.70	18.93	18.55
	RMSE	3.61	3.53	9.58	9.37	13.41	13.17	19.17	18.88
RVoG Forest Height using GEDI DTM	r^2	0.48	0.39	0.52	0.38	0.51	0.37	0.48	0.35
	bias	1.82	2.33	9.25	9.77	13.22	13.74	19.08	19.59
	RMSE	4.59	5.36	10.06	10.87	13.78	14.54	19.46	20.14

Note that even though the other RH metrics (25, 50 and 75) are highly correlated with the outputs from RVoG, the RMSE values become increasingly larger and are comprised almost entirely of bias. This is consistent because the RH metrics are usually correlated with each other (that is they are almost never independent). So the RVoG results for top height (RH100) will be correlated with the other metrics, but

these metrics must be biased given that the RH metrics occur at increasingly large distances below the canopy. This last point explains why the SPH results for RH100 using the full LVIS DTM had an equally high r^2 to the results from RVoG but a much larger RMSE, again because of bias. SPH is most closely approximating RH50 and will therefore be biased low relative to the canopy top.

There is a seemingly inconsistent result of achieving better results (a larger r^2 and smaller RMSE) using RVoG to predict height without an ancillary DTM to correct the ground phase vs. using a GEDI-derived DTM. This occurs because of DTM inaccuracies in the spaces between points where GEDI heights are available during the gridding process. But even though the RMSE may be lower, the bias is improved when a GEDI DTM exists, discussed further below.

2.5 Discussion

The underlying assumption of our research was that the inclusion of external topographic data should improve estimates of canopy height from TDX. Such data are a requirement if an approach using scattering phase height is used, as there is no way to translate the height above the ellipsoid (that is an elevation) into a canopy height otherwise. However, given that the RVoG model does not explicitly require external topographic data, but can instead uses an assumption about the extinction coefficient, there is considerable interest to understand the accuracies achievable as we move from no external topographic data to a high-resolution external DTM. In between the two are coarser DTM data that will be available globally from GEDI.

Our first objective in this process was to examine the responsiveness of TDX coherence to variations in canopy height using two acquisitions from different environmental conditions. We accomplished this by simulating coherence using lidar canopy heights in the inverted RVoG model and found strong agreement with actual coherences for both acquisitions, indicating the robustness of sensitivity of TDX coherence to canopy height variation under the two distinct forest conditions. We then explored the use of the RVoG model with and without ancillary DTM data to correct the ground phase. Using TDX acquisitions our best results were found using the full-resolution DTM from LVIS. While it is true that such data are not available globally, airborne lidar data are continually being collected and it is likely that coverage will expand rapidly. For example in the United States it is estimated that over 50% of the land surface has lidar coverage (NOAA 2015). In developing countries, the focus on carbon markets and treaty compliances, such as REDD (Agrawal et al. 2011) will also provide more coverage. However, complete wall-to-wall data is likely decades away while data from space-based sensors such as GEDI should appear within a few years, which while globally extensive, will be lower resolution given likely sampling patterns.

Summarizing our results using the RVoG model, we generated canopy heights with the best accuracy and highest correlation to the reference heights for both acquisitions using the full LVIS DTM, with (in the best case of the two) an r^2 of around 0.7 and 3.5 m RMSE with a bias of about 1.3 m at 90 m resolution. Similar

correlations between lidar canopy height and TDX-derived heights were reported using a high-resolution lidar DTM over a temperate test site, with an r^2 of 0.80-0.86 at stand level over homogeneous areas with a mean size of 3-5 ha (Kugler et al. 2014). Using the simulated GEDI DTM resulted in an r^2 of around 0.5 and 4.6 m RMSE with a bias of about 1.8 m at 90 m resolution during leaf-on condition. Thus, the improvement in accuracy was not large when the full DTM was available, but was significant, reducing RMSE and bias relative to the GEDI DTM. In both cases, however, a positive bias remained; that is, TDX was always higher than the LVIS canopy height. The bias can possibly be corrected or reduced once transects of GEDI relative height metrics are explored for the fusion.

Somewhat surprisingly, the RVoG model performed almost as well when no external DTM data were available for ground phase correction, showing strong correlations with LVIS-derived heights and an RMSE of about 4.3 m and bias of about 2.5 m. However, an important caveat is that the inversion accuracy relies on choosing the fixed extinction value (σ) within an appropriate range (0 - 0.04 dB/m in this study) which in general is representative for the overall forest condition. We chose ours based on our simulations from GEDI canopy height. It may not be unreasonable that in future applications, GEDI or limited airborne lidar data could be used from similar forests to help determine an appropriate extinction when study areas do not have GEDI measurements due to clouds or other issues.

Are the improvements seen in going from no DTM data to a full-resolution DTM significant enough to warrant the development of a coherent strategy based on using TDX and a mission such as GEDI? To answer this question requires knowledge of the scientific uses and requirements for canopy height data. One of the main application of height data is for biomass estimation (Goetz and Dubayah 2011; Hall et al. 2011; Sun et al. 2011). Studies done as part of the DESDynI mission (Hall et al. 2011) as well as for carbon modeling using ecosystem models (Hurtt et al. 2010) have argued that a height accuracy of about 1 – 2 m at spatial scales of between 100 m – 1000 m (with finer resolution more desirable) is required to keep biomass errors sufficiently small. The TDX data by themselves did not meet this requirement at HBEF at 90 m resolution (the RMSE was about 4.3 m). Averaging to a coarser resolution may reduce this RMSE but cannot reduce the bias, which is too large (~2.5 m). The bias was reduced to < 2 m by using the GEDI DTM. Averaging to a coarser resolution brings the RMSE estimates even lower. For example, at 500 m, the best GEDI RMSE estimate was ~ 2.7 m with a bias of ~ 1.7 m while the TDX data by themselves was ~ 3.3 m with a bias of ~ 2.5 m. In future studies, introducing forest height measurements from GEDI data to reduce the coherence estimation bias and to provide a low-resolution spatial map of the mean extinction may help achieve required accuracies.

There is also the question of how generalizable our results are. HBEF has a diverse topography with steep slopes and mix of broad leaf and conifer forests.

Results are likely to be improved for less topographically diverse regions. However, HBEF does not have particularly high canopy cover, with a mean cover of about 80%, and the X-band SAR signal in general had a good penetration into the HBEF canopy (as illustrated by its scattering phase height negative biases). Though lower penetration will occur in tropical forests, similar relationships can be expected because unlike dual-pol inversion methods, DTM-aided single-pol inversion does not rely on penetration of the signal through the canopy to the ground to derive forest height (as it does not require the estimation of ground phase from the TDX data). A recent study has reported that in tropical forests, correlations as high as 0.97 – 0.98 with RMSE of 1.8 m – 3.2 m can be established using the single-pol inversion at stand level over homogeneous forest areas with a mean size of ~ 3-5 ha, though much lower correlations of 0.54 – 0.69 with RMSE of 5.7 – 6.3 m were found using the dual-pol inversion method (Kugler et al. 2014).

We did not model any of the expected errors from the GEDI data themselves. GEDI data are similar to LVIS waveform data but will differ mainly in terms of signal to noise ratio (which will impact the ability to detect the ground), with GEDI having lower SNR. GEDI is designed to reliably detect the ground through canopy covers between 95% and 98%, for the coverage and strong beams, respectively, and to provide canopy height to better than 1 m accuracy. Secondly, GEDI geolocation errors are larger than LVIS, being about 7 – 10 m, while LVIS data are geolocated to tens of centimeters. This may increase the error of the underlying gridded DTM.

Lastly, there are issues of spatial coverage and clouds. The ISS is in a non-repeating orbit and the exact placement of tracks cannot be predicted *a priori*. Furthermore, some areas may get less tracks than anticipated due to clouds, resulting in different numbers and distribution of GEDI footprints. All of these factors may decrease the accuracy and grid resolution of the final DTM made from GEDI.

Our results improved as the lidar shot density increased. In areas with less rugged terrain, the coverage needed to reduce bias may be reduced. For flat areas, e.g. mangrove forests (Lee and Fatoyinbo 2015), it may be that even a severely reduced density of observations from space, sparse but spatially extensive, may provide enough data to improve RVoG model predictions.

Lastly, as noted above, we did not make full use of the actual canopy height observations from GEDI in our estimation process using the RVoG model. GEDI is expected to provide close to 15 billion observations of canopy height per year. Future work should consider how to explicitly fuse these high quality canopy observations in the estimation process.

2.6 Conclusion

This study has investigated forest structure mapping from TDX InSAR data over a mountainous temperate forest - Hubbard Brook Experimental Forest (HBEF). Our results have demonstrated that an approach which uses Tandem-X observations with ancillary topographic data produced from a spaceborne lidar such as GEDI can potentially achieve higher accuracy or more complete spatial coverage than using

either alone. The ability of the RVoG model to recover canopy structure was impressive, even without ancillary topographic data for ground phase correction. However, reducing biases and other errors to a sufficient level for applications such as biomass modeling should be facilitated through the use of GEDI data.

As currently implemented, GEDI will produce gridded data sets of canopy height and biomass at 1 km resolution. However, GEDI will also produce experimental products at much higher resolutions using both topographic and canopy observations in fusion with Tandem-X. The contours of such fusion algorithms are being developed as part of a collaboration between the German Aerospace Center (DLR) and the GEDI mission. Work such as presented here heralds a trend towards a new era of active remote sensing, where diverse missions such as Tandem-X, GEDI, NASA's ICESat-2 and NISAR, ESA BIOMASS and others are used in concert to derive key aspects of ecosystem structure.

2.7 Acknowledgements

This work was supported by NASA's Earth and Space Science Fellowship Program (Grant # 5268930) and NASA's Global Ecosystem Dynamics Investigation mission. We thank the Microwaves and Radar Institute of the German Aerospace Center (DLR) for providing the TerraSAR-X/TanDEM-X data and for the assistance in processing these data. We also thank Dr. Kostas Papathanasiou for useful comments on the manuscript and research direction. Thanks to Dr. Chengquan Huang for providing the forest disturbance data product.

Chapter 3: Improved Forest Height Estimation by Fusion of Simulated GEDI Lidar Data and TanDEM-X InSAR Data

3.1 Introduction

Forest Above-Ground Biomass (AGB) has been identified as a key parameter for assessing the role of forests in the global carbon cycle and for analyzing ecosystem productivity. However, current quantification of forest AGB worldwide and associated biomass changes remain uncertain (CEOS 2014; Pan et al. 2011). Forest inventory methods have been widely used to estimate AGB at field scales, either through destructive sampling or by measurement of various biomass-related forest structural properties and a subsequent employment of allometric equations. However, these methods are often labor-intensive and time-consuming, and do not yield continuous AGB maps over the landscape (Clark and Kellner 2012; Duncanson et al. 2015a; Duncanson et al. 2015b; Keller et al. 2001). Therefore, there is an interest to capitalize on field-scale biomass and remotely measured forest parameters (particularly height) to provide more cost-effective AGB mappings at large areas (Goetz and Dubayah 2011; Huang et al. 2012).

Lidar and Interferometric Synthetic Aperture Radar (InSAR) remote sensing techniques are playing increasingly important roles in forest structure estimation (Anderson et al. 2006; Dubayah and Drake 2000; Dubayah et al. 2010; Garestier et al. 2008; Hajnsek et al. 2009; Papathanassiou and Cloude 2001). However, taken

individually, each technique has particular limitations and difficulties to deliver large-area forest structure dataset for reducing the uncertainty of forest AGB quantification (Goetz and Dubayah 2011; Hall et al. 2011). Lidar observations of forest structure are mainly restricted to local regions where airborne lidar campaigns were conducted (Anderson et al. 2006; Drake et al. 2002; Dubayah et al. 2010). The sole spaceborne Earth observation lidar instrument, Geoscience Laser Altimeter System (GLAS) onboard ICESat, has enabled global coverage and produced consistent forest structure measurements. However, it had low sampling density over mid-latitude and tropical forests (Dolan et al. 2011; Dolan et al. 2009; Fatoyinbo and Simard 2013).

InSAR has been widely used to generate wall-to-wall forest structure maps (Askne et al. 2013; Schlund et al. 2015; Soja et al. 2014; Solberg et al. 2013). However, accuracies of InSAR products are often reduced by temporal decorrelation effect which occurs when the SAR images forming the interferometric coherence are acquired at different times. This temporal decorrelation limits the accuracy of repeat-pass interformetry (Lavalle and Hensley 2015; Lee et al. 2013; Papathanassiou and Cloude 2003). To address this problem, the German Aerospace Center (DLR) launched the first dual-satellite (bistatic) SAR spaceborne mission – TanDEM-X (TDX). There is no temporal decorrelation using TDX because the data from each satellite are obtained at the same time, allowing more accurate estimation of forest height (Caicoya et al. 2016; Kugler et al. 2014; Treuhaft et al. 2015). A simple forest scattering model, the Random Volume over Ground (RVoG) model, has been widely

used to produce forest height maps from TDX coherence under a variety of terrain conditions and forest types. However, because TDX images are generally acquired at a single polarization, determination of forest height using the RVoG model must assume known canopy extinction and topographic parameters (Hajnsek et al. 2009; Kugler et al. 2014; Qi and Dubayah 2016).

The aforementioned issues of lidar and InSAR can be potentially addressed by combining their complementary observations, where lidar data may be used to constrain the forest scattering model and to validate InSAR height inversion while InSAR images are exploited to extend lidar observations (Bergen et al. 2009; Goetz and Dubayah 2011; Hall et al. 2011; Qi and Dubayah 2016; Sun et al. 2011). For example, previous studies have used airborne lidar elevation data to provide the needed external DTM to estimate forest height from TDX single-polarization (single-pol) coherence (Cloude et al. 2013; Kugler et al. 2014; Schlund et al. 2015; Soja and Ulander 2013; Solberg et al. 2013). Accurate airborne lidar observations of forest vertical structure have also been used to enhance parameterization of the forest scattering models for improved forest height estimation (Brolly et al. 2016). The elevation data derived from the first spaceborne InSAR mission - Shuttle Radar Topography Mission (SRTM) - has been calibrated and validated with local height measurements from GLAS to produce continuous canopy height and AGB maps over Mangrove forests (Fatoyinbo and Simard 2013). These studies have demonstrated the

potential advantages in combining lidar and InSAR to map forest structure at better accuracy and coverage.

An unprecedented opportunity of global forest structure and biomass mapping from lidar/InSAR fusion has emerged with the upcoming launch of the Global Ecosystem Dynamics Investigation (GEDI) mission (Qi and Dubayah 2016). GEDI is a full-waveform lidar system to be deployed on the International Space Station (ISS) by NASA in 2018 (Stysley et al. 2015). During its nominal two-year mission, GEDI will provide about 15 billion accurate ground elevation and forest vertical structure measurements at a footprint size of ~25 m in diameter. Aided by these GEDI observations, TDX data can potentially provide wall-to-wall forest height maps, which in turn can be used to extend GEDI observations for forest structure and biomass estimation at finer resolution, accuracy and coverage (Qi and Dubayah 2016). However, the effects of using different elements of forest vertical structure observed by GEDI on TDX height inversion are still unclear and largely unexplored. Also, the performance of GEDI/TDX fusion needs to be investigated for different forest structural types and environmental conditions.

The goal of this study is to develop lidar/InSAR fusion methods for improved TDX height estimates using simulated GEDI observations. GEDI data are simulated using airborne laser scanning (ALS) data and combined with single-pol TDX InSAR data. Our test sites include three contrasting forest types: Hubbard Brook Experimental Forest (HBEF), a temperate mixed broadleaf deciduous and conifer

forest; Teakettle Experimental Forest (TEF), a mountainous conifer forest; and La Selva Biological Station (LSBS), a tropical broadleaf rainforest. Specifically we perform three sets of analyses to explore the impact on height derivations using fusion. First, we establish a baseline accuracy for our study sites by using only TDX data and simple assumptions of RVoG parameters. Next, we utilize an external DTM derived from simulated GEDI-derived DTM in the RVoG model. Lastly, we investigate the impact of using both a simulated GEDI DTM and GEDI-derived canopy heights within the RVoG model. In each case, we also examine the impact of environmental conditions on the fusion results by comparison of GEDI tracks under cloud-free vs. 50%-cloud (which is the mean of global cloud coverage) conditions, and comparison of leaf-on vs. leaf-off TDX acquisitions. Results from this study will help to broaden the application scale of both lidar and InSAR remote sensing on forest height and biomass mapping.

3.2 Test Sites and Data

3.2.1 Test sites

Three sites were chosen, representing a range of forest characteristics (Figure *3-1*). Hubbard Brook Experimental Forest (HBEF) (43°56′12″N, 71°45′01″W) is a closed-canopy broadleaf-dominated forest located in the White Mountain National Forest, New Hampshire, USA, and is typical of temperate forest conditions. Covering an area of 3,100 ha the topography is rugged, with steep slopes occuring within a

bowl-shaped watershed. Elevations range from about 150 m to 1000 m. It is a managed forest consisting of mostly deciduous northern hardwoods and a small percentage (10-20%) of spruce-fir. Measured forest heights mainly range from ~2 to ~42 m, with a mean of ~24 m and a standard deviation of ~5 m (Campbell and Bailey 2014; Qi and Dubayah 2016; Schwarz et al. 2001; Siccama et al. 2007; Swatantran et al. 2012; Thomas et al. 2008; Whitehurst et al. 2013). HBEF has a moderate amount of above ground biomass, with a mean of 216 Mg/ha in 2001 (Siccama et al. 2007).

Teakettle Experimental Forest (TEF) (36°57′60″N, 119°01′0″W) is a conifer-dominated forest located along the western slopes of Sierra Nevada Mountain Range, USA. The study site is a mountainous region covering an area of around 1,300 ha, with elevations ranging from about 1,800 m to 2,500 m. It is an old-growth forest with mature and complex structure. Tree heights mainly range from ~3 m to ~68 m, with a mean height of ~39 m and a standard deviation of ~11 m. Major tree types include White fir (*Abies concolor*), Ponderosa pine (*Pinus ponderosa*), Red fir (*Abies magnifica*) and California black oak (*Quercus kellogi*) (Pierce et al. 2002). The averaged aboveground biomass is about 200 Mg/ha with individual tree values up to 20 Mg per tree (Duncanson et al. 2015a; Smith et al. 2005; Swatantran et al. 2011).

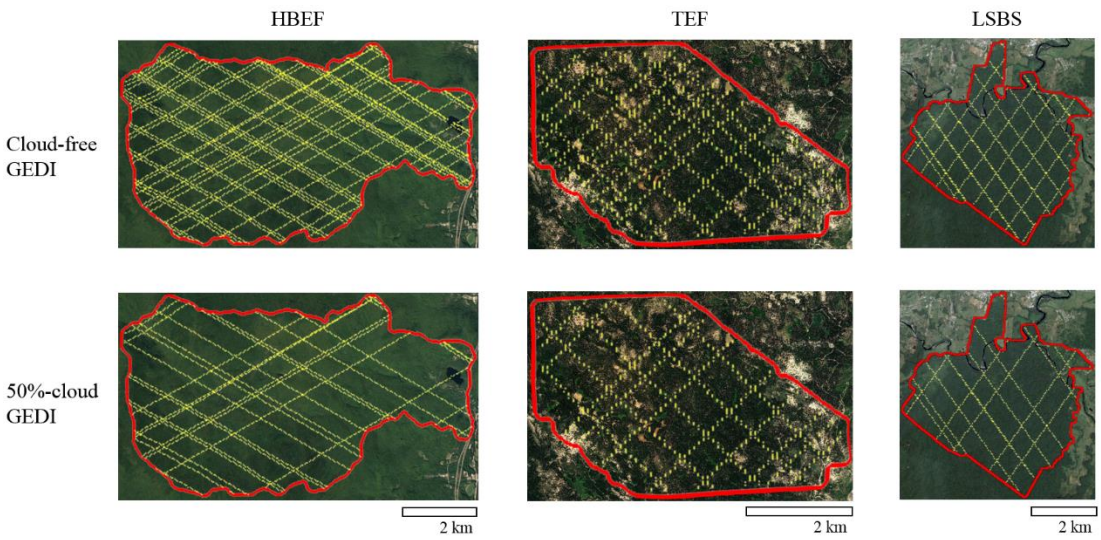


Figure 3-1. Simulated two-year GEDI tracks over HBEF, TEF and LSBS test sites respectively based on cloud free and 50% cloud cover conditions.

La Selva Biological Station (LSBS) (10°25′44″N, 84°00′29″W) is a low-land (elevation <150 m) tropical rain forest in northeastern Costa Rica. The site is a protected region covering about 1,600 ha, and contains a mixture of old-growth, secondary and selectively logged forests as well as agroforestry plantations, developed areas, and abandoned pastures. Tree height ranges from ~3 to ~59 m, with a mean of ~28 m and a standard deviation of ~11 m. Estimate of aboveground biomass spans from 0 to 279 Mg/ha, and averaged biomass of old-growth forest, which is the major components of total LSBS biomass, is around 169 Mg/ha (Clark et al. 2011). Detailed site characteristics can be found in (Clark et al. 2008; Dubayah et al. 2010; Tang et al. 2014; Tang et al. 2012).

3.2.2 Datasets

A. Airborne Lidar data and simulation of GEDI

Airborne Lidar Scanning (ALS) data were collected in 09/2009, 09/2008 and 09-10/2009 over HBEF, TEF and LSBS, respectively. The data were processed to simulate GEDI full-waveform data following the method of Blair and Hofton (1999) with measurement noise added following Hancock et al. (2011). In the simulation process, discrete ALS returns were weighted by the laser footprint shape (Gaussian with σ (standard deviation) = 5.5m), binned into vertical histograms at GEDI's resolution (15 cm) and convolved with the laser pulse shape (Gaussian with FWHM (full width at half maximum) = 15 ns). White Gaussian noise was added to give the expected signal-to-noise ratio (Davidson and Sun 1988; Hancock et al. 2011). The ground was identified from the high-resolution ALS data (Isenberg 2011) and canopy height was calculated relative to that ground surface as the 98^{th} percentile (RH98) (Drake et al. 2002), after denoising the waveform by subtracting the mean noise and smoothing by a Gaussian of FWHM 11ns (Hancock et al. 2017).

We then simulated the expected ISS ground track distribution to extract GEDI footprints from the ALS-derived pseudo-waveforms. GEDI comprises three identical lasers (Coyle et al. 2015; Stysley et al. 2015). One laser is not split, whereas each of the other two lasers is split into two beams (coverage beams), with power reduced to slightly less than half that of the strong beam. These five beams are dithered across-track on every other line to produce 10 parallel ground tracks with approximately 600

m spacing across track and 60 m spacing along track. The GEDI across-track ground swath width (the distance from beams 1 to 10) is therefore approximately 5.4 km. The inclination of tracks relative to north, determined by the inclination of the ISS orbit (Qi and Dubayah 2016), is latitude-dependent and thus different for each site (see Figure *3-1*). After simulating the likely number of times the 10-beam pattern of GEDI would cross each site after the full two-year period, we obtained the track patterns under cloud-free condition as shown in the top row of Figure 3-1.

Since future local cloud conditions for each GEDI orbital pass are unknown, an estimate of ~50% for the mean global cloud cover (Downs and Day 2005) was applied to obtain the track patterns under cloudy condition for all sites (see lower row in Figure 3-1). Specifically, the impact of data losses due to cloud cover was simulated by removing complete GEDI tracks. For each track, a random number (0-1) was selected and if that number was greater than the cloud over (0.5), it was used. If it was less, all GEDI footprints in that track were rejected. This assumes that the cloud length scale was large enough to remove a complete track, but not so large that adjacent tracks were affected. Both track patterns under cloud-free and 50%-cloud conditions were then used as templates for the extraction of ALS-derived waveforms to simulate GEDI observations (see Figure 3-2).

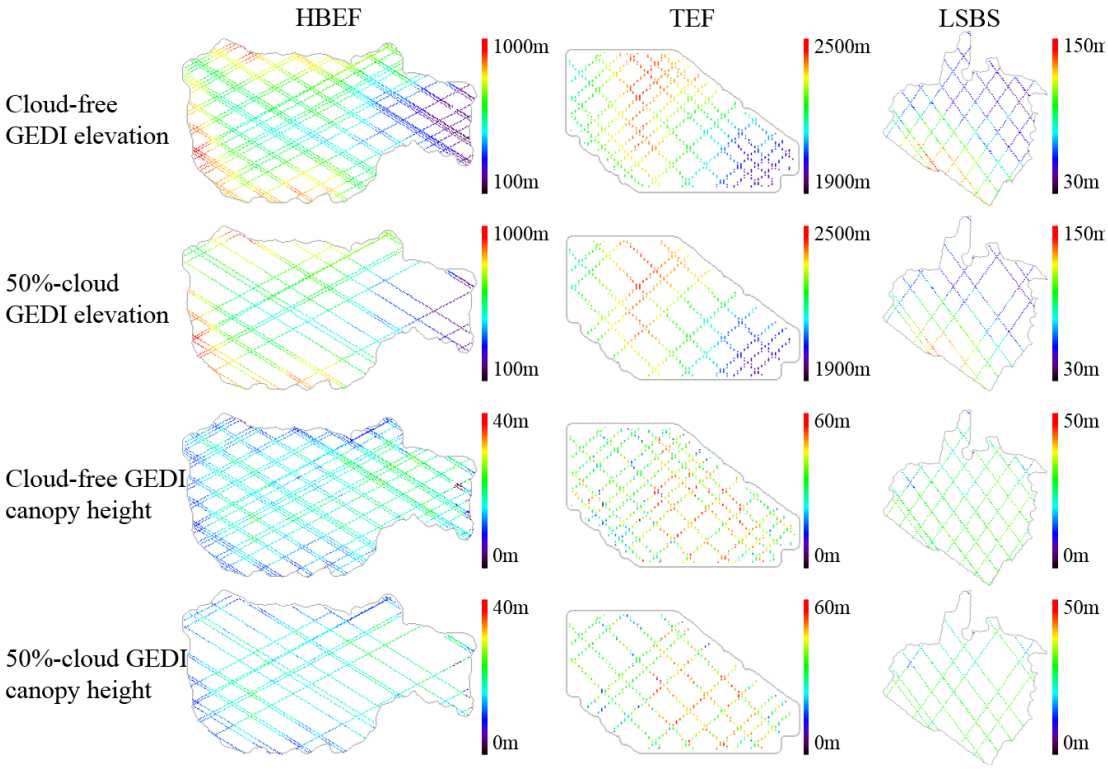


Figure 3-2. Simulated GEDI observations of elevation and canopy height over nominal two-year period based on cloud-free and 50%-cloud cover conditions.

B. TDX data

The simulated GEDI data were based on ALS data acquired pre-TDX-launch. We therefore used TDX acquisitions closest in time, to minimize temporal discrepancies. Specifically, TDX acquisitions in 2011 were used for TEF and LSBS test sites; for HBEF, both 2011 (leaf-on) and 2012 (leaf-off) acquisitions were used (Table 3-1). Selection of TDX data within the desired temporal windows was further refined based on their Height of Ambiguity (*HoA*) values. *HoA* can be calculated as $2\pi/\kappa_z$ (see Section 3.3.1 for the definition of κ_z and its use in RVoG model) and

defines the maximum height retrieval allowed by a specific acquisition geometry (Kugler et al. 2015). In terms of polarization state, we only explored *HH* data because of its availability at the global scale (Krieger et al. 2007; Kugler et al. 2014). All data were acquired in bistatic mode, where one satellite was transmitting and both satellites were simultaneously receiving the returned signal, and thus had no temporal decorrelation effect (Abdullahi et al. 2016; Kugler et al. 2014; Lee and Fatoyinbo 2015).

The time difference between the ALS and TDX acquisitions was 2-3 years for HBEF, 3 years for TEF and 2 years for LSBS. The magnitude of forest change was minor over most undisturbed places within these time intervals (Dubayah et al. 2010; Duncanson et al. 2015a; North et al. 2010; Siccama et al. 2007; Smith et al. 2005; Van Doorn et al. 2011; Whitehurst et al. 2013). All areas disturbed between the acquisition dates of ALS and TDX data were removed using ancillary disturbance product from Landsat images (Huang et al. 2010).

Table 3-1. Summary of TDX acquisitions over the study areas.

	Date	Eff. Bsl. (m)	HoA (m)	Rg. Res. (m)	Az. Res. (m)	Pol.
HBEF (temperate mixed forest)	2011/10/21 (Leaf-on)	121.42	-47.43	2.99	3.30	HH
	2012/01/28 (Leaf-off)	85.37	-68.12	2.99	3.30	HH

TEF (mountainous conifer forest)	2011/12/10	103.59	-64.47	2.71	3.30	HH
LSBS (tropical broadleaf forest)	2011/12/05	89.43	67.79	1.93	6.60	HH, VV

*Dates are listed in Year/Month/Date format. Eff. Bsl. - Effective Baseline; HoA - Height of ambiguity; Rg. Res. - Range Resolution; Az. Res. - Azimuth Resolution; Pol – Polarization.

3.3 Single-polarization RVoG Inversion and Combination with GEDI Data

3.3.1 RVoG model and height inversion from single-pol InSAR data

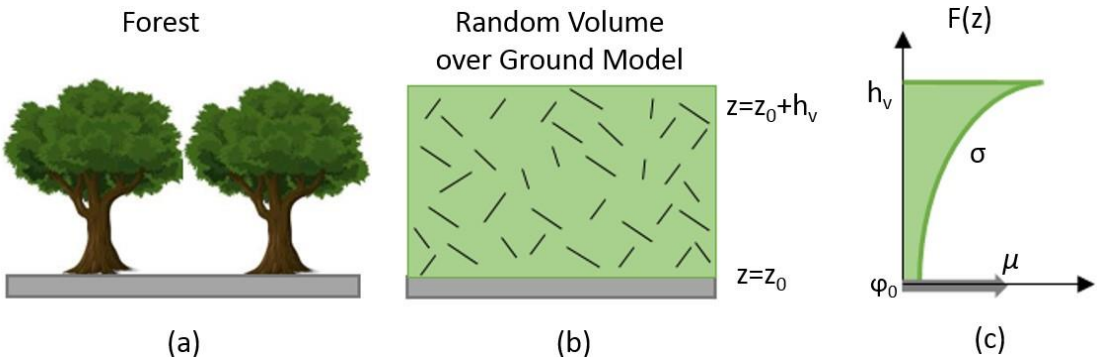


Figure 3-3. The basis of the RVoG model. Forest structure in (a) is modeled using the two-layer scattering model in (b) with ground elevation z_0 and volume height (h_V). Scatterers are randomly distributed and oriented inside the forest volume (Cloude and Papathanassiou 2003). *F(z)*, radar reflectivity of forest scatterers at different height *z*, decays as a function of extinction coefficient (σ) as shown in (c). The term φ_0 denotes

the ground phase ($e^{i\varphi_0} = e^{i\kappa_z z_0}$) and μ is the ground-to-volume amplitude ratio (Cloude and Papathanassiou 2003).

Random Volume over Ground (RVoG) model is a widely used two-layer scattering model (see Figure 3-3) that enables the inversion of physical forest parameters from InSAR coherences. Based on the RVoG model, the complex interferometric coherence $\tilde{\gamma}(\vec{\omega})$ at a polarization ($\vec{\omega}$), after compensating system and geometry induced decorrelation effects (Kugler et al. 2015), can be simply represented by equation (2-1),

$$\tilde{\gamma}(\vec{\omega}) = e^{i\varphi_0} \frac{\tilde{\gamma}_V + \mu(\vec{\omega})}{1 + \mu(\vec{\omega})} \tag{2-1}$$

where φ_0 is the phase corresponding to the ground elevation z_0; and $\mu(\vec{\omega})$ denotes the (polarization-dependent) ratio of powers echoed from ground and forest volume (Hajnsek et al. 2009; Kugler et al. 2014; Papathanassiou and Cloude 2001) (Figure 3-3). Since this study only works with *HH* TDX data ($\vec{\omega} = HH$), $\tilde{\gamma}(\vec{\omega})$ will be written as $\tilde{\gamma}$ and $\mu(\vec{\omega})$ as μ hereafter. $\tilde{\gamma}_V$ represents volume coherence and can be described by a Fourier relationship of the vertical profile of the radar reflectivity *F(z)* and the volume height h_V,

$$\tilde{\gamma}_V = \frac{\int_0^{h_V} \mathrm{F(z')}\, \mathrm{e}^{i\kappa_z z'}\, dz'}{\int_0^{h_V} \mathrm{F(z')} dz'} \tag{2-2}$$

where κ_z is the effective vertical wavenumber. Therefore, when the ground-to-volume amplitude ratio (μ) was zero, the correlation coefficient $|\tilde{\gamma}| = |\tilde{\gamma}_V|$. The estimation of

h_V requires the parameterization of *F(z)*. A widely and successfully employed approach is to assume that the distribution of scatterers decreases exponentially from the volume top downward, i.e., $F(z) = e^{\frac{2\sigma}{\cos\theta}z}$ where θ describes the incidence angle and σ, the mean extinction coefficient, represents the attenuation rate of the wave within the volume (see Figure 3-3) (Cloude and Papathanassiou 2003; Papathanassiou and Cloude 2001).

Single-pol InSAR inversion is an underdetermined problem, meaning that the number of observables from the interferometric coherence is smaller than the number of the unknown parameters. Previous studies solved this problem with two constraints: 1) using an external digital terrain model (DTM) to estimate ground phase (φ_0) or using a fixed mean extinction coefficient (σ) for the entire study site; and 2) assuming a zero ground-to-volume amplitude ratio (μ=0) at the polarization state of the acquisition. However, external DTMs are often unavailable over large areas, and the accuracy of height inversion may be compromised when a fixed σ (as opposed to one that varies spatially from using an external DTM) is used (Hajnsek et al. 2009; Kugler et al. 2014). Also, ground scattering may be present (i.e. $\mu \neq 0$) in areas with low forest density or low vegetation water content (Kugler et al. 2014). To overcome these issues, we assess the efficacy of using GEDI-derived DTM and canopy heights to provide the needed prerequisite information for TDX single-pol inversion.

3.3.2 Combining RVoG single-pol inversion with GEDI data

We perform a set of three analyses (Figure 3-4) where DTM and tree height variables derived from simulated GEDI lidar data were added progressively as inputs to improve the parameterization of single-pol RVoG inversion, enabling an examination of the respective performance gain on height estimation.

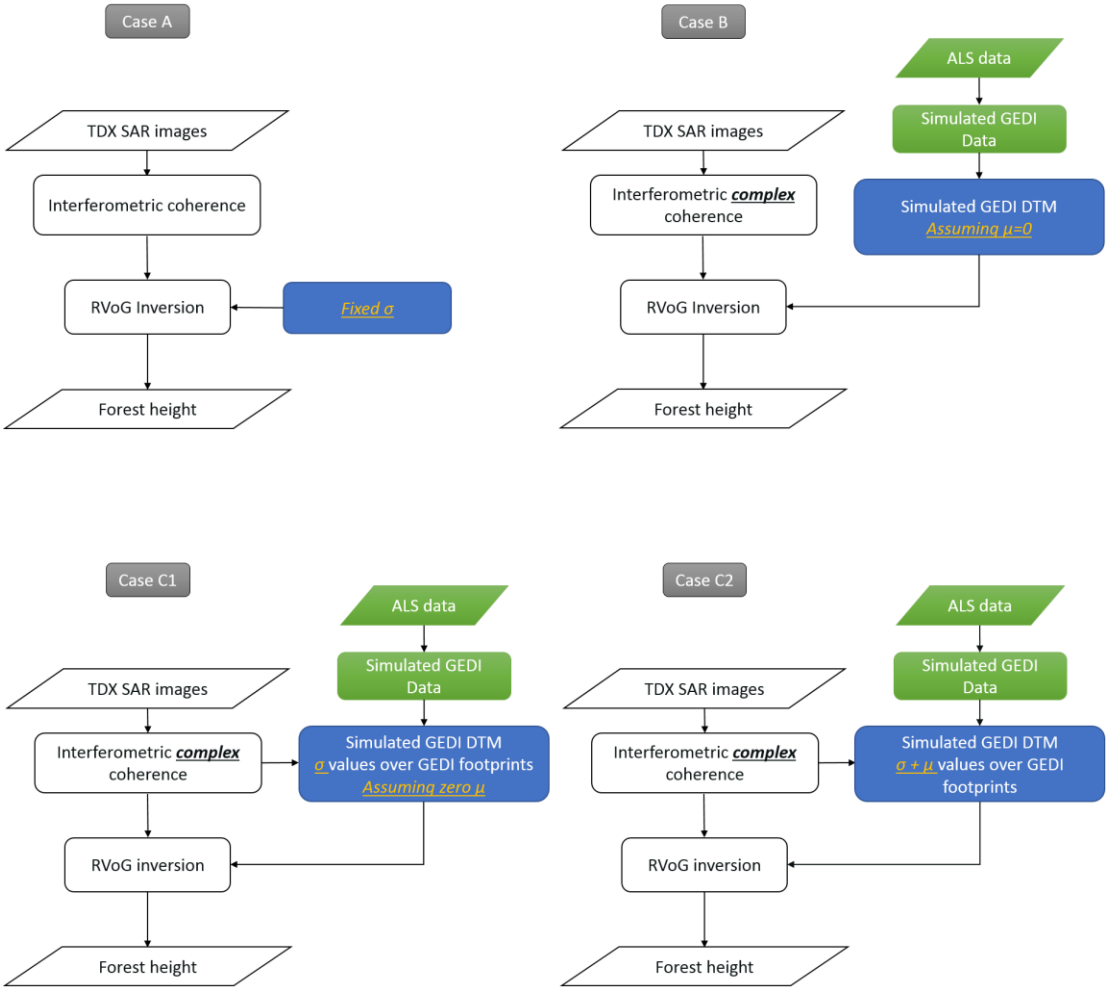


Figure 3-4. Main procedures for the four different fusion approaches – cases A, B, C1 and C2.

Table 3-2. RVoG parameter summary for different cases performed in this study.

Cases	**Used TDX observables**	**Known parameters for RVoG**
A Only TDX	Magnitude of coherence	σ = *constant*; assuming μ = 0
B Using GEDI DTM	Complex coherence	Ground phase φ_0 from GEDI DTM; assuming μ = 0
C1 & C2 Using GEDI height and DTM	Complex coherence	C1) Ground phase φ_0 from GEDI DTM; σ values over GEDI footprints; assuming μ = 0
		C2) Ground phase φ_0 from GEDI DTM; Both σ and μ values over GEDI footprints

***Case A - Only TDX*.** This case served as a baseline to assess what improvements in canopy height accuracy, if any, would be achieved from the addition of data derived from GEDI. Here, forest height was derived solely from the magnitude of TDX interferometric coherence $|\tilde{\gamma}|$ (i.e. interferometric correlation coefficient) by using a constant value of extinction coefficient (σ) and a zero ground-to-volume amplitude ratio (μ=0) assumption (see Table 3-2 case A). A key step of this method is to determine an appropriate σ value that in general represents forest density and dielectric constant for the entire study site. For a particular acquisition

with *HoA* larger than forest height, a σ value higher than optimum often leads to an overestimation of h_V whereas a σ value lower than optimum may result in an underestimation of h_V (see Figure 3-5) (Caicoya et al. 2012; Hajnsek et al. 2009). The presence of ground scattering that violates the μ=0 assumption may also lead to increased errors in tree height estimation (Kugler et al. 2014). Previous studies found a variation of 0.3 dB/m – 1 dB/m for σ values in temperate leaf-on broadleaf forest (Kugler et al. 2010), 0 – 0.4 dB/m in conifer forest (Caicoya et al. 2012) and 0.1 dB – 0.9 dB/m in tropical broadleaf forest (Hajnsek et al. 2009). Forest heights had been retrieved using a constant σ value of 0.3 dB/m for a tropical forest and a leaf-on temperate forest (Hajnsek et al. 2009; Kugler et al. 2010), and a value of 0.2 dB/m for a conifer forest (Caicoya et al. 2012). In this study, we applied similar σ values of 0.3 dB/m for leaf-on HBEF and LSBS, and 0.2 dB/m for TEF. The relatively smaller extinction of 0.2 dB/m was applied for leaf-off HBEF as better penetration capability of TDX was usually observed for leaf-off deciduous forest due to the relatively lower canopy cover and forest density (Abdullahi et al. 2016; Olesk et al. 2015).

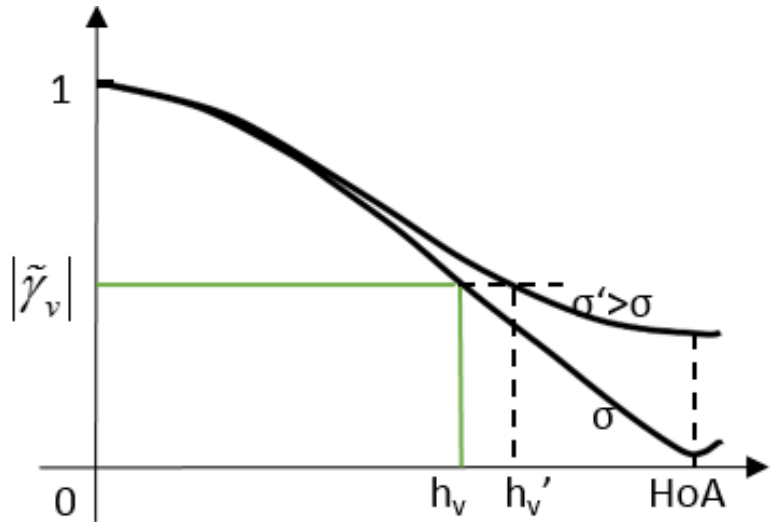


Figure 3-5. An increase of h_V up to *HoA* corresponds to the decrease of $|\tilde{\gamma}_V|$ for a fixed σ. For the same $|\tilde{\gamma}_V|$ value, a higher σ' value derived a larger forest height h_V'.

Case B – Using a simulated GEDI DTM. This case was designed to examine the impact of adding a GEDI-derived DTM on the TDX RVoG inversion. The DTM was created at 30 m resolution using simulated GEDI elevation data and a kriging-interpolation method based on a spherical semivariogram model. Ground phase (φ_0) was estimated from this DTM and subsequently used to derive forest scattering phase (φ); φ equals TDX interferogram ($\varphi_{\_interf}$) subtracted by flat-earth-phase ($\varphi_{\_flat}$) and ground phase φ_0; both $\varphi_{\_interf}$ and $\varphi_{\_flat}$ can be calculated from TDX acquisitions. The scattering phase (φ), combined with the interferometric correlation coefficient ($|\tilde{\gamma}|$), allowed the establishment of a balanced single-pol RVoG inversion after using the μ=0 assumption to derive h_V and σ (see Table 3-2 case B). Compared to the fixed σ value employed in case A, data-driven σ values may reflect better the variation of

forest environment, such as volume density and vegetation water content, and thus enhanced the h_V inversion. However, because the accuracy of scattering phase (φ) estimation is sensitive to the accuracy of the GEDI-derived DTM that derived φ_0, efficacy of this method is impacted by the available lidar shot density and the local topography variation, i.e. if the DTM is created from sparse data and the topography has large variation at local scales, the kriged DTM may not capture this variation accurately.

Case C1 and C2 – Using a simulated GEDI DTM and GEDI canopy heights. In these two cases, we assessed the effect of using both DTM and canopy height from simulated GEDI data on RVoG height inversion. The auxiliary information from lidar enabled the determination of two RVoG parameters over GEDI tracks (σ and μ), to constrain the inversion. To quantify the performance gain, we designed cases C1 and C2 to parameterize these variables progressively with the added lidar inputs. First, for case C1, we calculated only σ values (kriging was then used to estimate σ for the entire study area), assuming μ=0 as in previous cases, and tested the improvement of h_V estimation (see Table 3-2 case C1). By constraining σ with the additional input of simulated GEDI canopy height, case C1 was expected to be less sensitive to errors of DTM estimation than case B. Second, to further evaluate the effect of using GEDI-based RVoG parameterization, we calculated both σ and μ values and applied their interpolated maps (based on kriging) to derive h_V (see Table 3-2 case C2). Over areas where ground scattering is present, we hypothesize that case

C2 can reflect better forest structure variation and thus should outperform case C1, which assumed no scattering from ground (i.e. μ=0). For all cases, the derived forest height maps were resampled at 30 m resolution and subsequently averaged to 90 m (using a 3 × 3 window) to compare against reference lidar canopy heights.

3.4 Results

3.4.1 Case A – Only TDX

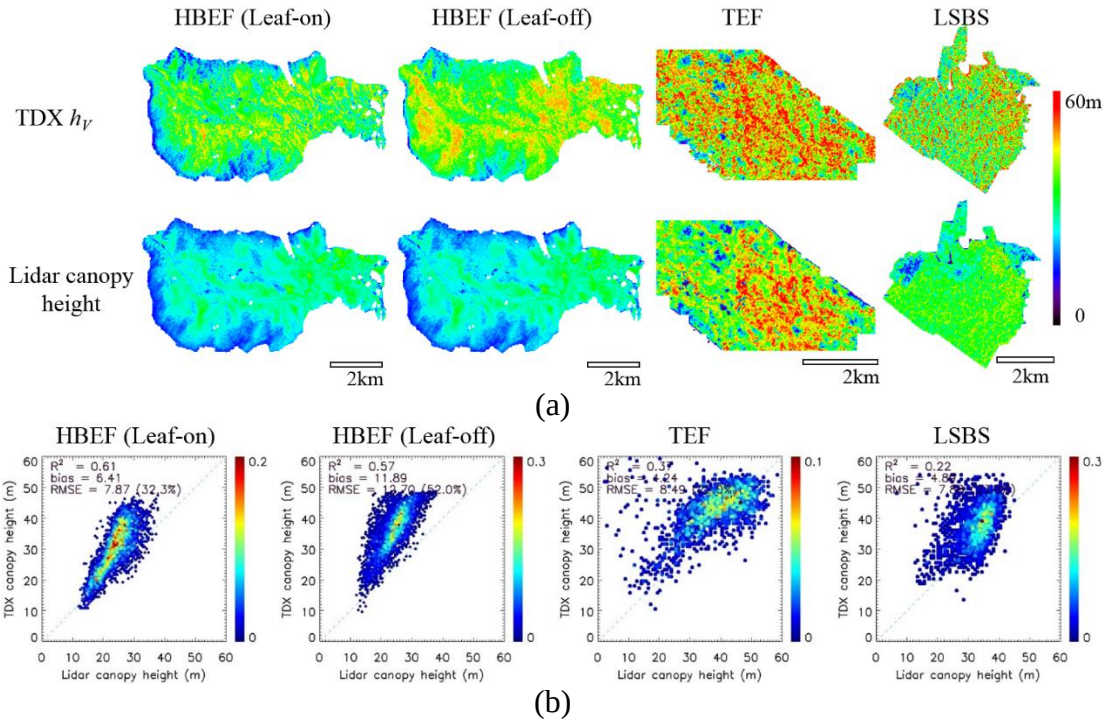


Figure 3-6. (Case A) (a) Forest heights derived using fixed extinction (σ) values of 0.3 dB/m (for leaf-off, 0.2 dB/m was used for leaf-on condition), 0.2 dB/m, and 0.3 dB/m respectively for HBEF, TEF, and LSBS. (b) Comparisons of the derived heights and reference lidar heights at 90 m resolution.

Following the method described in Section 3.3.2 (Case A), we derived forest heights directly from TDX correlation coefficient ($|\tilde{\gamma}|$) using fixed σ values of 0.3 dB/m (for leaf-off, 0.2 dB/m was used for leaf-on condition), 0.2 dB/m, and 0.3 dB/m respectively for HBEF, TEF, and LSBS (Figure 3-6). As mentioned earlier, for a particular acquisition with *HoA* larger than forest height, the use of a σ value that is too high may result in an overestimation of h_V, and vice versa. Biases of 6.4 m (leaf-on)/11.9 m (leaf-off), 4.2 m and 4.9 m were respectively found for HBEF, TEF and LSBS. These results indicated that optimum σ values may be smaller than those were used here. The particularly large bias at HBEF during leaf-off season may also be caused by the violation of the μ=0 assumption as there are areas of low canopy cover that could lead to ground scattering.

Despite the large biases over HBEF, good correlations were observed between the derived heights and reference heights with r^2 of 0.61 for leaf-on and 0.57 for leaf-off conditions, indicating an overall homogeneous forest structure (because only one σ was given) and good explanatory power of TDX $|\tilde{\gamma}|$ at this site. In contrast, lower correlations were found at TEF (r^2=0.37) and LSBS (r^2=0.22), probably resulting from the lower explanatory power of the used TDX coherences, given that TDX signal is expected to have less penetration capability over areas with taller trees and higher forest density. In addition, these sites have a somewhat heterogeneous forest structures, and therefore an expectation that σ may have a larger spatial variation and thus is less suitable for using the fixed value assumption.

3.4.2 Case B – Using simulated GEDI DTM

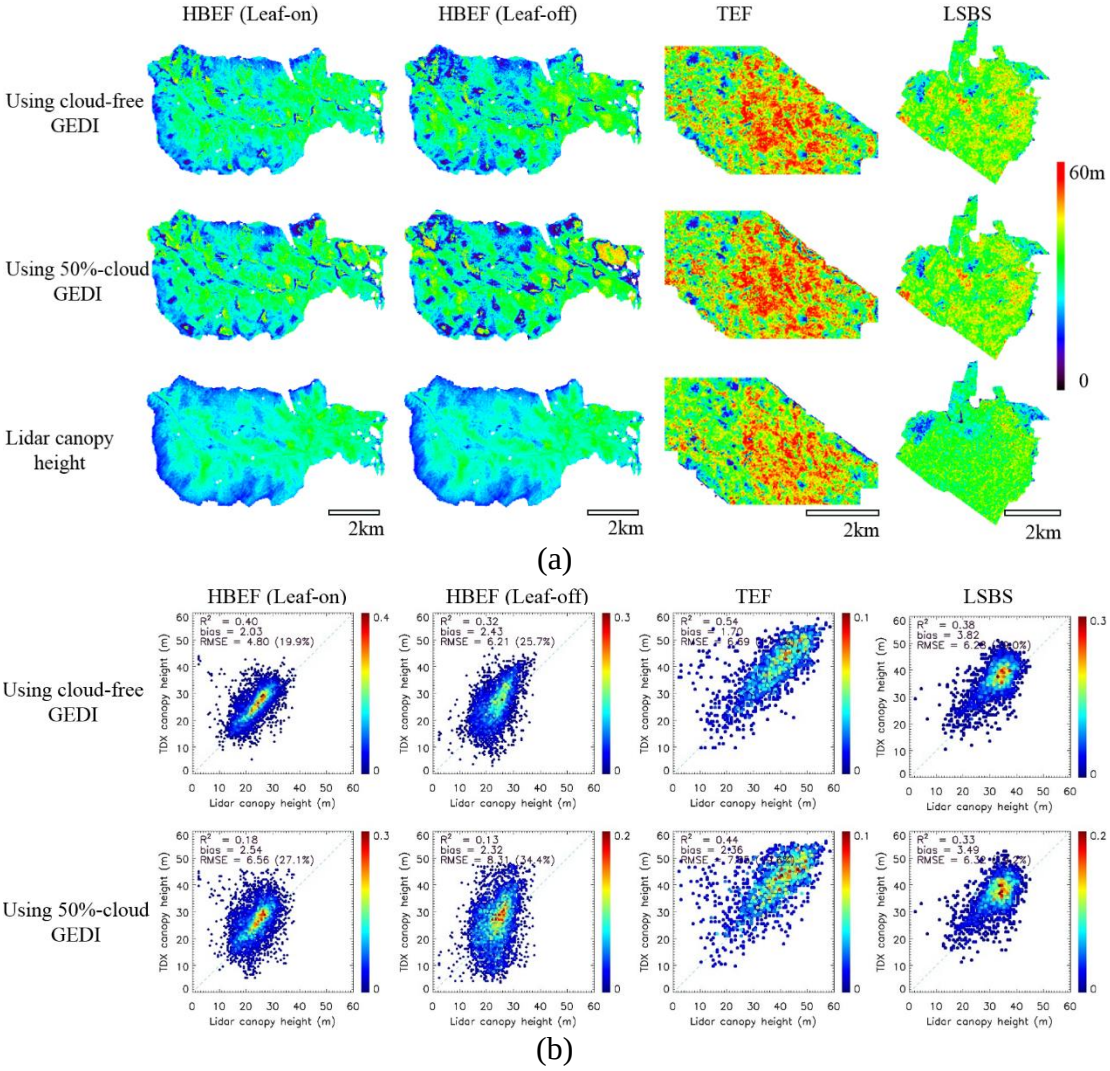


Figure 3-7. (Case B) (a) Forest heights derived from complex TDX coherence using simulated GEDI DTM, based on cloud-free and 50%-cloud cover conditions. (b) Comparisons of the derived heights and reference lidar heights at 90 m resolution.

We estimated a scattering phase (φ) map for each site using the external DTM derived from simulated GEDI elevation data. Forest height (h_V) as well as extinction

(σ) were then derived from the RVoG model using φ and correlation coefficient ($|\tilde{\gamma}|$) as inputs. Moderate agreement was found between the heights derived using cloud-free GEDI vs. lidar canopy heights, with r^2 of 0.40 (leaf-on) / 0.32 (leaf-off), 0.54 and 0.38 respectively at HBEF, TEF and LSBS (see Figure 3-7 and Table 3-3). Biases were reduced to 2.0 m (leaf-on) / 2.4 m (leaf-off) for HBEF, 1.7 m for TEF and 3.8 m for LSBS. Relatively lower agreement was found when using GEDI under 50% cloud cover, with r^2 of 0.18 (leaf-on) / 0.13 (leaf-off), 0.44 and 0.33, and biases of 2.5 m (leaf-on) / 2.3 m (leaf-off), 2.4 m and 3.5 m respectively for HBEF, TEF and LSBS. As mentioned earlier, σ represents the attenuation rate of the microwave signal inside the forest volume and reflects the variation of forest scatterer density and dielectric constant. Therefore, compared to case A which used a fixed σ, case B provided improved height estimates by exploiting a spatially varying σ, providing a better fit to the environmental condition at the time of acquisition.

Table 3-3. Validation results of RVoG heights from all cases at 90 m resolution.

Cases	Validation Parameters	HBEF				TEF		LSBS	
		Leaf-on		Leaf-off		-		-	
		Cloud-free	50%-cloud	Cloud-free	50%-cloud	Cloud-free	50%-cloud	Cloud-free	50%-cloud
Case A	r^2	0.61		0.57		0.37		0.22	
	Bias (m)	6.41		11.89		4.24		4.89	
	RMSE (m)	7.87		12.70		8.49		7.88	

	RMSE (%)	32.3		52.0		22.0		23.9	
Case B	r^2	0.40	0.18	0.32	0.13	0.54	0.44	0.38	0.33
	Bias (m)	2.03	2.54	2.43	2.32	1.70	2.36	3.82	3.49
	RMSE (m)	4.80	6.56	6.21	8.31	6.69	7.65	6.28	6.32
	RMSE (%)	19.9	27.1	25.7	34.4	17.0	19.6	19.0	19.2
Case C1	r^2	0.53	0.27	0.50	0.22	0.60	0.50	0.44	0.38
	Bias (m)	0.68	0.93	-0.66	-0.80	-0.84	-0.47	0.59	0.10
	RMSE (m)	3.83	5.51	4.16	6.49	5.95	6.77	4.65	4.94
	RMSE (%)	15.9	22.8	17.3	26.9	15.2	17.3	14.1	15.0
Case C2	r^2	0.71	0.38	0.69	0.32	0.58	0.43	0.45	0.40
	Bias (m)	-0.33	0.27	0.29	1.14	-0.62	-0.50	0.35	0.02
	RMSE (m)	2.58	4.02	2.66	4.90	5.99	7.07	4.27	4.51
	RMSE (%)	10.7	16.7	11.0	20.3	15.3	18.1	12.9	13.7

3.4.3 Cases C1 and C2 – Using simulated GEDI DTM and canopy heights

As described in 3.3.2, simulated GEDI observations of DTM and canopy height were combined to assist the parameterization of RVoG based on two approaches, either refining σ alone (case C1) or σ and μ combined (case C2).

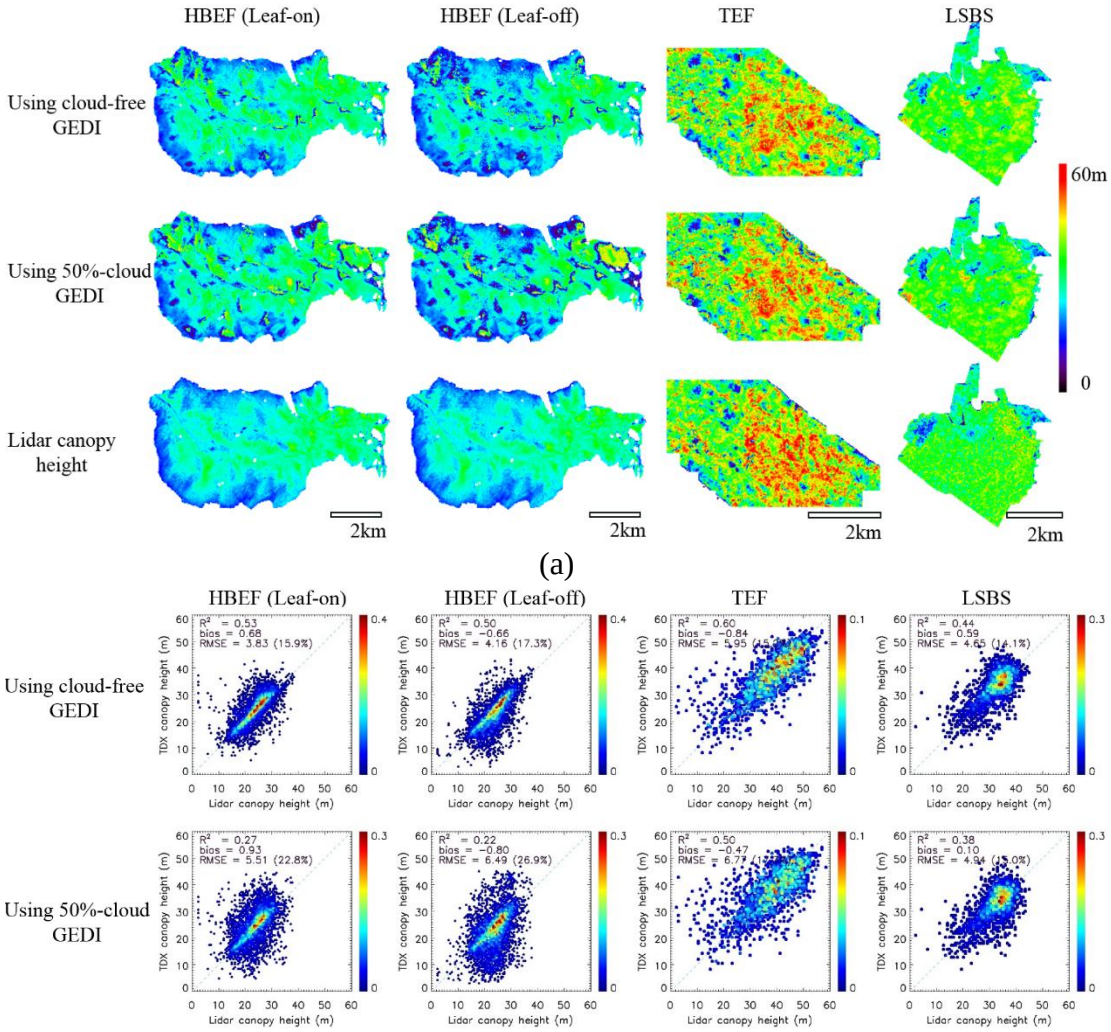
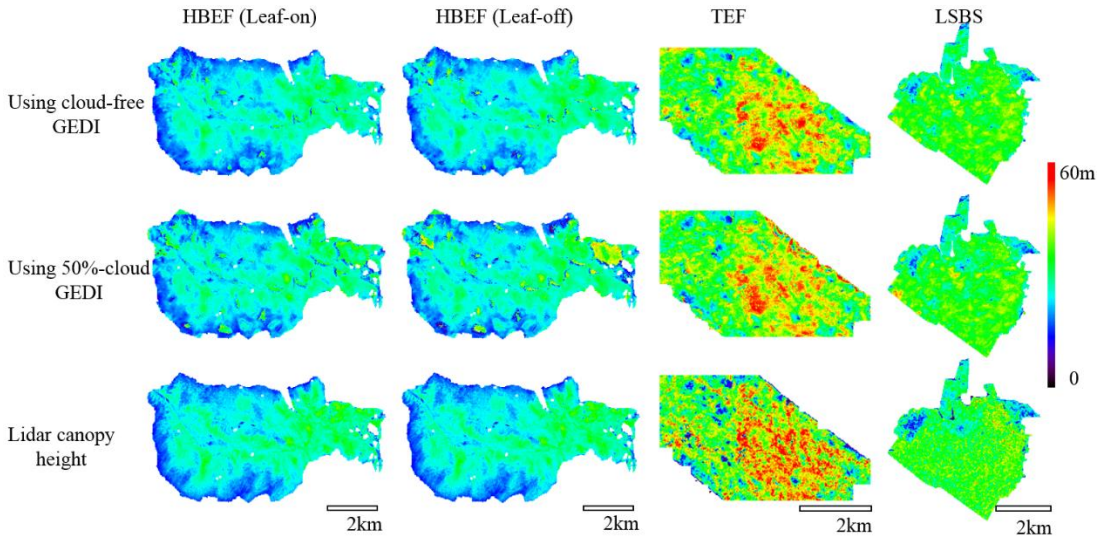


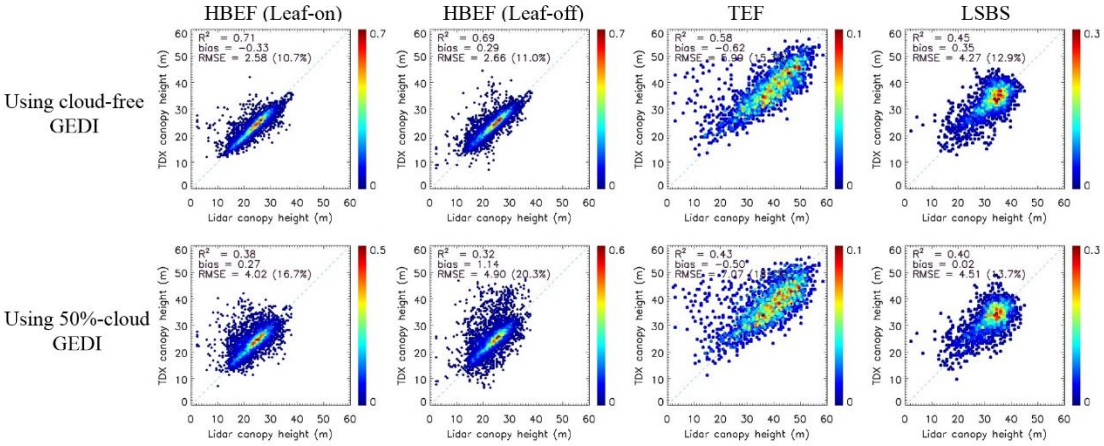
Figure 3-8. (Case C1) (a) Forest heights derived from complex TDX coherence using DTM and canopy height derived from simulated GEDI observations, respectively based on cloud-free and 50%-cloud conditions, to constrain σ. (b) Comparisons of the derived heights and reference lidar heights at 90 m resolution.

For case C1, simulated GEDI-derived extinction coefficient (σ) values were employed. Compared to case B which just used the external GEDI-derived DTM, the

constraining of σ from additional GEDI height information enhanced the single-pol inversion as expected. When using cloud-free GEDI observations, improved correlation between estimated height and lidar reference was obtained, with r^2 of 0.53 (leaf-on) / 0.50 (leaf-off), 0.60 and 0.44 respectively at HBEF, TEF and LSBS. Biases were reduced to 0.7 m (leaf-on) / -0.7 m (leaf-off) for HBEF, -0.8 m for TEF and 0.6 m for LSBS. The RMSEs were 3.8 m (relative error of 16% for leaf-on) / 4.2 m (17% for leaf-off), 6.0 m (15%), and 4.7 m (14%) respectively (see Figure 3-8, Table 3-3). When using GEDI under 50% cloud cover, height estimates were also improved compared to case B under the same GEDI coverage, with r^2 of 0.27 (leaf-on) / 0.22 (leaf-off), 0.50 and 0.38 respectively at HBEF, TEF and LSBS. Biases of 0.9 m (leaf-on) / -0.8 m (leaf-off), -0.5 m and 0.1 m, and RMSEs of 5.5 m (relative error of 23% for leaf-on) / 6.5 m (27% for leaf-off), 6.8 m (17%) and 4.9 m (15%) were found for the three sites (see Figure 3-8, Table 3-3). These results showed that by constraining σ estimation alone using local tree height information from GEDI, single-pol RVoG height inversion is significantly improved under a μ=0 assumption.



(a)



(b)

Figure 3-9. (Case C2) (a) Forest heights derived from complex TDX coherence using DTM and canopy height derived from simulated GEDI observations, respectively based on cloud-free and 50%-cloud cover conditions, to constrain σ and μ. (b) Comparisons of the derived heights and reference lidar heights at 90 m resolution.

The second method (case C2) was the only approach that derived both extinction coefficient (σ) and ground-to-volume amplitude ratio (μ) values to improve

the height inversion. This was made possible by adding both simulated GEDI canopy height and GEDI-derived DTM as inputs. These σ and μ values were interpolated and used jointly to calculate forest height from the complex coherence ($\tilde{\gamma}$). The estimated heights were improved at HBEF relative to all other cases. When using cloud-free GEDI data, we found r^2 values of 0.71 (leaf-on) / 0.69 (leaf-off), biases of -0.3 m (leaf-on) / 0.3 m (leaf-off) and RMSEs of 2.6 m (11% for leaf-on)/ 2.7 m (11% for leaf-off). When using 50% cloud-cover GEDI data, r^2 of 0.38 (leaf-on) / 0.32 (leaf-off), biases of 0.3 m (leaf-on) / 1.1 m (leaf-off), and RMSEs of 4.0 m (17% for leaf-on)/ 4.9 m (20% for leaf-off) were observed. For each specific case, leaf-on TDX-derived heights had stronger agreement with reference lidar heights than leaf-off TDX-heights. Somewhat paradoxically, greater improvements were observed from case A to case C2 using leaf-off data at HBEF. This is mainly because leaf-off forests have relatively lower volume scattering and higher ground scattering, and are more likely to violate the μ=0 assumption; therefore, the RVoG model using leaf-off data had greater reliance on GEDI inputs (particularly canopy height) for constraining the σ and μ parameters to accurately invert forest heights.

At TEF and LSBS, case C2 derived heights with r^2 values of 0.45-0.58 (cloud-free)/0.40-0.43 (50%-cloud), biases of -0.6 to 0.4 m (cloud-free)/-0.5 to -0.02 m (50%-cloud), and RMSEs of 4.3-6.0 m (13%-15%, cloud-free)/4.5-7.1 m (14%-18%, 50%-cloud). The overall improvements from case C1 to case C2 were not seen (at TEF) or only marginal (at LSBS). This indicated a lower utility of constraining the μ

values in improving height estimation over areas where taller trees, higher canopy cover or heterogeneous forest structure prevailed (Figure 3-9 and Figure 3-10; Table 3-3).

3.5 Discussion

There is the potential to combine the sparse, footprint level estimates of GEDI with wall-to-wall SAR measurements from TDX to provide continuous estimates of canopy height at much finer spatial resolution than what can be obtained by GEDI alone. Indeed, as currently planned, GEDI will grid its height observations to a resolution of between 500 m – 1000 m. Our work presented here provides a realistic pathway towards the goal of improved height mapping at these finer resolutions.

Our study explored the efficacy of using simulated GEDI observations in improving TDX estimate of canopy heights. The utility of two GEDI-aided RVoG parameters – extinction coefficient (σ) and ground-to-volume amplitude ratio (μ) – in improving forest height estimation was assessed. These two parameters are related to forest height, density, canopy cover, as well as the dielectric constant of scatterers in a forest, and vary across the landscape in different forest environments. In previous studies, these were mainly derived using full-polarimetric InSAR data at longer-wavelength (such as L-band), which are currently unavailable at the global scale (Hajnsek et al. 2009; Kugler et al. 2015; Neumann et al. 2012). Our study demonstrated that these RVoG parameters can be effectively derived from single-pol TDX data by adding simulated GEDI observations of terrain elevation and canopy

height as model inputs, and can be applied to improve forest height estimation over a wide range of forest types and terrain conditions.

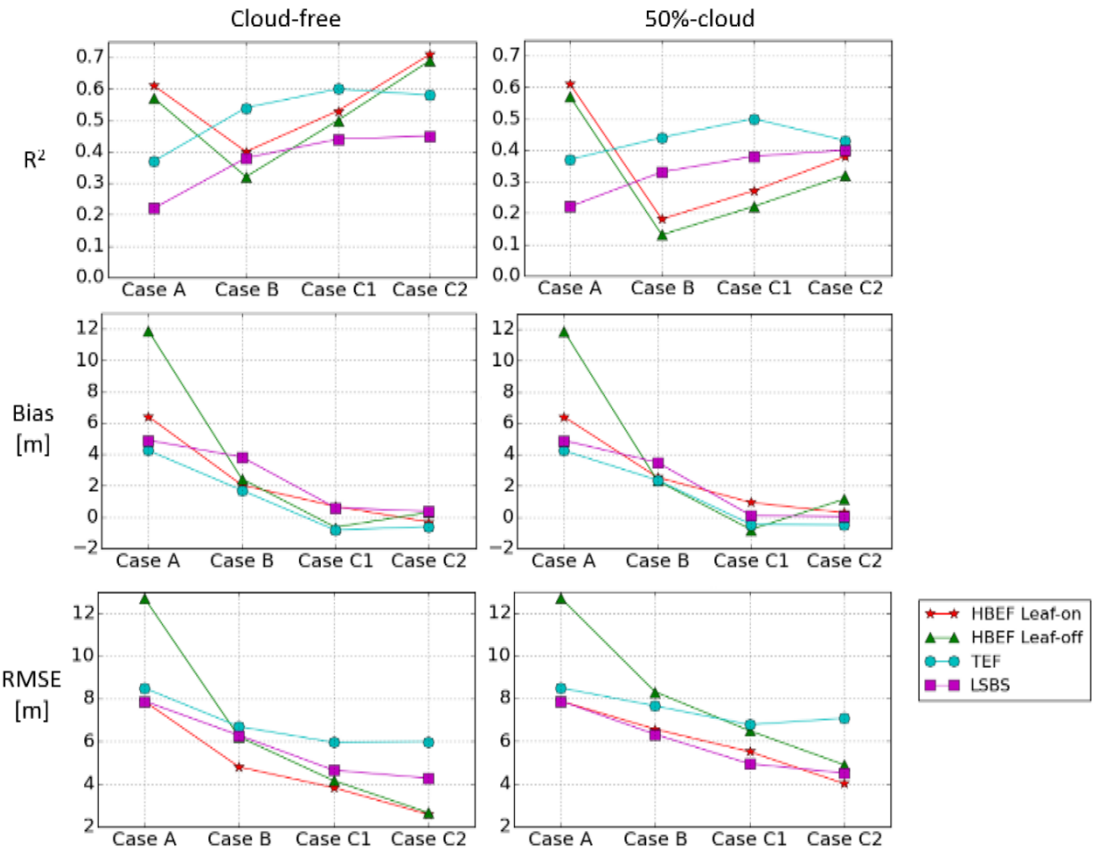


Figure 3-10. Different model performance corresponding to the four different cases over HBEF, TEF and LSBS under both cloud-free and 50%-cloud cover conditions.

In general, height estimates improved as more information was used from GEDI to parameterize the RVoG model (Figure 3-10 and Table 3-3). Our results also demonstrated that height estimation using TDX data acquired in leaf-off conditions could be significantly improved through inclusion of GEDI data, opening up the

possibility of using a much broader range of TDX acquisitions in temperate deciduous forests. We did not, however, evaluate the impact of using leaf-off GEDI data, which is the focus of a future study.

The fidelity of the GEDI-derived DTM had a significant impact on the efficacy of GEDI/TDX fusion. A key step to providing more accurate height products may be to enhance the GEDI DTM (below canopy topography) using, for example, DEM (surface elevation) products from TDX (Bräutigam et al. 2014), SRTM (Rodriguez et al. 2006), ASTER (Abrams et al. 2010), or from future missions such as ICESat-2 (Abdalati et al. 2010) and NISAR (Hoffman et al. 2016). In particular, the combination of the transect sampling lidar observations from ICESat-2 and GEDI, when combined with continuous, but less accurate below canopy elevation measurements from other missions, within an improved spatial interpolation/kriging framework is a promising avenue for future research.

Related to this, is the fact that GEDI is limited by cloud-cover and the vagaries of the ISS precessing orbit, which may limit the number of observations available for a given region (and thus lead to an inaccurate DTM, for example, in those areas). When there are insufficient GEDI observations for a given study site, parameters derived over limited GEDI footprints may fail to cover the whole spectrum of forest structure and topographic conditions and may smooth through spatial discontinuities in forest structure (see results in Case C2 at TEF and LSBS). GEDI tries to overcome some of this issue by pointing to acquire a more uniform

track coverage. Since σ and μ are related to forest structural characteristics, an alternative approach may be to input σ and μ derived from the same TDX acquisition or those with similar geometries (particularly baselines) over similar forest types and environmental conditions, based on TDX coherence and other continuous fields (e.g. canopy cover maps from Landsat). This can be done using segmentation and clustering algorithms to group segments with similar expected σ and μ values (Clewley et al. 2014). Such fusion approaches are being developed as part of a collaboration between the German Aerospace Center (DLR) and the GEDI mission.

Extrapolation of our results to real data derived from GEDI should be done carefully. Our simulated GEDI data is based on using ALS data, along with the expected ISS track patterns under cloud-free and 50%-cloud cover conditions within an end-end simulator (Hancock et al. 2016). While the simulator has been validated, the on-orbit data from the GEDI instrument may differ from our simulations. Likewise, the location of the orbital tracks, and their density, while based on realistic scenarios using real orbital data from ISS may differ. Another source of potential error is geo-location uncertainty. The geolocation accuracy of GEDI footprints is estimated to be around 7 m at 1-sigma level. This geolocation uncertainty was not modeled in our experiments. Such error may lead to a less-accurate DTM generation for scattering phase (φ) estimation over sloping surfaces and or less-representative RVoG parameters of σ (and μ) over heterogeneous forest structure, and thereby lower the inversion accuracy. Minimizing geolocation uncertainty for GEDI has been a

priority during mission development precisely so we may preserve our ability to do fusion at fine spatial scales with other data.

One major application of height estimates (and the main driver behind the GEDI mission) is forest AGB estimation. Previous studies have identified a height accuracy requirement of about 1 m to 2 m at 100 m to 1000 m resolution (with finer resolution more favorable) for effective biomass estimates (Hall et al. 2011; Hurtt et al. 2010; Qi and Dubayah 2016). Our fusion results at 90 m resolution do not quite meet that requirement. However, one of the most important results is that fusion greatly reduces bias. This is key because if bias can be get low, the fused heights can be aggregated to a coarser resolution until the desired height accuracies are achieved. For example, starting from the 30 m resolution at which the GEDI/TDX fusion was conducted, our height products from case C2 agreed with reference lidar heights at RMSEs of 2.8 m (leaf-on) / 3.0 m (leaf-off) at HBEF, 7.5 m at TEF and 5.5 m at LSBS under cloud-free condition, and 4.1 m (leaf-on) / 5.0 (leaf-off) m at HBEF, 8.8 m at TEF, 5.8 m at LSBS under 50%-cloud condition; after averaging to 200 m, RMSEs were improved to 1.9 m (leaf-on) / 2.1 m (leaf-off) m at HBEF, 3.7 m at TEF and 3.3 m at LSBS under cloud-free condition and 3.0 m (leaf-on) / 3.9 m (leaf-off) at HBEF, 4.6 m at TEF and 3.7 m at LSBS under 50%-cloud condition; at 1000 m resolution, RMSEs were further improved to 1.0 m (leaf-on) / 1.1 m (leaf-off) m at HBEF, 2.1 m at TEF and 1.7 m at LSBS under cloud-free condition, and 1.5 m (leaf-on) / 2.0 m (leaf-off) at HBEF, 2.4 m at TEF and 1.8 m at LSBS under 50%-cloud

cover. These results demonstrated that aided by GEDI observations, TDX single-pol InSAR data has the potential for large-scale forest biomass estimation.

3.6 Conclusions

We investigated the fusion of Global Ecosystem Dynamics Investigation (GEDI) lidar data and TanDEM-X (TDX) InSAR data to improve forest structure mapping over three contrasting forest types covering a wide range of heights, canopy covers and topographies. Our results showed that forest height retrievals from TDX single-polarization InSAR acquisitions based on the widely Random Volume over Ground (RVoG) were significantly improved using GEDI observations of bare-ground topography and canopy top height as inputs to constrain the model parameterization. Improving TDX height estimates with the aid of GEDI measurements is a meaningful step towards deriving blended height products from the two missions with better accuracy and coverage than using either data source alone. These height products, if sufficiently accurate, should improve the potential use of these data for applications such as biomass modeling and biodiversity.

3.7 Acknowledgements

This study was funded by NASA's Earth and Space Science Fellowship Program (Grant # 5268930) and NASA's Global Ecosystem Dynamics Investigation mission. We appreciate the Microwaves and Radar Institute of the German Aerospace Center (DLR) for providing the TanDEM-X data and for the assistance in data

processing. We are especially grateful to Dr. Kostas Papathanasiou and Dr. Matteo Pardini for their useful comments on this research. We also thank Dr. Chengquan Huang, Dr. Min Feng and Dr. Feng Zhao for providing the forest disturbance data products.

Chapter 4: Forest Biomass Estimation over Three Distinct Forest Types Using Tandem-X InSAR Data and GEDI Lidar Data

4.1 Introduction

Improving estimates of forest above-ground biomass (hereafter, biomass) is critical for reducing the great uncertainties associated with the quantification of terrestrial carbon dynamics and the implications of such results to changes in climate, habitat and biodiversity (Goetz and Dubayah 2011; Hall et al. 2011; Le Toan et al. 2011). The ecosystem community has suggested a desired mapping scale of 1 ha (i.e., a resolution of 100 m) that matches important scales of vegetation dynamics to minimize the biomass and flux estimation errors. However, given the large uncertainties in the existing coarse-resolution global estimates, biomass maps created at resolutions of close to or finer than 1 km, with properly characterized uncertainties and sufficient accuracy, would provide superior information in comparison to current knowledge (Hall et al. 2011).

A promising contributor to solving the biomass mapping challenge is lidar remote sensing technology (Goetz and Dubayah 2011). Lidar data have been collected to retrieve important forest structural attributes including heights, canopy cover, and crown volume that can be closely related to forest biomass. However, lidar-derived biomass maps have been mainly restricted to regional scales where

airborne lidar campaigns were conducted. Global lidar coverage has been enabled by the Geoscience Laser Altimeter System (GLAS) onboard the Ice, Cloud and land Elevation Satellite (ICESat) (2003-2009); however, the mission had low sampling density over low- and mid-latitude areas (Lefsky et al. 2005; Simard et al. 2011). The acquisition of a consistent and comprehensive set of lidar measurements on forest structure is thus a pressing need for improving the current estimates of forest biomass and terrestrial carbon (Hall et al. 2011).

NASA's Global Ecosystem Dynamics Investigation (GEDI) mission has been specifically designed to provide the required lidar observations and is scheduled to be onboard the International Space Station (ISS) in late 2018. After the completion of its nominal 2-year mission, GEDI will have obtained unprecedented lidar observations between the latitudes of 51.5° north and south (Coyle et al. 2015). The footprints (~25 m in diameter) will be separated with an along-track distance of 60 m and an across-track spacing of 600 m. These GEDI observations by themselves will be used to create gridded data sets of forest structure and biomass at 1 km resolution using statistical methods (Healey and Patterson 2018). However, due to gaps between expected GEDI samples and possible impacts from cloud cover, it is of interest to develop algorithms that appropriately fuse GEDI lidar data with ancillary satellite data of forest structural properties to produce contiguous forest biomass maps at finer resolution and accuracy.

One source of such ancillary data can be from spaceborne Interferometric Synthetic Aperture Radar (InSAR) missions. InSAR technology was developed to retrieve the Earth's surface structure following the acquisition of two or more images at slightly different viewing angles, from either multiple or single orbital passes. To estimate forest attributes with higher accuracy, simultaneous InSAR acquisitions are preferred to construct coherences without the temporal decorrelation effect (Cloude and Papathanassiou 2003; Lavalle and Hensley 2015; Lee et al. 2013). With an aim of obtaining such acquisitions globally, the German Aerospace Center (DLR) launched the first dual-platform (bistatic) InSAR mission–TerraSAR-X in 2007 and TanDEM-X in 2010 respectively (with the joint mission abbreviated as TDX). Forest structural properties, in particular height, have been successfully derived from TDX data over a variety of terrain and environmental conditions using the well-acknowledged Random Volume over Ground (RVoG) model (Cloude and Papathanassiou 2003; Kugler et al. 2014; Lee and Fatoyinbo 2015) or other SAR scattering models (Soja et al. 2017). Yet accuracies of these estimated parameters are subject to the availability and accuracy of *a priori* information on forest structure and ground elevation, particularly when single-polarization (the mode at which TDX data were collected at the global scale) data were used (Kugler et al. 2014).

While multiple data sources may be available to provide the *a priori* information for TDX, GEDI would certainly yield one of the most accurate estimates of ground topography and forest structure at its footprint level globally. Recent

studies have demonstrated the improvement of height estimation from TDX using simulated GEDI sampled observations to enhance the RVoG parameterizations (Qi and Dubayah 2016; Qi et al. Submitted). Since TDX data are available wall-to-wall (i.e. spatially complete), the GEDI-aided TDX height estimates in turn can be used as ancillary data for GEDI to potentially provide biomass maps at finer resolutions and accuracies. If characterized with appropriate uncertainties, these biomass products (particularly when produced at a fine resolution of within a few hundred meters) may help improve our understanding of the terrestrial carbon budget and to assess its impact on biodiversity and habitat at both regional and global scales.

This study investigates the effectiveness of using TDX InSAR heights to improve biomass estimations from GEDI lidar observations. We estimate biomass at two mapping scales (100 m and 1 km) by integrating GEDI and TDX data, in contrast to the use of GEDI data alone. Statistical uncertainty analyses are conducted for both the products at the two spatial scales, with the aim of addressing the following two objectives: 1) Evaluate the improvement in the accuracy of biomass estimation at 1 km by fusing GEDI data with TDX data; 2) Assess the capability of GEDI/TDX fusion to produce biomass products at 100 m.

Experiments are conducted at three major forest types: mixed temperate (conifer and broadleaf), mountainous conifer, and tropical rainforest. We first simulate pseudo-GEDI waveform data using airborne laser scanning (ALS) data and develop regression models linking the biomass response to lidar height metrics, i.e.

biomass–lidar models, at ground plots. Based on these models, GEDI footprint-level biomass are estimated over the expected track patterns for a full mission period of 2 years under an assumption of 50% cloud cover. These GEDI-based biomass estimates are then averaged to provide biomass products at 100 m (with a large proportion of gaps) and 1 km. We characterize uncertainty by accounting for both the model parameter- and sampling-induced errors according to the hybrid inference method described in Saarela et al. (2016). The TDX heights co-located with GEDI footprints are subsequently extracted to ultimately develop regression models linking the biomass response to TDX observations as predictors, i.e. GEDI–TDX models, for final biomass product generation at both mapping scales. Uncertainties are assessed by propagating errors induced by the biomass–lidar model parameters to the errors estimated for the GEDI–TDX models, following the hierarchical model-based inference described in Saarela et al. (2016). The study is the first attempt to produce improved forest biomass maps based on GEDI and TDX data fusion.

4.2 Study Sites and Datasets

4.2.1 Study sites and field data

Our study sites included the Hubbard Brook Experimental Forest (HBEF) in the White Mountain National Forest in New Hampshire, Teakettle Experimental Forest (TEF) in the Western Sierra Nevada Mountain Range, and La Selva Biological Station (LSBS) in the Atlantic lowlands of Costa Rica (Figure 4-1). HBEF is a typical

mixed temperate forest that mainly consists of deciduous northern hardwoods and a small percentage (10–20%) of spruce-fir, TEF is a high-biomass conifer-dominated forest with dominant species of white fir, ponderosa pine, red fir, and California black oak, and LSBS is largely comprised of a mixture of old growth and secondary lowland tropical rainforests, with mostly evergreen canopy and a small number of deciduous trees.

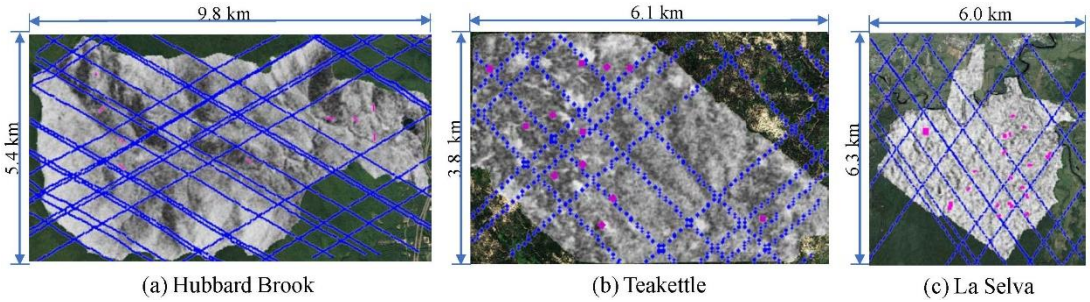


Figure 4-1. TDX coherence maps overlaid by in situ plots (red) and simulated GEDI tracks (blue) for the nominal 2-year period after cloud decimation based on 50% cloud cover over (a) Hubbard Brook, (b) Teakettle and (c) La Selva test sites.

Table 4-1. Summary of acquisition time for the used field, lidar, and InSAR data.

Sites	**In situ data**	**ALS data**	**InSAR data**
HBEF	2009	2009	2011
TEF	2008	2008	2011
LSBS	2009	2009	2011

At HBEF, forest height primarily ranges from 5–48 m and has a mean value of ~24 m. Mean above-ground biomass was estimated to be around 216 Mg/ha in 2001 (Siccama et al. 2007). Eighteen field plots were surveyed in New England (nine at HBEF and nine at BF (Bartlett Forest, which shares similar forest types as HBEF)) by investigators from federal and university laboratories during the NASA DESDynI (Freeman et al. 2009; Hall et al. 2011) field campaigns in 2009 (Table 4-1). This campaign was carried out to obtain estimates of useful forest biophysical attributes. These inventories were conducted in field plots of either 1 ha (200 m × 50 m, n=17), each of which consisted of sixteen 25 m × 25 m subplots (Table 4-2), or 0.5 ha (100 m × 50 m, n=1), which was comprised of eight 25 m × 25 m subplots. Various forest structural elements, including the diameters, heights, and species information, were recorded for trees with diameter at breast height (dbh, at 1.37 m) of 10 cm and larger. Biomass was then estimated for each tree following the method described in Jenkins et al. (2003), which includes 10 different equations depending on the tree species. These estimates were then aggregated to calculate the total above-ground biomass for each subplot and plot, and divided by the area of the subplot/plot to obtain biomass density.

TEF has a mature and complex structure with mean biomass values of around 200 Mg/ha, which may reach up to around 1000 Mg/ha in Giant Sequoia (*Sequoiadendron giganteum*) stands (Duncanson et al. 2015a; Smith et al. 2005; Swatantran et al. 2011). Field data were acquired in July 2008 (Table 4-1). The

sampled areas covered 25 square plots (12 overlapped the study site) with a size of 100 m × 100 m (1 ha), and each was divided into nine subplots of 33 m × 33 m (Table 4-2). The sample layout was oriented in azimuth so that plot sides followed the primary directions downslope and across slope. Characteristics including *dbh*, crown form, height, species, and location of each tree were all measured for trees with a *dbh* of 10 cm and larger, and these data were used for biomass estimations (Jenkins et al. 2003). Subplot- and plot-level total above-ground biomass and above-ground biomass density were subsequently calculated using these tree biomass data.

Table 4-2. Summary of scales of field data used for each forest type.

Forest Types	**Subplot-size**	**Plot-size**	**Number of plots**
Temperate mixed (HBEF and BF)	25 m × 25 m (0.0625 ha)	200 m × 50 m (1.0 ha); 100 m × 50 m (0.5 ha)	9 plots (8 at 1.0 ha, 1 at 0.5 ha) at HBEF and 9 plots at BF
Mountainous conifer (TEF)	33 m × 33 m (0.11 ha)	100 m × 100 m (1.0 ha)	12 plots overlapped TEF
Tropical rain (LSBS)	25 m × 25 m (0.0625 ha)	100 m × 50 m (0.5 ha); 100 m × 100 m (1.0 ha)	20 plots at 0.5 ha; 1 plot at 1.0 ha

LSBS consists of a mixture of lowland primary and secondary tropical rain forest, abandoned pasture, current and abandoned plantations, and agroforestry plots (Clark et al. 2008; Dubayah et al. 2010; Tang et al. 2014; Tang et al. 2012). As part of a landscape-scale carbon storage and flux study, the primary tropical forest data were

collected during 1997–2009 in eighteen 100 m × 50 m (0.5 ha) plots (Table 4-1), each of which was divided into eight 25 m × 25 m (0.0625 ha) subplots (Table 4-2) (Clark and Clark 2000). The plots were stratified over three edaphic conditions, namely, six plots on flat inceptisols, six on flat ultisols, and six on steep ultisol slopes. The secondary forest data were collected in three different areas that were approximately 25, 33, and 42 years old as of 2009. The two younger secondary forest plots were 100 m × 50 m each (0.5 ha; same as those in the primary forest), while the oldest one was 100 m × 100 m (1.0 ha; including 16 subplots). In each plot, all stems with *dbh* (at 1.37 m) or diameter above buttressing greater than 10 cm (in primary and 42-year-old secondary forests) or 5 cm (in 25- and 33-year-old secondary forests) were labelled and measured (Clark and Clark 2000). Biomass reference data were then generated based on stem diameter measurements using allometric equations for tropical wet forests (Rejou-Mechain et al. 2017).

4.2.2 Lidar data

Airborne laser scanning (ALS) data were collected at HBEF, TEF, and LSBS in September 2009, September 2008, and September to October 2009, respectively. These data were first reprocessed to medium-footprint (around 25 m diameter) pseudo-waveforms with signal to noise ratio (SNR), footprint shape and vertical resolution equivalent to those expected from GEDI data (Hancock et al. 2011). RH metrics, i.e., height at which a percentage of laser energy is returned, were then derived at 2% intervals from these pseudo-waveforms. Among the metrics, RH100 is

a noisy metric because it is associated with the first return and dependent on the SNR setup in the measurements; RH98 has been found to be more robust and is therefore used to represent canopy top height in this study; RH50–the height of median energy (HOME)–has been demonstrated to be sensitive to changes in the vertical arrangement of canopy elements and tree density (Huang et al. 2013). Ground elevation (Zg) has been calculated as the center of the lowest mode in the pseudo-waveform greater than mean signal noise (Drake et al. 2002). Other forest properties, such as canopy cover, can also be retrieved from the pseudo-waveforms. However, we used only RH metrics for estimating biomass in this study.

After simulating the number of times the International Space Station (ISS) was expected to pass each study site within a 2-year period, we obtained the GEDI ground-track patterns as shown in Figure 4-1, after applying a mean cloud cover of 50% to randomly decimate whole tracks. By using such track patterns as templates, simulated GEDI RH metrics were extracted to derive the GEDI-based biomass products as well as to fuse with TDX for improved biomass mapping.

4.2.3 TDX data

Because ALS data were collected in 2008–2009, we used only TDX data acquired in 2011 to minimize their temporal differences. The TDX acquisitions used for each site were further selected so that their HoAs (defined as one full phase cycle 2π divided by the effective vertical wavenumber κ_z) exceed maximum tree height to ensure the height of all trees are within measurable range. However, because

increasing HoA leads to higher levels of noise sensitivity (Kugler et al. 2015), acquisitions with HoAs slightly larger than the maximum tree height were preferred (see Table 4-3). These TDX data were obtained in bistatic mode, where one satellite was transmitting and both were simultaneously receiving the backscattered signal, so that coherences can be generated without temporal decorrelation effects. Also, the selected images were acquired in strip-map imaging mode (Abdullahi et al. 2016; Kugler et al. 2014; Lee and Fatoyinbo 2015), which enabled high resolutions of 1.93–2.99 m at ground range and 3.30–6.60 m at azimuth over the three study sites (see Table 4-3). Although both HH- (horizontal transmit and horizontal receive) and VV- (vertical transmit and vertical receive) polarization data were available at LSBS, only HH data were used for all sites because of their availability at the global scale.

Table 4-3. Summary of TDX acquisitions over the study areas.

Study Sites	**Date**	**Eff. Bsl. (m)**	**HoA (m)**	**Rg. Res. (m)**	**Az. Res. (m)**	**Pol.**
HBEF (temperate mixed)	2011/10/21	121.42	-47.43	2.99	3.30	HH
TEF (mountainous conifer)	2011/12/10	103.59	-64.47	2.71	3.30	HH
LSBS (tropical rainforest)	2011/12/05	89.43	67.79	1.93	6.60	HH, VV

Field inventory data and lidar data were collected in the same years over all sites (Table 4-1). However, both datasets were collected two (at HBEF and LSBS) or three (at TEF) years earlier than the TDX acquisitions. We removed places that were disturbed (0–7.8% of total area for each site) between the dates of field/lidar data collection and TDX acquisitions for more effective fusion based on ancillary Landsat disturbance products (Huang et al. 2010). Although biomass changes were small over most undisturbed forests within the time span of analysis, such changes may still have produced minor impacts on the TDX biomass estimates as compared to those from lidar, discussed later in sections 4.5 and 4.6 (Dubayah et al. 2010; Siccama et al. 2007; Smith et al. 2005; Van Doorn et al. 2011).

4.3 Methods

With the use of forest inventory data and co-located simulated GEDI RH metrics, biomass–lidar relationship models were developed and subsequently applied to estimate biomass respectively using wall-to-wall lidar data (as a reference) and simulated GEDI lidar data. Their means within each 100 m or 1 km grid cell were calculated to provide the respective biomass products (section 4.3.1). The GEDI track estimates, combined with co-located TDX heights, were applied to derive the GEDI biomass–TDX height relationship models for mapping TDX biomass at 25 m wall-to-wall. These estimates were then aggregated to provide contiguous biomass products at 100 m and 1 km, and the results were compared against those from wall-to-wall

lidar observations (section 4.3.2). Lastly, we characterized and analyzed uncertainties for all products statistically (section 4.3.3; see flow chart in Figure 4-2).

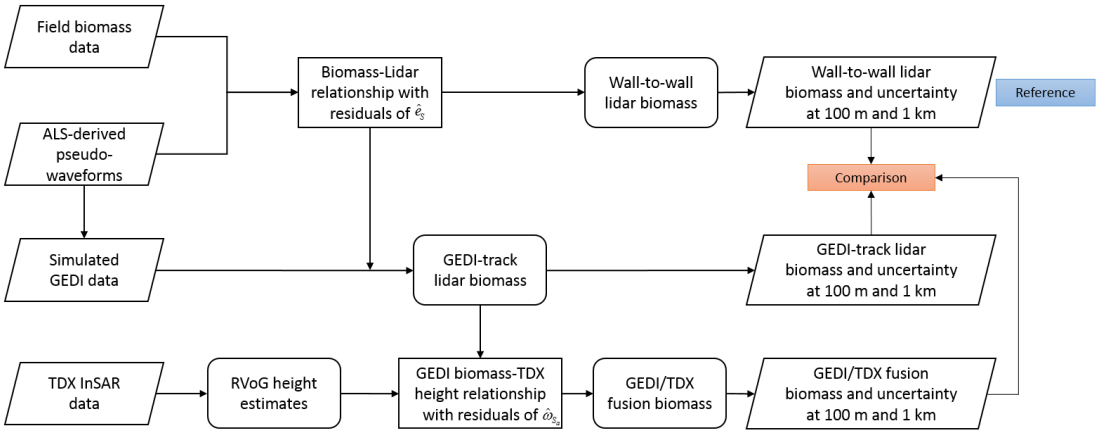


Figure 4-2. Methods designed to estimate biomass and uncertainty maps respectively using wall-to-wall (reference) and GEDI-track lidar data, and fused GEDI/TDX data.

4.3.1 Establishment of relationships between field biomass and lidar data

Site-specific biomass–lidar relationships were first established using the subplot field biomass (see Table 4-2 for the size and number of subplot field data) and co-located simulated GEDI metrics. Multiple linear regression was applied to relate the biomass (response) and lidar (independent) variables. Because relationships between biomass and each single lidar metric term may be non-linear (Giannico et al. 2016; Hudak et al. 2006; Zhao et al. 2013), we included RH metrics, squared RH metrics and products of canopy height (RH98) with other RH metrics as inputs, from which the best combination of these variables for biomass prediction were chosen for each site.

To select the best subsets, we conducted an exhaustive search among the inputs by measuring how good each subset is using information criterion such as AIC (Akaike Information Criterion) (see "regsubsets" function in the R package "leaps" (R Core Team 2014)). At each subset size, one model with the best correlation with field biomass was selected. Equivalent correlations may be achieved with different subset sizes (i.e. using a different number of variables); here we selected the model with the smallest amount of predictors to increase the degree of freedom in the regressions and to avoid the overfitting issues as discussed in previous studies (Valbuena et al. 2017). Linear equations were then developed for each site using the selected variables (Table 4-4), to facilitate uncertainty estimation following the methods described in Saarela et al. (2016). The selected variables and established models were then applied to the pseudo-waveform-derived metrics both wall-to-wall and along simulated GEDI tracks for biomass estimation. These estimates were subsequently averaged to 100 m and 1 km to generate reference-lidar and GEDI-lidar-biomass products. Note that gaps will present in the GEDI products, particularly in the 100 m estimates, because of the spacing between adjacent tracks.

4.3.2 Establishment of relationships between GEDI biomass and TDX heights

The TDX InSAR heights were then related to GEDI-derived biomass to fill the gaps between adjacent GEDI tracks and extend the estimates wall-to-wall. Unlike biomass–lidar models that were developed for each entire site, GEDI biomass–TDX height relationship models were established for every 1 km grid cell (which contains

1600 pixel cells with 25 m resolution), since biomass samples available from GEDI data were more widely distributed and had larger quantity than that from inventory plots, i.e. each 1 km grid cell had its own equation.

To be more specific, depending on the availability of GEDI observations in each 1 km grid cell, we applied two different methods to derive TDX heights for GEDI–TDX relationship establishment and biomass prediction. The first method was applied for grid cells containing GEDI observations; here TDX heights were estimated with the aid of GEDI observations of elevation and canopy height (Qi et al. Submitted). The second method was designed for grid cells without GEDI observations; here TDX heights were estimated solely from coherence data by making assumptions about certain parameters of the height inversion model (Qi and Dubayah 2016; Qi et al. Submitted).

Since in the second method, GEDI data were unavailable to train GEDI biomass–TDX height relationship, we enlarged the 1 km area to 2 km (or further to 3 km if necessary) to include GEDI observations for model establishment (see Figure 4-3). For each grid cell, we subsequently examined the correlations and RMSDs between GEDI-biomass and biomass derived from TDX height (h_t) or h_t^2, and chose the variable with better performance as model inputs. We still applied linear regression, using h_t or h_t^2 as input variables, to facilitate the uncertainty analysis following methods described in Saarela et al. (2016). Based on the established models (multiple models (one model for each grid cell) have been used for TDX heights),

TDX biomass maps were estimated at 25 m resolution. We then aggregated these estimates to 100 m and 1 km, with uncertainties characterized in a statistical manner as described below.

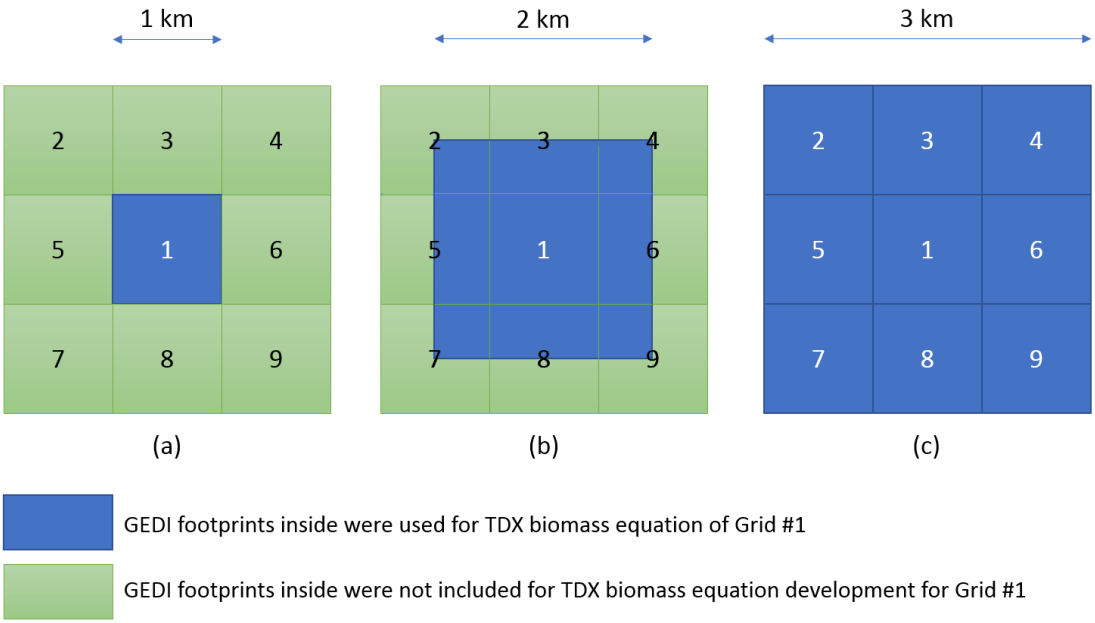


Figure 4-3. Area selection for establishing regression models for Grid cell #1 (each grid has a size of 1 km × 1 km). (a) When Grid cell #1 contained GEDI tracks, co-located GEDI biomass and TDX height (derived with the aid of GEDI observations) in Grid #1 were used to train the regression model. (b) When Grid cell #1 did not contain GEDI observations, co-located GEDI biomass and TDX height (derived using TDX coherence alone) in an enlarged area around Grid #1 (2 km × 2 km) were used to develop the regression model. (c) When the enlarged area in (b) did not contain enough GEDI observations, the area was enlarged to 3 km × 3 km for establishing the GEDI biomass–TDX height model.

4.3.3 Uncertainty analysis for the biomass products at mapping scales of 100 m and 1 km

Statistical uncertainty analyses were then performed for the derived biomass products at the two mapping scales (100 m and 1 km). Following the methods used in Saarela et al. (2016), uncertainties associated with the GEDI-based biomass products were estimated by considering both the sampling error and biomass–lidar model parameter error (see section 4.3.3 A), whereas uncertainties associated with GEDI/TDX fusion biomass products were estimated by propagating the modeling errors from the biomass–lidar relationship to the GEDI biomass–TDX height relationship (see section 4.3.3 B). Since both ALS and TDX data are spatially complete, they were assumed to have no sampling error. Residual error (also refers to residual prediction error) was currently not considered for all products. For large scale, this error has a minor contribution to the total uncertainty and is often ignored due to the residual (spatial) autocorrelation, discussed later in 4.5.

A. Estimation of wall-to-wall- and GEDI-lidar biomass uncertainty

Using the co-located biomass data and lidar predictors, a linear model was fitted for each study site with the following relationship:

$$\boldsymbol{y}_S = \boldsymbol{X}_S \boldsymbol{\alpha}_S + \boldsymbol{e}_S, \tag{4-1}$$

where *S* describes the field samples of size *m* at the subplot level, which has an equivalent scale to the GEDI footprints, $\boldsymbol{y}_S$ is a column vector of the field-based

biomass, $\boldsymbol{X}_S$ is a matrix formed by the lidar predictors (RH metrics and variables derived from these metrics) over the subplot samples, $\boldsymbol{\alpha}_S$ is the column vector of the linear model parameters, and $\boldsymbol{e}_S$ is the column vector of random errors with zero expectation. We applied the ordinary least squares (OLS) estimator to estimate model parameters, i.e. $\hat{\boldsymbol{\alpha}}_S = (\boldsymbol{X}_S^T \boldsymbol{X}_S)^{-1} \boldsymbol{X}_S^T \boldsymbol{y}_S$ (Davidson and MacKinnon 1993).

The variance of the biomass estimator for each 100 m or 1 km grid cell can then be described as follows (Saarela et al. 2016):

$$\begin{aligned} V\left[E\left(\mu\right)_{Lidar} \right] &= V_I + V_{II} \\ &= (1 - M / N) \frac{1}{M} {\omega_M}^2 + \boldsymbol{\iota}_{S_a}^T \boldsymbol{X}_{S_a} Cov\left(\hat{\boldsymbol{\alpha}}_S\right) \boldsymbol{X}_{s_a}^T \boldsymbol{\iota}_{S_a}, \end{aligned} \tag{4-2}$$

where V_I and V_{II} respectively account for the variances related to the sampling error and modeling error. Specifically, for V_I, S_a represents the available lidar samples of size M in each grid and N is the 25 m-cell population size within that grid, i.e., N equals 16 for each 100 m grid and 1600 for each 1 km grid. ${\omega_M}^2$ corresponds to the sample-based population variance from the M-length column vector of $\hat{\boldsymbol{y}}_{Sa} = \boldsymbol{X}_{Sa} \hat{\boldsymbol{\alpha}}_S$ values. Therefore, the sampling error decreases with the increase of the ratio of available lidar samples to the total population size (M/N), which leads to its minimum value of zero when lidar data are available wall-to-wall (M=N). For V_{II}, $\boldsymbol{\iota}_{S_a}$ denotes the M-length column vector of entities 1/M and $\boldsymbol{X}_{Sa}$ represents an $M \times (p+1)$ matrix formed by the lidar predictors, where p is the number of the used input lidar variables.

The covariance matrix of the estimated model parameters $\hat{\boldsymbol{\alpha}}_S$ is estimated in a heteroscedasticity-consistent manner, because of the violation of the ordinary least squares assumptions by using predictors correlated to each other (Saarela et al. 2016), and can be calculated as:

$$\boldsymbol{Cov}(\hat{\boldsymbol{\alpha}}_S) = \left(\boldsymbol{X}_S^T \boldsymbol{X}_S\right)^{-1} [\sum_{i=1}^{m} \hat{e}_i^2 \boldsymbol{x}_i^T \boldsymbol{x}_i] \left(\boldsymbol{X}_S^T \boldsymbol{X}_S\right)^{-1}, \quad (4\text{-}3)$$

where $\hat{e}_i$ is the residual and x_i the (p+1)-length row vector of lidar predictors for the i[th] observation from sample S. To overcome the issue of the squared residual $\hat{e}_i^2$ yielding biased estimators of the squared errors e_i^2, a correction of $\frac{m}{m-p-1}\hat{e}_i^2$ was applied.

B. Estimation of GEDI/TDX-fusion biomass uncertainty

After estimating the GEDI-lidar biomass based on the biomass–lidar relationship, GEDI biomass was used as the response variable for estimating model parameters linking biomass and TDX height (or variables derived from the height) over the lidar sample S_a assuming the following:

$$\boldsymbol{X}_{S_a} \hat{\boldsymbol{\alpha}}_{S_a} = \boldsymbol{Z}_{S_a} \boldsymbol{\beta}_{S_a} + \boldsymbol{\omega}_{S_a}, \quad (4\text{-}4)$$

where $\boldsymbol{X}_{Sa}\hat{\boldsymbol{\alpha}}_{Sa}$ is an M-length column vector of GEDI biomass estimates, $\boldsymbol{Z}_{Sa}$ is the TDX height (variables derived from the height) co-located with GEDI tracks, and $\boldsymbol{\beta}_{Sa}$ is the column vector of model parameters linking the estimated GEDI biomass values

and TDX predictors; $\boldsymbol{\omega}_{Sa}$ denotes the *M*-length column vector of random errors with zero expectation.

The variance of the biomass estimator can then be represented by:

$$\begin{aligned} V\left[E\left(\mu\right)_{TDX} \right] &= V_I + V_{II} \\ &= (1-N/N)\frac{1}{N}\omega_N{}^2 + \boldsymbol{\iota}_U^T \boldsymbol{Z}_U Cov(\hat{\boldsymbol{\beta}}_{S_a}) \boldsymbol{Z}_U^T \boldsymbol{\iota}_U . \end{aligned} \tag{4-5}$$

Similar to 4.3.3 A, V_I and V_{II} are variances of the mean biomass associated with sampling error and modeling error, respectively. Here V_I, the sampling error, is assumed to be zero as the TDX observations are available wall-to-wall. In V_{II}, $\boldsymbol{\iota}_U$ is the *N*-length column vector of entities 1/*N* (as mentioned earlier, *N* equals 16 for each 100 m grid and 1600 for each 1 km grid). $\boldsymbol{Z}_U$ is the $N \times (q + 1)$ matrix formed by TDX height input variables, where *q* is the number of TDX predictors. $\boldsymbol{Cov}(\hat{\beta}_{S_a})$ is the covariance matrix of $\hat{\boldsymbol{\beta}}_{s_a}$ and can be estimated with the following equation:

$$\begin{aligned} \boldsymbol{Cov}\left(\boldsymbol{\beta}_{S_a}\right) = &\left(\boldsymbol{Z}_{S_a}^T \boldsymbol{Z}_{S_a}\right)^{-1} \sum_{i=1}^{M} \hat{\omega}_i^2 z_i^T z_i \left(\boldsymbol{Z}_{S_a}^T \boldsymbol{Z}_{S_a}\right)^{-1} \\ &+ \left(\boldsymbol{Z}_{S_a}^T \boldsymbol{Z}_{S_a}\right)^{-1} \boldsymbol{Z}_{S_a}^T \left[\boldsymbol{X}_{S_a} \boldsymbol{Cov}\left(\boldsymbol{\alpha}_S\right) \boldsymbol{X}_{S_a}^T \right] \boldsymbol{Z}_{S_a} \left(\boldsymbol{Z}_{S_a}^T \boldsymbol{Z}_{S_a}\right)^{-1}, \end{aligned} \tag{4-6}$$

where $\mathbf{Z}_{Sa}$ is a $M \times (q + 1)$ matrix of TDX predictors over sample S_a (i.e., the GEDI samples within each grid where the grid-specific GEDI biomass–TDX relationship was established); $\hat{\omega}_i^2$ is a squared residual for the i^{th} observation in sample S_a, and the correction $\frac{M}{M-q-1}\hat{\omega}_i^2$ was applied, as done in earlier studies.

After estimating V for the biomass products at both 100 m and 1 km, the prediction interval (PI) can be estimated as $PI = E \pm t\sqrt{V}$, where E is the mean biomass in each grid at the respective mapping scale and t is the Student's distribution critical value for a given level of confidence and degree of freedom. For a large number of observations, the Student's and normal distributions give almost identical critical values. Here, we used the normal distribution since estimating the degrees of freedom for the Student's distribution with a two-phase sampling design is complex. Therefore, the following formula was used to estimate the prediction interval: $PI = E \pm Z\sqrt{V}$ (e.g., Z = 1.96 at 95% confidence), where $Z\sqrt{V}$ is the estimated uncertainty for each grid.

4.4 Results

4.4.1 Biomass estimation

A. Wall-to-wall- and GEDI-lidar biomass

Biomass–lidar relationships were developed respectively for HBEF, TEF and LSBS based on the method described in section 4.3 (see Table 4-4 for the selected metrics and the established equations). Moderately good performances were found for these models, with r^2 values of respectively 0.59, 0.52 and 0.37, and RMSEs of respectively 66.45 (31.6%) Mg/ha, 185.45 (43.4%) Mg/ha and 80.76 (34.9%) Mg/ha) between the field biomass and lidar estimates (Figure 4-4). Note that these relationships were established at the smaller subplot level (25×25 m at HBEF and

LSBS, and 33×33 m at TEF) at which crown geometry effects and random geolocation errors may lead to higher model uncertainties compared to those developed at larger plot levels.

Table 4-4. Relationship between field biomass and lidar RH metrics for each site.

Sites	Equations	R^2	RMSE (Mg/ha)
HBEF	-19.22 + 8.50×RH80 + 0.18×RH98×RH80	0.59	66.45 (31.6%)
TEF	97.04 + 12.42×RH60 + 0.25×RH98×RH80	0.52	185.45 (43.4%)
LSBS	53.13 + 0.22×RH40×RH40 + 0.14×RH98×RH40	0.37	80.76 (34.9%)

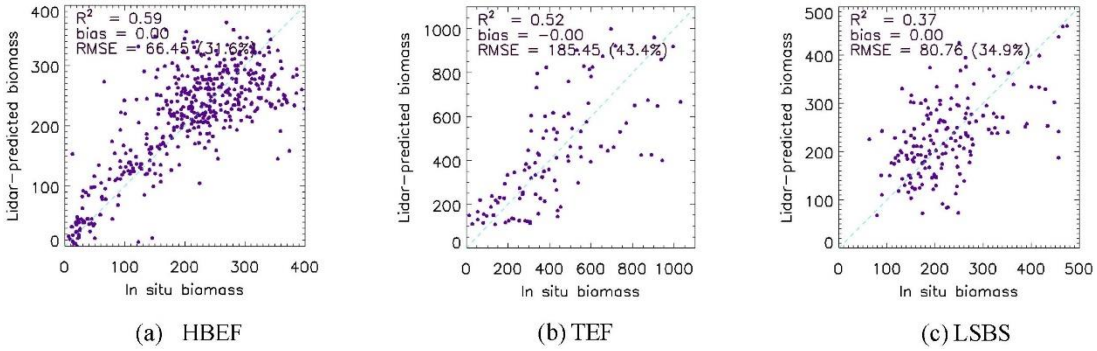
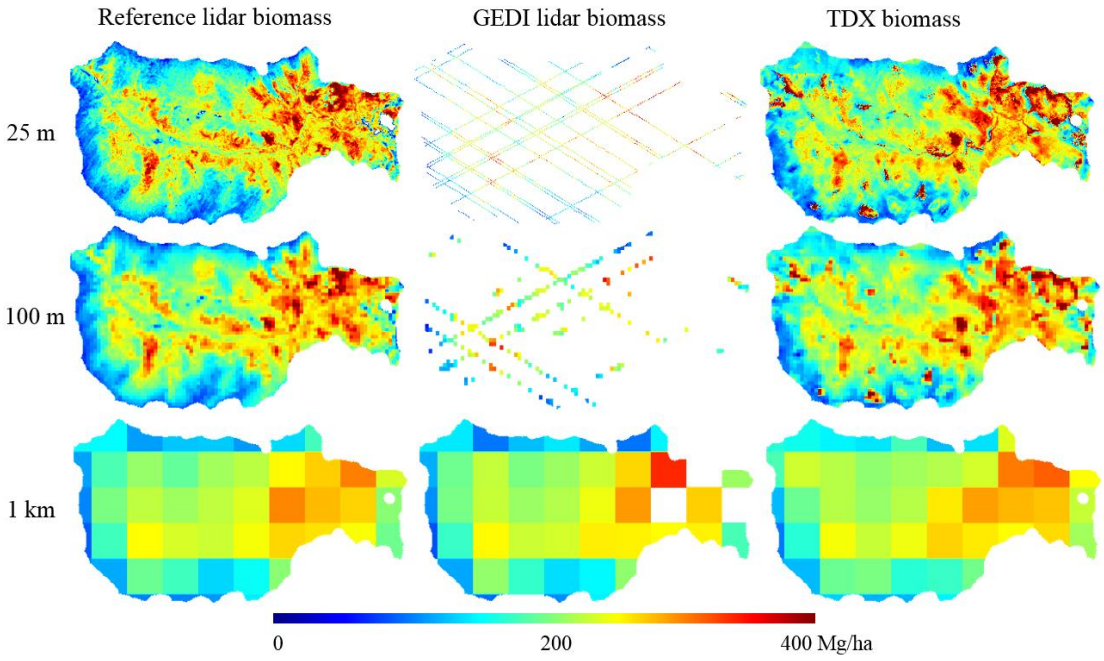


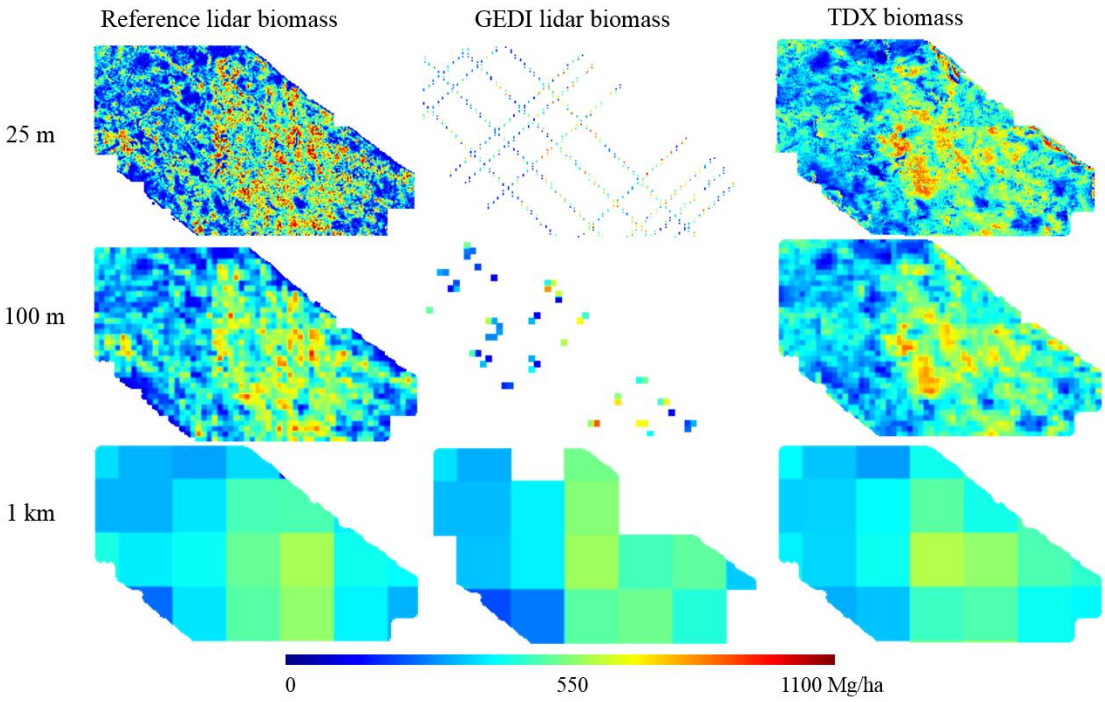
Figure 4-4. Lidar-predicted biomass vs. in situ biomass at the subplot level (25 m over HBEF and LSBS, or 33 m over TEF).

The established biomass–lidar models were then applied to predict reference-lidar biomass and GEDI-lidar biomass at 25 m using wall-to-wall- and GEDI track-height metrics derived from the simulated pseudo-waveforms. These estimates were

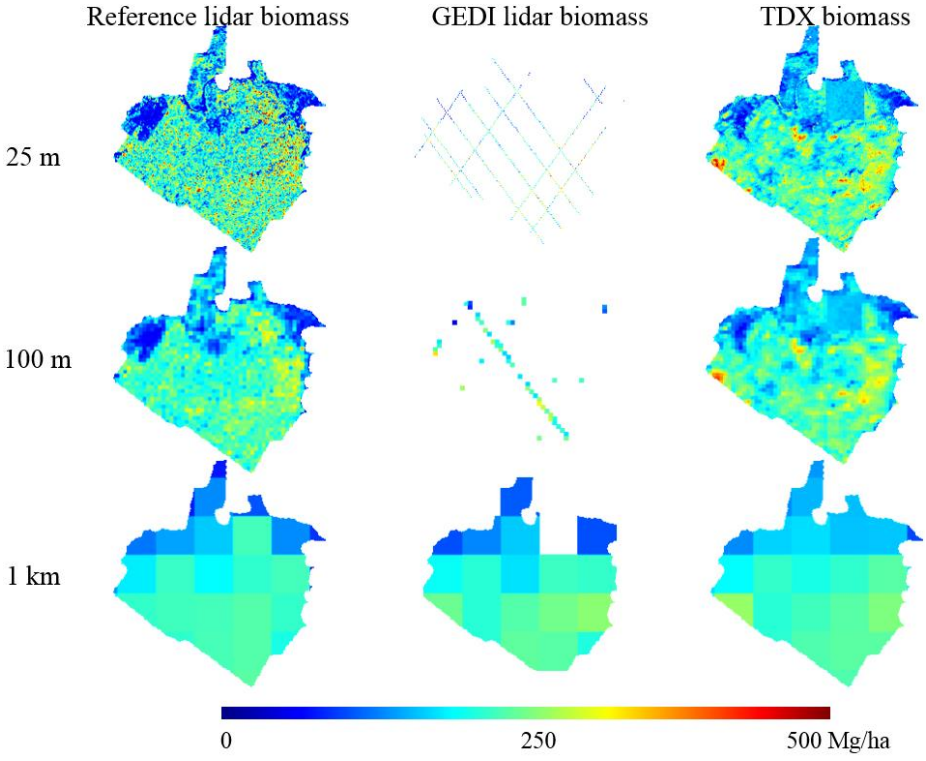
then averaged at 100 m and 1 km, and characterized with uncertainties for their means at the respective mapping scale. At least two independent tracks were required in each mapping grid to calculate the cross-track variance for estimating the uncertainties (Ståhl et al. 2010). Therefore, we removed grid cells that contained less than two tracks and derived variance of the biomass mean only in the remaining grid cells (Figure 4-5). Estimates over the invalid grid cells can be provided once the contiguous TDX forest heights are fused with GEDI, as illustrated in the next section.



(a) HBEF


Reference lidar biomass
GEDI lidar biomass
TDX biomass
25 m
100 m
1 km
0
550
1100 Mg/ha

(b) TEF


Reference lidar biomass
GEDI lidar biomass
TDX biomass
25 m
100 m
1 km
0
250
500 Mg/ha

(c) LSBS

Figure 4-5. Biomass maps predicted from the wall-to-wall- (left) and GEDI- (middle) lidar data, and TDX heights (right) respectively at 25 m, 100 m and 1 km over (a) HBEF, (b) TEF and (c) LSBS.

B. GEDI/TDX-fusion biomass

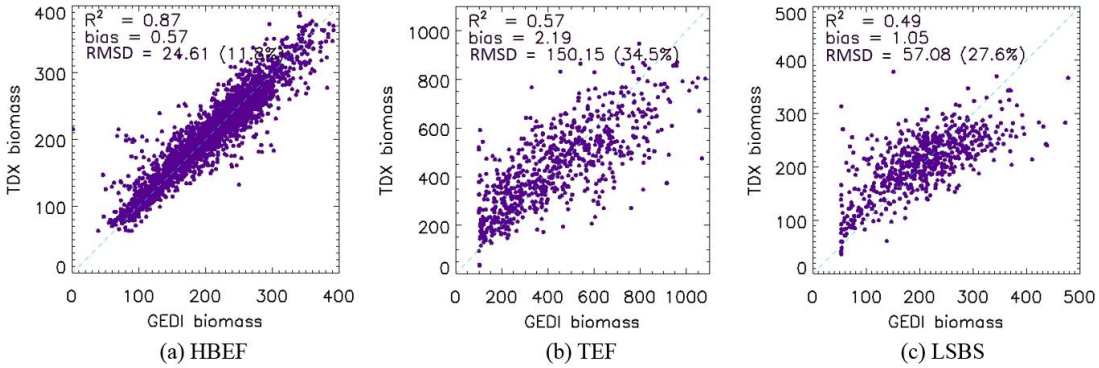


Figure 4-6. GEDI biomass vs. co-located TDX biomass at 25 m respectively over (a) HBEF, (b) TEF, and (c) LSBS.

Based on the method described in section 4.3, we extracted the 25 m GEDI biomass estimates within each 1 km grid (or enlarged areas; see 4.3.2) and related these estimates to the co-located TDX heights to develop grid-specific linear models. Each 1 km grid corresponded to a linear model to be used for the TDX biomass prediction at 25 m. Note that for each grid containing no GEDI tracks, we applied the TDX heights derived using only coherence to develop equation for biomass prediction. After applying TDX heights to the respective relationship model of each grid cell, we obtained r^2 values of 0.87, 0.57 and 0.49, and RMSDs of 24.61 (11.8%) Mg/ha, 150.15 (34.5%) Mg/ha and 57.08 (27.6%) Mg/ha at 25 m between the co-

located TDX biomass and GEDI biomass, respectively over HBEF, TEF, and LSBS (Figure 4-6).

The estimated TDX biomass maps were then averaged to 100 m and 1 km (Figure 4-5), with uncertainties described in 4.4.2. To increase the accuracy of the predicted biomass from the TDX height, pixels with coherence values lower than 0.3 and effective vertical wavenumbers (which vary as a function of terrain slopes and interferometric baseline) out of the range 0.04 – 0.35 (Kugler et al. 2015; Qi and Dubayah 2016) were all removed (less than 3% of the total pixels). Just to clarify, in this study TDX biomass also means GEDI/TDX-fusion biomass as it was derived by fusing GEDI and TDX data instead of using TDX data alone.

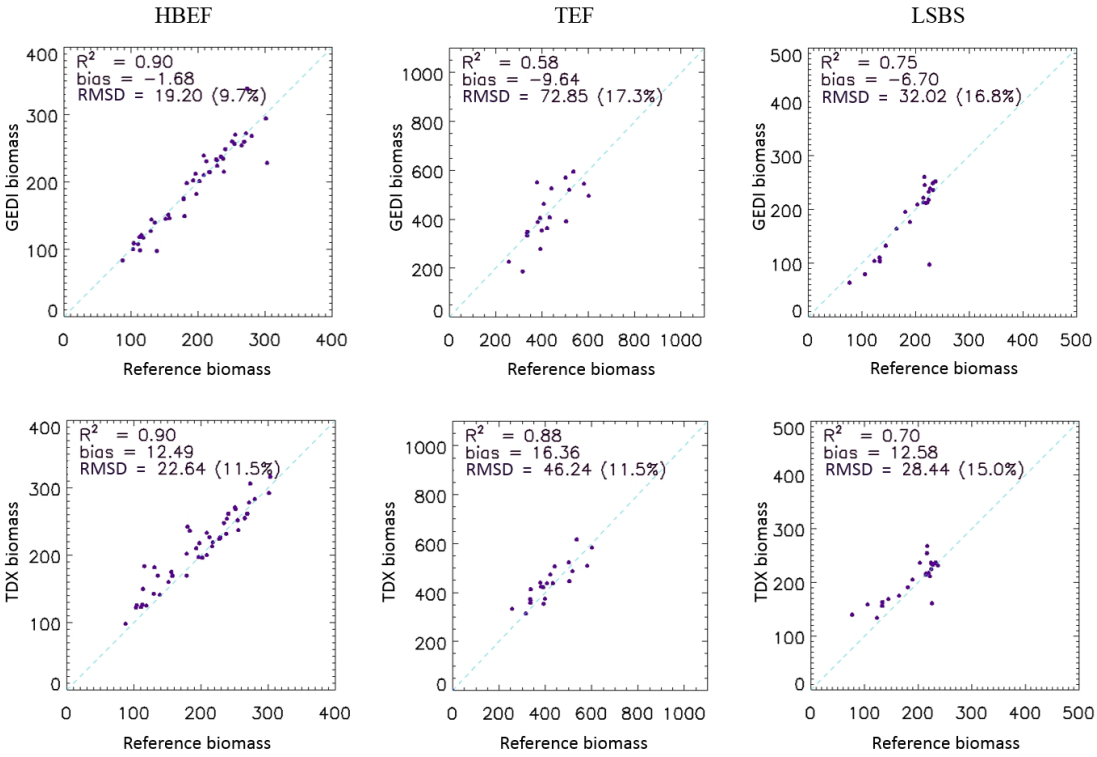


Figure 4-7. GEDI vs. Reference (wall-to-wall-lidar) biomass (top row), and TDX vs. Reference biomass (bottom row) at 1 km.

Using the wall-to-wall TDX heights, we derived contiguous biomass maps, including in places where no GEDI observations were acquired (Figure 4-5). At 1 km, the biases of the predicted TDX biomass against the reference-lidar biomass were estimated to be 12.5 Mg/ha, 16.4 Mg/ha and 12.6 Mg/ha respectively for HBEF, TEF and LSBS, which may be related to real biomass changes due to the temporal discrepancy of between the field/lidar data and TDX data. RMSDs between TDX and reference-lidar biomass were 22.6 (11.5%) Mg/ha, 46.2 (11.5%) Mg/ha and 28.4 (15.0%) Mg/ha respectively for the three sites, compared to RMSDs of 19.2 (9.7%) Mg/ha, 72.9 (17.3%) Mg/ha and 32.0 (16.8%) Mg/ha between GEDI-lidar and reference-lidar biomass (Figure 4-7). Since reference (wall-to-wall)-lidar biomass are not truth data, we subsequently performed uncertainty analysis for these mean estimates in a statistical manner at both 100 m and 1 km mapping scales, as discussed below.

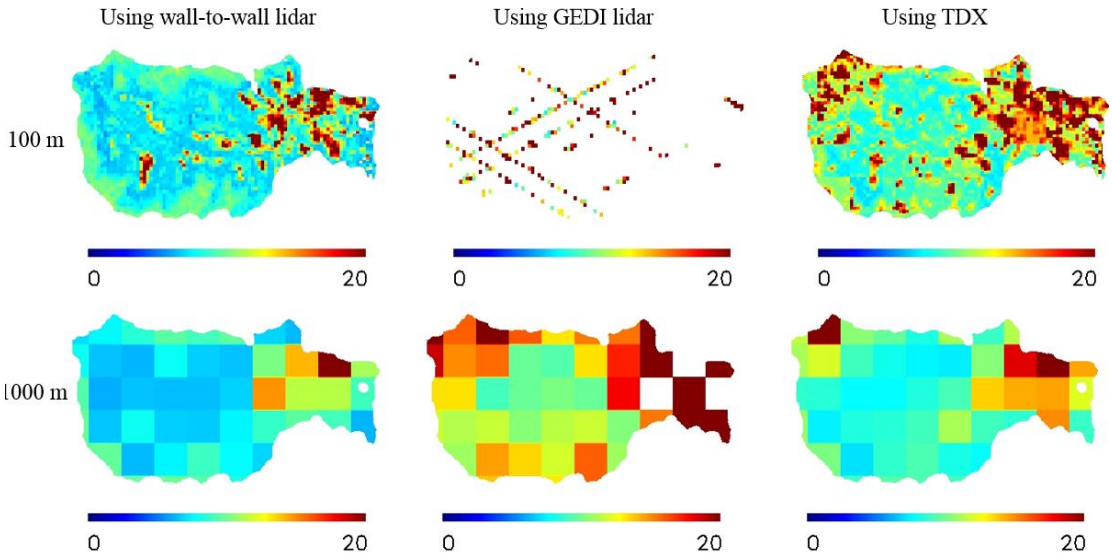
4.4.2 Uncertainty estimation

A. Wall-to-wall- and GEDI-lidar biomass uncertainty

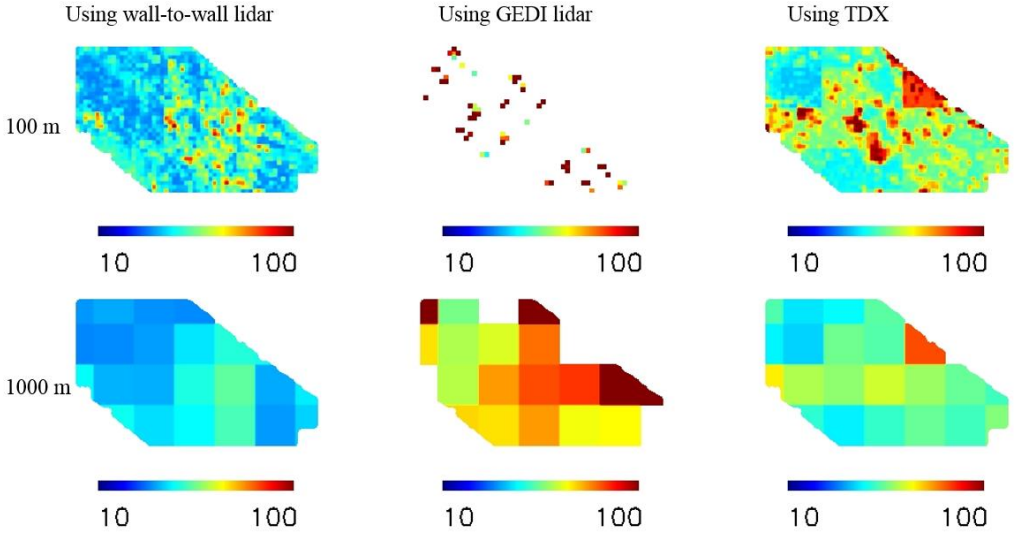
As discussed in section 4.3.3, we derived only V_{II} (variance associated with the modeling error) and ignored V_I (variance associated with the sampling error) for the wall-to-wall-(reference) lidar biomass, and estimated both V_I and V_{II} for the GEDI

sampling-based biomass. Uncertainties (Mg/ha) at a 95% confidence level were subsequently calculated from these variance estimates (see left two columns of Figure 4-8).

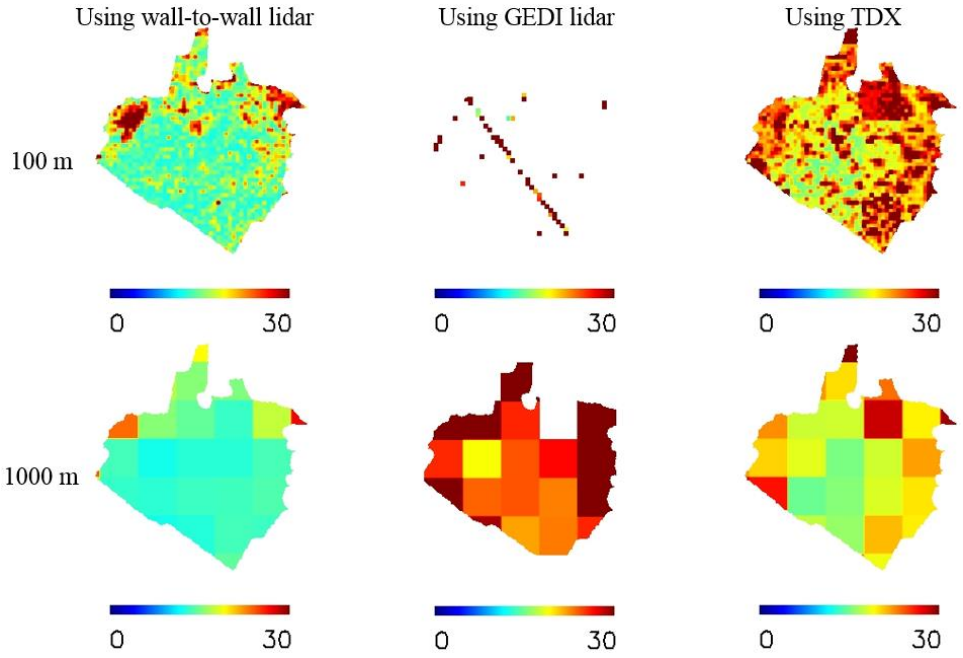
Among the three sites, uncertainties were the largest over TEF and the smallest over HBEF, perhaps related to the increased structural complexity observed by TDX data at TEF (LSBS may be more structurally complex than TEF but the complex understory of LSBS might be less seen by X-band signal due to the higher canopy cover at this site) and the opposite condition at HBEF. We observed reduced uncertainty levels for all sites from 100 m to 1 km. Note the large voids at 100 m resolution using GEDI lidar data (and a small proportion at 1 km) (Figure 4-8). Again, this arises when a cell has less than two tracks for cross-track variance calculation so that we cannot provide uncertainties over the cell (see 4.3.3).



(a) HBEF


Using wall-to-wall lidar
Using GEDI lidar
Using TDX
100 m
1000 m
10
100

(b) TEF


Using wall-to-wall lidar
Using GEDI lidar
Using TDX
100 m
1000 m
0
30

(c) LSBS

Figure 4-8. Uncertainties (Mg/ha) for biomass mean estimates at 100 m and 1 km scales respectively using the wall-to-wall- (left) and GEDI- (middle) lidar data, and TDX heights (right) over (a) HBEF, (b) TEF, and (c) LSBS.

B. GEDI/TDX-fusion biomass uncertainty

For all sites, uncertainties were reduced as resolution coarsened; for both mapping scales of 100 m and 1 km, we observed uncertainty levels of wall-to-wall-lidar biomass < GEDI/TDX-fusion biomass < GEDI-lidar biomass (Figure 4-8 and Figure 4-9). The results demonstrated that fusing GEDI data with wall-to-wall TDX data reduced uncertainties of GEDI biomass estimation at 1 km over the three forest sites. The fusion also showed the potential of GEDI/TDX fusion to provide contiguous biomass maps at a fine resolution of 100 m.

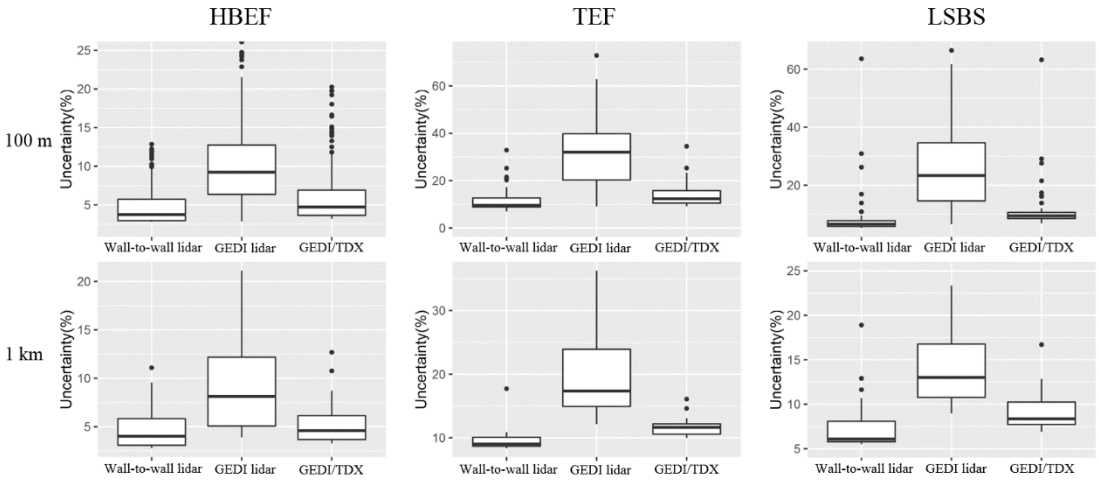


Figure 4-9. Uncertainty (%) of wall-to-wall- and GEDI-lidar biomass, and GEDI/TDX-fusion biomass at 100 m (top row) and 1 km (bottom row). Uncertainty

(%) was calculated as uncertainty (Mg/ha) divided by mean biomass (Mg/ha) of each grid cell. For consistency, uncertainties at 100 m were only compared over grid cells with valid GEDI estimates.

4.5 Discussion

With the launch of GEDI mission, more than 10 billion lidar observations of forest vertical structure and ground topography will be collected. These data will be used by themselves to provide biomass maps globally at a moderate resolution of 1 km using appropriate empirical models (Dubayah et al. 2014). However, since gaps are expected between adjacent GEDI ground tracks, it is of significant interest to utilize ancillary remote sensing datasets for fusing with GEDI so that biomass could be mapped contiguously at a finer resolution or with better accuracy at a moderate resolution especially for areas without GEDI coverage.

Previous studies have used passive optical and polarimetric radar observations for biomass prediction by correlating the spectral reflectance or radar backscatter signals to lidar-derived biomass, or directly to field-based data. However, passive optical and radar observations often fail to capture the whole spectrum of forest structure and saturate at the successional state (Mitchard et al. 2012; Saatchi et al. 2011). InSAR technology, on the other hand, has the capability to retrieve biomass-related structural properties, particularly height, to predict biomass beyond the saturation threshold of passive optical and traditional radar sensors, as demonstrated

in recent studies (Askne et al. 2017; Caicoya et al. 2016; Soja et al. 2014). Using forest height derived from TDX InSAR data, we have filled in gaps of GEDI biomass products at 100 m and at 1 km where GEDI observations were not available. Even when there were no gaps at 1 km, uncertainties were significantly reduced by using the wall-to-wall TDX data.

At 1 km, the mean uncertainties of wall-to-wall-lidar biomass products were estimated to be 4.8% at HBEF, 9.7% at TEF and 7.8% at LSBS, compared to uncertainties of 9.0% at HBEF, 19.9% at TEF and 16.1% at LSBS for biomass derived using GEDI sampled observations. The increased uncertainty was related to the sampling errors induced by using unrepresentative lidar samples from GEDI to estimate the mean biomass. Using the far larger sample size of GEDI shots compared to field plots, we were able to establish grid cell-specific GEDI biomass–TDX height relationships for predicting TDX biomass. Lower uncertainty levels of TDX-derived biomass were found compared to GEDI-based biomass, with uncertainties of 5.2% at HBEF, 11.7% at TEF and 9.3% at LSBS characterized at 1 km. At both 100 m and 1 km, estimated uncertainties were ranked as GEDI-lidar > TDX/GEDI fusion > wall-to-wall-lidar for all sites. These results demonstrated that the InSAR observations from TDX can be of great use to GEDI for providing contiguous biomass products at finer-resolutions and reducing the sampling-induced uncertainties.

We did not include the residual error (see 4.3.3) (Peters 2001) analysis in this study. Such error may be minor at 1 km but could be significant at smaller spatial

scales. Methods to estimate residual errors at smaller scales, such as 100 m, are currently under development. Furthermore, we did not consider uncertainties of the field biomass derived from the allometric equations and assumed the inventory data as truth. Ignoring this error should not affect the comparison of uncertainties characterized for our wall-to-wall-lidar, GEDI-lidar and GEDI/TDX fusion biomass products, but clearly will impact the absolute uncertainties.

Here, we developed our fusion algorithms over three contrasting forest types for improved forest biomass estimation. These methods used GEDI data with a variety of sampling density across each study site (i.e. the number of tracks varied widely from 0 to 7 for different 1 km grid cells) and were implemented with consideration of larger-scale applications. Areas with sufficient GEDI coverage can be provided with biomass estimates from GEDI alone and improved through fusion with TDX data; whereas for areas with few GEDI observations, biomass can be predicted from TDX height derived using coherence alone (Qi and Dubayah 2016; Qi et al. Submitted) based on GEDI biomass–TDX height relationships developed over places with similar forest types and conditions.

4.6 Conclusions

Our study used wall-to-wall forest heights estimated from TerraSAR-X/TanDEM-X (TDX) acquisitions to improve biomass derived from Global Ecosystem Dynamics Investigation (GEDI) lidar observations. It was demonstrated that TDX heights were useful to fill gaps of GEDI biomass products at 100 m and at 1

km where GEDI observations were not available. Better accuracy was also achieved through fusion with TDX heights when there were GEDI observations, with uncertainties of 5.2-11.7% from the fusion across the three sites, as compared to 9.0-19.9% achieved from the use of GEDI data alone. There is a great potential to fuse GEDI and TDX data to generate high-quality forest biomass maps over large spatial scales.

4.7 Acknowledgements

This work was funded by NASA's Earth and Space Science Fellowship Program (Grant # 5268930) and NASA's Global Ecosystem Dynamics Investigation mission. We would like to thank the researchers conducting the field work and biomass calculations. The used TerraSAR-X/TanDEM-X data were provided and pre-processed by Microwaves and Radar Institute of the German Aerospace Center (DLR). We are also grateful to Dr. Chengquan Huang, Dr. Min Feng and Dr. Feng Zhao for providing the forest disturbance data product.

Chapter 5: Conclusions and Future Research

5.1 *Summary of the Research*

Large-scale mapping of forest biomass at fine spatial resolutions (100–1000 m) is challenging, yet it is a critical step in improving quantifications of the terrestrial carbon budget and reducing the large uncertainties in predictions of future climate. One major obstacle to achieving this goal is the lack of a sufficiently accurate and spatially comprehensive global dataset of forest vertical structure. Important estimates of forest structure and biomass have been carried out using lidar remote sensing technology; however, there are limitations (e.g. restricted spatial coverage) in mapping biomass using lidar data alone. The launch of NASA's GEDI mission in 2018 will provide the most complete set of global-scale lidar observations of forest structure and ground elevation to date. Nevertheless, the nature of lidar sampling and possible impacts from cloud cover indicate that there will be gaps between adjacent ground tracks, which will limit GEDI's capability to provide biomass data at the desired resolution and accuracy. This study is a first effort at investigating the potential of fusing GEDI lidar data with ancillary data acquired by the TDX satellite mission to provide forest structure and biomass maps at the needed resolution and accuracy.

The research mainly focused on improving estimates of forest height and biomass by fusing GEDI and TDX observations. As the first long-term single-pass

InSAR mission in space, TDX generates coherence images for height estimations without any temporal decorrelation effects. The data also possess a high spatial resolution, complete global coverage and multiple spatial baselines. The results of this study demonstrated that GEDI and TDX observations are complementary, and that fusing them can potentially provide large-scale forest height and biomass data at a finer resolution and with higher accuracy than is currently possible. Adding GEDI lidar data to TDX InSAR acquisitions substantially reduced biases and errors of height estimation from the latter, whereas the wall-to-wall TDX heights have helped to significantly reduce uncertainties in GEDI's sampling-based biomass estimates. The fusion algorithms developed in this study are innovative and could potentially be applied to provide global biomass products at resolutions of 100–1000 m.

5.1.1 Evaluating the effectiveness of GEDI elevation data to improve forest height estimations from TDX

The investigation in Chapter 2 focused on combining the simulated GEDI elevation data and TDX data to improve height estimation under different foliage conditions and achieved the following results. The responsiveness of TDX coherence to the variation of lidar-derived canopy height was evaluated; this was important as their relationship formed the basis for the GEDI/TDX fusion – if there was no or little correlation between them, it would be unlikely that heights can be derived from TDX with an acceptable accuracy for an effective fusion. The chosen study area, Hubbard Brook Experimental Forest, is a temperate mixed deciduous and conifer forest with a

large gradient of forest height, canopy density and topography. Results demonstrated that there were consistently strong and robust correlations between lidar-derived forest height and TDX coherence at varied TDX spatial baselines and environment (leaf on vs. leaf off) conditions, providing meaningful evidence for the efficacy of fusing GEDI/TDX data on a large scale.

This study also verified the utility of using GEDI elevation data in TDX height inversions. Previous studies using single-pol TDX InSAR acquisitions often required external DTMs derived from airborne lidar elevation data to estimate height, because otherwise height accuracies would be compromised by using assumptions (e.g. a fixed extinction coefficient (σ) and a zero ground-to-volume amplitude ratio (μ=0)) for model parameters. The results of this study indicated that using GEDI elevation data effectively improved TDX height estimation compared to results using TDX data alone. There were good correlations between TDX-derived height and lidar canopy height (RH100) as well as other lidar height metrics commonly used for biomass estimation (e.g. RH75, RH50, and RH25). This demonstrated the great potential of using TDX height, particularly when derived with aid from GEDI lidar observations of bare ground elevation, to map large-scale forest biomass.

5.1.2 Improving the parameterization of TDX inversion model to enhance forest height estimation using GEDI canopy height and elevation data

Chapter 3, with its focus on fusing GEDI/TDX data for height estimation, was somewhat similar to Chapter 2. However, Chapter 3 also investigated the utility of

GEDI-derived canopy height, in addition to the DTM, to enhance RVoG model parameterizations for improved height estimations.

Previous studies had mainly explored external DTMs derived from airborne lidar observations for TDX height inversion. As mentioned in Chapter 2, the input of external DTMs allowed the RVoG to use a spatially varied extinction coefficient (σ) for estimating height; however, the ground-to-volume-amplitude ratio (μ) was still assumed zero in the inversion process. Exploring full-polarimetric InSAR data at longer-wavelength (L- or P-band) would allow the derivation of both σ and μ for better height estimation, but these data were still restricted to areas where airborne (in many cases multiple-baseline) PolInSAR campaigns were conducted. This study developed an approach to incorporate GEDI lidar observations into the RVoG model to estimate both parameters. Results demonstrated performance gain from calculating μ for height estimations in certain environmental conditions (for instance, over leaf-off forested areas).

Chapter 3 also evaluated the utility of each refined RVoG parameter over different forest conditions. Three major representative forest types – mixed temperate deciduous and conifer, mountainous conifer, and tropical rainforest – were investigated. For all places, the estimated heights improved from using no GEDI data (only TDX) to using GEDI elevation, and further to using GEDI elevation and canopy height combined. Including both GEDI elevation and canopy height as RVoG inputs allowed the derivation of both σ and μ parameters; μ was found to be less useful over

areas with high canopy covers and more useful over areas where more ground scattering was present.

Furthermore, this study assessed the impact of cloud cover on the effectiveness of GEDI/TDX fusion for height estimation. Results derived under the cloud-free condition (i.e. with no GEDI tracks lost due to clouds) were compared with those derived under the 50% cloud cover condition (i.e. with 50% GEDI observations lost due to clouds). For each fusion method, forest heights derived using the simulated GEDI data under cloud-cover condition were less accurate because of the sparser data available for local calibrations performed on RVoG parameters. Nevertheless, GEDI/TDX data fusion was effective under both conditions to improve forest height, particularly when both simulated GEDI canopy height and elevation data were used.

5.1.3 Assessing the efficacy of TDX and GEDI data fusion to improve estimations of forest biomass

While Chapters 2 and 3 focused on enhancing forest height estimation from TDX using GEDI lidar data as inputs to the InSAR inversion model, Chapter 4 focused on improving GEDI-based biomass estimations through the use of wall-to-wall TDX height maps. The following major issues were addressed in this study.

Simulated GEDI data were used to demonstrate the creation of biomass products at a map scale of 1 km over the three representative forest types. My research also showed that we can create biomass products at finer resolutions from

GEDI/TDX fusion. TDX has the potential to provide the information for GEDI with its global coverage, high spatial resolution, and simultaneous interferometric acquisitions (i.e. without temporal decorrelation effect). This study provided detailed procedures on how to fuse the complementary GEDI and TDX data to improve height estimations and use them to upscale GEDI-derived biomass data. Contiguous biomass maps were produced not only at a moderate resolution of 1 km, but also at a finer resolution of 100 m. At 1 km, the fused biomass values were found to have equivalent (at HBEF and LSBS) or better (at TEF) correlations with reference (i.e. wall-to-wall)-lidar biomass, when compared to correlations between GEDI- and reference-lidar biomass at this scale, further manifesting the advantages of the GEDI/TDX fusion.

This study also implemented new methods to characterize the uncertainties of the GEDI-based and GEDI/TDX fusion-based biomass products at varied mapping scales. Estimating uncertainty has been difficult for mapping forest biomass from remote sensing data, especially at scales (e.g. 1 km) not sufficiently covered by field inventories. Uncertainties associated with GEDI biomass at its native footprints could be easily estimated by comparing subplot field biomass and co-located GEDI-derived biomass. However, when GEDI biomass is aggregated to larger grids or related to TDX observations to provide spatially contiguous biomass, there were questions regarding the propagation of errors from the native footprint to larger grids as well as from GEDI biomass to TDX-derived biomass. Based on previous work on estimating variances for hybrid and hierarchical-modeling inferences, this study addressed

uncertainties for both the GEDI-based (by considering both the modeling and sampling errors) and GEDI/TDX fusion-based (by properly propagating the modeling errors from the biomass–lidar relationship to the GEDI biomass–TDX height relationship) products. The fusion products had lower uncertainties at both 100 m and 1 km resolutions, showcasing the effectiveness of fusing the data to map forest biomass at better resolutions and accuracies.

This research has added insights into approaches to fuse lidar and InSAR data, and has also demonstrated the potential of fusing GEDI data with complementary satellite images such as those from TDX in order to produce large-scale biomass maps at resolutions and accuracies desired to reduce high uncertainties regarding the global carbon budget.

5.2 Future Research

The findings in this study indicate several avenues for future research. The first is to explore approaches to improve the accuracy of the GEDI-derived DTM for fusion, as results shown in Chapters 2 and 3 suggest an impact of the DTM on the fusion efficacy. Factors contributing to a less-accurate GEDI DTM may include a lower lidar shot density (e.g. due to higher cloud covers or fewer orbital passes over a certain site), less accurate ground detection (e.g. because of signal saturation over areas with higher canopy cover), or geolocation errors over topographically complex terrain. Based on a less accurate GEDI DTM, a less accurate ground phase would be

simulated to estimate the scattering phase and retrieve less accurate forest height from the fusion. To address these issues, future work will focus on the generation of more-accurate DTMs by integrating bare-ground GEDI elevation data with surface DEMs derived from, for example, TDX (12 m), SRTM (30 m) and the Advanced Spaceborne Thermal Emission and Reflection Radiometer (30 m), or from future missions such as ICESat-2 and NISAR, with an aim to produce reliable ground DTMs at a fine resolution over sloped surfaces and other areas with less GEDI coverage. Preliminary studies are being conducted on merging GEDI DTMs with TDX DEMs using wavelet transformations (Lee et al. 2018), demonstrating the effectiveness of developing fusion algorithms to produce a bare-ground DTM at better resolution and accuracy than from either mission alone.

Results in Chapter 3 also suggested the importance of developing alternate methods to provide improved maps of RVoG parameters (e.g. extinction coefficient σ and ground-to-volume amplitude ratio μ). In this study, RVoG parameters were derived at native simulated GEDI footprints and interpolated into spatial maps for height estimation. These maps were used to greatly improve height estimations compared to those derived from TDX coherence alone or a joint use with GEDI-derived DTM. However, such results were based on the use of GEDI observations with sufficient coverage – the estimated RVoG parameters were representative enough of the structural characteristics across the three forest sites, as indicated by the fusion results. As noted, GEDI lidar density would vary from site to site due to the

influence of cloud cover, variable orbital passes, and topography variations. This requires the development of approaches that reliably estimate RVoG parameters, which is particularly necessary for areas with less GEDI coverage, and greater structural and topographical complexity. Future studies may explore alternate ways that input RVoG parameters derived from areas with similar expected parameter values. This might be done by using segmentation and clustering algorithms to group segments from similar forest types and environmental conditions based on ancillary maps of, for example, TDX coherence or Landsat canopy cover (Clewley et al. 2014).

Furthermore, the waveforms (or pseudo-waveforms) from lidar were not fully exploited and related to TDX observations. This was mainly because X-band radar signals often have limited penetration capabilities over dense forests and does not represent the full range of forest vertical structure, particularly over the understories. This is the primary reason why the underlying topography from GEDI is essential for TDX InSAR studies – for more reliable ground detection and height inversion. To overcome this limitation, DLR has planned another spaceborne full-polarimetric tandem SAR mission–TanDEM-L–for launch in 2023. This mission will work at L-band and have a better canopy penetration capability, which will allow the retrieval of more detailed structural properties and the establishment of better relationships with GEDI-derived profiles. It would be of great significance to explore the fusion of GEDI and TanDEM-L datasets, when those become available, for improved biomass estimations.

A key element of mapping biomass with GEDI/TDX fusion is the accuracy of biomass estimation at the native GEDI footprints, a step depending on the quality of the empirical biomass–lidar relationship models (Duncanson et al. 2015b). Here multiple linear regression equations were established for simplicity and efficiency as it was not the primary focus of this study. It is possible to further reduce the total biomass estimation uncertainty by exploring the establishment of non-linear or other complex relationships based on machine learning algorithms. However, the propagation of modeling errors based on these complex relationships remains a challenging issue to be addressed. Attempts in this study formed only a prototype for characterizing this process at varied mapping scales under a rigorous statistical framework.

To conclude, the forthcoming GEDI mission–among other spaceborne Earth-observation lidar missions–can be a game changer for large scale biomass estimation and terrestrial carbon stock quantification in the next few years. This study is among the pioneering efforts to develop an algorithm for generating contiguous biomass maps from GEDI and TDX missions. Its efficacy has been largely demonstrated over representative biomes, and are expected to be globally applicable given sufficient GEDI coverage and availability of appropriate TDX acquisitions. Among many other similar studies, this research presents a trend towards a new era of active remote sensing, where data from missions as diverse as the TDX, GEDI, DLR's TanDEM-L, NASA's ICESat-2 and NISAR, and ESA's BIOMASS, among others, are used jointly

to deepen our understanding of ecosystem structures and improve the quantification of terrestrial carbon stocks and fluxes.

Bibliography

Abdalati, W., Zwally, H.J., Bindschadler, R., Csatho, B., Farrell, S.L., Fricker, H.A., Harding, D., Kwok, R., Lefsky, M., Markus, T., Marshak, A., Neumann, T., Palm, S., Schutz, B., Smith, B., Spinhirne, J., & Webb, C. (2010). The ICESat-2 laser altimetry mission. *Proceedings of the IEEE, 98*, 735-751

Abdullahi, S., Kugler, F., & Pretzsch, H. (2016). Prediction of stem volume in complex temperate forest stands using TanDEM-X SAR data. *Remote Sensing of Environment, 174*, 197-211

Abrams, M., Bailey, B., Tsu, H., & Hato, M. (2010). The aster global dem. *Photogrammetric Engineering and Remote Sensing, 76*, 344-348

Adam, E., Mutanga, O., & Rugege, D. (2010). Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management, 18*, 281-296

Agrawal, A., Nepstad, D., & Chhatre, A. (2011). Reducing emissions from deforestation and forest degradation. *Annual Review of Environment and Resources, 36*, 373-396

Anderson, J.E., Martin, M.E., Smith, M.L., Dubayah, R.O., Hofton, M.A., Hyde, P., Peterson, B.E., Blair, J.B., & Knox, R.G. (2006). The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. *Remote Sensing of Environment, 105*, 248-261

Anderson, J.E., Plourde, L.C., Martin, M.E., Braswell, B.H., Smith, M.L., Dubayah, R.O., Hofton, M.A., & Blair, J.B. (2008). Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sensing of Environment, 112*, 1856-1870

Askne, J.I.H., Fransson, J.E.S., Santoro, M., Soja, M.J., & Ulander, L.M.H. (2013). Model-based biomass estimation of a hemi-boreal forest from multitemporal TanDEM-X acquisitions. *Remote Sensing, 5*, 5574-5597

Askne, J.I.H., & Santoro, M. (2005). Multitemporal repeat pass SAR interferometry of boreal forests. *IEEE Transactions on Geoscience and Remote Sensing, 43*, 1219-1228

Askne, J.I.H., Soja, M.J., & Ulander, L.M.H. (2017). Biomass estimation in a boreal forest from TanDEM-X data, lidar DTM, and the interferometric water cloud model. *Remote Sensing of Environment, 196,* 265-278

Benson, M.L., Pierce, L.E., Bergen, K.M., Sarabandi, K., Zhang, K., & Ryan, C.E. (2010). Extrapolation of LiDAR for forest structure estimation using SAR, InSAR, and optical data. In, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1633-1636)

Bergen, K.M., Goetz, S.J., Dubayah, R.O., Henebry, G.M., Hunsaker, C.T., Imhoff, M.L., Nelson, R.F., Parker, G.G., & Radeloff, V.C. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions. *Journal of Geophysical Research: Biogeosciences, 114*, 1-13

Blackard, J.A., Finco, M.V., Helmer, E.H., Holden, G.R., Hoppus, M.L., Jacobs, D.M., Lister, A.J., Moisen, G.G., Nelson, M.D., Riemann, R., Ruefenacht, B., Salajanu, D., Weyermann, D.L., Winterberger, K.C., Brandeis, T.J., Czaplewski, R.L., McRoberts, R.E., Patterson, P.L., & Tymcio, R.P. (2008). Mapping US forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sensing of Environment, 112,* 1658-1677

Blair, J.B., & Hofton, M.A. (1999). Modeling laser altimeter return waveforms over complex vegetation using high-resolution elevation data. *Geophysical Research Letters, 26*, 2509-2512

Blair, J.B., Rabine, D.L., & Hofton, M.A. (1999). The Laser Vegetation Imaging Sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. *ISPRS Journal of Photogrammetry and Remote Sensing, 54*, 115-122

Brolly, M., Simard, M., Tang, H., Dubayah, R.O., & Fisk, J.P. (2016). A Lidar-Radar Framework to Assess the Impact of Vertical Forest Structure on Interferometric Coherence. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9,* 5830-5841

Brown, C.G., & Sarabandi, K. (2001). Model-based estimation of forest canopy parameters using polarimetric and interferometric SAR. In, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 357-359)

Bräutigam, B., Bachmann, M., Schulze, D., Tridon, D.B., Rizzoli, P., Martone, M., Gonzalez, C., Zink, M., & Krieger, G. (2014). TanDEM-X global DEM

quality status and acquisition completion. In, *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International* (pp. 3390-3393): IEEE

Caicoya, A.T., Kugler, F., Hajnsek, I., & Papathanassiou, K. (2012). Boreal forest biomass classification with TanDEM-X. In, *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International* (pp. 3439-3442): IEEE

Caicoya, A.T., Kugler, F., Hajnsek, I., & Papathanassiou, K.P. (2016). Large-scale biomass classification in boreal forests with TanDEM-X data. *IEEE Transactions on Geoscience and Remote Sensing, 54*, 5935-5951

Campbell, J.L., & Bailey, A.S. (2014). Hubbard Brook Experimental Forest (USDA Forest Service): Daily Mean Temperature Data, 1955 - present. In: Hubbard Brook Data Archive

Canadell, J.G., Le Quere, C., Raupach, M.R., Field, C.B., Buitenhuis, E.T., Ciais, P., Conway, T.J., Gillett, N.P., Houghton, R.A., & Marland, G. (2007). Contributions to accelerating atmospheric CO2 growth from economic activity, carbon intensity, and efficiency of natural sinks. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 18866-18870

Carver, K.R. (1988). SAR Synthetic Aperture RADAR–Earth Observing System, Vol. In. Washington DC: IIf, NASA Instrument Panel Report

CEOS (2014). *CEOS Strategy for Carbon Observations from Space. The Committee on Earth Observation Satellites (CEOS) Response to the Group on Earth Observations (GEO) Carbon Strategy.*: JAXA and I&A Corporation

Chen, W., Blain, D., Li, J., Keohler, K., Fraser, R., Zhang, Y., Leblanc, S., Olthof, I., Wang, J., & McGovern, M. (2009). Biomass measurements and relationships with Landsat-7/ETM+ and JERS-1/SAR data over Canada's western sub-arctic and low arctic. *International Journal of Remote Sensing, 30*, 2355-2376

Chowdhury, T.A., Thiel, C., & Schmullius, C. (2014). Growing stock volume estimation from L-band ALOS PALSAR polarimetric coherence in Siberian forest. *Remote Sensing of Environment, 155*, 129-144

Clark, D.B., & Clark, D.A. (2000). Landscape-scale variation in forest structure and biomass in a tropical rain forest. *Forest Ecology and Management, 137*, 185-198

Clark, D.B., & Kellner, J.R. (2012). Tropical forest biomass estimation and the fallacy of misplaced concreteness. *Journal of Vegetation Science, 23*, 1191-1196

Clark, D.B., Olivas, P.C., Oberbauer, S.F., Clark, D.A., & Ryan, M.G. (2008). First direct landscape-scale measurement of tropical rain forest Leaf Area Index, a key driver of global primary productivity. *Ecology Letters, 11,* 163-172

Clark, M.L., Roberts, D.A., Ewel, J.J., & Clark, D.B. (2011). Estimation of tropical rain forest aboveground biomass with small-footprint lidar and hyperspectral sensors. *Remote Sensing of Environment, 115,* 2931-2942

Clewley, D., Bunting, P., Shepherd, J., Gillingham, S., Flood, N., Dymond, J., Lucas, R., Armston, J., & Moghaddam, M. (2014). A Python-Based Open Source System for Geographic Object-Based Image Analysis (GEOBIA) Utilizing Raster Attribute Tables. *Remote Sensing, 6,* 6111

Cloude, S.R., Chen, H., & Goodenough, D.G. (2013). Forest height estimation and validation using Tandem-X polinsar. In, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1889-1892)

Cloude, S.R., & Papathanassiou, K.P. (1998). Polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing, 36,* 1551-1565

Cloude, S.R., & Papathanassiou, K.P. (2003). Three-stage inversion process for polarimetric SAR interferometry. *IEE Proceedings-Radar, Sonar and Navigation, 150,* 125-134

Coyle, D.B., Stysley, P.R., Poulios, D., Clarke, G.B., & Kay, R.B. (2015). Laser transmitter development for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar. In, *SPIE Optical Engineering+ Applications* (pp. 961208-961208): International Society for Optics and Photonics

Davidson, F.M., & Sun, X.L. (1988). Gaussian approximation versus nearly exact performance analysis of optical communication systems with PPM signaling and APD receivers. *IEEE Transactions on Communications, 36,* 1185-1192

Davidson, R., & MacKinnon, J.G. (1993). *Estimation and inference in econometrics*. *OUP Catalogue*

Dolan, K.A., Hurtt, G.C., Chambers, J.Q., Dubayah, R.O., Frolking, S., & Masek, J.G. (2011). Using ICESat's Geoscience Laser Altimeter System (GLAS) to assess large-scale forest disturbance caused by hurricane Katrina. *Remote Sensing of Environment, 115,* 86-96

Dolan, K.A., Masek, J.G., Huang, C., & Sun, G. (2009). Regional forest growth rates measured by combining ICESat GLAS and Landsat data. *Journal of Geophysical Research: Biogeosciences, 114*, 7

Drake, J.B., Dubayah, R.O., Clark, D.B., Knox, R.G., Blair, J.B., Hofton, M.A., Chazdon, R.L., Weishampel, J.F., & Prince, S.D. (2002). Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment, 79*, 305-319

Drake, J.B., Knox, R.G., Dubayah, R.O., Clark, D.B., Condit, R., Blair, J.B., & Hofton, M. (2003). Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. *Global Ecology and Biogeography, 12*, 147-159

Dubayah, R.O., & Drake, J.B. (2000). Lidar remote sensing for forestry. *Journal of Forestry, 98*, 44-46

Dubayah, R.O., Goetz, S.J., Blair, J.B., Fatoyinbo, T.E., Hansen, M., Healey, S.P., Hofton, M.A., Hurtt, G.C., Kellner, J., Luthcke, S.B., & Swatantran, A. (2014). The global ecosystem dynamics investigation. In, *AGU Fall Meeting Abstracts*

Dubayah, R.O., Sheldon, S.L., Clark, D.B., Hofton, M.A., Blair, J.B., Hurtt, G.C., & Chazdon, R.L. (2010). Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. *Journal of Geophysical Research: Biogeosciences, 115*, 17

Duncanson, L.I., Dubayah, R.O., Cook, B.D., Rosette, J., & Parker, G. (2015a). The importance of spatial detail: Assessing the utility of individual crown information and scaling approaches for lidar-based biomass density estimation. *Remote Sensing of Environment, 168*, 102-112

Duncanson, L.I., Dubayah, R.O., & Enquist, B.J. (2015b). Assessing the general patterns of forest structure: quantifying tree and forest allometric scaling relationships in the United States. *Global Ecology and Biogeography, 24*, 1465-1475

Englhart, S., Keuck, V., & Siegert, F. (2011). Aboveground biomass retrieval in tropical forests - The potential of combined X- and L-band SAR data use. *Remote Sensing of Environment, 115*, 1260-1271

Fassnacht, F.E., Hartig, F., Latifi, H., Berger, C., Hernandez, J., Corvalan, P., & Koch, B. (2014). Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment, 154*, 102-114

Fatoyinbo, T.E., & Simard, M. (2013). Height and biomass of mangroves in Africa from ICESat/GLAS and SRTM. *International Journal of Remote Sensing, 34*, 668-681

Files, C. (2012). NACP aboveground biomass and carbon baseline data, V. 2 (NBCD 2000), USA, 2000.

Freeman, A., Rosen, P., Jordan, R., Johnson, W., Hensley, S., Sweetser, T., Loverro, A., Smith, J., Sprague, G., & Shen, Y. (2009). DESDynI–A NASA Mission for Ecosystems, Solid Earth, and Cryosphere Science. In

Frolking, S., Palace, M.W., Clark, D.B., Chambers, J.Q., Shugart, H.H., & Hurtt, G.C. (2009). Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure. *Journal of Geophysical Research: Biogeosciences, 114*, 27

Garestier, F., Dubois-Fernandez, P.C., & Papathanassiou, K.P. (2008). Pine forest height inversion using single-pass X-band PolInSAR data. *IEEE Transactions on Geoscience and Remote Sensing, 46*, 59-68

Giannico, V., Lafortezza, R., John, R., Sanesi, G., Pesola, L., & Chen, J.Q. (2016). Estimating stand volume and above-ground biomass of urban forests using LiDAR. *Remote Sensing, 8*, 339-352

Goetz, S.J., & Dubayah, R.O. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Management, 2*, 231-244

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*, 481-493

Hall, F.G., Bergen, K., Blair, J.B., Dubayah, R.O., Houghton, R., Hurtt, G., Kellndorfer, J., Lefsky, M., Ranson, J., Saatchi, S., Shugart, H.H., & Wickland, D. (2011). Characterizing 3D vegetation structure from space: Mission requirements. *Remote Sensing of Environment, 115*, 2753-2775

Hancock, S., Anderson, K., Disney, M., & Gaston, K.J. (2017). Measurement of fine-spatial-resolution 3D vegetation structure with airborne waveform lidar: Calibration and validation with voxelised terrestrial lidar. *Remote Sensing of Environment, 188*, 37-50

Hancock, S., Disney, M., Muller, J.P., Lewis, P., & Foster, M. (2011). A threshold insensitive method for locating the forest canopy top with waveform lidar. *Remote Sensing of Environment, 115*, 3286-3297

Hancock, S., Tang, H., Blair, J.B., Dubayah, R., Armston, J., Sun, X., Hofton, M.A., Duncanson, L., Kellner, J.R., & Marselis, S. (2016). Assessing the performance of the GEDI spaceborne lidar instrument using a waveform lidar simulator. In, *American Geophysical Union*. San Francisco

Hayashi, M., Saigusa, N., Borjigin, H., Sawada, Y., Yamagata, Y., Hirano, T., & Ichii, K. (2015a). Estimation of regional-scale forest resources using ICESat/GLAS spaceborne LiDAR. In, *SPIE Remote Sensing* (pp. 96440V-96440V): International Society for Optics and Photonics

Hayashi, M., Saigusa, N., Oguma, H., Yamagata, Y., & Takao, G. (2015b). Quantitative assessment of the impact of typhoon disturbance on a Japanese forest using satellite laser altimetry. *Remote Sensing of Environment, 156*, 216-225

Healey, S., & Patterson, P. (2018). Algorithm Theoretical Basis Document (ATBD) for GEDI Level 4B Gridded Aboveground Biomass Product. In

Hoffman, J.P., Shaffer, S., & Perkovic-Martin, D. (2016). NASA L-SAR instrument for the NISAR (NASA-ISRO) Synthetic Aperture Radar mission. In, *Earth Observing Missions and Sensors: Development, Implementation, and Characterization IV* (p. 988108): International Society for Optics and Photonics

Houghton, R.A., & Goetz, S.J. (2008). New satellites help quantify carbon sources and sinks. *Eos, Transactions American Geophysical Union, 89*, 417-418

Houghton, R.A., Hall, F., & Goetz, S.J. (2009). Importance of biomass in the global carbon cycle. *Journal of Geophysical Research, 114*, 2156-2202

Houghton, R.A., House, J.I., Pongratz, J., van der Werf, G.R., DeFries, R.S., Hansen, M.C., Le Quere, C., & Ramankutty, N. (2012). Carbon emissions from land use and land-cover change. *Biogeosciences, 9*, 5125-5142

Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., & Vogelmann, J.E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment, 114*, 183-198

Huang, W., Sun, G., Dubayah, R.O., Cook, B., Montesano, P., Ni, W., & Zhang, Z. (2013). Mapping biomass change after forest disturbance: Applying

LiDAR footprint-derived models at key map scales. *Remote Sensing of Environment, 134*, 319-332

Huang, W., Sun, G., Dubayah, R.O., Zhang, Z., & Ni, W. (2012). Mapping forest above-ground biomass and its changes from LVIS waveform data. In, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 6561-6564)

Hudak, A.T., Crookston, N.L., Evans, J.S., Falkowski, M.J., Smith, A.M.S., Gessler, P.E., & Morgan, P. (2006). Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Canadian Journal of Remote Sensing, 32*, 126-138

Hurtt, G.C., Fisk, J., Thomas, R.Q., Dubayah, R.O., Moorcroft, P.R., & Shugart, H.H. (2010). Linking models and data on vegetation structure. *Journal of Geophysical Research: Biogeosciences, 115*, 11

Hyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M., & Hunsaker, C. (2006). Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM plus , Quickbird) synergy. *Remote Sensing of Environment, 102*, 63-73

IPCC (2013). Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change

Isenberg, M. (2011). LAStools—Efficient Tools for LiDAR Processing. In. http://www.cs.unc.edu/~isenburg/lastools/

Jenkins, J.C., Chojnacky, D.C., Heath, L.S., & Birdsey, R.A. (2003). National-scale biomass estimators for United States tree species. *Forest Science, 49*, 12-35

Jensen, M.A., Arnold, D.V., & Crockett, D.E. (2000). System-level microwave design: Radar-based laboratory projects. *IEEE Transactions on Education, 43*, 414-419

Joshi, N., Mitchard, E.T., Brolly, M., Schumacher, J., Fernández-Landa, A., Johannsen, V.K., Marchamalo, M., & Fensholt, R. (2017). Understanding 'saturation'of radar signals over forests. *Scientific Reports, 7*, 3505

Keller, M., Palace, M., & Hurtt, G. (2001). Biomass estimation in the Tapajos National Forest, Brazil - Examination of sampling and allometric uncertainties. *Forest Ecology and Management, 154*, 371-382

Kellndorfer, J.M., Walker, W.S., LaPoint, E., Kirsch, K., Bishop, J., & Fiske, G. (2010). Statistical fusion of lidar, InSAR, and optical remote sensing data for forest stand height characterization: A regional-scale method based on LVIS, SRTM, Landsat ETM plus , and ancillary data sets. *Journal of Geophysical Research: Biogeosciences, 115*, 1-10

Kellndorfer, J.M., Walker, W.S., 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, 93*, 339-358

Kenyi, L.W., Dubayah, R.O., Hofton, M., & Schardt, M. (2009). Comparative analysis of SRTM-NED vegetation canopy height to LIDAR-derived vegetation canopy metrics. *International Journal of Remote Sensing, 30*, 2797-2811

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 Transactions on Geoscience and Remote Sensing, 45*, 3317-3341

Kugler, F., Lee, S.K., Hajnsek, I., & Papathanassiou, K.P. (2015). Forest height estimation by means of Pol-InSAR data inversion: The role of the vertical wavenumber. *IEEE Transactions on Geoscience and Remote Sensing, 53*, 5294-5311

Kugler, F., Sauer, S., Lee, S.-K., Papathanassiou, K., & Hajnsek, I. (2010). Potential of TanDEM-X for forest parameter estimation. In, *Synthetic Aperture Radar (EUSAR), 2010 8th European Conference on* (pp. 1-4): VDE

Kugler, F., Schulze, D., Hajnsek, I., Pretzsch, H., & Papathanassiou, K.P. (2014). TanDEM-X Pol-InSAR performance for forest height estimation. *IEEE Transactions on Geoscience and Remote Sensing, 52*, 6404-6422

Langner, A., Achard, F., & Grassi, G. (2014). Can recent pan-tropical biomass maps be used to derive alternative Tier 1 values for reporting REDD plus activities under UNFCCC? *Environmental Research Letters, 9*, 124008-124019

Lavalle, M., & Hensley, S. (2015). Extraction of structural and dynamic properties of forests from polarimetric-interferometric SAR data affected by temporal decorrelation. *IEEE Transactions on Geoscience and Remote Sensing, 53*, 4752-4767

Le Toan, T., Quegan, S., Davidson, M.W.J., Balzter, H., Paillou, P., Papathanassiou, K., Plummer, S., Rocca, F., Saatchi, S., Shugart, H., & Ulander, L.

(2011). The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sensing of Environment, 115*, 2850-2860

Lee, S.-K., Fatoyinbo, T., Qi, W., Hancock, S., Armston, J., & Dubayah, R. (2018). GEDI AND TANDEM-X FUSION FOR 3D FOREST STRUCTURE PARAMETER RETRIEVAL. In, *2018 International Geoscience and Remote Sensing Symposium*. Valencia, Spain: IEEE GRSS

Lee, S.K., & Fatoyinbo, T.E. (2015). TanDEM-X Pol-InSAR inversion for mangrove canopy height estimation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*, 3608-3618

Lee, S.K., Kugler, F., Papathanassiou, K.P., & Hajnsek, I. (2013). Quantification of temporal decorrelation effects at L-band for polarimetric SAR interferometry applications. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6*, 1351-1367

Lefsky, M.A., Cohen, W.B., Harding, D.J., Parker, G.G., Acker, S.A., & Gower, S.T. (2002). Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology and Biogeography, 11*, 393-399

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Del Bom Espirito‐Santo, F., Hunter, M.O., & de Oliveira, R. (2005). Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical research letters, 32*

Li, A., Glenn, N.F., Olsoy, P.J., Mitchell, J.J., & Shrestha, R. (2015). Aboveground biomass estimates of sagebrush using terrestrial and airborne LiDAR data in a dryland ecosystem. *Agricultural and Forest Meteorology, 213*, 138-147

Lu, H., & Liu, G. (2014). Modeling the impacts of policy interventions from REDD+ in Southeast Asia: A case study in Indonesia. *Science China Earth Sciences, 57*, 2374-2385

Malhi, Y., Baldocchi, D.D., & Jarvis, P.G. (1999). The carbon balance of tropical, temperate and boreal forests. *Plant Cell and Environment, 22*, 715-740

Millennium Ecosystem Assessment (2005). Ecosystems and Human Well-Being: Biodiversity Synthesis. In. Island Press, Washington, DC

Mitchard, E.T.A., Saatchi, S.S., White, L.J.T., Abernethy, K.A., Jeffery, K.J., Lewis, S.L., Collins, M., Lefsky, M.A., Leal, M.E., Woodhouse, I.H., & Meir, P. (2012). Mapping tropical forest biomass with radar and spaceborne LiDAR in Lope

National Park, Gabon: overcoming problems of high biomass and persistent cloud. *Biogeosciences, 9*, 179-191

Montesano, P.M., Cook, B.D., Sun, G., Simard, M., Nelson, R.F., Ranson, K.J., Zhang, Z., & Luthcke, S. (2013). Achieving accuracy requirements for forest biomass mapping: A spaceborne data fusion method for estimating forest biomass and LiDAR sampling error. *Remote Sensing of Environment, 130*, 153-170

Mu, Q., Heinsch, F.A., Zhao, M., & Running, S.W. (2007). Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote sensing of Environment, 111*, 519-536

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., & Smith, G. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote sensing of environment, 83*, 214-231

Naesset, E., Gobakken, T., Bollandsas, O.M., Gregoire, T.G., Nelson, R., & Stahl, G. (2013). Comparison of precision of biomass estimates in regional field sample surveys and airborne LiDAR-assisted surveys in Hedmark County, Norway. *Remote Sensing of Environment, 130*, 108-120

Nagendra, H. (2001). Using remote sensing to assess biodiversity. *International journal of remote sensing, 22*, 2377-2400

Neumann, M., Saatchi, S.S., Ulander, L.M.H., & Fransson, J.E.S. (2012). Assessing performance of L-and P-band polarimetric interferometric SAR data in estimating boreal forest above-ground biomass. *IEEE Transactions on Geoscience and Remote Sensing, 50*, 714-726

NOAA (2011-2012). Daily Record of Climatological Observations. In: National Oceanic and Atmospheric Administration

NOAA (2015). United States Interagency Elevation Inventory (USIEI). In: NOAA's Ocean Service, Office for Coastal Management (OCM)

North, M., Oakley, B., Chen, J., Erickson, H., Gray, A., Izzo, A., Johnson, D., Ma, S., Marra, J., Meyers, M., & Purcell, K. (2010). *Vegetation and ecological characteristics of mixed-conifer and Red Fir forests at the Teakettle Experiemental Forest*Vegetation and ecological characteristics of mixed-conifer and Red Fir forests at the Teakettle Experiemental Forest

Nunery, J.S., & Keeton, W.S. (2010). Forest carbon storage in the northeastern United States: Net effects of harvesting frequency, post-harvest retention, and wood products. *Forest Ecology and Management, 259*, 1363-1375

Olesk, A., Voormansik, K., Vain, A., Noorma, M., & Praks, J. (2015). Seasonal differences in forest height estimation from interferometric TanDEM-X coherence data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*, 5565-5572

Oliver, M.A., & Webster, R. (1990). Kriging: a method of interpolation for geographical information systems. *International Journal of Geographical Information System, 4*, 313-332

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., & Canadell, J.G. (2011). A large and persistent carbon sink in the world's forests. *Science, 333*, 988-993

Papathanassiou, K.P., & Cloude, S.R. (2001). Single-baseline polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing, 39*, 2352-2363

Papathanassiou, K.P., & Cloude, S.R. (2003). The effect of temporal decorrelation on the inversion of forest parameters from Pol-InSAR data. In, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1429-1431)

Pardini, M., Qi, W.L., Dubayah, R., Papathanassiou, K.P., & Vde Verlag, G. (2016). Exploiting TanDEM-X Pol-InSAR Data for Forest Structure Observation and Potential Synergies with NASA's Global Ecosystem Dynamics Investigation Lidar (GEDI) Mission. In, *Proceedings of European Conference on Synthetic Aperture Radar (EUSAR)* (pp. 1-6). Hamburg, GERMANY: Vde Verlag Gmbh

Peters, C.A. (2001). Statistics for analysis of experimental data. *Environmental engineering processes laboratory manual*, 1-25

Pierce, L.E., Walker, W.S., Dobson, M.C., Hunsaker, C.T., Fites-Kaufman, J.A., & Dubayah, R.O. (2002). Fusion of Optical and SAR data for forestry applications in the Sierra Nevada of California. In, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1771-1773)

Pollard, J.E., Westfall, J.A., Patterson, P.L., Gartner, D.L., Hansen, M., & Kuegler, O. (2006). Forest inventory and analysis national data quality assessment report for 2000 to 2003. *Gen. Tech. Rep. RMRS-GTR-181. Fort Collins, CO: US*

Department of Agriculture, Forest Service, Rocky Mountain Research Station. 43 p., 181

Popkin, G. (2015). The hunt for the world's missing carbon. *Nature News, 523*, 20

Praks, J., Antropov, O., & Hallikainen, M.T. (2012). LIDAR-aided SAR interferometry studies in boreal forest: Scattering phase center and extinction coefficient at X-and L-band. *IEEE Transactions on Geoscience and Remote Sensing, 50*, 3831-3843

Praks, J., Kugler, F., Papathanassiou, K.P., Hajnsek, I., & Hallikainen, M. (2007). Height estimation of boreal forest: Interferometric model-based inversion at L- and X-band versus HUTSCAT profiling scatterometer. *IEEE Geoscience and Remote Sensing Letters, 4*, 466-470

Qi, W., & Dubayah, R.O. (2016). Combining Tandem-X InSAR and simulated GEDI lidar observations for forest structure mapping. *Remote Sensing of Environment, 187*, 253-266

Qi, W., Lee, S.-K., Hancock, S., Tang, H., Armston, J., Dubayah, R., style="text-align:justify, p.c.B., text-justify:inter-ideograph, & line-height:200%"> (Submitted). Improved forest height estimation by fusion of simulated GEDI Lidar data and TanDEM-X InSAR data. *Remote Sensing of Environment*

Quéré, C.L., Andrew, R.M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A.C., Korsbakken, J.I., Peters, G.P., Canadell, J.G., Jackson, R.B., Boden, T.A., Tans, P.P., Andrews, O.D., ..., & Zhu, D. (2018). Global Carbon Budget 2017

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing

Ranson, K.J., Saatchi, S., & Sun, G. (1995). Boreal forest ecosystem characterization with SIR-C/XSAR. *IEEE Transactions on Geoscience and Remote Sensing, 33*, 867-876

Ranson, K.J., & Sun, G. (2000). Effects of environmental conditions on boreal forest classification and biomass estimates with SAR. *IEEE Transactions on Geoscience and Remote Sensing, 38*, 1242-1252

Ranson, K.J., & Sun, G.Q. (1994). Northern forest classification using temporal multifrequency and multipolarimetric SAR images. *Remote Sensing of Environment, 47*, 142-153

Rejou-Mechain, M., Tanguy, A., Piponiot, C., Chave, J., & Herault, B. (2017). BIOMASS: an R package for estimating above-ground biomass and its uncertainty in tropical forests. *Methods in Ecology and Evolution, 8*, 1163-1167

Rodriguez, E., Morris, C.S., & Belz, J.E. (2006). A global assessment of the SRTM performance. *Photogrammetric Engineering & Remote Sensing, 72*, 249-260

Rosen, P.A., Hensley, S., Wheeler, K., Sadowy, G., Miller, T., Shaffer, S., Muellerschoen, R., Jones, C., Zebker, H., & Madsen, S. (2006). UAVSAR: A new NASA airborne SAR system for science and technology research. In, *Radar, 2006 IEEE Conference on* (p. 8 pp.): IEEE

Saarela, S., Holm, S., Grafstrom, A., Schnell, S., Naesset, E., Gregoire, T.G., Nelson, R.F., & Stahl, G. (2016). Hierarchical model-based inference for forest inventory utilizing three sources of information. *Annals of Forest Science, 73*, 895-910

Saatchi, S.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., & Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences, 108*, 9899-9904

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

Schwarz, P.A., Fahey, T.J., Martin, C.W., Siccama, T.G., & Bailey, A. (2001). Structure and composition of three northern hardwood-conifer forests with differing disturbance histories. *Forest Ecology and Management, 144*, 201-212

Seymour, M.S., & Cumming, I.G. (1994). Maximum likelihood estimation for SAR interferometry. In, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 2272-2275)

Siccama, T.G., Fahey, T.J., Johnson, C.E., Sherry, T.W., Denny, E.G., Girdler, E.B., Likens, G.E., & Schwarz, P.A. (2007). Population and biomass dynamics of trees in a northern hardwood forest at Hubbard Brook. *Canadian Journal of Forest Research, 37*, 737-749

Simard, M., Pinto, N., Fisher, J.B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research-Biogeosciences, 116*, 12

Smith, T.F., Rizzo, D.M., & North, M. (2005). Patterns of mortality in an old-growth mixed-conifer forest of the southern Sierra Nevada, California. *Forest Science, 51*, 266-275

Soja, M.J., Askne, J.I.H., & Ulander, L.M.H. (2017). Estimation of boreal forest properties from TanDEM-X data using inversion of the interferometric water cloud model. *IEEE Geoscience and Remote Sensing Letters, 14*, 997-1001

Soja, M.J., Persson, H., & Ulander, L.M.H. (2014). Estimation of boreal forest biomass from two-level model inversion of interferometric TanDEM-X data. In, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 3398-3401)

Soja, M.J., & Ulander, L.M.H. (2013). Digital canopy model estimation from TanDEM-X interferometry using high-resolution lidar DEM. In, *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 165-168)

Solberg, S., Astrup, R., Breidenbach, J., Nilsen, B., & Weydahl, D. (2013). Monitoring spruce volume and biomass with InSAR data from TanDEM-X. *Remote Sensing of Environment, 139*, 60-67

Stysley, P.R., Coyle, D.B., Kay, R.B., Frederickson, R., Poulios, 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. *Optics & Laser Technology, 68*, 67-72

Ståhl, G., Holm, S., Gregoire, T.G., Gobakken, T., Næsset, E., & Nelson, R. (2010). Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, NorwayThis article is one of a selection of papers from Extending Forest Inventory and Monitoring over Space and Time. *Canadian Journal of Forest Research, 41*, 96-107

Sun, G., Ranson, K.J., Guo, Z., Zhang, Z., Montesano, P., & Kimes, D. (2011). Forest biomass mapping from lidar and radar synergies. *Remote Sensing of Environment, 115*, 2906-2916

Swatantran, A., Dubayah, R.O., Goetz, S.J., Hofton, M.A., Betts, M.G., Sun, M., Simard, M., & Holmes, R. (2012). Mapping migratory bird prevalence using remote sensing data fusion. *PloS one, 7*, e28922

Swatantran, A., Dubayah, R.O., Roberts, D., Hofton, M.A., & Blair, J.B. (2011). Mapping biomass and stress in the Sierra Nevada using lidar and hyperspectral data fusion. *Remote Sensing of Environment, 115*, 2917-2930

Tang, H., Brolly, M., Zhao, F., Strahler, A.H., Schaaf, C.L., Ganguly, S., Zhang, G., & Dubayah, R.O. (2014). Deriving and validating Leaf Area Index (LAI) at multiple spatial scales through lidar remote sensing: A case study in Sierra National Forest, CA. *Remote Sensing of Environment, 143*, 131-141

Tang, H., Dubayah, R.O., Swatantran, A., Hofton, M., Sheldon, S., Clark, D.B., & Blair, B. (2012). Retrieval of vertical LAI profiles over tropical rain forests using waveform lidar at La Selva, Costa Rica. *Remote Sensing of Environment, 124*, 242-250

Thomas, R.Q., Hurtt, G.C., Dubayah, R.O., & Schilz, M.H. (2008). Using lidar data and a height-structured ecosystem model to estimate forest carbon stocks and fluxes over mountainous terrain. *Canadian Journal of Remote Sensing, 34*, S351-S363

Treuhaft, R.N., Chapman, B.D., dos Santos, J.R., Goncalves, F.G., Dutra, L.V., Graca, P., & Drake, J.B. (2009). Vegetation profiles in tropical forests from multibaseline interferometric synthetic aperture radar, field, and lidar measurements. *Journal of Geophysical Research: Atmospheres, 114*, 16

Treuhaft, R.N., Goncalves, F., dos Santos, J.R., Keller, M., Palace, M., Madsen, S.N., Sullivan, F., & Graca, P. (2015). Tropical-forest biomass estimation at X-band from the spaceborne TanDEM-X interferometer. *IEEE Geoscience and Remote Sensing Letters, 12*, 239-243

Tsui, O.W., Coops, N.C., Wulder, M.A., & Marshall, P.L. (2013). Integrating airborne LiDAR and space-borne radar via multivariate kriging to estimate above-ground biomass. *Remote Sensing of Environment, 139*, 340-352

Valbuena, R., Hernando, A., Manzanera, J., Görgens, E., Almeida, D., Mauro, F., & García-Abril, A. (2017). Enhancing of accuracy assessment for forest above-ground biomass estimates obtained from remote sensing via hypothesis testing and overfitting evaluation. *Ecological Modelling, 366*, 15-26

Van Doorn, N.S., Battles, J.J., Fahey, T.J., Siccama, T.G., & Schwarz, P.A. (2011). Links between biomass and tree demography in a northern hardwood forest: a decade of stability and change in Hubbard Brook Valley, New Hampshire. *Canadian Journal of Forest Research, 41*, 1369-1379

Whitehead, D. (2011). Forests as carbon sinks—benefits and consequences. *Tree physiology, 31*, 893-902

Whitehurst, A.S., Swatantran, A., Blair, J.B., Hofton, M.A., & Dubayah, R.O. (2013). Characterization of canopy layering in forested ecosystems using full waveform lidar. *Remote Sensing, 5*, 2014-2036

Wilson, B.T., Woodall, C.W., & Griffith, D.M. (2013). Imputing forest carbon stock estimates from inventory plots to a nationally continuous coverage. *Carbon Balance and Management, 8*, 1-15

Yang, X., Strahler, A.H., Schaaf, C.B., Jupp, D.L.B., Yao, T., Zhao, F., Wang, Z., Culvenor, D.S., Newnham, G.J., Lovell, J.L., Dubayah, R.O., Woodcock, C.E., & Ni-Meister, W. (2013). Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidn (R)). *Remote Sensing of Environment, 135*, 36-51

Zhao, F., Guo, Q., & Kelly, M. (2012). Allometric equation choice impacts lidar-based forest biomass estimates: A case study from the Sierra National Forest, CA. *Agricultural and Forest Meteorology, 165*, 64-72

Zhao, F., Yang, X., Strahler, A.H., Schaaf, C.L., Yao, T., Wang, Z., Roman, M.O., Woodcock, C.E., Ni-Meister, W., Jupp, D.L.B., Lovell, J.L., Culvenor, D.S., Newnham, G.J., Tang, H., & Dubayah, R.O. (2013). A comparison of foliage profiles in the Sierra National Forest obtained with a full-waveform under-canopy EVI lidar system with the foliage profiles obtained with an airborne full-waveform LVIS lidar system. *Remote Sensing of Environment, 136*, 330-341

Zhao, M., & Zhou, G.-S. (2005). Estimation of biomass and net primary productivity of major planted forests in China based on forest inventory data. *Forest Ecology and Management, 207*, 295-313

Álvarez-Dávila, E., Cayuela, L., González-Caro, S., Aldana, A.M., Stevenson, P.R., Phillips, O., Cogollo, Á., Peñuela, M.C., Von Hildebrand, P., & Jiménez, E. (2017). Forest biomass density across large climate gradients in northern South America is related to water availability but not with temperature. *PloS one, 12*, e0171072