Forschungsbericht 2023-16

Combining TanDEM-X Interferometric SAR and GEDI Lidar Measurements for Improving Forest Height, Structure and Biomass Estimates

Changhyun Choi

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Hochfrequenztechnik und Radarsysteme

Oberpfaffenhofen

DLR Deutsches Zentrum für Luft- und Raumfahrt

Forschungsbericht 2023-16

Combining TanDEM-X Interferometric SAR and GEDI Lidar Measurements for Improving Forest Height, Structure and Biomass Estimates

Changhyun Choi

Deutsches Zentrum für Luft- und Raumfahrt
Institut für Hochfrequenztechnik und Radarsysteme

Oberpfaffenhofen

130 Seiten
46 Bilder
3 Tabellen
228 Literaturstellen

Deutsches Zentrum
DLR für Luft- und Raumfahrt

Herausgeber:

Deutsches Zentrum
für Luft- und Raumfahrt e. V.
Wissenschaftliche Information
Linder Höhe
D-51147 Köln

ISSN 1434-8454
ISRN DLR-FB-2023-16
Erscheinungsjahr 2023

DOI: 10.57676/v2fg-0z95

This version reproduces the manuscript submitted to ETH Zurich revised for typos.

SAR, interferometrischen SAR, Lidar, TanDEM-X, GEDI, Wald, Waldhöhe, Waldstruktur, Waldbiomasse

Changhyun Choi
DLR, Institut für Hochfrequenztechnik und Radarsysteme, Oberpfaffenhofen

Kombination von TanDEM-X interferometrischen SAR- und GEDI Lidar-Messungen zur Verbesserung von Waldhöhe, -struktur und Biomasse-Schätzungen

ETH Zürich

Die Parameter der Waldhöhe, Waldstruktur und Waldbiomasse sind weithin als die drei essenziellen Elemente zur Beschreibung und Überwachung von Wäldern bekannt. Für jegliche Schätzung dieser Parameter auf einer annähernd globalen Skala, braucht es weltraumgestützte Fernerkundungssysteme. Nur durch die Anwendung des optischen *Light Dectection and Ranging* (LiDar) und dem auf Mikrowellen basierten *Synthetic Aperture Radar* (SAR) können Wälder dreidimensional, und mit einer räumlichen Auflösung weniger Meter wiederholt betrachtet werden. Beide Technologien spielen dabei entscheidende Rollen. Von der gemeinsamen Nutzung der LiDAR Mission GEDI der nordamerikanischen Raumfahrbehörde NASA, wie auch der interferometrischen SAR Mission TanDEM-X des Deutschen Zentrums für Luft- und Raumfahrt (DLR e.V.) werden erhebliche Synergieeffekte erwartet. Diese begründen sich auf einer gemeinsamen Sensitivität zur geometrischen Struktur der Baumkronen bei gleichzeitig hoher räumlicher und zeitlicher Auflösung, wie auch der Möglichkeit zur Aufnahme komplementärer Geometrien und Aufnahmearten.

Aufgrund dieses einzigartigen Potentials, beruht die Zielsetzung dieser Arbeit darin, einen synergetischen und systematischen Ansatz zur Kombination der räumlich diskreten GEDI LiDAR *waveforms* und den räumlich kontinuierlichen TanDEM-X Kohärenzen zu entwickeln. Zusätzlich dazu, werden die benötigten Algorithmen zur räumlich-kontinuierlichen Schätzung der Waldhöhe (1) bei einer räumlichen Auflösung besser 1 ha, wie auch eine Waldhöhe-zu-Waldbiomasse Allometrie (2) entwickelt. Die in (1) abgeleitete kontinuierliche Waldhöhe wird zusammen mit einem aus TanDEM-X Daten berechneten Waldstruktur-Index zur Verbesserung der allometrischen Beziehung verwendet. Im Rahmen der Entwicklung zur synergetischen Nutzung, werden die GEDI *waveforms* zur Initialisierung der Waldhöhen-Invertierung aus TanDEM-X Daten genutzt. Gleichzeitig, werden die GEDI Waldhöhe und Waldbiomasse Daten auf Aufnahmeskala dazu verwendet, die Waldhöhe-zu-Waldbiomasse Allometrie aufzustellen. Ein neuartiger horizontaler Waldstrukturindex auf Basis von TanDEM-X InSAR Daten wurde entwickelt, um die Bestandsdichte zu approximieren. Dazu wird die Variabilität der Baumkronen-Höhe durch die InSAR Phasenzentren geschätzt. Weniger dichte Bestände zeichnen sich durch eine hohe Variabilität in der Baumkronen-Höhe aus, während dichte(re) Bestände eine wesentlich homogenere Struktur aufweisen. Darauf aufbauend, wurde zusätzlich eine Methode entwickelt, bei der diese Strukturinformation zur Kompensierung der durch Penetration induzierten Unterschätzung der Waldhöhe verwendet wird, auf Kosten einer verschlechterten räumlichen Auflösung.

Die experimentelle Analyse zur Beurteilung der Ergebnisse deutet auf eine hohe Effektivität des kombinierten Ansatzes hin. Die Ergebnisse aus den tasmanischen Wäldern zeigen, dass die Modelparametrisierung für GEDI und die empirische Profilkorrektur robuste Waldhöhen Daten auf großen Skalen liefert. Die großskalige Implementierung der Waldhöhenschätzung wird überwiegend durch die lokale Variabilität in der Waldstruktur und dem Profil begrenzt. Parallel dazu, legen die Ergebnisse über den Untersuchungsgebieten in Gabun dar, dass der TanDEM-X Strukturindex die Fehler in der Waldhöhenschätzung größtenteils kompensieren kann und die Höhe-zu-Biomasse Allometrie lokal anpasst. Dadurch konnte die Biomasse-Schätzung im Vergleich zu einer singulären Allometrie verbessert werden. Die größte verbleibende Unsicherheit (Fehler/Varianz) findet sich in der Fehlerfortführung der übrigen Unsicherheit in der Waldhöhen-Schätzung.

SAR, SAR Interferometry, Lidar, TanDEM-X, GEDI, Forest, Forest height, Forest structure, Forest biomass

Changhyun Choi
German Aerospace Center (DLR), Microwaves and Radar Institute, Oberpfaffenhofen

Combining TanDEM-X Interferometric SAR and GEDI Lidar Measurements for Improving Forest Height, Structure and Biomass Estimates

ETH Zurich

Forest height, structure and biomass are widely recognized as three important elements to monitor as they describe state and development. Any global scale estimation of forest height, structure and biomass requires the support of spaceborne remote sensing techniques. Only Light Detection And Ranging (lidar) and Synthetic Aperture Radar (SAR) configurations can map forests in 3D with metric resolution and repeatedly in time, and contribute critically to this challenge. The NASA GEDI lidar mission and the DLR TanDEM-X interferometric SAR (InSAR) mission are expected to possess a great synergy potential in this regard due to the common sensitivity to the geometric architecture of the forest canopy and high resolution, and to the complementary acquisition geometries and measurement approaches.

With this motivation, the objective of this thesis is to propose a synergetic and systematic combination of spatially discrete GEDI lidar waveforms and continuous TanDEM-X InSAR coherences, and to develope the necessary algorithms for (i) estimating spatially continuous and unbiased large scale forest height maps with resolution of 1 ha or less, and (ii) using the estimated height in a (stand-level) height-to-biomass allometric relationship adapted and improved at local scales using a horizontal structure index derived from the TanDEM-X data. In the developed combination framework, the GEDI waveforms are used to initialize the inversion of forest height from TanDEM-X data. A set of performance criteria are implemented throughout the inversion to ensure a certain estimation quality. At the same time, the footprint-level height and biomass are used to define the height-to-biomass allometry. A new horizontal structure index from TanDEM-X InSAR data has been developed as a proxy to forest density by characterizing the variability of the top forest height through the variability of the InSAR phase center heights. Sparse stands are characterized by a high variability, which decreases for dense(r) stands. A methodology has also been devised to use the structure information to compensate for (penetration-induced) height underestimation biases at the cost of a reduced spatial resolution.

The experimental performance analysis points to the effectiveness of the developed combination framework. Height inversion results in the Tasmanian forests demonstrate that the appropriate model parameterization provided by GEDI and the devised implementation can provide robust estimates at large scale. The estimation performance is mainly limited by the uncompenstaed mismatch between the lidar reflectance and the X-band reflectivity profiles induced by local variabilities of forest structure and slopes. In parallel, the results over tropical forest test sites in Gabon confirm that the developed TanDEM-X structure index can compensate height estimation biases at a large extent, and is able to adapt the height-to-biomass allometry locally and to improve the biomass estimation performance when compared to the use of a single allometry. The largest remaining uncertainty contribution (in terms of bias and / or variance) is attributed to the propagation of the remaining height estimation uncertainty.

DISS. ETH NO. 28769

COMBINING TANDEM-X INTERFEROMETRIC SAR AND GEDI LIDAR MEASUREMENTS FOR IMPROVING FOREST HEIGHT, STRUCTURE AND BIOMASS ESTIMATES

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

CHANGHYUN CHOI

M.Sc. Earth and Environmental Sciences
Seoul National University (SNU)

born on 27.01.1994
citizen of Republic of Korea

accepted on the recommendation of

Prof. Dr. Irena Hajnsek
Prof. Dr. Andreas Huth
Prof. Dr. Juan M. Lopez-Sanchez
Prof. Dr. Ralph Dubayah

2023

ABSTRACT

Forest height, structure and biomass are widely recognized as three important elements to monitor as they describe state and development. Any global scale estimation of forest height, structure and biomass requires the support of spaceborne remote sensing techniques. Only Light Detection And Ranging (lidar) and Synthetic Aperture Radar (SAR) configurations can map forests in 3D with metric resolution and repeatedly in time, and contribute critically to this challenge. The NASA GEDI lidar mission and the DLR TanDEM-X interferometric SAR (InSAR) mission are expected to possess a great synergy potential in this regard due to the common sensitivity to the geometric architecture of the forest canopy and high resolution, and to the complementary acquisition geometries and measurement approaches.

With this motivation, the objective of this thesis is to propose a synergetic and systematic combination of spatially discrete GEDI lidar waveforms and continuous TanDEM-X InSAR coherences, and to develope the necessary algorithms for (i) estimating spatially continuous and unbiased large scale forest height maps with resolution of 1 ha or less, and (ii) using the estimated height in a (stand-level) height-to-biomass allometric relationship adapted and improved at local scales using a horizontal structure index derived from the TanDEM-X data. In the developed combination framework, the GEDI waveforms are used to initialize the inversion of forest height from TanDEM-X data. A set of performance criteria are implemented throughout the inversion to ensure a certain estimation quality. At the same time, the footprint-level height and biomass are used to define the height-to-biomass allometry. A new horizontal structure index from TanDEM-X InSAR data has been developed as a proxy to forest density by characterizing the variability of the top forest height through the variability of the InSAR phase center heights. Sparse stands are characterized by a high variability, which decreases for dense(r) stands. A methodology has also been devised to use the structure information to compensate for (penetration-induced) height underestimation biases at the cost of a reduced spatial resolution.

The experimental performance analysis points to the effectiveness of the developed combination framework. Height inversion results in the Tasmanian forests demonstrate that the appropriate model parameterization provided by GEDI and the devised implementation can provide robust estimates at large scale. The estimation performance is mainly limited by the uncompenstaed mismatch between the lidar reflectance and the X-band reflectivity profiles induced by local variabilities of forest structure and slopes. In parallel, the results over tropical forest test sites in Gabon confirm that the developed TanDEM-X structure index can compensate height estimation biases at a large extent, and is able to adapt the height-to-biomass allometry locally and to improve the biomass estimation performance when compared to the use of a single allometry. The largest remaining uncertainty contribution (in terms of bias and / or variance) is attributed to the propagation of the remaining height estimation uncertainty.

KURZFASSUNG

Die Parameter der Waldhöhe, Waldstruktur und Waldbiomasse sind weithin als die drei essenziellen Elemente zur Beschreibung und Überwachung von Wäldern bekannt. Für jegliche Schätzung dieser Parameter auf einer annähernd globalen Skala, braucht es weltraumgestützte Fernerkundungssysteme. Nur durch die Anwendung des optischen *Light Dectection and Ranging* (LiDar) und dem auf Mikrowellen basierten *Synthetic Aperture Radar* (SAR) können Wälder dreidimensional, und mit einer räumlichen Auflösung weniger Meter wiederholt betrachtet werden. Beide Technologien spielen dabei entscheidende Rollen. Von der gemeinsamen Nutzung der LiDAR Mission GEDI der nordamerikanischen Raumfahrbehörde NASA, wie auch der interferometrischen SAR Mission TanDEM-X des Deutschen Zentrums für Luft- und Raumfahrt (DLR e.V.) werden erhebliche Synergieeffekte erwartet. Diese begründen sich auf einer gemeinsamen Sensitivität zur geometrischen Struktur der Baumkronen bei gleichzeitig hoher räumlicher und zeitlicher Auflösung, wie auch der Möglichkeit zur Aufnahme komplementärer Geometrien und Aufnahmearten.

Aufgrund dieses einzigartigen Potentials, beruht die Zielsetzung dieser Arbeit darin, einen synergetischen und systematischen Ansatz zur Kombination der räumlich diskreten GEDI LiDAR *waveforms* und den räumlich kontinuierlichen TanDEM-X Kohärenzen zu entwickeln. Zusätzlich dazu, werden die benötigten Algorithmen zur räumlich-kontinuierlichen Schätzung der Waldhöhe (1) bei einer räumlichen Auflösung besser 1 ha, wie auch eine Waldhöhe-zu-Waldbiomasse Allometrie (2) entwickelt. Die in (1) abgeleitete kontinuierliche Waldhöhe wird zusammen mit einem aus TanDEM-X Daten berechneten Waldstruktur-Index zur Verbesserung der allometrischen Beziehung verwendet. Im Rahmen der Entwicklung zur synergetischen Nutzung, werden die GEDI *waveforms* zur Initialisierung der Waldhöhen-Invertierung aus TanDEM-X Daten genutzt. Gleichzeitig, werden die GEDI Waldhöhe und Waldbiomasse Daten auf Aufnahmeskala dazu verwendet, die Waldhöhe-zu-Waldbiomasse Allometrie aufzustellen. Ein neuartiger horizontaler Waldstrukturindex auf Basis von TanDEM-X InSAR Daten wurde entwickelt, um die Bestandsdichte zu approximieren. Dazu wird die Variabilität der Baumkronen-Höhe durch die InSAR Phasenzentren geschätzt. Weniger dichte Bestände zeichnen sich durch eine hohe Variabilität in der Baumkronen-Höhe aus, während dichte(re) Bestände eine wesentlich homogenere Struktur aufweisen. Darauf aufbauend, wurde zusätzlich eine Methode entwickelt, bei der diese Strukturinformation zur Kompensierung der durch Penetration induzierten Unterschätzung der Waldhöhe verwendet wird, auf Kosten einer verschlechterten räumlichen Auflösung.

Die experimentelle Analyse zur Beurteilung der Ergebnisse deutet auf eine hohe Effektivität des kombinierten Ansatzes hin. Die Ergebnisse aus den tasmanischen Wäldern zeigen, dass die Modelparametrisierung für GEDI und die empirische Profilkorrektur robuste Waldhöhen Daten auf großen Skalen liefert. Die großskalige Implementierung der Waldhöhenschätzung wird überwiegend durch die lokale Variabilität in der Waldstruktur und dem Profil begrenzt. Parallel

dazu, legen die Ergebnisse über den Untersuchungsgebieten in Gabun dar, dass der TanDEM-X Strukturindex die Fehler in der Waldhöhenschätzung größtenteils kompensieren kann und die Höhe-zu-Biomasse Allometrie lokal anpasst. Dadurch konnte die Biomasse-Schätzung im Vergleich zu einer singulären Allometrie verbessert werden. Die größte verbleibende Unsicherheit (Fehler/Varianz) findet sich in der Fehlerfortführung der übrigen Unsicherheit in der Waldhöhen-Schätzung.

ACKNOWLEDGEMENTS

I will never forget the first day joining DLR. The day was cloudy, typical German weather, and a bit chilly. I was excited to start with new topic but nervous for the challenge at the same time. In the end of PhD, I can say that I have learned and achieved a lot more than I could have ever imagined. I should say there were so many great people directly and indirectly contributed to the accomplishment of my PhD.

First of all, I'm deeply grateful to Kostas, who gave me the opportunity to pursue a PhD with an incredible topic and project in DLR. I thank you for all your guidance and the discussions to develop my research. I consider it was very special that you trusted me in proceeding of GEDI-TDX project which was one of the great experiences during my PhD. You always asked me the correct and the challenging questions, as well as finding ways to improve the outline of researches and manuscripts. Additionally, I have really happy memories of "Running" in many countries.

At the same time, I really appreciate all the things Matteo did for me since the beginning of my PhD. He has been my closest mentor during my PhD. I have learned so many things from him including ways developing research and thinking and his attitude to research. I believed I was lucky to have such a great supervisor.

I would like to thank Irena for supporting and trusting me to pursue PhD at ETH. Thanks to Alberto Moreira and Gerhard Krieger for giving me the possibility to work in the Microwave and Radar institute and always try to improve life of PhD students. Many thanks to Ralph Dubayah, Andreas Huth, and Juan M. Lopez-Sanchez for accepting being in the committee and reviewing this dissertation.

I want to express particularly that I thank you all colleagues in the Pol-InSAR and Information Retrieval groups at DLR. Victor, it was great to share the office with you. The time of sharing the office was so happy and warm. Jun Su, thank you for helping me in various ways as a senior Korean DLR researcher. Giuseppe, you were my Italian and coffee teacher. Roman, it was nice to share part of our journey towards PhD and as a friend. Sibylle and Gabi, thank you for your best supports. Georg, thank you for uncountable helps to solve many language problems. I enjoyed many activities with you as well. Alberto, Thomas, Kristina, and many past colleges, you gave me an amazing experience. Ben (thank you for your help in abstract!), Islam (happy to share the office), Paloma, Lea, Nikita, and Noelia, I wish your happier next days. You were massive parts of my PhD-life.

I really want express my gratitude to my Korean friends, Juno, Min Jae, Dae Geon, Ji Hoon, Simchan, Hyeong-oh, Taekyung. Thank you so much for contacts and supports from Korea. Hyun Choung and Seokkyun, and all my Korean friends in Munich, maybe you guys donn't realize but you helped me a lot during my hard times of being in Germany. A special thanks to my family – Father, Mother, and Brother - who always believed in me and encouraged me I couldn't take a step further without your supports. It is impossible to list all the things you all contributed to me being able to achieve my PhD.

Most importantly, I acknowledge my fiancé Heasun with your supports from Korea and Germany. Without you, my life journey cannot be completed. Thank you for being such a wonderful way to me. Love you.

CONTENTS

LIST OF ABBREVIATIONS

AGB	AboveGround Biomass
BIOMASS	P-band SAR mission currently developed by European Space Agency
CHM	Canopy Height Model
DBH	Diameter of tree at Breast Height
DEM	Digital Elevation Model
DTM	Digital Terrain Model
DLR	Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center)
ESA	European Space Agency
GEDI	Global Ecosystem Dynamics Investigation NASA mission
HH	Horizontal receive, horizontal transmit
HOME	Height Of Median Energy
HS	Horizontal Structure index
ICESat	NASA’s Ice, Cloud and land Elevation Satellite
InSAR	Synthetic Aperture Radar Interferometry
ISS	International Space Station
Lidar	Light Imaging Detection and Ranging
LVIS	NASA's full waveform lidar Land, Vegetation and Ice Sensor
NASA	National Aeronautics and Space Administration
PAI	Plant Area Index
P-band	Frequency of 435 MHz, wavelength of 69 cm in case of BIOMASS
Radar	Radio Detection and Ranging
RH	Relative Height Metric
RVoG	Random-Volume-over-Ground
SAR	Synthetic Aperture Radar
SLA	Shuttle Laser Altimeter
SLC	Single Look Complex
SNR	Signal To Noise ratio
SRTM	Shuttle Radar Topography Mission
TanDEM-X	X-band single-pass SAR mission operated by DLR
X-band	Frequency of 9.65 GHz, wavelength of 3.1 cm in case of TanDEM-X

LIST OF SYMBOLS

α	Allometric coefficient
β	Allometric exponent
$\tilde{\gamma}(\kappa_z)$	Measured complex interferometric coherence
γ_{Scat}	Baseline separation induced decorrelation
γ_{Sys}	Systematic decorrelation
γ_{tmp}	Temporal decorrelation
γ_{Vol}	Volume coherence
$\sigma(x, y, z)$	Volumetric reflectivity of the vegetation layer
κ_z	Interferometric vertical wavenumber [rad/m]
λ	Wavelength [m]
HoA	Height of Ambiguity [m]
H_V	Top forest height
i	Imaginary unit
$P(x, y, z)$	Received lidar signal power at a position (x,y,z)
S_j	SAR acquisition j
$f_V(z)$	Radar vertical reflectivity function
θ	Incidence angle [rad]
θ_{local}	Local incidence angle [rad]
z	Height axis [m]
z_0	Underlying ground location [m], mostly defined as $z_0 = 0\ m$
$(\cdot)^*$	Complex conjugate
$\langle\cdot\rangle$	Expectation value
$\lvert\cdot\rvert$	Magnitude

1 INTRODUCTION

1.1 MOTIVATION

Forests are one of the most important terrestrial ecosystems and play a significant role in the global carbon and water cycle. Besides their importance, forests are today more than ever before under the pressure of a growing world population and climate change. Globally, forests are characterized by complex 3D patterns which are the result of natural, climatic and anthropogenic processes. In this context, strategies for characterizing forest height, structure and biomass, and their dynamics in time, are of significant importance. Mapping forest height provides information on stand condition and site index. It allows characterizing the successional state of the forest and thus can be used to describe forest dynamics [1]. Forest height is an indicator for the site dependent timber production potential of a stand and is closely related to forest biomass [2-3]. The distribution of forest heights within a stand may be used to assess the disturbance regime. High spatial and temporal resolution forest height maps can be used to detect logging activities. At the same time, mapping forest structure (intended as "the organization in space and time, including the position, extent, quantity, type and connectivity, of the aboveground components of vegetation" [4-8]) is critical for understanding the history, function and future of forest ecosystems. Indeed, forest structure expresses forest state, functionality, biodiversity and evolution, and is an indicator of the successional stage and development as well as sustainability and habitability [9-11]. Due to this, it is an important parameter for assessing forest productivity, biomass and biodiversity [12-15]. Finally, the knowledge of both height and structure and their dynamics can be used to constrain model estimates of above-ground biomass (AGB – intended as "living vegetation above the soil, including stem, stump, branches, bark, seeds, and foliage" [16] in ecology and carbon stock estimates, but more focused on the dominant stem contribution in forestry) and associated carbon flux components between the vegetation and the atmosphere. While forest characterization traditionally relies on sampling techniques by means of field inventory plots or more recently by using terrestrial laser scanning techniques, any global scale estimation of forest height, structure and biomass requires the support of (spaceborne) remote sensing techniques.

Today, only Light Detection and Ranging (lidar) and Synthetic Aperture Radar (SAR) configurations can monitor forests in 3D with metric resolution and repeatedly in time, and contribute critically to the observation and quantitative characterization of forest height, structure and biomass at large scales [17-18]. Full-waveform lidar systems measure directly vegetation reflectance profiles (or waveforms) with a fairly high vertical resolution over footprint samples along rather narrow stripes in a nadir-looking geometry. This configuration is implemented by the NASA's / University of Maryland Global Ecosystem Dynamics Investigation lidar (GEDI) mission [19], operating on the International Space Station (ISS) since 2018 with the scientific

objective of mapping forest height, structure and biomass of temperate and tropical forests globally. Despite meeting the required measurement accuracy and resolution, GEDI can only sample forest structure over a grid of footprints on ground. The resulting sparsity of the spatial coverage (possibly reduced even further by cloud cover) may not be adequate to represent actual structure variations at small scale without possibility of interpolation (gridding) below 1 km × 1 km resolution. In contrast, SAR measurements are continuous with high spatial resolution, and can be implemented by means of wide swaths in a side-looking geometry allowing large scale coverage and short revisit time. Depending on the configuration, SAR systems can measure backscattering amplitudes, interferometric coherences or even radar vertical reflectivity profiles. However, these profiles are not always straightforward to be interpreted in terms of physical forest structure, partially due to their limited vertical resolution. The DLR TerraSAR-X add-on for Digital Elevation Measurement (TanDEM-X) mission [20-21] has acquired single-pass interferometric SAR (InSAR) data globally without temporal decorrelation at X-band (wavelength 3.1 cm) since 2010 for the generation of a global digital elevation model. Despite the short wavelength, and the consequential limitations in the penetration of forest volumes, the exploration of TanDEM-X InSAR data for forestry has been increasing in time especially concerning forest height [22-26] and biomass estimation and classification [27-30]. More recent experiments have demonstrated that TanDEM-X data can contribute to the quantification of forest structure [31-35]. These application potentials and the obtained performances were certainly unexpected before the mission launch, but not all of them are today systematically understood and characterized.

The common sensitivity to forest structure, and the complementary measurement configurations, motivate the combined and synergistic use of lidar and SAR data. This combination is expected to enhance the quality of forest characterization by improving the accuracy and / or the resolution, spatial continuity and coverage of physical forest parameter estimates, like forest height, structure, and biomass. This topic is addressed in this thesis with particular reference to the combination of spatially discrete GEDI and continuous TanDEM-X measurements for obtaining (1) spatially continuous forest height estimates at large scale and resolution at 1 ha or even below, and (2) spatially continuous biomass estimates through a (stand-level) height-to-biomass allometry adapted and improved at local scales using structure information. While the differences in wavelength and observation geometry make GEDI waveforms and X-band reflectivity profiles different, the high attenuation makes both of them sensitive to the geometric architecture of the canopy with a similar information content. Exactly this common information content allows GEDI samples to initialize parameter inversions from the TanDEM-X data over the complementary continuous coverage, and both the coverage continuity and high resolution of TanDEM-X measurements make possible the exploration of a set of relevant scales in order to provide (horizontal) structural information that the sampled GEDI nature cannot achieve. The ability to link GEDI and TanDEM-X measurements at the different complementary coverage (sampled vs. continuous) and scales / resolution performed defines in this case the combination. Its realization requires both a deeper understanding of the measurements and the development of novel estimation algorithms.

1.2 STATE-OF-THE-ART

Forest remote sensing applications have been developed with both passive and active sensors. Passive sensors include in particular optical systems, which record spectral responses from the interaction between the solar radiance and forest stand canopies in the nanometers wavelength range. Optical data have been predominant in forest remote sensing since decades due to their relatively easy processing, interpretation and continuity of observations since the 1970s [36] for forest/non-forest mapping [37-38], forest disturbances monitoring [39], forest structure indices such the leaf area index [40] or the normalised difference vegetation index. However, cloud coverage limits the availability of data for large regions and/or amount of time. Additionally, the limited sensitivity to the vertical distribution of vegetation elements especially after canopy closure introduces large uncertainties in structural and biomass maps at scales lower than a few kilometers [41-42].

In contrast, measurements from active remote sensing systems like lidar and SAR (briefly reviewd in this Section) are inherently sensitive to the full 3D forest attributes, and because of this they have been widely used for height, structure and biomass mapping [17-18]. It is worth mentioning that even if height- and structure-related parameters are more or less directly retrievable at stand level, the typical resolution of lidar and SAR measurements – especially in spaceborne implementation – does not allow the extraction of single-tree information. Thus, well-defined concepts to estimate forest height, structure, and biomass for the specific sensor measurements, resolutions and acquisition configurations are required to link data to (conventional) physical parameters at the scales / resolutions performed. The state-of-the-art of these concepts are briefly reviewed and discussed in the next Sections with particular reference to lidar and SAR configurations like GEDI and TanDEM-X alone and in combination.

1.2.1 FULL WAVEFORM LIDAR MEASUREMENTS: THE GEDI MISSION

Vegetation lidar instruments are active configurations usually operating at the wavelength of 1064 nm or 1550 nm as they propagate well through the atmosphere, they are eye-safe, and the reflectance of vegetation elements is very high resulting into a high signal-to-noise ratio. Laser pulses are transmitted in a nadir-looking geometry, and intercept the vegetation elements within the illuminated footprint on the ground, get attenuated and reflected back to the receiver [5], [7], [43-44]. In full waveform lidars the receiver records the reflected distribution of (light) energy (i.e., the waveform) as a function of time, which can be converted in range (distance) through the speed of light. As a consequence, the received waveform shape depends directly on the 3D geometric distribution of the intercepted vegetation elements within the footprint, as well as their scattering properties and the actual atmospheric conditions. The footprint diameter can range between decimeter in small-footprint lidars and tens of meters in large-footprint lidars. While the former is typically designed for fine topographic mapping (and only the discretized returns may be recorded), the latter is instead preferred for vegetation mapping. Indeed, footprints comparable to the crown diameter (10 – 25 m) are expected to contain both tree top and ground returns. In

addition, they increase coverage with a lower cost [43]. Ignoring for simplicity the propagation through the atmosphere, and assuming a penetration until the terrain height z_0 and a constant weighting within the footprint, the received waveform $P(x, y, z)$ at a position (x, y) on ground from a vegetation volume of thickness H_V can be written as

$$P(x, y, z) = \int_{z_0}^{z_0 + H_V} p(z' - z) \iint \sigma(x' - x, y' - x, z') \cdot e^{-\tau(z' - z_0)} dy' dx' dz' \qquad (1.1)$$

where $\sigma(x, y, z)$ is the volumetric reflection of the vegetation layer, τ the volumetric extinction coefficient that accounts for the two-way attenuation within the vegetation layer, and $p(z)$ is the system pulse shape.

The first and last peaks of a waveform are typically associated with reflecting surfaces at the highest and lowest heights within the footprint, respectively. An accurate identification of these peaks above the noise level is required on the one hand to precisely geolocate the corresponding reflecting surfaces, and on the other hand to calculate statistical quantile metrics from the integrated waveform between these ranging points. These statistical quantiles are typically referred to as relative height (RH) metrics [45]. Each RH metric expresses the height above the ground at which a certain percentage of the total waveform energy is received. For instance, the RH100 represents to the height above the ground in which the total energy is received. The RH98 is often used as a measurement of top canopy height as it is less affected by noise than the RH100. RH metrics are typically used alone or in combination to describe the canopy vertical structure [46]. The distribution of the waveform peaks has been considered as well [45], [47], under the assumption that a peak corresponds to an accumulation of reflected energy induced by an accumulation of canopy material. The RH metrics have also been related to biomass. In particular, the so-called HOME (Height Of Median Energy), i.e. the RH50, is usually considered [47-48] as a proxy to structure, and for instance can locally adapt height-to-biomass allometric relationship parameterized e.g. by the RH98.

The derivation of vertical structure of vegetation canopies from lidar waveform data depends on the knowledge of the relationship between lidar waveforms and the spatial structure and optical properties of vegetation canopies. Indeed, in the absence of multiple scattering, a received waveform can be expressed as a function of the projected area of canopy / background materials along the path of the laser pulse [7], [44], which is often expressed in terms of a canopy gap probability. This parameter links lidar measurements, vertical and horizontal canopy structure, the radiation regime of a plant canopy [44], [49], and enables the estimation of Plant Area Index (PAI), cover and their vertical profiles [44], [50].

NASA's large-footprint airborne instrument LVIS (Land, Vegetation, and Ice Sensor) [51] is continuously being used in many experimental campaigns, and has demonstrated the value of lidar observation for forest mapping. In parallel, there have been only a few spaceborne lidar systems dedicated to the observation of land surfaces, like the Shuttle Laser Altimeter (SLA) [52], the NASA's Ice, Cloud and land Elevation Satellite (ICESat) [53] and the current ICESat2 missions [54]. In 2018, the GEDI lidar was deployed on the International Space Station (ISS) and is

currently collecting waveform measurements (beyond the two-year mission length) over the Earth's surface between 51.6° N and 51.6° S, i.e. covering tropical and temperate forests. In contrast to airborne scanning systems such as LVIS, GEDI can only sample forest structure in a more or less dense grid which is anyway optimized to maximize geographic coverage in contrast to previous missions. For each ISS pass, the three GEDI lasers generate four beams which are dithered to collect waveform data along 8 tracks separated by about 600 m in the across-track direction. Along each track, every 60 m, waveforms with a footprint of approximately 25 m diameter are measured.

In order to achieve its scientific objectives, the GEDI mission produces a number of science data products including footprint-level and gridded data sets. A comprehensive overview is provided in [19]. The following data products are of particular interest in the context of this thesis:

- The GEDI waveforms after geolocation and correct positioning with respect to the Earth ellipsoid (Level L1B data). A representative transect of GEDI waveforms in the Lopé National Park is shown in Fig. 1.1;
- The RH metrics (Level L2A data) obtained from the waveforms after identification of the ground elevation and top canopy height. Fig. 1.2 shows the RH98 in correspondence of the footiprints in Lopé within the coverage of a TanDEM-X image acquired on Jan. 25, 2016;
- Footprint-level above ground biomass (AGB, level L4A).

The L4A footprint-level AGB is obtained from models linking the GEDI L2 waveform height metrics (especially RH98) to AGB estimated from field measurements. The candidate models change with the plant functional types (deciduous broad-leaf trees, evergreen broad-leaf trees, evergreen and deciduous needle-leaf trees, and combinations of woodlands, grasslands and shrubs) and regions (all continents except Antarctica); these models are discussed and described in detail in [55]. The model selection is then driven by the data. Operating with a relatively small number of locations available for the training and testing, and ensuring the transferability of the derived relationships outside of each location are two main challenges to be faced in the development of the AGB models [19], [55]. The first challenge has been solved by simulating thousands of large footprint training samples from discrete-return airborne lidar in correspondence of the field measurements. To overcome the second challenge, models are updated including new GEDI observations in a specific framework [19].

A gridded (continuous) mean AGB product (L4B) is then produced from the (sparse) footprint-level L4A at 1 km × 1 km resolution, together with the related estimated variance. The aggregation approach in each 1 km cell changes depending on the density of L4A AGB values. It is noted in [19], [56] that the resolution of 1 km is the coarsest resolution required by the ecosystem science community. However, the sampling nature of the GEDI mission constrain the spatial resolution of the grids that can be produced. Finer resolutions in the order of 100 m × 100 m (1 ha), which are commonly understood (and more often desired) to render a meaningful estimate of the complexity of structure and biomass variability in space, can be realized only by combining the

GEDI data with other remote sensing data characterized by high resolution (comparable to the GEDI footprint diameter), spatial acquisition continuity, and sensitivity to forest structure, like SAR data in general and TanDEM-X InSAR data in particular.

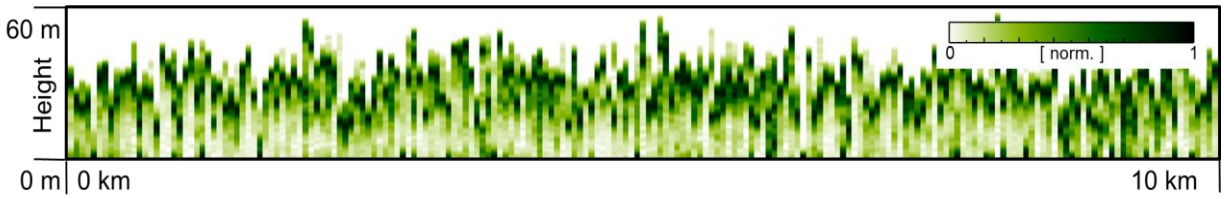


Fig. 1.1. GEDI full waveforms along a representative transect within the Lopé National Park in Gabon along the transect indicated in Fig. 1.2. The distance between consecutive profiles is about 60 m. The height axis is referred to the ground (0 m). Each profile is normalized by its maximum amplitude. Dark green indicates higher reflectance and white indicates lower reflectance.

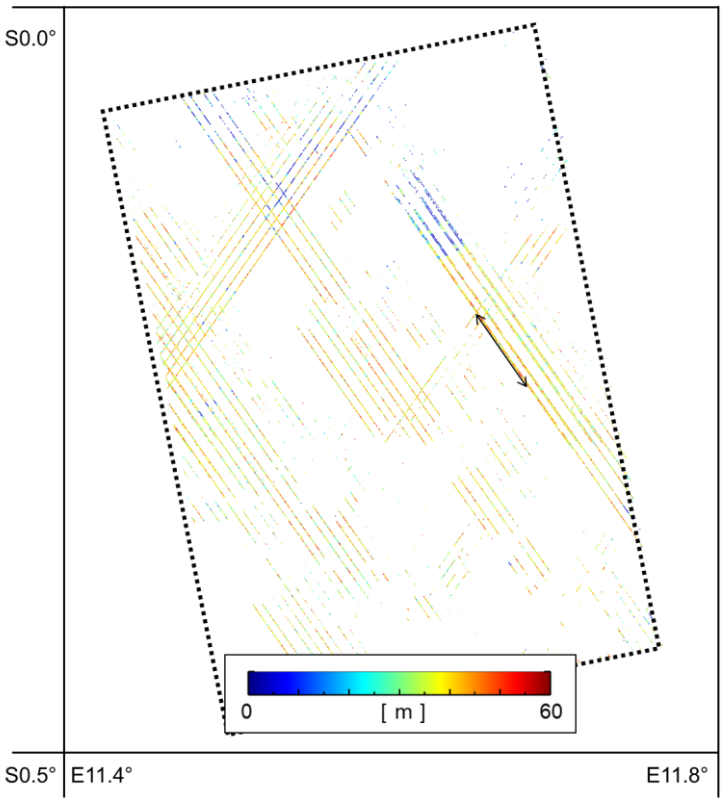


Fig. 1.2. Lopé: GEDI RH98 map in lat-lon coordinates in correspondence of footprint positions of the available GEDI measurements with 25 m resolution (footprint diameter), covering around 45 km × 60 km, corresponding to the area of the TanDEM-X scene acquired on Jan. 25, 2016 (dotted border). The black arrow indicates the representative transect whose waforms are shown in Fig. 1.1.

1.2.2 INTERFEROMETRIC SAR MEASUREMENTS: THE TANDEM-X MISSION

Similar to lidar sensors, SAR sensors transmit electromagnetic pulses, but in the microwave region of the electromagnetic spectrum corresponding to wavelengths in the order of cm to dm, and in a side-looking geometry. The amplitude and phase of the received backscattered pulses is measured with resolution on ground often between 1 m and 10 m after the image formation process. Both amplitude and phase depend on the geometric and dielectric properties of the physical scatterers on ground, their distribution within the antenna footprint (i.e. the illuminated area/volume), and the transmit-receive polarization. The side-looking geometry complicates additionally the interpretation of a single SAR image, but at the same time it allows the realization of wide footprints, i.e. swath widths up to hundreds of kilometers, and the separation of the scatterers in ground range direction. Spaceborne SAR configurations of the latest generation can therefore realize revisit times between 1 and 2 weeks [18].

Radar pulses penetrate more and more through vegetation layers at the decrease of frequency (increase of the wavelength), thus interact with vegetation elements at different heights and with the underlying ground. However, a single SAR image (or a set of SAR images in different polarimetric channels), even if it results from the interaction of the transmitted pulse(s) with the whole 3D forest structure, does not allow the reconstruction of the properties of the 3D distribution of scatterers within the illuminated volume. Fig. 1.3 shows TanDEM-X SAR image amplitudes of a scene over Lopé acquired on Jan. 25, 2016, at X-band (wavelength 3.1 cm). In this case, the saturation of the backscattered amplitudes does not allow to distinguish among forest stands with different characteristics, and even between forested and bare areas. To remove this limitation, a set of SAR images acquired under (slightly) different angular directions (i.e. incidence angles) is required in the context of InSAR and tomographic SAR measurements [18].

The main InSAR measurement is a complex interferometric coherence $\tilde{\gamma}(\kappa_z)$ formed by using the two images S_1 and S_2 acquired at a given polarization with a given spatial, i.e. baseline, and temporal separation [22]. It is a measure of similarity of the two images and can be written as:

$$\tilde{\gamma}(\kappa_z) = \frac{< S_1 \cdot S_2^* >}{\sqrt{< S_1 \cdot S_1^* >< S_2 \cdot S_2^* >}} \tag{1.2}$$

where <…> denotes the expected value. The dependency on range and azimuth in (1.2) and in the following has been dropped for simplicity. κ_z is the so-called vertical wavenumber and is linearly proportional to the change of incidence angle induced by the spatial separation (baseline) between the two images.

In conventional interferometric applications, κ_z expresses the sensitivity of the interferometric phase to (terrain) height variations. $\tilde{\gamma}(\kappa_z)$ includes several decorrelation contributions and can be re-written as [22]:

$$\tilde{\gamma}(\kappa_z) = \tilde{\gamma}_{Tmp} \cdot \tilde{\gamma}_{Sys} \cdot \tilde{\gamma}_{Scat}(\kappa_z). \tag{1.3}$$

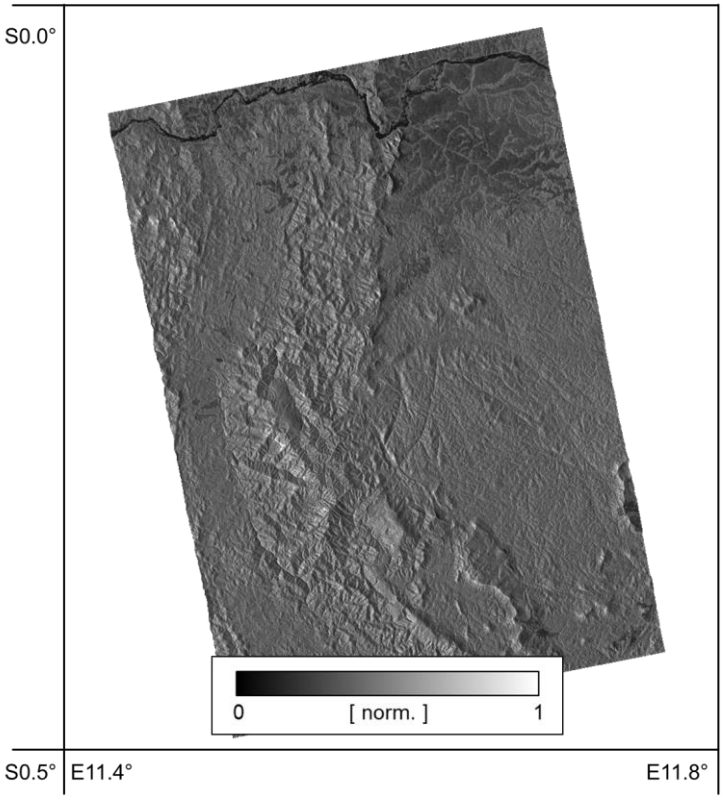


Fig. 1.3. Lopé: TanDEM-X (normalized) SLC amplitude image acquired on Jan. 25 2016. The map is in lat-lon coordinates, with resolution 25 m × 25 m (13 × 12 looks in range and azimuth), and with coverage around 45 km × 60 km.

$\tilde{\gamma}_{Tmp}$ arises from geometric and/or dielectric changes of the scatterers in the time interval between the two interferometric acquisitions. $\tilde{\gamma}_{Sys}$ comprises a wide range of decorrelation effects related to the non-ideality of the SAR system and processing implementations. $\tilde{\gamma}_{Scat}(\kappa_z)$ is originated by the phase stability of the scatterers under the different incidence angles induced by the InSAR baseline. After range and azimuth spectral filtering [34] $\tilde{\gamma}_{Scat}(\kappa_z)$ becomes the volume decorrelation contribution $\tilde{\gamma}_{Vol}(\kappa_z)$ [57]:

$$\tilde{\gamma}_{Vol}(\kappa_z) = \exp(i\kappa_z z_0) \cdot \frac{\int_0^{H_V} f_V(z) \exp(i\kappa_Z z)\, dz}{\int_0^{H_V} f_V(z)\, dz} \tag{1.4}$$

where once again z_0 is the ground height, H_V is the (top) forest height, and $f_V(z)$ is the vertical reflectivity profile, expressing the vertical distribution of scatterers seen by the interferometer. In (1.4), the phase of $\tilde{\gamma}_{Vol}(\kappa_z)$ converted to height, i.e. the so-called phase center height, corresponds to the "center of mass" of $f_V(z)$. The shape of $f_V(z)$ effects also the absolute value $|\tilde{\gamma}_{Vol}(\kappa_z)|$: the concentration of $f_V(z)$ around a height makes a $|\tilde{\gamma}_{Vol}(\kappa_z)|$ very close to 1 and almost constant at the increase of κ_z, while a more extended $f_V(z)$ makes $|\tilde{\gamma}_{Vol}(\kappa_z)|$ lower than 1, and typically

decreasing at the increase of κ_z. This sensitivity enables the retrieval of structural information from a single (or a limited number of) coherence(s) by means of a model parameterized in terms of geometric and scattering parameters. Obviously, the validity of the model is critical for the physical significance of the extracted parameters. Fig. 1.4 shows the TanDEM-X interferometric coherences over Lopé obtained over the same scene in Fig. 1.3. If many $\tilde{\gamma}_{Vol}(\kappa_z)$ are available with an appropriate diversity and distribution of κ_z, a direct tomographic inversion of the full $f_V(z)$ can be attempted without models [18].

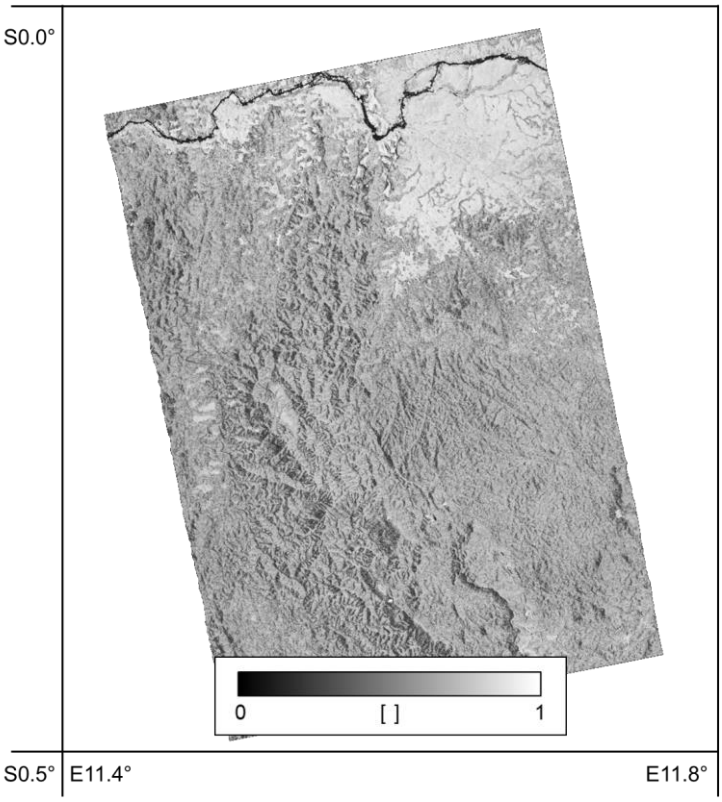


Fig. 1.4. Lopé: TanDEM-X coherence image obtained over the same scene of Fig. 1.3. The map is in lat-lon coordinates, with resolution 25 m × 25 m (13 × 12 looks in range and azimuth), and with coverage around 45 km × 60 km.

Since its launch in 2010 and still today, the DLR's TanDEM-X mission (with resolution around 2 m in range and azimuth, respectively, and swath width of 30 to 50 km) allows single-pass InSAR measurements at X-band of $\tilde{\gamma}(\kappa_z)$ from space in a bistatic configuration [20-21]. In the TanDEM-X bistatic mode, one of the two satellites transmits and both satellites receive the scattered signal quasi simultaneously (with temporal differences in the order of a fraction of a second). For this reason, $\tilde{\gamma}_{Tmp} = 1$, and after compensation of the $\tilde{\gamma}_{Sys}$, $\tilde{\gamma}_{Scat}(\kappa_z)$ and in turn $\tilde{\gamma}_{Vol}(\kappa_z)$ are immediately available. Fig. 1.6 shows the distribution of $|\tilde{\gamma}_{Vol}(\kappa_z)|$ over the Amazon forest as measured by TanDEM-X at κ_z values with no significant variations. Compared to Fig 1.5, it is apparent that lower values of $|\tilde{\gamma}_{Vol}(\kappa_z)|$ in forested areas reflect at the first order a larger extent of $f_V(z)$ in height as a consequence of a larger top height.

The estimation of forest height [i.e. H_V in (1.4)] is by far the most assessed application of TanDEM-X data in forestry [25], [58-60]. Because of the high TanDEM-X resolution, height maps can be obtained already with resolution in the order of 10 m × 10 m. Dual-polarised TanDEM-X InSAR measurements enable forest height estimation without any a priori information [22]. The achieved estimation error is between 10% and 20% if sufficient capabilities to "see" the entire volume extent in height are guaranteed by the penetration at X-band. If this condition is not met, taller stands are systematically underestimated. However, dual-pol data are available only over selected test sites. If the underlying topography is known, forest height can be estimated even from single-polarimetric InSAR coherences [22]. Despite this inversion scheme is constrained to flat areas or where a digital terrain model (DTM) is available, it can be implemented using data of the standard TanDEM-X acquisition mode, enabling large scale applications [23]. Remarkably, the knowledge of the terrain height leads to unbiased height estimates and the visibility of the full volume extent is not required anymore. But a DTM is not available for most of the forest regions. In this case, remembering equation (1.4) the inversion of forest height can still be performed for the absolute value of the measured volume coherence, which does not depend anymore on the interferometric phase induced by the terrain height. In order to obtain a determined inversion problem (meaning one unknown, H_V, for one measurement, $|\tilde{\gamma}_{Vol}(\kappa_z)|$), the full $f_V(z)$ must now be assumed or approximated from other sources of information. The choice of a suitable shape is critical for controlling the inversion bias (primarily for the taller stands). A typical (but not the only one) assumption is a uniform $f_V(z)$ ("box"-shaped) from the ground up to H_V [61-63]: the inversion can be performed even by means of closed-from expressions, however the obtained values can show a bias (underestimation) even when X-band penetration is not problematic.

In contrast, the characterization of forest structure is not as advanced as the estimation of forest height, mainly because any physical interpretations of InSAR coherences is in an early stage of development. Scattering models can include descriptors of structural properties [64]. For instance, the application of the interferometric water cloud model as well as a two-level model has been shown in boreal forest to allow the retrieval of a forest density index [25], [65], stem volume [65-67], and to describe growth and management actions [68] from TanDEM-X data. Again, the low dimensionality of the observation space (one complex InSAR coherence corresponds to two measurements – the amplitude and the phase) and the introduction of a structure parameter additional to height require the knowledge of the terrain height to obtain a determined inversion problem. On the other hand, the projection of the 3D structure and dielectric characteristics onto one single parameter implied by the InSAR models may also lead to ambiguous physical interpretations. A second approach relies on the use of allometric relationships to forest height only. An attempt in this context regards e.g. the estimation of stem volume at plot level (500 m^2) in a temperate forest [69]. But the parameterization of this kind of relationships requires the availability of an appropriate amount of specific ground measurements sampling uniformly all the different conditions. A third investigated approach aims at using the estimated top heights to characterize horizontal structure by quantifying the variability in height of the top canopy "surface" at a certain resolution (e.g. 100 m × 100 m) [70]. A high variability indicates

(horizontally) sparse stands, while a low(er) variability indicates dense(r) ones. It is worth noting that this structural information is complementary to the one in the vertical direction provided by the GEDI waveform samples. The horizontal structure definition based on top height variability, together with an evaluation of the mean height within a stand, has been used in [33] to identify growth stages in a test site (Tapajos) in the Amazon forest with a resolution in the order of 50 m × 50 m. Importantly, the sensitivity of X-band to the top canopy layer makes possible the quantification of the top canopy height variability just by using the TanDEM-X phase center height after compensation of the terrain height. This observation is at the basis of the approach in [32], in which the top canopy variability is expressed by extending a horizontal structure index originally developed for tomographically reconstructed profiles [70] at longer wavelengths to the TanDEM-X phase center heights. Interestingly, this index has been found to be in a close relationship with the well-established stand density index [70-71]. The quantification of forest structure is now fully independent of specific scattering models and allometric relationships, and is expected to make a generalization and suitability to different ecosystems easier. Two-point statistics can be employed as well. Indeed, in [31] a wavelet analysis at different scales was used to separate stands with different structural characteristics within a tropical test site. The analysis in [31] also points out that if the terrain height is not compensated from the TanDEM-X phase center heights, the retrieved top canopy height variations are not affected by terrain height variations for scales (resolutions) smaller than 10 m × 10 m (order of magnitude). However, measurements of the phase center height at these fine scales are affected by the phase variance induced by the interferometric decorrelation, therefore operating at larger scales (resolutions) is preferred. In this case, the compensation of the terrain topography is needed to distinguish a wider range of structure types. Independently of the approach, strategies not relying on the knowledge of terrain height are not available in the literature, and they are a critical step to extend structure characterization to wide coverages, accepting in this case some performance degradation.

Finally, the use of InSAR measurements for estimating biomass is critical in order to overcome backscatter saturation occurring already for very low biomass levels at X-band. Several attempts have been reported in the literature with TanDEM-X data. Change of digital elevation model heights (correlating to top height changes) have been related to biomass changes [59-60], [72]. In parallel, allometric relationships at stand level relating top canopy height and biomass (which are also common in forestry) have been used in [24], [28-29], [73] in boreal forests. In [29], [73], the forest density as a result of the inversion of a (simplified) coherence model has been included to account for the spatial variability of the allometric coefficients, as already mentioned under the assumption of a known terrain height. An alternative approach to overcome the inversion of a model considers empirical direct relationships between the complex coherence [30] or its absolute value [74] and biomass. The general argument for this kind of relationships is that the coherence contains information of both the volume height and its (vertical) structure. While it is understood that height is the main contribution to volume coherence [26], the contribution of (vertical) structure is maybe lower. A simple coherence-to-biomass relationship does not distinguish between the two contributions, and its physical interpretability is not straightforward. The observed overall correlations between the biomass estimated from the TanDEM-X data and the

one estimated from field measurements are promising, but the errors are higher than in model-based physical approaches. Yet, both model-based and empirical approaches have the common issues of the initialization of the relationship parameters, and of their adaptation to local scales.

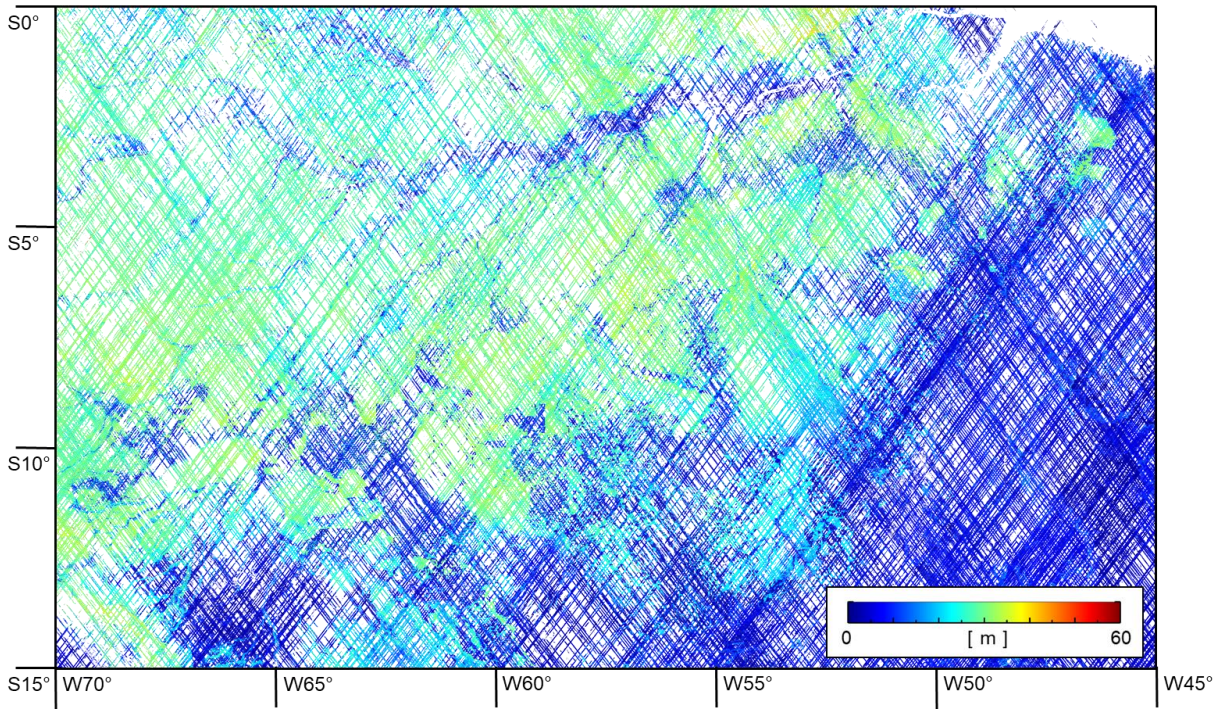


Fig. 1.5. GEDI RH98 map in meters (m) over the Brazilian Amazon forest. The map is in geographic coordinates in lat-lon, with resolution of 25 m × 25 m. The acquisition period is from 19th mission week (2019) to 130th mission week (2021).

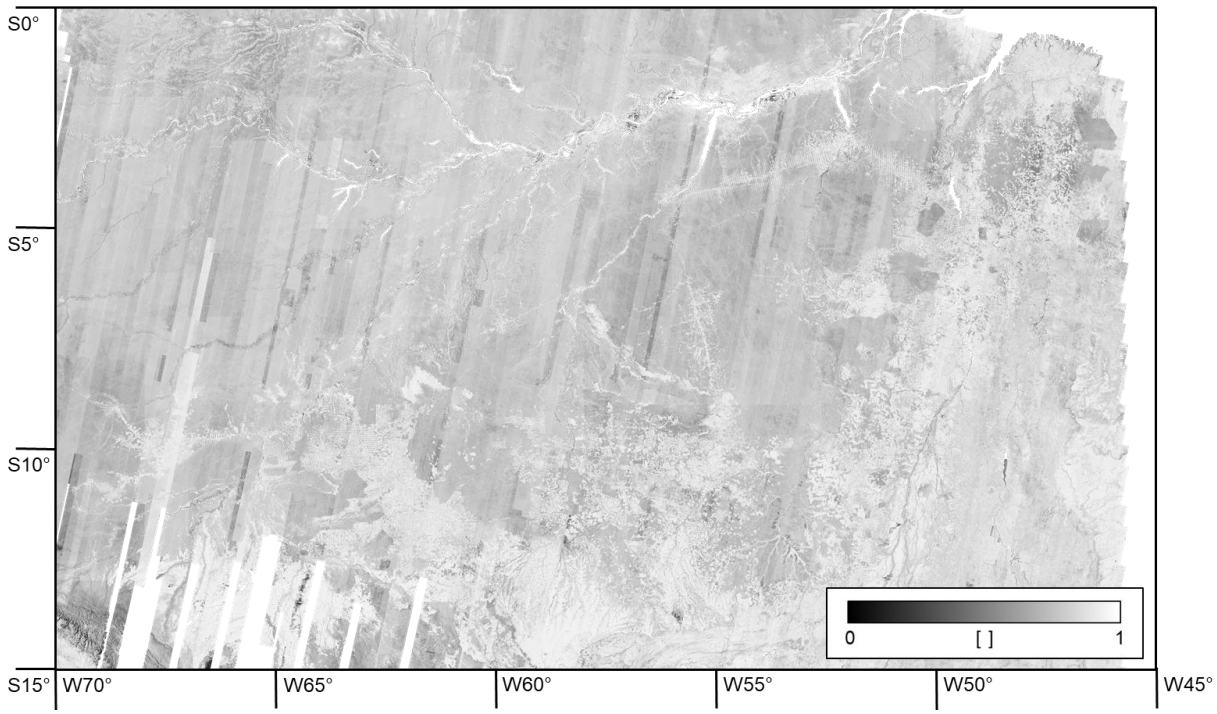


Fig. 1.6. TanDEM-X volume coherence $\tilde{\gamma}_{Vol}(\kappa_z)$ map over the Brazilian Amazon forest with the same extension as in Fig. 1.5. The acquisition period is between 2010 and 2020.

1.2.3 LIDAR – SAR COMBINATION

The development of strategies for the combination of lidar and SAR measurements over forest scenarios critically depends on the ability to relate the lidar reflectance profiles and the radar reflectivity profiles for the same physical distribution of vegetation elements. This is in general not established today at all the radar frequencies. Nevertheless, a few approaches from space have been tested using SRTM data in the past, and TanDEM-X data more recently.

Statistical regressions are an obvious way to overcome the lack of modelling understanding. Several examples for such an approach have been reported in the literature, for example aiming at biomass estimation by combining lidar heights and SAR backscatter and / or interferometric parameters [see e.g. 75-79].

On the other hand, lidar and interferometric SAR measurements can be linked more systematically by means of a scattering model. For instance, lidar data and/or lidar derived products have been used to solve the inherent underdetermination of the forest height inversion problem from single-polarimetric TanDEM-X InSAR coherences. The most basic combination consists in using the lidar-derived DTM to directly enable the forest height inversion, see Section 1.2.2 [22-25], [73]. Such a combination can exploit even the temporal [23] complementarity of lidar and SAR measurements to obtain continuous estimates of forest height and of its changes over large scales. A more advanced combination uses lidar data to constrain the forest height inversion problem reducing in this way its dimensionality [80-82]. Such approaches are supported by a certain similarity of the structural information content of lidar and TanDEM-X measurements, induced by the high sensitivity to the geometrical architecture of the canopy, the high attenuation rates and the high spatial resolution common to both configurations. In this context, very recently the use of lidar waveforms to directly approximate the TanDEM-X reflectivity in (1.4) has been proposed [26], [83] or to train models [84]. Experimental results demonstrate an adequate performance achieved with respect to fixed reflectivity profiles (see the discussion in Section 1.2.2) by using the absolute value of the InSAR coherence, therefore not needing the knowledge of the terrain height. For such application, lidar waveforms appear a sufficiently good approximation of TanDEM-X reflectivity profiles, even if there might exist significant devaitions. Clearly, the underestimation of forest heights as a result of a limited X-band penetration remains as a serious error contribution for tall and/or dense forest conditions.

Combination strategies based on the exchange of structural characterization between lidar and SAR, and in particular TanDEM-X, measurements have not been developed so far. A reason for this is certainly the fact that the interpretation of the SAR reflectivity profiles underlying the InSAR coherences in terms of physical forest structure attributes is not as advanced as for lidar profiles. Nevertheless, structural characterizations could be used to better select the lidar profile to be used for height inversion from TanDEM-X data if only a limited set of lidar profiles is available for a TanDEM-X scene (like in the GEDI case), or to initialize the compensation of penetration biases in the height estimates, or even to guide / adapt the creation of allometric relationships for biomass estimation.

1.3 RESEARCH OBJECTIVES

Given the scientific motivation described in Section 1.1, and the current challenges described in the state-of-the-art in Section 1.2, this thesis has the objective of proposing a novel, systematic and synergetic combination of spatially discrete GEDI lidar waveforms and continuous TanDEM-X InSAR coherences, by developing the necessary algorithms for

- estimating spatially continuous and unbiased large scale forest height maps with resolution of 1 ha or less,
- using the estimated height in a (stand-level) height-to-biomass allometric relationship adapted and improved at local scales using a horizontal structure index derived from the TanDEM-X data.

The proposed combination framework does not rely on any terrain height information. Thus, it can be used to derive forest height, structure and biomass estimates at large scales, contributing to fill an important application gap as well.

Forest heights are initially estimated from the absolute value of the TanDEM-X coherences using the approach in [26] by initializing the inversion using a “mean” profile within a TanDEM-X scene derived from the corresponding set of GEDI waveforms. After the inversion, the resulting height map is at a fine resolution (25 m) equivalent to the GEDI footprint diameter. Estimation biases are counteracted in two steps. First, a remaining profile mismatch is accounted for by compensating at fine resolution any residual global bias by linear regression of vertical wavenumber – height products (TanDEM-X estimated vs GEDI measured). Second, a TanDEM-X horizontal structure index is used to compensate for a local residual penetration-induced bias by deriving a mean height at a coarser resolution (100 m). This index expresses top forest height variations (at 100 m resolutions), and requires the spatially continuous and high-resolution measurements that only TanDEM-X can provide in the combination. It is worth remarking that while the first compensation makes use of a global relationship built up using the GEDI heights, the subsequent averaging process is carried out at smaller scales and relies on TanDEM-X data only.

Next, biomass is estimated (at 100 m resolution) from a height-to-biomass allometric relationship in which an allometric factor is multiplied by an exponential function of height parameterized by the allometric exponent. The allometric exponent depends on species composition and growth conditions, and is reasonably assumed constant even at large scales. The allometric factor depends on anthropogenic or natural variations in stand density resulting from differences in basal area, age composition but also thinning operations, or disturbance effects, and changes at significantly smaller scales. The values of both the allometric exponent and factor are here derived by using the GEDI heights and biomass at footprint level. The height-to-biomass relationship is then improved by making the allometric factor vary with the TanDEM-X horizontal structure index defined above, while the exponent is kept constant. The GEDI heights and biomass are then used in this case to derive the relationship linking the allometric factor to the horizontal

structure index. The use of such an index is motivated by the fact that it indeed reflects a horizontal variability of forest heights expressing density.

The research goals associated with the thesis objective can be stated by means of six relevant questions that define both the novel aspects and the proposed solutions in this thesis:

Q1. How far can a horizontal forest structure index derived from TanDEM-X data be used to locally adapt the height-to-biomass allometry in heterogeneous forests, and improve biomass estimation performance?

Q2. In which way can a horizontal forest structure index be estimated from TanDEM-X data in the absence of a DTM?

Q3. What is the role of a horizontal forest structure index in compensating the forest height estimation bias in dense (tropical) forests in the absence of a DTM?

Q4. Which is the ability of the GEDI waveform sampling in the parameterization of a height-to-biomass allometry as a function of resolution and sampling density?

Q5. In which way can GEDI waveforms and heights initialize and/or correct TanDEM-X forest height inversion in the absence of a DTM?

Q6. Which performance can be achieved in height estimation by the implemented GEDI-TanDEM-X combination for wide areas (e.g. country-wide) applications?

1.4 ORGANIZATION OF THE THESIS

The research questions presented in Section 1.3 have been addressed in Chapters 2, 3, and 4. These Chapters consist of peer-reviewed published papers. Each of them can be read and understood separately.

Chapter 2 “Improving Forest Height-To-Biomass Allometry with Structure Information: A TanDEM-X Study” answers question Q1 in the ideal case in which a DTM is available. Under this assumption, a horizontal forest structure index previously derived for tomographic SAR profiles [70] is extended to the TanDEM-X InSAR case following the first experiences in [32]. Then, a continuous relationship between the allometric factor defining the forest height-to-biomass allometry and the structure index is reconstructed from the available lidar data. The experimental results show the appropriateness of TanDEM-X data for characterizing structure and in this way improving the biomass estimation performance. The experimental analysis has been carried out by processing LVIS lidar data and TanDEM-X data acquired during the AfriSAR campaign over three test sites (Lopé, Mondah, Mabounie) in Gabon.

In Chapter 3 “Forest Biomass Mapping Using Continuous InSAR and Discrete Waveform Lidar Measurements: A TanDEM-X / GEDI Test Study”, the questions Q1 in absence of a DTM, Q2, Q3, and Q4 are addressed. The estimation of forest height and horizontal structure and further biomass from TanDEM-X data in the absence of a DTM is discussed. The possibility of estimating top canopy height variations independently of topographic height variations is investigated using a wavelet-based scale analysis. The gained understanding is used to define a new horizontal forest structure index with physical meaning similar to the one presented in Chapter 2. This index is used, first, to compensate for the underestimation of forest height in dense stands, and second to adapt the height-to-biomass allometry. The ability of the available GEDI sampling to parameterize this (adaptive) relationship is evaluated as well. TanDEM-X and GEDI data over the Lopé test site in Gabon have been processed in the experimental analysis. LVIS data have been used to validate the results.

Chapter 4 “Large Scale Forest Height Mapping by Combining TanDEM-X and GEDI data” answers Q5 and Q6. This Chapter addresses the potential of the GEDI-TanDEM-X combination for large scale high-resolution forest height mapping. The factors affecting the height estimation performance are discussed in the context of the combination, and proper performance criteria are used to ensure a certain estimation quality thoughout the large scale inversion. A way to compensate for the height estimation bias is also proposed using the vertical wavenumber – height product. The experimental analysis has been carried out over the forests of Tasmania, and the height estimation performance is quantitatively assessed by comparison with discrete-return lidar height measurements.

Finally, Chapter 5 draws the conclusions of this thesis by summarizing the answers to the research questions. Possible future research directions originated by the work presented in this thesis are also discussed.

1.5 REFERENCES

[1] H. Pretzsch, "Forest Dynamics, Growth and Yield: From Measurement to Model," Berlin, Germany: *Springer*, 2009.

[2] T. Mette, "Forest biomass estimation from polarimetric SAR interferometry," Ph.D. dissertation, *Tech. Univ. München*, 2007.

[3] T. Mette, K. Papathanassiou, I. Hajnsek, H. Pretzsch, and P. Biber, "Applying a common allometric equation to convert forest height from Pol-InSAR data to forest biomass," *IEEE International Geoscience and Remote Sensing Symposium*, Anchorage, USA, pp.269–272, 2004.

[4] G. G. Parker, "Structure and microclimate of forest canopies," *In Forest Canopies - A Review of Research on a Biological Frontier*, Elsevier Academic Press, pp.73–106, 2004.

[5] M. A. Lefsky, D. J. Harding, W. B. Cohen, G. G. Parker, H. H. Shugart, "Surface lidar remote sensing of basal area and biomass in deciduous forests of eastern Maryland, USA," *Remote Sens Environ*, vol.67, pp.83–98, 1999.

[6] T. A. Spies, "Forest structure: A key to the ecosystem," *Northwest Sci.*, vol.72, pp.34–36, 1998.

[7] D. J. Harding, M. A. Lefsky, G. G. Parker, and J. B. Blair, "Laser altimeter canopy height profiles—Methods and validation for closed-canopy, broadleaf forest," *Remote Sens. Environ.*, vol.76, pp.283–297, 2001.

[8] M. Snyder, "What is forest stand structure and how is it measured?" *North. Woodl.*, vol.64, 2010.

[9] F. G. Hall, K. Bergen, J. B. Blair, R. O. Dubayah, R. Houghton, G. Hurtt, J. Kellndorfer, M. Lefsky, J. Ranson, S. Saatchi, H. H. Shugart, and D. Wickland, "Characterizing 3D vegetation structure from space: Mission requirements," *Remote Sens. Environ.*, vol.115, pp.2753–2775, 2011.

[10] N. Brokaw and R. Lent, "Vertical Structure. In Maintaining Biodiversity in Forest Ecosystems," *Cambridge University Press*, pp.373–399, 1999.

[11] C. McElhinny, P. Gibbons, C. Brack, and J. Bauhus, "Forest and woodland stand structural complexity: Its definition and measurement," *For. Ecol. Manag.*, vol.218, pp.1–24, 2005.

[12] F. J. Bohn, A. Huth, "The importance of forest structure to biodiversity-productivity relationships," *R. Soc. Open Sci.*, 2017

[13] K. M. Bergen, S. J. Goetz, R. O. Dubayah, G. M. Henebry, C. T. Hunsaker, M. L. Imhoff, R. F. Nelson, G. G. Parker, and V. C. Radeloff, "Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions," *J. Geophys. Res.*, vol.114, 2009.

[14] S. J. Goetz, D. Steinberg, R. O. Dubayah, and B. Blair, "Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest," *Remote Sens. Environ.*, vol.108, pp.254–263, 2007.

[15] W. Turner, S. Spector, N. Gardiner, M. Fladeland, E. Sterling, and M. Steininger, "Remote sensing for biodiversity science and conservation," *Trend Ecol. Evol.*, vol.18, pp.306–314, 2003.

[16] M. S. Ashton, M. L. Ytrrell, and D. Spalding, "Managing Forest Carbon in a Changing Climate," *Springer Science & Business Media*, 2012.

[17] M. Pardini, J. Armston, W. Qi, S. K. Lee, M. Tello-Alonso, V. Cazcarra-Bes, C. Choi, K. Papathanassiou, R. Dubayah, and L. E. Fatoyinbo, "Early lessons on combining lidar and multi-baseline SAR Measurements for forest structure characterization," *Surv. Geophys.*, vol.40, no.4, pp.803–837, 2019

[18] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geosci. Remote Sens. Mag.* vol.1, pp.6–43, 2013.

[19] R. Dubayah, J. B. Blair, S. Goetz, L. E. Fatoyinbo, M. Hansen, S. Healey, M. Hofton, G. Hurtt, J. Kellner, S. Luthcke, J. Armston, H. Tang, L. Duncanson, S. Hancock, P. Jantz, S. Marselis, P. L. Patterson, W. Qi, and C. Silva, "The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography," *Science of Remote Sensing*, vol.1, 2020.

[20] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, and M. Zink, "TanDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol.45, pp.3317-3341, 2007.

[21] G. Krieger, M. Zink, M. Bachmann, B. Bräutigam, D. Schulze, M. Martone, P. Rizzoli, U. Steinbrecher, J. Antony, F. De Zan, I. Hajnsek, K. Papathanassiou, F. Kugler, M. Rodriguez -Cassola, M. Younis, and S. Baumgartner, P. López-Dekker, P. Prats, A. Moreira, "Tan-DEM-X: a radar interferometer with two formation-flying satellites," *Acta Astr.* vol.89 pp.83-98, 2013

[22] F. Kugler, D. Schulze, I. Hajnsek, H. Pretzsch, and K. Papathanassiou, "TanDEM-X Pol-InSAR Performance for Forest Height Estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6404–6422, 2014.

[23] H. J. Persson, H. Olsson, M.J. Soja, L.M. Ulander, and J.E. Fransson, "Experiences from large-scale forest mapping of Sweden using TanDEM-X data," *Remote sens.*, vol.9 no.12, p.1253, 2017.

[24] M. Schlund, S. Erasmi, and K. Scipal, "Comparison of aboveground biomass estimation from InSAR and LiDAR canopy height models in tropical forests," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 367–371, 2020.

[25] M. J. Soja, H. J. Persson, and L. M. H. Ulander, "Estimation of forest height and canopy density from a single InSAR correlation coefficient," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 646–650, 2015.

[26] R. Guliaev, V. Cazcarra-Bes, M. Pardini, and K. Papathanassiou, "Forest height estimation by means of TanDEM-X InSAR and waveform lidar data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3084–3094, 2021.

[27] S. Solberg, E. H. Hansen, T. Gobakken, E. Naessset, and E. Zahabu, "Biomass and InSAR height relationship in a dense tropical forest," *Remote Sens. Environ.*, vol. 192, pp. 166-175, 2017.

[28] A. Torano-Caicoya, F. Kugler, I. Hajnsek, and K. Papathanassiou, "Large-Scale Biomass Classification in Boreal Forests with TanDEM-X Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5935–5951, 2016.

[29] M. J. Soja, H. J. Persson, and L. M. H. Ulander, "Estimation of forest biomass from two-level model inversion of single-pass InSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 5083–5099, Sep. 2015.

[30] R. Treuhaft, Y. Lei, F. Gonçalves, M. Keller, J. R. D. Santos, M. Neumann, and A. Almeida, "Tropical-forest structure and biomass dynamics from TanDEM-X radar interferometry," *Forests*, vol. 8, no. 8, Jul, 2017.

[31] E. C. De Grandi, E. Mitchard, and D. Hoekman, "Wavelet based analysis of TanDEM-X and LiDAR DEMs across a tropical vegetation heterogeneity gradient driven by fire disturbance in Indonesia," *Remote Sens.*, vol. 8, no. 8, pp. 641–667, 2016.

[32] A. Pulella, P. C. Bispo, M. Pardini, F. Kugler, V. Cazcarra, M. Tello, K. Papathanassiou, H. Balzter, I. Rizaev, M. Santos, and J. R. Dos Santos, "Tropical forest structure observation with TanDEM-X data," *IEEE International Geoscience and Remote Sensing Symposium*, pp. 918-921, 2017.

[33] P. C. Bispo, M. Pardini, K. P. Papathanassiou, F. Kugler, H. Balzter, D. Rains, J. R. dos Santos, I. G. Rizaev, K. Tansey, M. N. dos Santos, and L. Spinelli Araujoi, "Mapping forest successional stages in the Brazilian Amazon using forest heights derived from TanDEM-X SAR interferometry," *Remote Sens. Environ.*, vol. 232, 2019.

[34] C. Choi, M. Pardini and K. Papathanassiou, "Quantification of horizontal forest structure from high resolution TanDEM-X interferometric coherences," *IEEE International Geoscience and Remote Sensing Symposium*, Valencia, pp. 376-379, 2018.

[35] M. Martone, P. Rizzoli, C. Wecklich, C. González, J. L. Bueso-Bello, P. Valdo, and A. Moreira, "The Global Forest/Non-Forest Map from TanDEMX Interferometric SAR data", *Remote Sens. Environ.*, vol. 205, 2018.

[36] J. Reiche, R. Lucas, A. L. Mitchell, J. Verbesselt, D. H. Hoekman, J. Haarpaintner, J. M. Kellndorfer, A. Rosenqvist, E. A. Lehmann, C. E. Woodcock, F. M. Seifert, and M. Herold, "Combining satellite data for better tropical forest monitoring", Nature Climate Change, vol. 6, no. 2, p. 120, 2016.

[37] M. Pax-Lenney, C. E. Woodcock, S. A. Macomber, S. Gopal, and C. Song, "Forest mapping with a generalized classifier and Landsat TM data," *Remote Sens. Environ.*, vol.77, pp.241-250, 2001.

[38] A. Pekkarinen, L. Reithmaier, and P. Strobl, "Pan-European forest/non-forest mapping with Landsat ETM+ and CORINE Land Cover 2000 data," *ISPRS J. Photogramm. Remote Sens.*, vol.64, pp.171-183, 2009.

[39] R. E. Kennedy, Z. Yang, and W. B. Cohen, "Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms," *Remote Sens. Environ.*, vol.114, pp.2897-2910, 2010.

[40] T. N. Carlson, and D. A. Ripley, "On the relation between NDVI, fractional vegetation cover, and leaf area index," *Remote Sens. Environ.*, vol.62, pp.241-252, 1997.

[41] H. G. Jones and R. A. Vaughan, "Remote sensing of vegetation: principles, techniques, and applications." *Oxford university press*, 2010

[42] P. S. Thenkabail and J. G. Lyon, "Hyperspectral remote sensing of vegetation." *CRC press*, 2016.

[43] R. O. Dubayah and J. B. Drake, "Lidar Remote Sensing for Forestry," *J. Forestry*, vol.98 pp.44-46, 2000.

[44] W. Ni-Meister, D. L. B. Jupp, and R. O. Dubayah, "Modeling lidar waveforms in heterogeneous and discrete canopies," *IEEE Trans. Geosci. Remote Sens.*, vol.39, pp.1943–1958, 2001.

[45] R. O. Dubayah, S. L. Sheldon, D. B. Clark, M. A. Hofton, J. B. Blair, G. C. Hurtt, and R. L. Chazdon, "Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica," *J. Geophys. Res.*, vol.115, 2010.

[46] A. S. Whitehurst, A. Swatantran, J. B. Blair, M. A. Hofton, and R. Dubayah, "Characterization of canopy layering in forested ecosystems using full waveform lidar," *Remote Sens.*, vol.5, pp.2014-2036, 2013.

[47] J. B. Drake, R. O. Dubayah, D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton, R. L. Chazdon, J. F. Weishampel, and S. D. Prince, "Estimation of tropical forest structure characteristics using large-footprint lidar," *Remote Sens. Environ.*, vol.79, pp.305-319, 2002.

[48] J. Armston, H. Tang, S. Hancock, S. Marselis, L. Duncanson, J. Kellner, M. Hofton, J.B. Blair, T. Fatoyinbo, and R.O. Dubayah, AfriSAR: Gridded Forest Biomass and Canopy Metrics Derived from LVIS, Gabon, ORNL DAAC, Oak Ridge, Tennessee, USA. 2016 doi: https://doi.org/10.3334/ORNLDAAC/1775.

[49] J. Armston, M. Disney, P. Lewis, P. Scarth, S. Phinn, R. Lucas, P. Bunting, and N. Goodwin, "Direct retrieval of canopy gap probability using airborne waveform lidar," *Remote Sens. Environ.*, vol.134, pp.24–38, 2013.

[50] H. Tang, R. Dubayah, A. Swatantran, M. Hofton, S. Sheldon, D. B. Clark, and B. Blair, "Retrieval of vertical LAI profiles over tropical rain forests using waveform lidar at La Selva, Costa Rica," *Remote Sens. Environ.*, vol.124, pp.242-250, 2012.

[51] J. Blair, D. Rabine, and M. Hofton, “The laser vegetation imaging sensor: a medium altitude, digitization-only, airborne laser altimeter for mapping vegetation and topography,” *ISPRS J. Photogramm. Remote Sens.*, vol.54, pp.115-122, 1999.

[52] J. Garvin, J. Bufton, J. Blair, D. Harding, S. Luthcke, J. Frawley, and D. Rowlands, “Observations of the earth’s topography from the Shuttle Laser Altimeter (SLA): laser-pulse echo-recovery measurements of terrestrial surfaces,” *Phys. Chem. Earth*, 1998.

[53] B. E. Schutz, H. J. Zwally, C. A. Shuman, D. Hancock, J. P. DiMarzio, “Overview of the ICESat mission,” *Geophys. Res. Lett.*, 2005.

[54] W. Abdalati, H.J. Zwally, R. Bindschadler, B. Csatho, S.L. Farrell, H.A. Fricker, D. Harding, R. Kwok, M. Lefsky, T. Markus, A. Marshak, T. Neumann, S. Palm, B. Schutz, B. Smith, J. Spinhirne, and C. Webb, “The ICESat-2 laser altimetry mission,” *Proc. IEEE*, 2010.

[55] L. Duncanson, J. R. Kellner, J. Armston, R. Dubayah, D. M. Minor, S. Hancock, S. P.Healey, P. L. Patterson, S. Saarela, S. Marselis, C. E. Silva, J. Bruening, S. J. Goetz, H. Tang, M. Hofton, B. Blair, S. Luthcke, L. Fatoyinbo, K. Abernethy, A. Alonso, H. Andersen, P.Aplin, T. R. Baker, N. Barbier, J. F. Bastin, P. Biber, P. Boeckx, J. Bogaert, L. Boschetti, P. B. Boucher, D. S. Boyd, D. F.R.P. Burslem, S. Calvo-Rodriguez, J. Chave, R. L. Chazdon, D. B. Clark, D. A. Clark, W. B. Cohen, D. A. Coomes, P. Corona, K. C. Cushman, M. E. J. Cutler, J. W. Dalling, M. Dalponte, J. Dash, S. de-Miguel, S. Deng, P. W. Ellis, B. Erasmus, P. A. Fekety, A. Fernandez-Landa, A. Ferraz, R. Fischer, A. G. Fisher, A. García-Abril, T. Gobakken, J. M. Hacker, M. Heurich, R. A. Hill, C. Hopkinson, H. Huang, S. P. Hubbell, A. T. Hudak, A. Huth, B. Imbach, K. J. Jeffery, M. Katoh, E. Kearsley, D. Kenfack, N. Kljun, N. Knapp, K. Král, M. Krůček, N. Labrière, S. L. Lewis, M. Longo, R. M. Lucas, R. Main, J. A. Manzanera, R. V. Martínez, R. Mathieu, H. Memiaghe, V. Meyer, A. M. Mendoza, A. Monerris, P. Montesano, F. Morsdorf, E. Næsset, L. Naidoo, R. Nilus, M. O’Brien, D. A. Orwig, K. Papathanassiou, G. Parker, C. Philipson, O. L. Phillips, J. Pisek, J. R. Poulsen, H. Pretzsch, C. Rüdiger, S. Saatchi, A. Sanchez-Azofeifa, N. Sanchez-Lopez, R. Scholes, C. A. Silva, M. Simard, A. Skidmore, K. Stereńczak, M. Tanase, C. Torresan, R. Valbuena, H. Verbeeck, T. Vrska, K. Wessels, J. C. White, L. J. T. White, E. Zahabu, and C. Zgraggen, “Aboveground biomass density models for NASA’s Global Ecosystem Dynamics Investigation (GEDI) lidar mission,” *Remote Sens. Environ.*, vol.270, 2022.

[56] R. O. Dubayah, J. Armston, J. R. Kellner, L. Duncanson, S.P. Healey, P.L. Patterson, S. Hancock, H. Tang, M.A. Hofton, J.B. Blair, and S.B. Luthcke. “GEDI L4A Footprint Level Aboveground Biomass Density, Version 1,” ORNL DAAC, Oak Ridge, Tennessee, USA, 2021. https://doi.org/10.3334/ORNLDAAC/1907.

[57] K. P. Papathanassiou and S. R. Cloude, "Single-baseline polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol.39, no.11, pp.2352-2363, 2001.

[58] H. J. Persson and J. E. S. Fransson, "Comparison between TanDEM-X and ALS based estimation of above ground biomass and tree height in boreal forests," *Scand. J. For. Res.*, pp.306–319, 2017.

[59] Yang Lei, Robert Treuhaft, and Fabio Gonçalves, “Automated estimation of forest height and underlying topography over a Brazilian tropical forest with single-baseline single-polarization TanDEM-X SAR interferometry,” *Remote Sens. Environ.*, vol.252, 2021

[60] S. Solberg, E. Næsset, T. Gobakken, and O. M. Bollandsås, “Forest biomass change estimated from height change in interferometric SAR height models,” *Carbon Balance and Management*, vol.9, pp.1-12, 2014.

[61] H. Chen, S. R. Cloude, D. G. Goodenough, D. A. Hill and A. Nesdoly, "Radar Forest Height Estimation in Mountainous Terrain Using Tandem-X Coherence Data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3443-3452, Oct. 2018

[62] H. Chen, S. R. Cloude and D. G. Goodenough, "Forest Canopy Height Estimation Using Tandem-X Coherence Data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol.9, pp.3177-3188, 2016.

[63] C. Gómez, J. M. Lopez-Sanchez, N. Romero-Puig, J. Zhu, H. Fu, W. He, Y. Xie, and Q. Xie, "Canopy Height Estimation in Mediterranean Forests of Spain With TanDEM-X Data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2956-2970, 2021.

[64] D. Hoekman and C. Varekamp, "Observation of tropical rain forest trees by airborne high-resolution interferometric radar," *IEEE Trans Geosci Rem Sens.*, vol.39, pp.584-594, 2001

[65] J. I. H. Askne and L. M. H. Ulander, "Boreal Forest Properties From TanDEM-X Data Using Interferometric Water Cloud Model and Implications for a Bistatic C-Band Mission," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol.14, pp.8627-8637, 2021

[66] J. I. H. Askne, H. J. Persson, and L.M.H. Ulander, "On the sensitivity of TanDEM-X observations to boreal forest structure," *Remote Sens.*, vol.11, 2019

[67] S. Erasmi, M. Semmler, P. Schall, and M. Schlund, "Sensitivity of bistatic TanDEM-X data to stand structural parameters in temperate forests," *Remote Sens.*, vol.11, 2019.

[68] M. J. Soja, H. J. Persson and L. M. H. Ulander, "Modeling and Detection of Deforestation and Forest Growth in Multitemporal TanDEM-X Data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol.11, pp.3548-3563, 2018.

[69] Sahra Abdullahi, Florian Kugler, Hans Pretzsch, "Prediction of stem volume in complex temperate forest stands using TanDEM-X SAR data," *Remote Sens. Environ.*, vol.174, pp.197-211, 2016.

[70] M. Tello Alonso, V. Cazcarra Bes, M. Pardini, and K. Papathanassiou, "Forest structure characterization from SAR tomography at L-band," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol.11, pp.3402–3414, 2018.

[71] L. Reineke, "Perfecting a stand-density index for even aged forests," *J. Agricultural Research*, vol.46, pp.627–638, 1933.

[72] N. Knapp, A. Huth, F. Kugler, K. Papathanassiou, R. Condit, S.P. Hubbell, and R. Fischer, "Model-Assisted Estimation of Tropical Forest Biomass Change: A Comparison of Approaches," *Remote Sens.*, vol.10, 2018.

[73] H. J. Persson, M. J. Soja, J. E. S. Fransson, and L. M. H. Ulander, "National Forest Biomass Mapping Using the Two-Level Model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol.13, pp.6391-6400, 2020.

[74] M. Schlund, F. von Poncet, S. Kuntz, C. Schmullius, and D. H. Hoekman, "TanDEM-X data for aboveground biomass retrieval in a tropical peat swamp forest, Remote Sensing of Environment, vol.158, pp.255-266, 2015.

[75] J. M. Kellndorfer, W. S. Walker, E. LaPoint, K. Kirsch, J. Bishop, and G. Fiske, "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+, and ancil-lary data sets," *J. Geophy.s Res.*, vol.115, 2010.

[76] S. Kaasalainen, M. Holopainen, M. Karjalainen, M. Vastaranta, V. Kankare, K. Karila, and B. Osmanoglu, "Combining lidar and synthet-ic aperture radar data to estimate forest biomass: Status and prospects," *Forests*, vol.6, pp.252-270, 2015.

[77] T. E. Fatoyinbo and M. Simard, "Height and biomass of mangroves in Africa from ICESat /GLAS and SRTM," *Int. J. Remote Sens.*, vol.34, pp.668-681, 2013.

[78] G. Sun, K. J. Ranson, Z. Guo, Z. Zhang, P. Montesano, and D. Kimes, "Forest biomass mapping from lidar and radar synergies," *Remote Sens. Environ.*, vol.115, pp.2906-2916, 2011.

[79] O. W. Tsui, N. C. Coops, M. A. Wulder, and P. Marshall, "Integrating airborne LiDAR and space-borne radar via multivariate kriging to estimate above-ground biomass," *Remote Sens. Environ.*, vol.139, pp.340-352, 2013.

[80] M. Brolly, M. Simard, H. Tang, R. O. Dubayah and J. P. Fisk, "A Lidar-Radar Framework to Assess the Impact of Vertical Forest Structure on Interferometric Coherence," *IEEE J. Sel. Top. Appl. Remote Sens. Earth Obs.*, vol.9, 2016.

[81] W. Qi, S. K. Lee, S. Hancock, S. Luthcke, H. Tang, J. Armston, and R. Dubayah, "Improved forest height estimation by fusion of simulated GEDI Lidar data and TanDEM-X InSAR data," *Remote Sens. Environ.*, vol.221, pp.621–634, 2019.
[82] W. Qi and R. O. Dubayah, "Combining Tandem-X InSAR and simulated GEDI lidar observations for forest structure mapping," *Remote Sens. Environ.*, vol.187, p.253–266, 2016.
[83] H. Chen, S. R. Cloude, and J. C. White, “Using GEDI Waveforms for Improved TanDEM-X Forest Height Mapping: A Combined SINC + Legendre Approach,” *Remote Sens.* 2021.
[84] M. Schlund, A. Wenzel, N. Camarretta, C. Stiegler, and S. Erasmi, “Vegetation canopy height estimation in dynamic tropical landscapes with TanDEM-X supported by GEDI data,” *Methods in Ecology and Evolution*, 2022.

2 IMPROVING FOREST HEIGHT-TO-BIOMASS ALLOMETRY WITH STRUCTURE INFORMATION: A TANDEM-X STUDY

C. Choi, M. Pardini, M. Heym, and K. P. Papathanassiou

IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING

Published in vol.14, pp. 10415 – 10427, September 2021. DOI: 10.1109/JSTARS.2021.3116443

This Chapter is a post-print, differing from the published paper only in terms of layout and formatting.

The author's contributions:

- Interferometric processing and analysis of the TanDEM-X data.
- Inversion of forest height and structure index from the TanDEM-X data.
- Development of biomass estimation method including the structure index from TanDEM-X
- Writing of the manuscript.

The co-authors' contributions:

- M. Pardini and K. P. Papathanassiou provided guidance throughout the research.
- M. Pardini and K. P. Papathanassiou contributed to the main ideas, the discussion of the results, and reviewed the manuscript.
- M. Heym contributed to the discussion of the forest AGB estimation from remote sensing data.

IMPROVING FOREST HEIGHT-TO-BIOMASS ALLOMETRY WITH STRUCTURE INFORMATION: A TANDEM-X STUDY

Changhyun Choi[1,2], Matteo Pardini[1], Michael Heym[3], and Konstantinos P. Papathanassiou[1]
[1] German Aerospace Center, Microwaves and Radar Institute, Wessling, Germany
[2] ETH Zurich, Institute of Environmental Engineering, Zurich, Switzerland
[3] Bavarian State Institute of Forestry, Freising, Germany

Abstract

Allometric relations that link forest above ground biomass to top forest (i.e. canopy) height are of particular significance in the context of lidar and interferometric synthetic aperture radar remote sensing, as both techniques allow accurate height measurements at ecologically relevant spatial scales. Besides the often unknown allometry itself, its spatial variation in heterogenous forest environments restricts the performance when using a single fixed height-to-biomass allometric relation. This paper addresses how forest structure information derived from interferometric TanDEM-X data can be used to locally adapt the height-to-biomass allometry in heterogeneous forests, and to improve biomass estimation performance. The analysis is carried out using TanDEM-X interferometric measurements in three tropical forest test sites in Gabon. A structure index expressing forest density is derived from the TanDEM-X data. Then, a continuous relationship between the structure index and the allometric level that defines the forest height-to-biomass allometry is reconstructed from the available lidar data, and used to vary the height-to-biomass relationship. Finally, the potential of the derived structure index to support an allometric relationship common to all sites is evaluated. The experimental results show the appropriateness of TanDEM-X data for characterizing structure and in this way improving the biomass estimation performance.

2.1 INTRODUCTION

Biomass has a direct relationship to carbon content and is a measure of forest and ecosystem productivity. Estimation of biomass is very inaccurate at local, regional and supra-regional scales. Ground measurements of biomass in natural forests often exhibit errors much greater than 20%. Particularly large are the deviations in tropical and natural forests due to their spatial heterogeneity. Dynamic changes of biomass and their spatial distribution are a direct measure of the exchange of carbon between the terrestrial ecosystem and the atmosphere [1]. At the same time, they characterize the variation of forest growth and productivity induced by water or climate stress [2].

The above ground biomass (AGB) B_T of a single tree can be expressed as the product of the tree volume V with its (species specific) wood density ρ [3], [4]. Different adequate standards of tree volume can be chosen depending on the individual application. In forestry, the focus on the dominant stem contribution suggests the use of stem volume that leads to

$$B_T = V \cdot \rho = F \cdot \left(\frac{\pi D^2}{4}\right) \cdot H \cdot \rho \tag{2.1}$$

where D is the stem diameter at breast height (also known as DBH), H is the tree height, and F is a factor that accounts for the shape of the stem [3], [4]. Accounting for the total tree volume including its leaves and branches is more common for ecological applications and carbon stock estimates.

However, the use of (2.1) in the context of remote sensing is rather limited to high resolution airborne implementations because the spatial resolution of conventional and especially spaceborne configurations does not allow us to measure single tree parameters. With a typical spatial resolution on the order of few to ten meters, such configurations can provide forest parameter estimates at some tenths of meters, representative of patches of trees or forest stands rather than of single trees. However, the transition of (2.1) from a single tree to a stand biomass relation by replacing the individual tree parameters with the stand means (e.g., mean forest height, mean diameter or alternatively basal area, and mean wood density) is not straightforward and strongly depends on the stand density and heterogeneity in terms of species and age composition [5], [6].

The potential of either spaceborne lidar or interferometric synthetic aperture radar (InSAR) configurations to measure forest height at spatial scales of 1 hectare (ha) or even below motivated the use of the so-called forest height-to-biomass allometry at stand level. The AGB B is expressed in terms of (top) height by means of an exponential relationship [4–8]

$$B = \alpha_0 \cdot H^{\beta_0} \tag{2.2}$$

where H is the maximum height within the stand area equivalent to the top canopy height, and (α_0, β_0), are the so-called allometric factor and reference (allometric) exponent, respectively. The reference exponent β_0 defines the underlying allometric relation accounting for species composition and growing conditions of a certain stand. The allometric factor accounts then for (anthropogenic and natural) density variations across stands of the same composition. When the reference exponent β_0 is fixed, α_0 is also known as the allometric level as it scales a set of height-to-biomass relations with the same reference exponent accounting for different stand densities (e.g., different basal areas at a given age) or stand ages. A successful implementation requires the two allometric parameters to be either a priori known or estimated from reference (for example inventory) measurements. Indeed, allometric relationships in form of (2.2) have been successfully used to derive biomass estimates from height estimates [6–11], however for rather homogeneous stand and forest conditions.

However, the spatial variability in stand density and/or structure reduces the performance of a single height-to-biomass relation with fixed α_0 and β_0. Accordingly, in spatially heterogenous (in terms of density and structure) forests a fixed height-to-biomass allometry leads to a poor performance or even breaks down [4–6], [9]. To visualize this insufficiency three stands with the same H are depicted in Fig. 2.1. The tree density decreases from left to right with the stand biomass decreasing accordingly. In this particular case, if only height is used and (α_0, β_0) are fixed, the same biomass is estimated across the three stands. Using allometric parameters (α_0, β_0) fitting the denser stand will overestimate the biomass in the sparser stand, and vice versa.

Fig. 2.1. The three depicted forest stands have the same top height H. Their horizontal density decreases from left to right. The height axis is referred to the ground, which is located in its origin (0 m).

Summarizing, two main factors limit the use of the height-to-biomass allometry in a wide context: (i) the large uncertainty in the knowledge of the allometric parameters (α_0, β_0) for the individual forest conditions (arising from the insufficiency or complete lack of appropriate reference measurements), and (ii) the inability to adapt (α_0, β_0) to the spatial variability of forests.

However, today both limitations appear less restrictive making a revisit of the height-to-biomass allometry attractive. Terrestrial lidar scanning (TLS) techniques have the potential to make plot inventory measurements more accurate and faster [12]. At the same time, spaceborne waveform lidar configurations sample forest height in a more or less dense grid, and provide a set of waveform metrics that allow us to estimate AGB (using empirically derived models) [13], [14]. TLS and/or spaceborne waveform lidars are able to provide enough forest height and biomass measurements to define a general height-to-biomass allometry at regional or even finer scales. Hence, they resolve, at least to a large extent, the first limitation. On the other hand, the spatial variability of allometry can be accounted with remote sensing configurations able to estimate not only forest height, but also forest structure information with a subhectare spatial resolution. Indeed, SAR interferometry [15–21] and especially SAR tomography have been proven to be able to characterize physical forest structure, the latter relying on the reconstruction of the three-dimensional (3-D) radar reflectivity [22–25].

The open question is how far the knowledge of forest structure can be used to adapt (and improve) the general height-to-biomass allometry to local scales. This paper addresses this

question in the context of TanDEM-X (i.e., bistatic interferometric acquisitions at X-band) and waveform lidar measurements. Height and biomass estimates from the waveform lidar measurements are used to establish a general height-to-biomass allometry. Then, the horizontal structure index HS as defined in [24] is reconstructed from TanDEM-X measurements and used to account for the spatial variability (by means of the allometric level) of the height-to-biomass allometry within the test site

$$B = \alpha(\mathrm{HS}) \cdot H^{\beta_0}. \tag{2.3}$$

There are two arguments for using HS to adapt α: first, its relation to the well-established stand density index [26] (and thus to basal area) as discussed and demonstrated in [24], and second, the ability to obtain HS estimates from TanDEM-X data.

Accordingly, in the following Sections, the estimation of H and HS from TanDEM-X data, the performance of conventional height-to-biomass allometry and the use of HS to improve the height-to-biomass allometry performance for different tropical forest types and conditions are discussed.

2.2 TEST SITES AND DATASET

2.2.1 TEST SITES

The AfriSAR campaign was carried out over tropical forest sites in Gabon in 2015 and 2016 [27], [28]. The objective of the campaign was to acquire air- and space-borne polarimetric SAR interferometric and tomographic data sets complemented by airborne waveform lidar [28] and field measurements [29] for the development and validation of forest height, structure and biomass estimation algorithms. Three of the AfriSAR sites are considered in this article: Lopé; Mabounie; and Mondah. Their locations are shown in Fig. 2.2.

The Lopé site is located within the Lopé National Park near the geographic center of Gabon. The site consists of a variety of structure types ranging from open savannas to undisturbed tall (sometimes exceeding 50 m) dense forest stands. Colonizing forest (sparse forest stands mixed up with savanna) or monodominant Okoume (dense, monolayered, tall and dense forest stands) are two particular cases [29]–[31]. Biomass ranges between 10 t/ha in savanna areas and 600 t/ha in dense forest areas. The terrain is hilly with many local slopes steeper than 20°.

The Mabounie site is part of the “Maboumine” mining project started in 2005. The site is covered by mature stands with canopy heights between 40 to 60 m and biomass levels up to 400 t/ha, but is signed by local degradation caused by roads, buildings, and other infrastructure. The site includes also partially flooded areas containing swamp mixed forest [28].

Finally, Mondah is a partially flooded area containing mangrove and mahogany woodlands in northwest Gabon. The western part is a primary forest taller than 50 m, with dense and

homogenous stands while the eastern part is secondary forest [28]. The topography is fairly flat within the site.

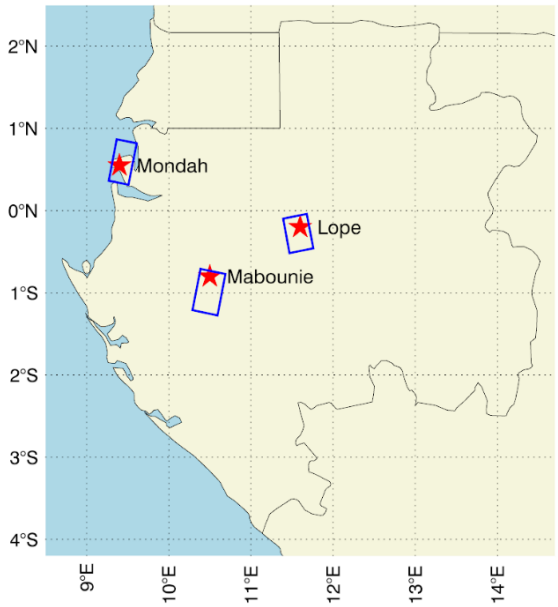


Fig. 2.2. Locations of the three selected AfriSAR test sites (red stars) and TanDEM-X acquisitions (blue rectangles) in Gabon (central Africa).

2.2.2 LIDAR ACQUISITIONS

Waveform lidar data were acquired by NASA's land and vegetation and ice sensor (LVIS) in February and March 2016 [28], [32], [33]. During AfriSAR, LVIS was operated at a nominal flight altitude of 24,000 ft (equivalent to 7315 m), and acquired data with partially overlapping footprints on ground with an average diameter around 22 m [28].

For each test site, the digital terrain model (DTM) and relative height (RH) metrics were derived from the waveforms [33]. Each RH metric expresses the height above ground at which a certain percentage of the total waveform energy is received [14]. For instance, the RH100 represents to the height above the ground in which the total energy is received. RH metrics have been used to estimate biomass and to describe the canopy vertical structure [13], [14].

In the following, the RH100 is considered as the top canopy height within a footprint, and is used as a reference for the validation of the heights estimated from TanDEM-X coherences. Consistently with the formulation in Section 2.1, the LVIS top height H_{LVIS} has been calculated by taking the maximum RH100 for every 1 ha resolution cell on ground.

Similarly, the waveform-derived AGB estimates B_{LVIS} and their uncertainty at 1 ha resolution, estimated by means of an allometric relationship linking height and RH metrics to the LVIS AGB and parameterized by using the available field inventory plots, have been used as reference [28], [34]. An overall root mean square error around 70 t/ha across all inventory plots was documented [28], [34]. LVIS RH and biomass data sets are ideal for the experiments in this article as they

cover large areas allowing to appreciate structure gradients, and provide a statistically (very) large number of samples for comparison at the desired 1 ha resolution.

Fig. 2.3 shows the maps of H_{LVIS} and B_{LVIS} for the three selected sites. The largest continuous LVIS coverage is over Lopé. H_{LVIS} is larger than 50 m in the northwestern part and lower in the southeastern part, but overall without relevant spatial gradients. At the same time, B_{LVIS} varies significantly stronger suggesting a likewise significant structure dependency. In Mabounie, many footprints are lost because of cloud coverage. In Mondah, the short stands (H_{LVIS}< 30 m) in the central part of the scene have a low biomass (B_{LVIS} < 100 t/ha). In contrast, in the western part B_{LVIS} can reach up to 700 t/ha for the taller primary stands (H_{LVIS} around 50 m).

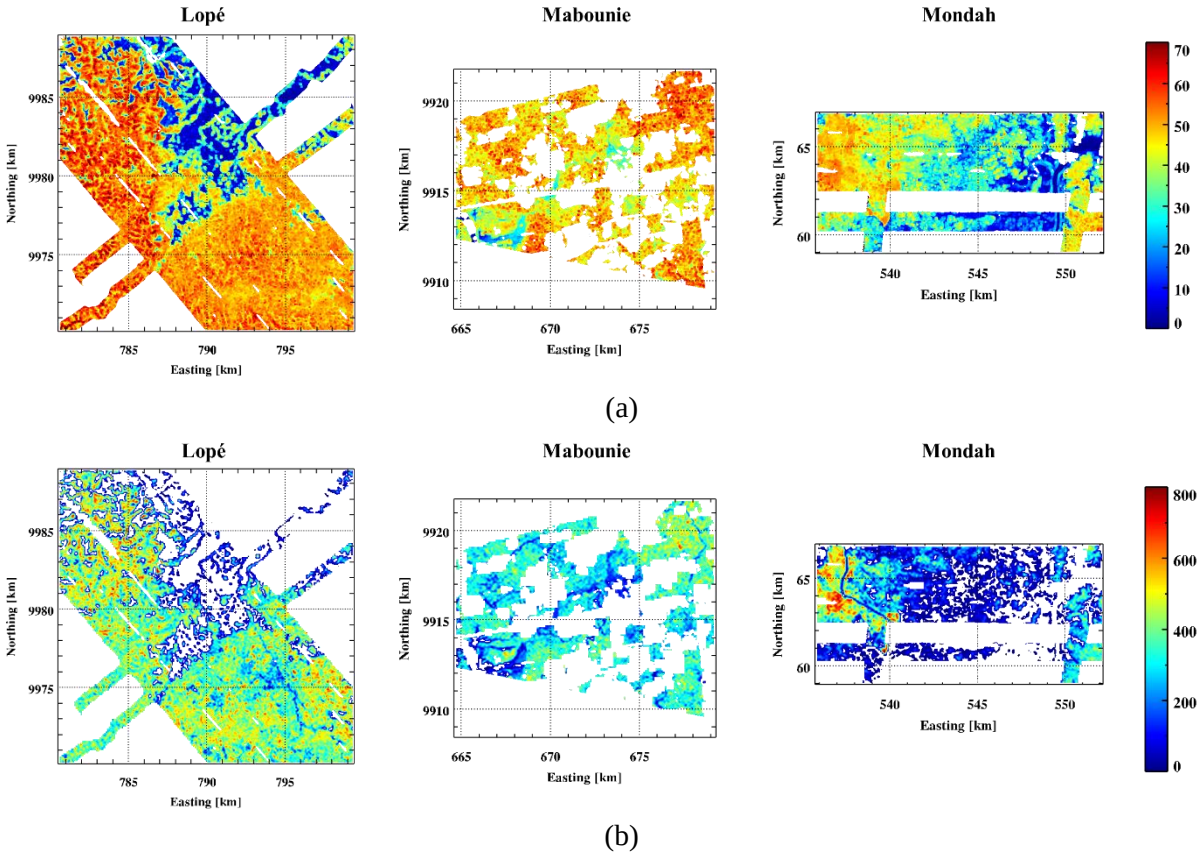


Fig. 2.3. (a) LVIS RH100 maps in meters (m), and (b) AGB maps in tons per hectare (t/ha) for Lopé, Mabounie, and Mondah. All maps are in UTM coordinates and their resolution is 100 m by 100 m in easting and northing direction. The coverage in Lopé is around 19 km by 19 km, in Mabounie is 15 km by 13 km, and in Mondah is 16 km by 8 km.

2.2.3 TanDEM-X Acquisitions

Single-polarimetric stripmap bistatic TanDEM-X data were acquired close in time to the LVIS flights. The relevant acquisition parameters are given in Table 2.1. For every data set the InSAR complex coherence has been calculated as

$$\gamma(\kappa_z) = \frac{<S_1 \cdot S_2^*>}{\sqrt{<S_1 \cdot S_1^*> \; <S_2 \cdot S_2^*>}} \tag{2.4}$$

where S_1 and S_2 are single-look complex amplitudes of the two images, and $(\cdot)^*$ and $\langle\cdot\rangle$ denote the complex conjugation and the spatial averaging operators. For forest height estimation, the spatial averaging has been performed using cells measuring 25 m × 25 m (corresponding to ∼150 independent looks) in ground range – azimuth, respectively. For the estimation of the horizontal structure index (see Section 2.4.1) a finer resolution of 10 m × 10 m (corresponding to 25 independent looks) has been chosen. Finally, biomass, horizontal structure index and top height have been calculated at the scale of 1 ha (100 m × 100 m), consistent to the reference lidar data.

TABLE 2.1
Summary of TanDEM-X Acquisitions Parameters

Test Site	Lopé	Mabounie	Mondah
Acquisition Date	Jan. 25, 2016	Oct. 5, 2015	Nov. 11, 2015
Frequency		X-band	
Polarization		HH	
Ground resolution (Range / Azimuth)	1.95 m / 1.99 m	2.26 m / 2.05 m	1.84 m / 1.87 m
Vert. wavenumber	~ 0.10 m^{-1}	~ 0.078 m^{-1}	~ 0.062 m^{-1}
HoA	~ 62.8 m	~ 80.5 m	~ 101.3 m
Incidence angle	~ 44.5°	~ 37.1°	~ 47.7°

In (2.4), κ_z represents the vertical wavenumber, which expresses the sensitivity (i.e., the derivative) of the InSAR phase difference with respect to (vertical) height [35], [36]. For the bistatic mode

$$\kappa_z = \frac{2\pi}{\mathrm{HoA}} \cong \frac{2\pi}{\lambda} \frac{B_\perp}{R \sin\theta_i} \tag{2.5}$$

where HoA is the InSAR Height Of Ambiguity, $B_\perp$ is the perpendicular InSAR baseline, λ is the radar wavelength, and R is the slant range distance. θ_i is the slope-corrected incidence angle, which corresponds to the difference between the incidence angle and the local terrain slope in range direction. The terrain slopes have been calculated using the TanDEM-X digital elevation model (DEM). The vertical wavenumber characterizes also the performance of parameter inversion from InSAR coherence measurements. In the case of forest height estimation, one single κ_z allows unbiased and accurate inversion only for a limited range of forest heights [36]. In Lopé a mean κ_z of 0.1 rad/m leads to an optimum performance range of 15 to 40 m appropriate to cover most of the forest height in the scene, however with some (significant) loss of performance for

the tallest stands with height close to the HoA (around 62 m), far above the optimum performance range. For Mabounie and Mondah, the κ_z values are lower, increasing further the HoA and shifting the optimum performance range to larger heights.

For bistatic acquisitions, the absence of temporal decorrelation leaves signal-to-noise ratio (SNR) decorrelation γ_{SNR}, range (spectral) decorrelation $\gamma_{rg}(\kappa_z)$, quantization decorrelation γ_Q and volume decorrelation $\gamma_{Vol}(\kappa_z)$ as the remaining decorrelation contributions [35]

$$\gamma(\kappa_z) = \gamma_{SNR} \cdot \gamma_{rg}(\kappa_z) \cdot \gamma_Q \cdot \gamma_{Vol}(\kappa_z). \tag{2.6}$$

Height is estimated from $\gamma_{Vol}(\kappa_z)$ obtained from $\gamma(\kappa_z)$ after compensating all the other contributions. γ_{SNR} and $\gamma_{rg}(\kappa_z)$ have been calculated and compensated using the procedure described in [35]. A fixed value of 0.97 was assumed for γ_Q according to the analysis in [37].

2.3 FOREST BIOMASS ESTIMATION FROM TANDEM-X HEIGHT

2.3.1 FOREST HEIGHT INVERSION

After coherence calibration (see Section 2.2.3), the vertical reflectivity function (i.e., the vertical distribution of scatterers) underlying $\gamma_{Vol}(\kappa_z)$ is usually represented by a two-layer model accounting for the ground and volume (back-) scattering contributions. However, at X-band and for dense(r) forest conditions, the ground contribution is often neglected so that [35], [36], [38]

$$\gamma_{Vol}(\kappa_z) = \exp(i\phi_0) \cdot \gamma_V(\kappa_z)\ . \tag{2.7}$$

The phase term $\varphi_0 = \kappa_z z_0$ is the InSAR phase corresponding to the ground height z_0 and

$$\gamma_V(\kappa_z) = \frac{\int_0^{h_V} f_V(z)\exp(i\kappa_Z z)\,dz}{\int_0^{h_V} f_V(z)\,dz} \tag{2.8}$$

where $f_V(z)$ is the volume-only vertical reflectivity function and h_V is the top volume (i.e., forest) height with respect to the ground topography. For $f_V(z)$ an exponential distribution of scatterers is widely used [35], [36], [38]

$$f_V(z) = \exp\left[\frac{2\sigma z}{\cos(\theta_i)}\right] \tag{2.9}$$

where σ is a coefficient defining the shape of the reflectivity function, interpreted as a mean extinction value.

The inversion of (2.7) with (2.8) and (2.9) using a single baseline is an underdetermined problem as the number of unknowns (φ_0, h_V, σ) exceeds the number of measurements ($\gamma_{Vol}(\kappa_z)$).

In this sense the use of an external DTM for calculating φ_0 has two advantages. First, it allows a balanced inversion. Second, it allows an unbiased estimation of h_V even if there is no penetration until the ground [35]. On the other hand, any offset (e.g., arising from residual TanDEM-X orbit errors) between φ_0 and $\gamma_{Vol}(\kappa_z)$ on surfaces must be compensated before the inversion [35].

The desired h_V is then obtained as the solution of [35]

$$\min_{h_V,\sigma} \ \|\gamma_{Vol}(\kappa_z) - \exp(i\phi_0) \cdot \gamma_V(\kappa_z; h_V, \sigma)\|. \tag{2.10}$$

Finally, similarly to the calculation of H_{LVIS}, the top height H_{TDX} is obtained by taking the maximum h_V for every 1 ha resolution cell on ground.

The inversion (2.10) was implemented and applied over the three sites using the LVIS DTMs to calculate φ_0. The inversion was carried out only for $|\gamma_{Vol}(\kappa_z)| > 0.25$ with $\kappa_z < 0.12$ rad/m (HoA > 55 m) corresponding to 85% of the forested areas in Lopé, 90% in Mabounie and 100% in Mondah. The obtained maps of H_{TDX} and the 2-D histograms validating H_{TDX} against H_{LVIS} are shown in Fig. 2.4. The overall RMSE amounts to 3.5, 3.5, and 4.1 m, and the bias amounts to −1.9, −0.1, and −2.9 m in Lopé, Mabounie, and Mondah, respectively. For all the sites, H_{TDX} covers the same height range of H_{LVIS}, and the two height maps are well correlated. The lower coherence level in taller stands especially in Lopé and Mabounie increases the standard deviation of H_{TDX}. H_{TDX} is clearly underestimated (i.e., has a negative bias) for $H_{\mathrm{LVIS}} < 30$ m in Lopé and even more in Mondah probably as a result of an actually present ground scattering contribution, constituting a model mismatch in (2.10), as these values are found in the more open canopy areas (e.g., colonizing forest) [7], [35], [36]. In Lopé, heights are slightly underestimated also for $H_{\mathrm{LVIS}} > 50$ m with a bias around −3 m. This is the effect of the relatively high κ_z that limits the sensitivity of the inversion for taller stands.

2.3.2 AGB Estimation Using Height

Here, the estimation of AGB by means of the height-to-biomass allometry is discussed. Using the LVIS height and biomass estimates the allometric level α_0 and the reference exponent β_0 of a height-to-biomass allometric relation with the form of (2.2) have been estimated by means of an ordinary least square regression

$$\min_{\alpha_0,\beta_0} \left\| \mathbf{b}_{LVIS} - \alpha_0 \mathbf{h}_{LVIS}^{\beta_0} \right\|^2, \tag{2.11}$$

where $\mathbf{b}_{LVIS}$ is a vector containing the biomass values b_{LVIS} and $\mathbf{h}_{LVIS}$ contains the associated h_{LVIS} values. For each site, the regression has been performed individually accounting for the fact that certain height ranges may be insufficiently represented leading to $\alpha_0 = 0.473,\ 1.25,\ 3.4 \times 10^{-5}$ and $\beta_0 = 1.72, 1.44, 4.73$ for Lopé, Mabounie, and Mondah, respectively. While the allometric exponents in Lopé and Mabounie are similar, both of them differ significantly from the

one in Mondah. This probably reflects the large structural heterogeneity and presence of secondary and disturbed (open) forest stands in Mondah in contrast to the other two test sites.

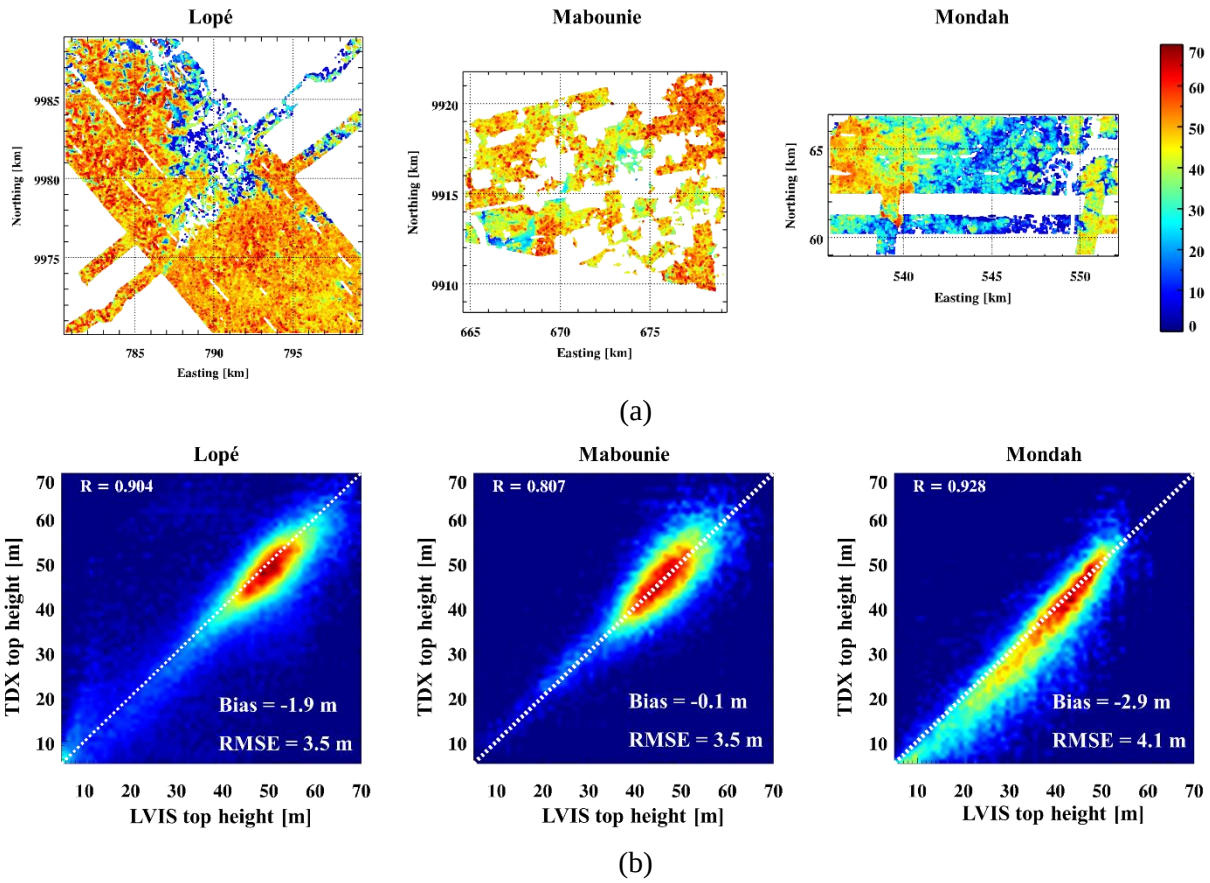
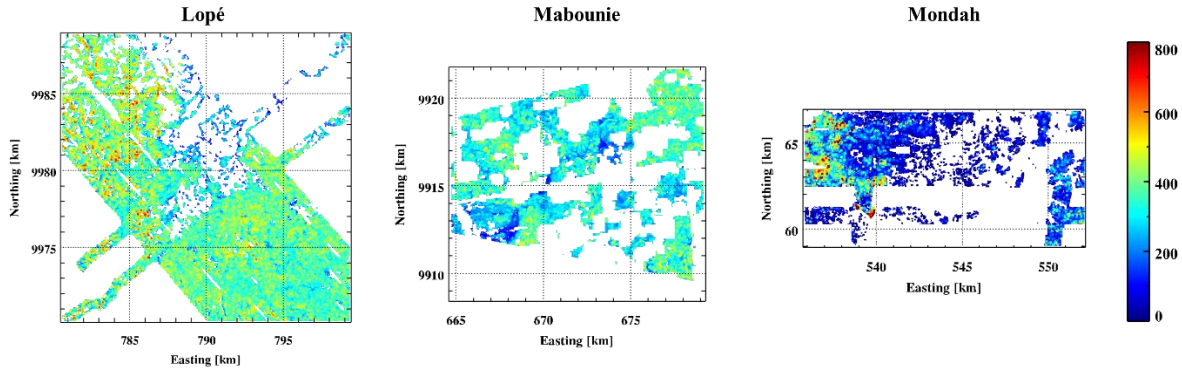


Fig. 2.4. (a) TanDEM-X top height maps at 1 ha resolution, and (b) 2D validation histograms against LVIS top heights for Lopé, Mabounie, and Mondah. The Pearson coefficient (R), bias, and RMSE are reported for each test sites.

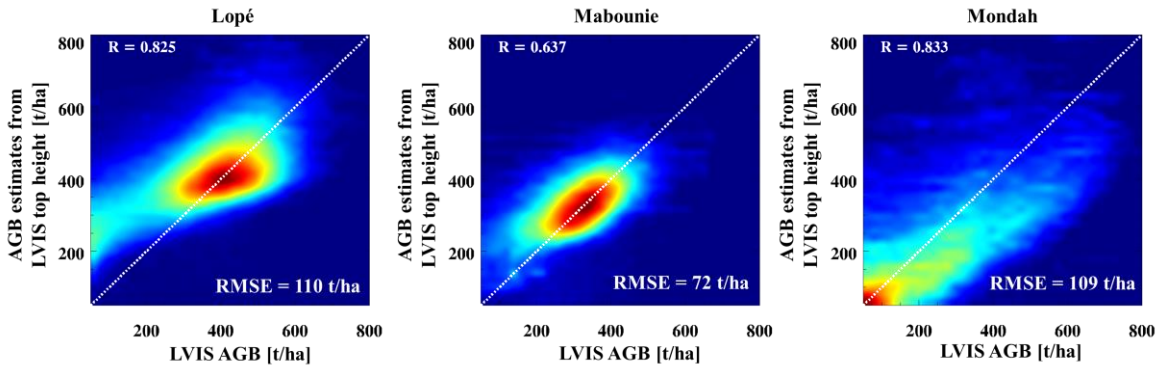
After the definition of the height-to-biomass allometry, forest height estimates, either H_{TDX} or H_{LVIS} can be transformed to AGB estimates. The obtained AGB maps at 1 ha resolution are shown in Fig. 2.5. For all three sites the estimated AGB range is smaller than the LVIS AGB range: high AGB levels are consistently underestimated (by 100 t/ha or even more), especially in Lopé and Mondah. In contrast, low AGB levels tend to be overestimated, especially in Lopé and Mabounie. This behavior is common whether TanDEM-X or LVIS heights are used.

The insufficiency of a single allometric relation to describe accurately the height to biomass relationship even within a single site becomes apparent in the 2-D histograms of B_{LVIS} against H_{LVIS} plotted for each site in Fig. 2.6. The middle one of the three dotted lines indicates the height-to-biomass allometry as obtained from (2.11). In Lopé the AGB ranges from 350 up to 500 t/ha at a height of 50 m. A single height-to-biomass relationship cannot describe this spread. Two additional forest height-to-biomass allometries are plotted defined by the same reference allometric exponent and an allometric level increased (or decreased) by 30% with respect to the original one. It becomes clear that a smaller α_0 is able to compensate the overestimation of lower

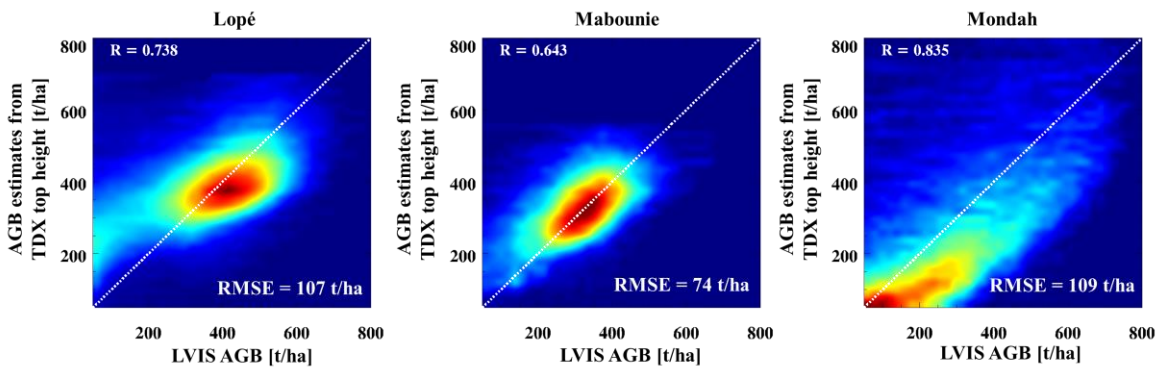
biomass levels (e.g., for sparse(r) stands) seen in Fig. 2.5. A larger α_0 fits better the higher AGB levels [e.g., for dense(r) stands] compensating their underestimation seen in Fig. 2.5. In conclusion, a single value of α_0 is not sufficient for an accurate biomass estimation. To account for this the use of a variable allometric level is attempted next.



(a)



(b)



(c)

Fig. 2.5. (a) AGB maps (1 ha resolution) estimated from TanDEM-X top heights, and 2D validation histograms of the biomass estimates against the reference biomass using (b) LVIS top heights (H_{LVIS}) and (c) TanDEM-X top heights (H_{TDX}) for Lopé, Mabounie, and Mondah. The overall estimation RMSE is reported.

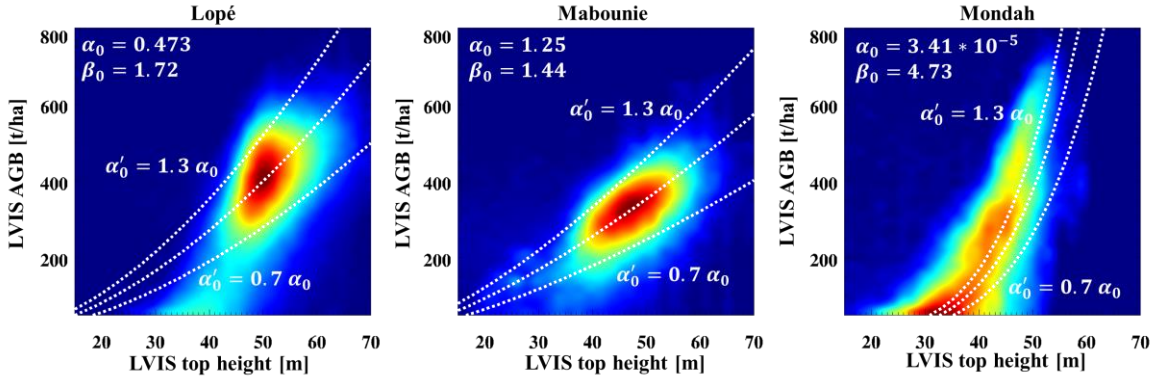


Fig. 2.6. 2D histograms relating the LVIS top height and reference biomass at 1 ha resolution. Each dashed line represents a different height-biomass allometric relationship. The estimated allometric coefficients are reported.

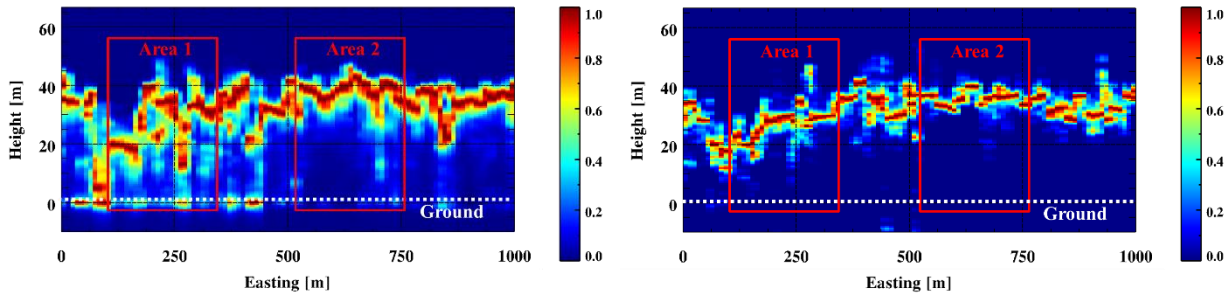
2.4 Forest Biomass Estimation From TanDEM-X Height and Horizontal Structure

2.4.1 Derivation of HS From TanDEM-X

The basic idea behind the structure estimation framework proposed in [23] and [24] is to quantify the variability in height of the top canopy "surface" from tomographically reconstructed 3-D reflectivity. In the case of TanDEM-X, where usually only one single-pass interferometric acquisition is available (with an appropriate vertical wavenumber), the reconstruction of a vertical reflectivity profile is not possible; at least not in a conventional tomographic way. However, the histogram of the interferometric phases (or alternatively of the converted phase center height) over a large enough area provides in many cases an approximation of the vertical reflectivity profiles [39]–[41]. Such profiles will be referred in the following as canopy height profiles (CHPs). Note that because of the different resolutions, the CHP can deviate significantly from the vertical reflectivity underlying the interferometric coherence. Nevertheless, the high attenuation at X-band combined with the high spatial resolution of the TanDEM-X interferograms support the correlation of phase center height variation to (top) canopy height variation allowing the use of the derived CHP to extract relevant horizontal structure information [41].

Interferometric coherences have been estimated with about 25 looks on a 10 m × 10 m ground range-azimuth cell (see Section 2.2). The LVIS DTM is converted to phase (by multiplying with the local vertical wavenumber) and subtracted from the interferometric phase in order to compensate any terrain-induced phase center variations. The obtained phase is then converted to phase center height (by dividing by the local vertical wavenumber). In this way, only the height variations induced by the canopy variability are relevant. Finally, vertical scattering profiles have been obtained within a 25 m × 25 m cell. An example of such CHPs along a 1 km transect across the Lopé site is shown in Fig. 2.7 and compared with the corresponding LVIS waveforms along

the same transect. The peaks of the CHP are distributed closer to the canopy top, while the waveform peaks are distributed much wider between canopy top and the ground. Despite this difference, it is apparent that just based on the top canopy variability, the CHPs allow us to distinguish a sparser area (marked as "Area 1") from a denser one (marked as "Area 2").



(a) (b)

Fig. 2.7. Representative transects in Lopé at a constant northing coordinate: (a) LVIS waveforms, (b) TanDEM-X CHPs. Each profile is normalized by its maximum. The height axis is referred to the ground, which is located in its origin (0 m).

With reference to Fig. 2.8, the calculation of the horizontal structure index HS proposed in [24] is based on "counting" the number of profile (i.e., reflectivity) peaks (associated with scattering contributions) located within a predefined top layer in a structure resolution cell. In practice, considering only the top layer allows us to separate the top canopy height variations from the (vertical) variations induced by volume scattering contributions [24]. As each CHP typically contributes with no more than a couple of peaks, the larger the number of peaks close to the canopy top, the more homogeneous is the forest height, and the higher is the local forest density. Referring the height of the highest peak in the structure resolution cell with H_P, the top height layer is defined as the fraction $\Delta \cdot H_P$ below H_P. The horizontal structure index is then calculated as [24]

$$HS = 1 - \frac{N_P}{N_{P,max}} \tag{2.12}$$

where N_P is the number of CHP peaks in the top layer within the structure resolution cell, and $N_{P,max}$ is a reference maximum often assumed as the maximum N_P within the site / scene. HS equals 0 in dense structure resolution cells, and 1 in sparse structure resolution cells corresponding to the cases illustrated in Fig. 2.8. The obtained HS value depends on the top layer extent: wider top layers contain more CHP peaks than thinner ones. As a consequence, the choice of an appropriate Δ is crucial for the ability of the index to reflect the physical structure (density) and its local variation.

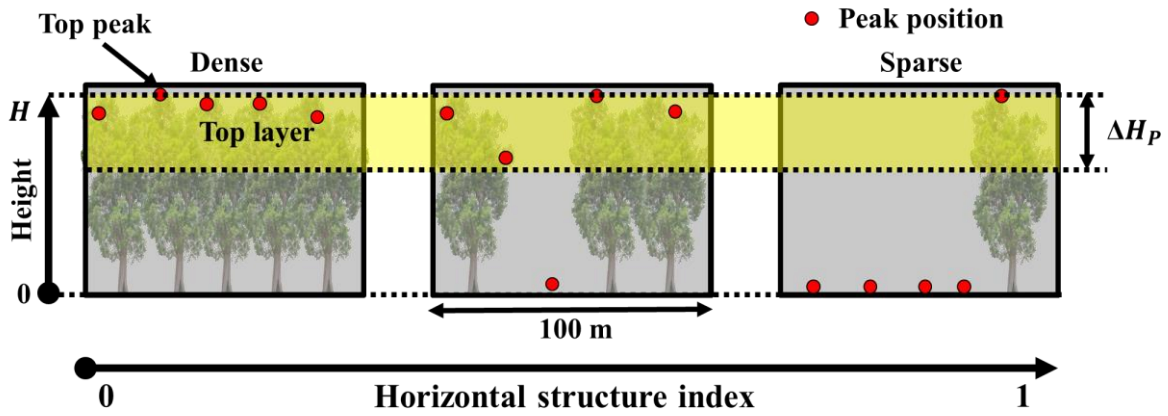


Fig. 2.8. Conceptual sketch related to the calculation of the horizontal structure index used in this paper depicted on the three stands of Fig. 2.1.

Horizontal structure indices have been derived from the TanDEM-X CHPs and the LVIS waveforms (using the same waveform maxima counting procedure and same top layer as with the CHPs) denoted as HS_{TDX} and HS_{LVIS} , respectively. The CHPs have been additionally "multilooked" in height by means of a 10 m Gaussian-shaped moving window in order to suppress peaks caused by the (interferometric phase) noise. A (sliding) 1 ha structure resolution cell has been used to aggregate a statistically relevant number of CHP's (or waveforms). Each structure resolution cell then contains 16 (4 × 4) profiles corresponding to non-overlapping areas on ground. A top layer extent $\Delta = 0.35$ has been chosen. This choice is rather empirical, although motivated by the analysis in [24]. However, both indices appear widely robust against the choice of the top layer extent: HS_{TDX} and HS_{LVIS} do not change significantly for Δ changing within 0.3 and 0.45, in accordance with the results reported in [41]. The derived HS_{TDX} and HS_{LVIS} maps for the three test sites are shown in Fig. 2.9 and reveal very similar structure patterns. In Lopé both indices distinguish the denser southeastern part from the sparser colonizing forest stands at the border to the savannah [30]. In Mabounie the mature forest stands are characterized as dense by both indices. Finally, Mondah's disturbed open forest stands in the central part of the scene (indicated as sparse) are clearly distinguished from the older tall stands in the western part (indicated as dense) by both indices.

Fig. 2.10 shows the 2-D scatterplots of B_{LVIS} against H_{LVIS} for each site. The color of each point corresponds to its HS_{LVIS} value. Despite the dispersion, the dependency of the height-to-biomass allometry on the horizontal structure index HS_{LVIS} becomes clearly visible: for a given height level B_{LVIS} increases with decreasing HS_{LVIS} (i.e., with increasing forest density) or degreases with increasing HS_{LVIS} (i.e., with decreasing forest density). Interestingly, tall stands with top heights of about 40 m appear to have very low biomass levels (< 100 t/ha). Recalling the maps in Fig. 2.9, these stands are composed by few isolated tall trees and shorter vegetation, like for example the colonizing forest stands in Lopé or the sparse forest stands in the center of the Mondah site. This agrees with Fig. 2.6, and supports the idea of using the horizontal structure index to adapt α_0 on local stand conditions.

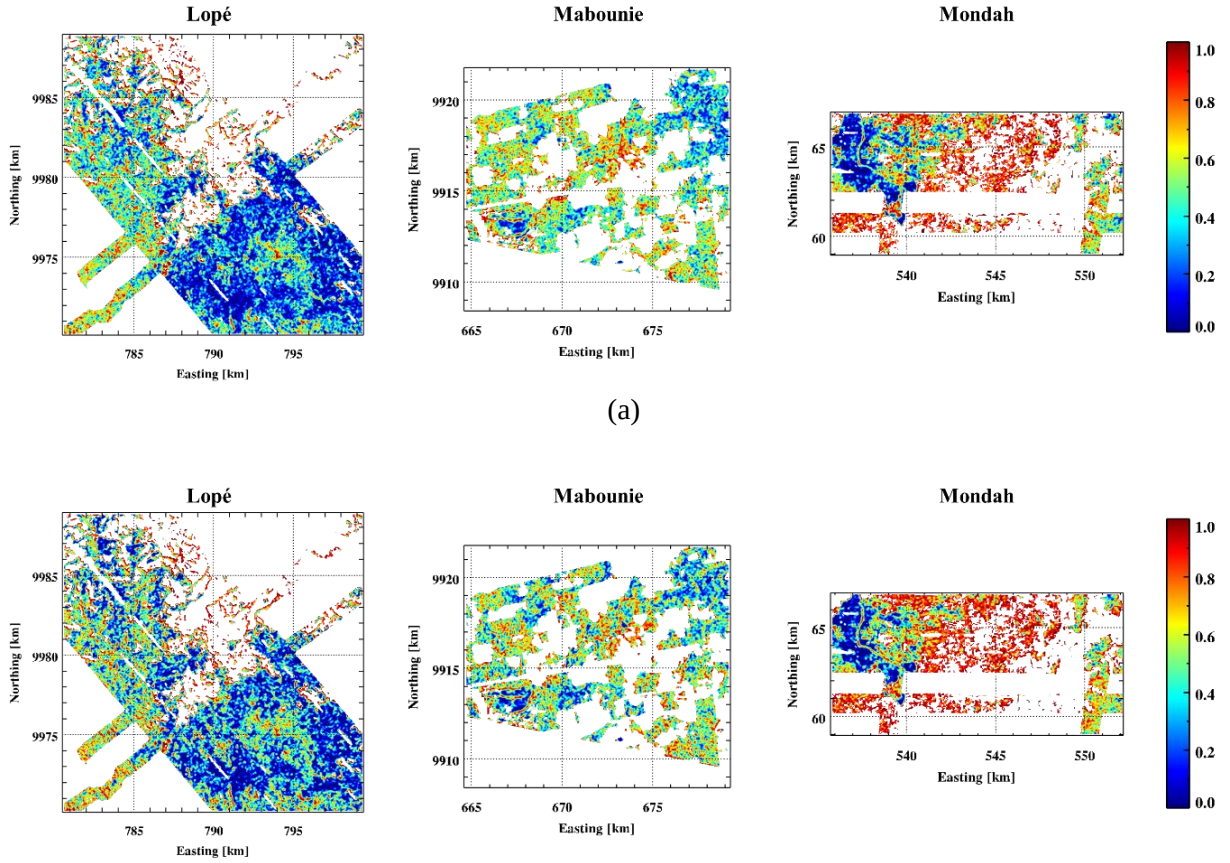


Fig. 2.9. Horizontal structure maps at 1 ha resolution (a) from LVIS waveforms and (b) from TanDEM-X CHPs for Lopé, Mabounie, and Mondah. Sparse forest stands appears in red; dense stands appear in blue.

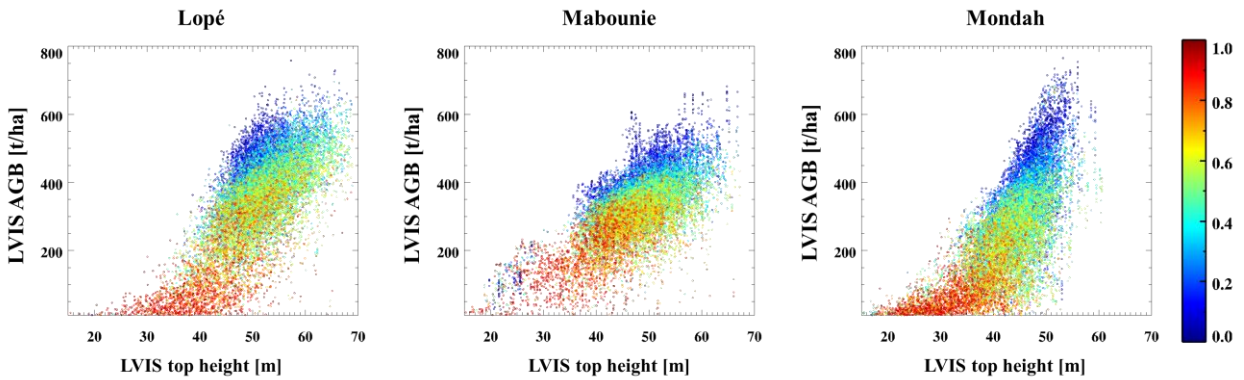


Fig. 2.10. Scatterplots between the reference LVIS top height and AGB for the three test sites. The color indicates the value of the LVIS horizontal structure index (HS_{LVIS}).

2.4.2 Biomass Estimation From TanDEM-X Forest Height and Horizontal Structure

In this Section, the improvement of the height-to-biomass allometry by exploring the dependency of the allometric level on the horizontal structure index is addressed. To define the relation of the allometric level α_0 with the horizontal structure index HS the LVIS estimates B_{LVIS}, H_{LVIS}, and $\mathrm{HS}_{\mathrm{LVIS}}$ are used. In the range from 0 to 1 $\mathrm{HS}_{\mathrm{LVIS}}$ has been segmented into N_α nonoverlapping and uniformly distributed intervals centered around the structure values $\{\mathrm{HS}_{\mathrm{LVIS,i}}\}_{i=1}^{N_\alpha}$. For each of these intervals the allometric level $\{\alpha_0(\mathrm{HS}_{\mathrm{LVIS,i}})\}_{i=1}^{N_\alpha}$ and a common reference allometric exponent β_0 are jointly estimated by means of a least squares optimization, similarly to (2.11)

$$\min_{\{\alpha(\mathrm{HS}_{LVIS,i})\}_{i=1}^{N_\alpha},\overline{\beta_0}} \left\| \begin{bmatrix} \mathbf{b}_{LVIS,1} \\ \vdots \\ \mathbf{b}_{LVIS,i} \\ \vdots \\ \mathbf{b}_{LVIS,N_\alpha} \end{bmatrix} - \begin{bmatrix} \alpha_0(\mathrm{HS}_{\mathrm{LVIS,1}})\mathbf{h}_{LVIS,1}^{\beta_0} \\ \vdots \\ \alpha_0(\mathrm{HS}_{\mathrm{LVIS,i}})\mathbf{h}_{LVIS,i}^{\beta_0} \\ \vdots \\ \alpha_0(\mathrm{HS}_{\mathrm{LVIS,N_\alpha}})\mathbf{h}_{LVIS,N_\alpha}^{\beta_0} \end{bmatrix} \right\|^2 \tag{2.13}$$

where $\mathbf{b}_{LVIS,i}$ and $\mathbf{h}_{LVIS,i}$ are the vectors containing the (reference) B_{LVIS} and H_{LVIS} values, for the generic i-th $\mathrm{HS}_{\mathrm{LVIS}}$ interval centered at $\mathrm{HS}_{\mathrm{LVIS,i}}$. The obtained allometric exponents for the three test sites Lopé, Mabounie and Mondah, are $\beta_0 = 1.8, 1.3$ and 2.5 respectively. Note that the obtained reference allometric exponent β_0 are very similar (at least in Lopé and Mabounie), but not identical to the ones obtained from (2.11) as the two optimisation problems are different.

The general behavior of $\alpha_0(\mathrm{HS}_{\mathrm{LVIS}})$ is obtained from (2.2) by

$$\alpha_0(\mathrm{HS}_{\mathrm{LVIS}}) = \frac{B_{LVIS}}{H_{LVIS}^{\beta_0}} \tag{2.14}$$

and is visualized by means of a 2-D histogram in Fig. 2.11 (a) for the Lopé site. The $\alpha_0(\mathrm{HS}_{\mathrm{LVIS}})$ obtained from (2.14) is indicated by the white dotted line. For all three sites a decreasing trend of $\alpha_0(\mathrm{HS}_{\mathrm{LVIS}})$for increasing $\mathrm{HS}_{\mathrm{LVIS}}$ is obtained. This agrees with Figs. 2.6 and 2.10, and confirms the relevance of the structure index for adjusting the allometric level. The distribution becomes wider for large $\mathrm{HS}_{\mathrm{LVIS}}$ values (0.75 in Lopé, 0.9 in Mabounie and 0.75 in Mondah). At larger $\mathrm{HS}_{\mathrm{LVIS}}$ levels, $\alpha_0(\mathrm{HS}_{\mathrm{LVIS}})$ jumps suddenly to very small values, and its distribution becomes asymmetrical. If $\mathrm{HS}_{\mathrm{TDX}}$ is used instead of $\mathrm{HS}_{\mathrm{LVIS}}$, the distribution of $\alpha_0(\mathrm{HS}_{\mathrm{TDX}})$ [see Fig. 2.11 (b)] becomes slightly wider, but its behavior does not change significantly.

AGB could be estimated by using H_{TDX} and $\alpha_0(\mathrm{HS}_{\mathrm{TDX}})$ in (2.3). The validation of the obtained AGB against the LVIS AGB are shown in Fig. 2.12. It becomes clear that the use of $\alpha_0(\mathrm{HS}_{\mathrm{TDX}})$ successfully compensates for the density and/or structure induced variation of the single allometry with fixed allometric level. Indeed, the estimation performance increases drastically: the RMSE decreases down to 15 - 25% while the correlation coefficient increases to

up to 0.7 - 0.9 for the different sites. The strong overestimation especially for low biomass values in Fig. 2.5 is widely compensated. The remaining residual overestimation, especially in Lopé, is a result of the ambiguity to relate α_0 to a single value of $\mathrm{HS_{TDX}}$ at large HS values (see Fig. 2.11). The optimization of (2.13) is equivalent to a minimization of the mean square error, including both the estimation bias and the standard deviation. Choosing $\alpha_0(\mathrm{HS})$ as the one minimizing the bias between the estimates and LVIS AGB compensates almost completely the bias in the low biomass regions at the cost of an overall increase of the RMSE by 10%. However, AGB can still be underestimated in the high biomass region as a result of the underestimation of forest height.

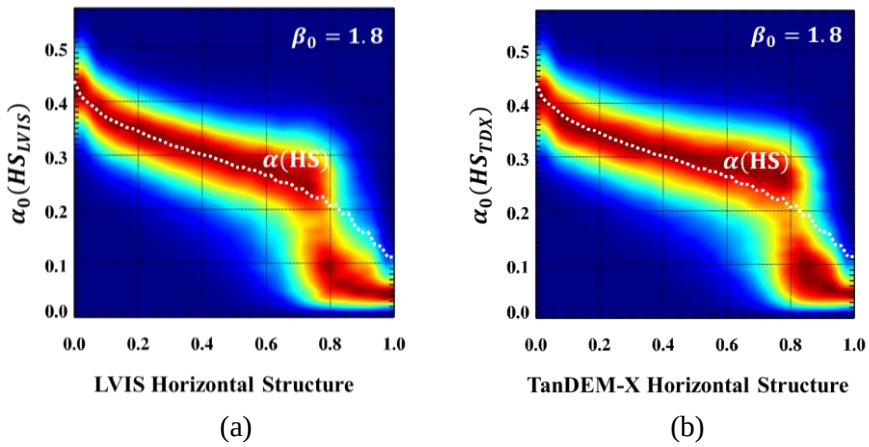


Fig. 2.11. 2D histograms relating α_0 and horizontal structure from (a) LVIS waveforms ($\mathrm{HS_{LVIS}}$), and (b) TanDEM-X CHPs ($\mathrm{HS_{TDXS}}$) in Lopé with $\beta_0 = 1.8$. The white dashed line represents the horizontal structure function $\alpha(\mathrm{HS})$ estimated from the reference biomass and height maps.

2.4.3 Common Height and Structure-to-Biomass Allometry Across Test Sites

The generalization of the height-to-biomass allometry across the three sites requires the estimation of a common (constant) β_{0C} and a common $\alpha_{0C}(\mathrm{HS_{TDX}})$ relation for all sites. For this, the procedure outlined in Section 2.4.2 has been applied to 50,000 B_{LVIS}, H_{LVIS} and $\mathrm{HS_{LVIS}}$ samples from all three sites. The samples are randomly selected and in order to equally represent each site. In this way, an allometric exponent of $\beta_{0C} = 1.7$ and the $\alpha_{0C}(\mathrm{HS_{TDX}})$ relationship shown in Fig. 2.13 (a) are obtained. The behavior of $\alpha_{0C}(\mathrm{HS_{TDX}})$ is consistent to the single-site relations obtained in the following Section. The AGB estimates obtained from the common allometry and its comparison with the reference AGB is shown in Fig. 2.13 (b). The overall RMSE is about 79 t/ha while the correlation is larger than 0.8. An underestimation of about 50 t/ha appears at around 300 t/ha. The overall performance is however convincing, keeping in mind the very different characteristics of the three sites.

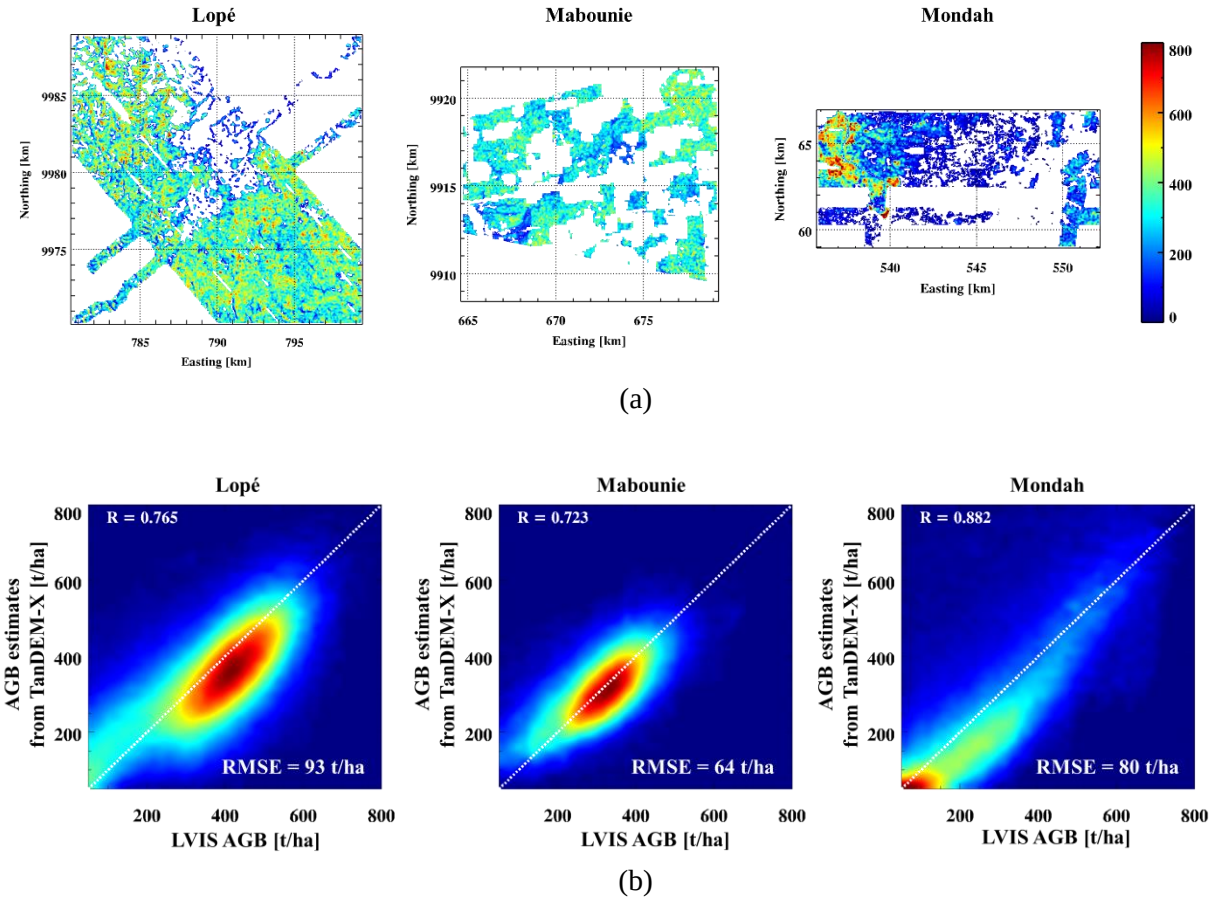


Fig. 2.12. (a) AGB maps from TanDEM-X top height and structure, and (b) 2D validation histograms of the TanDEM-X AGB estimates against the reference biomass for Lopé, Mabounie, and Mondah. The overall estimation RMSE is reported for each test sites.

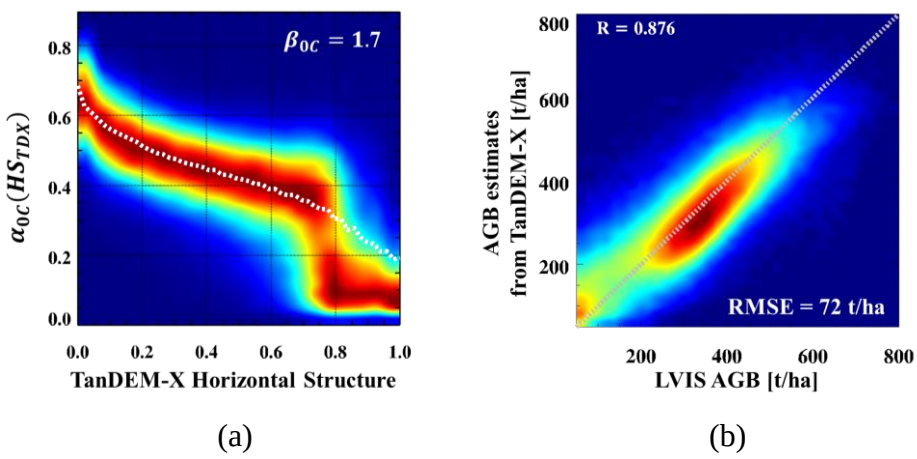


Fig. 2.13. Common height and structure-to-biomass allometry across test sites: (a) Horizontal structure function $\alpha_{0C}(HS)$ with $\beta_{0C} = 1.7$, and (b) 2D validation histogram of the TanDEM-X AGB estimates against reference biomass.

2.4.4 EFFECT OF REDUCED PENETRATION ON STRUCTURE CALCULATION

TanDEM-X forest height estimation is limited by the insufficient penetration capability at X-band due to the high attenuation rates especially in tall/dense/wet forest conditions. This limitation no longer exists if the ground topography is known (e.g., an external DTM is available). In this sense, the lack of a DTM becomes critical for forest height estimation in tall/dense/wet forest conditions. Different height estimation algorithms can be followed [42]. Any additional height estimation inaccuracy resulting from this is directly reflected into an additional inaccuracy of B_{TDX}. By differentiating (2.3), it is readily found that the sensitivity of a biomass estimation error is directly proportional to $\alpha_0(\mathrm{HS})$. From Figs. 2.11 or 2.13, it is apparent that a height error induces a larger biomass error for smaller $\mathrm{HS}_{\mathrm{TDX}}$ (i.e., denser forest stands) for a fixed allometric exponent.

With respect to the horizontal structure index, the high attenuation rates become an advantage increasing the sensitivity to the top-canopy height variations. Therefore, it is expected that the unavailability of an external DTM is less critical for its estimation [15], [41]. The external DTM allows us to separate the top-canopy height variations from the topographic induced height variations. However, as the two contributions are expected to occur at different spatial scales this may allows us to separate them from each other. As reported in an earlier study a low resolution TanDEM-X DEM can be used to compensate topographic induced height variations with sufficient accuracy [43]. A spatial resolution in the order of 100 m appears optimum [15], [43]. After DEM phase compensation both the CHP and the $\mathrm{HS}_{\mathrm{TDX}}$ are estimated as described.

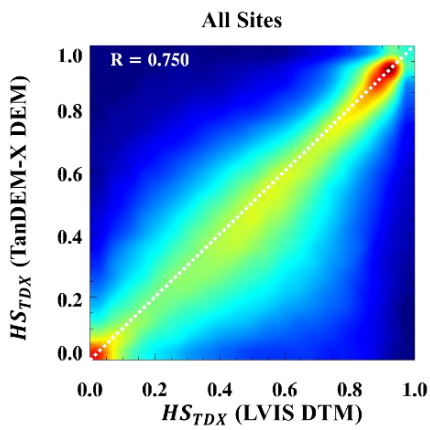


Fig. 2.14. 2D histograms relating $\mathrm{HS}_{\mathrm{TDX}}$ across test sites obtained using the LVIS DTM (horizontal axis) and the TanDEM-X low resolution DEM (vertical axis) as references for the phase center heights.

In order to test the validity of these expectations $\mathrm{HS}_{\mathrm{TDX}}$ calculated with and without DTM are compared in Fig. 2.14. In order to increase the significance of this comparison the same top layer width has been used. The good agreement obtained in all sites confirms that the reduced penetration is the key feature for the retrieval of structure information. The availability of a DTM

plays only a secondary role. However, it can still lead to a loss of performance especially in areas characterized by small-scale topographic variations as it has been seen in Lopé and Mabounie.

It is worth remarking that in the used structure estimation framework an error in the top layer width becomes an (relative) error in HS_{TDX}. HS_{TDX} increases for an underestimated width as the top layer includes more CHP peaks, and vice versa. In this sense, compared to the knowledge of the ground topography, the misidentification of the top layer can cause a larger performance degradation. However, as discussed above, the performance remains comparable even for significant variations (50%) of the selected top layer.

2.5 CONCLUSION

The potential of using forest structure information to adapt a general height-to-biomass allometry to local stand conditions for improving biomass estimation performance is investigated. It is addressed in terms of forest height and structure indices derived from remote sensing measurements particularly in the TanDEM-X context.

The discussed concepts have been demonstrated and validated with TanDEM-X data acquired in 2015–2016 over three different test sites in Gabon in the frame of the AfriSAR campaign. Reference forest height and biomass data as well as (horizontal) forest structure indices have been derived from airborne (LVIS) full waveform data acquired almost at the same time as the TanDEM-X data. TanDEM-X forest height, structure index and biomass estimates have been derived at a spatial scale of 1 ha and compared against the lidar reference data. The general height-to-biomass allometry used to transform forest height estimates to biomass has been established using the lidar height and biomass estimates. The lidar (LVIS) DTM has been used to compensate for the interferometric phase component induced by the terrain topography supporting in this way both the forest height inversion and the estimation of the horizontal structure index from TanDEM-X data.

Forest top height has been estimated with an RMSE well within 20%. In contrast, the biomass values obtained from a single height-to-biomass allometry is affected by large biases independently of the height estimation error. A continuous relationship between the allometric level and the horizontal structure index was derived by means of a least squares minimization of the biomass RMSE. The achieved results indicate that the horizontal structure index is able to adapt a more general height-to-biomass relation to local forest (density) conditions by changing the allometric level. The estimation biases appearing in the conventional single height-to-biomass allometry are widely compensated improving the overall biomass RMSE up to 30%. In addition, the ability to establish a common height-to-biomass allometry supported by the horizontal structure index for all three sites with a reasonable performance is an important result. The identification of the top canopy layer in which the height variability is evaluated remains a critical choice in the estimation of the horizontal structure index.

Regarding the role of the lidar data, they are indispensable for the definition of the forest height-to-biomass allometry. However, their role in the estimation of forest height and the horizontal structure index is less critical as both can be performed even in the absence of any lidar

measurements. More affected from the lack of lidar measurements is the estimation of forest height in tall/dense/wet forest conditions due to the high attenuation at X-band [42], rather than the estimation of the horizontal structure index itself.

There are three important features that make TanDEM-X especially appropriate for the quantitative characterization of the horizontal forest structure: the high attenuation rates at X-band and the associated limited penetration into the forest volume that maximize the sensitivity of the TanDEM-X (complex) coherence to the spatial variability of the top canopy layer; the high interferometric accuracy of TanDEM-X (driven by the single-pass implementation) that allows us to capture the structure induced (complex) coherence variation; and the high spatial resolution of the TanDEM-X (complex) coherence measurements and its continuous measurement nature allows the estimation of structure induced variations at spatial scales relevant for the characterization of the horizontal forest structure.

2.6 ACKNOWLEDGMENT

The lidar AfriSAR datasets were provided by the land, vegetation, and ice sensor team in Code 61A at NASA Goddard Space Flight Center with support from the University of Maryland, College Park, MD, USA. This work was performed in the frame of a joint research project between DLR and NASAs carbon monitoring system (CMS) program in support of NASA Grant #80NSSC20K0023 to the University of Maryland.

2.7 REFERENCES

[1] R. A. Houghton, F. Hall, and S. J. Goetz, “Importance of biomass in the global carbon cycle,” *J. Geophys. Res.-Biogeo.*, vol. 114, p. G00E03, 2009.

[2] M. Reichstein, and N. Carvalhais, "Aspects of forest biomass in the earth system: Its role and major unknowns, *Surveys in Geophysics*, vol. 40, no. 4, pp. 693–707, Jul. 2019.

[3] J. Chave, C. Andalo, S. Brown, M. A. Cairns, J. Q. Chambers, D. Eamus, H. Fölster, F. Fromard, N. Higuchi, T. Kira, J.-P. Lescure, B. W. Nelson, H. Ogawa, H. Puig, B. Riéra, and T. Yamakura, “Tree allometry and improved estimation of carbon stocks and balance in tropical forests,” *Oecologia*, vol. 145, no. 1, pp. 87–99, 2005.

[4] T. Mette, K. Papathanassiou, I. Hajnsek, H. Pretzsch and P. Biber, “Applying a common allometric equation to convert forest height from Pol-InSAR data to forest biomass,” *IGARSS 2004 - 2004 IEEE International Geoscience and Remote Sensing Symposium*, Anchorage, AK, 2004, pp. 269–272.

[5] G. P. Asner, J. Mascaro, H. C. Muller-Landau, G. Vieilledent, R. Vaudry, M. Rasamoelina, J. S. Hall, and M. van Breugel, “A universal airborne LiDAR approach for tropical forest carbon mapping,” *Oecologia*, vol. 168, no. 4, pp. 1147–1160, Apr. 2012.

[6] T. Mette, K. Papathanassiou and I. Hajnsek, “Biomass estimation from polarimetric SAR interferometry over heterogeneous forest terrain,” *IGARSS 2004 - 2004 IEEE International Geoscience and Remote Sensing Symposium*, Anchorage, AK, 2004, pp. 511–514.

[7] A. Torano Caicoya, F. Kugler, I. Hajnsek, and K. Papathanassiou, “Large-Scale Biomass Classification in Boreal Forests with TanDEM-X Data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5935–5951, Oct. 2016.

[8] N. Knapp, R. Fischer, and A. Huth, "Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states," *Remote Sens. Environ.*, vol. 205, Feb. 2020, 111597.

[9] N. Knapp, R. Fischer, V. Cazcarra-Bes and A. Huth, "Structure metrics to generalize biomass estimation from lidar across forest types from different continents," *Remote Sens. Environ.*, vol. 237, pp. 199–209, Feb. 2020.

[10] M. Schlund, S. Erasmi and K. Scipal, "Comparison of Aboveground Biomass Estimation From InSAR and LiDAR Canopy Height Models in Tropical Forests*," IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 3, pp. 367–371, Mar. 2020.

[11] M. J. Soja, H. J. Persson, and L. M. H. Ulander, "Estimation of Forest Biomass from Two-level Model Inversion of Single-pass InSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 5083–5099, 2015.

[12] A. E. L. Stovall, H. H. Shugart, "Improved biomass calibration and validation with terrestrial lidar: implications for future LiDAR and SAR missions," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 10, pp. 3527–3537, Oct. 2018.

[13] J. B. Drake, R. O. Dubayah. D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton, R. L. Chazdon, J. F. Weishampel, and S. Prince, "Estimation of tropical forest structure characteristics using large-footprint lidar," *Remote Sens. Environ.*, vol. 79, no. 2-3, pp. 305–319, Feb. 2002.

[14] R. O. Dubayah, S. L. Sheldon, D. B. Clark, M. A. Hofton, J. B. Blair, G. C. Hurtt, and R. L. Chazdon, "Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica," *J. Geophys. Res.*, vol. 115, 2010.

[15] E. C. De Grandi, E. Mitchard, and D. Hoekman, "Wavelet based analysis of TanDEM-X and LiDAR DEMs across a tropical vegetation heterogeneity gradient driven by fire disturbance in Indonesia," *Remote Sens.*, vol. 8, no. 8, pp. 641–667, Aug. 2016.

[16] P. Da Conceição Bispo, M. Pardini, K. P. Papathanassiou, F. Kugler, H. Balzter, D. Rains, J. R. dos Santos, I. G. Rizaev, K. Tansey, M. N. dos Santos, and L. Spinelli Araujoi, "Mapping forest successional stages in the Brazilian Amazon using forest heights derived from TanDEM-X SAR interferometry," *Remote Sens. Environ.*, vol. 232, Jul. 2019, 111194.

[17] S. Erasmi, M. Semmler, P. Schall, and M. Schlund, "Sensitivity of bistatic TanDEM-X data to stand structural parameters in temperate forests," *Remote Sens.*, vol. 11, 2966, Dec. 2019.

[18] J. I. H. Askne, H. J. Persson, and L.M.H. Ulander, "On the sensitivity of TanDEM-X observations to boreal forest structure," Remote Sens., vol. 11, 1644, Dec. 2019.

[19] M. J. Soja, H. J. Persson, and L. M. H. Ulander, "Estimation of Forest Height and Canopy Density from a Single InSAR Correlation Coefficient*," IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 3, pp. 646-650, 2014.

[20] S. Solberg, E. H. Hansen, T. Gobakken, E. Naessset, and E. Zahabu, "Biomass and InSAR height relationship in a dense tropical forest," *Remote Sens. Environ.*, vol. 192, pp. 166-175, Apr. 2017

[21] R. Treuhaft, Y. Lei, F. Gonçalves, M. Keller, J. R. D. Santos, M. Neumann, and A. Almeida, "Tropical-forest structure and biomass dynamics from TanDEM-X radar interferometry," *Forests*, vol. 8, no. 8, Jul, 2017.

[22] A. Toraño Caicoya, M. Pardini, I. Hajnsek and K. Papathanassiou, "Forest Above-Ground Biomass Estimation from Vertical Reflectivity Profiles at L-Band*," IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 12, pp. 2379–2383, Dec. 2015.

[23] V. Cazcarra Bes, M. Tello Alonso, R. Fischer, M. Heym, and K. Papathanassiou, "Monitoring of forest structure dynamics by means of L-band SAR tomography," *Remote Sens.*, vol. 9, no. 12, pp. 1229–1250, Nov. 2017.

[24] M. Tello Alonso, V. Cazcarra Bes, M. Pardini, and K. Papathanassiou, "Forest structure characterization from SAR tomography at L-band*," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 10, pp. 3402–3414, Oct. 2018.

[25] M. Pardini, J. Armston, W. Qi, S. K. Lee, M. Tello, V. Cazcarra Bes, C. Choi, K. P. Papathanassiou, R. O. Dubayah, and L. E. Fatoyinbo, "Early lessons on combining lidar and multi-baseline SAR Measurements for forest structure characterization," *Surveys in Geophysics,* vol. 40, no. 4, pp. 803–837, Jul. 2019.

[26] L. Reineke, "Perfecting a stand-density index for even aged forests," *J. Agricultural Research,* vol. 46, No. 7, pp. 627–638, 1933.

[27] I. Hajnsek, M. Pardini, M. Jäger, R. Horn, J. S. Kim, H. Jörg, and K. Papathanassiou, Technical assistance for the development of airborne SAR and geophysical measurements during the AfriSAR campaign, Final technical report, ESA contract no. 4000114293/15/NL/CT. Available at: https://earth.esa.int/documents/10174/134665/AfriSAR-Final-Report.

[28] L. Fatoyinbo, J. Armston, M. Simard, S. Saatchi, M. Denbina, M. Lavalle, M. Hofton, N. Pinto, S. Hancock, H. Tang, S. Marselis, B. Hawkins, L. Duncanson, B. Blair, C. Hansen, Y. Lou, R. Dubayah, S. Hensley, C. Silva, J. Poulsen, N. Labriere, N. Barbier, K. Jeffery, D. Kenfack, A. Alonso, G. Moussavou, S. Lewis, K Hibbard, "The NASA AfriSAR Campaign: Airborne SAR and Lidar Measurements of Tropical Forest Structure and Biomass in Support of Future Space Missions," *Remote Sens. Environ.,* vol. 264, 2021.

[29] N. Labriere, S. Tao, J. Chave, K. Scipal, T. L. Toan, K. Abernethy, A. Alonso, N. Barbier, P. Bissiengou, T. Casal, S. J. Davies, A. Ferraz, B. Hérault, G. Jaouen, K. J. Jeffery, D. Kenfack, L. Korte, S. L. Lewis, Y. Malhi, H. R. Memiaghe, J. R. Poulsen, M. Réjou-Méchain, L. Villard, G. Vincent, L. J. T. White, and S. S. Saatchi, "In Situ Reference Datasets from the TropiSAR and AfriSAR Campaigns in Support of Upcoming Spaceborne Biomass Missions," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.,* vol. 11, no. 10, pp. 3617–3627, Oct. 2018.

[30] S. M. Marselis, H. Tang, J. D. Armston, K. Calders, N. Labrière, and R. Dubayah, "Distinguishing vegetation types with airborne waveform lidar data in a tropical forest-savanna mosaic: A case study in Lopé National Park, Gabon," *Remote Sens. Environ.,* vol. 216, pp. 626–634, Oct. 2018.

[31] E. T. A. Mitchard, S. S. Saatchi, L. J. T. White, K. A. Abernethy, K. J. Jeffery, S. L. Lewis, M. Collins, M. A. Lefsky, M. E. Leal, I. H. Woodhouse, and P. Meir, "Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: overcoming problems of high biomass and persistent cloud," *Biogeosciences,* vol. 9, no. 1, pp. 179–191, Jan. 212.

[32] J. B. Blair and M. Hofton. 2018. AfriSAR LVIS L1B Geolocated Return Energy Waveforms, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: https://doi.org/10.5067/ED5IYGVTB50Z.

[33] J. B. Blair and M. Hofton. 2018. AfriSAR LVIS L2 Geolocated Surface Elevation Product, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: https://doi.org/10.5067/A0PMUXXVUYNH.

[34] J. Armston, H. Tang, S. Hancock, S. Marselis, L. Duncanson, J. Kellner, M. Hofton, J.B. Blair, T. Fatoyinbo, and R.O. Dubayah, AfriSAR: Gridded Forest Biomass and Canopy Metrics Derived from LVIS, Gabon, 2016. ORNL DAAC, Oak Ridge, Tennessee, USA. doi: https://doi.org/10.3334/ORNLDAAC/1775.

[35] F. Kugler, D. Schulze, I. Hajnsek, H. Pretzsch, and K. Papathanassiou, "TanDEM-X Pol-InSAR Performance for Forest Height Estimation," *IEEE Trans. Geosci. Remote Sens.,* vol. 52, no. 10, pp. 6404–6422, Oct. 2014.

[36] F. Kugler, S. K. Lee, I. Hajnsek, and K. P. Papathanassiou, "Forest Height Estimation by Means of Pol-InSAR Data Inversion: The Role of the Vertical Wavenumber*," IEEE Trans. Geosci. Remote Sens.,* vol. 53, no. 10, pp. 5294–5311, Oct. 2015.

[37] M. Martone, B. Bräutigam, P. Rizzoli, C. Gonzalez, M. Bachmann, and G. Krieger, "Coherence evaluation of TanDEM-X interferometric data," *ISPRS J. Photogramm. Remote Sens.*, vol. 73, pp. 21–29, Sep. 2012.

[38] K. Papathanassiou and S. R. Cloude, "Single-baseline polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 11, pp. 2352–2363, Nov. 2001.

[39] R. N. Treuhaft, B. D. Chapman, J. R. dos Santos, F. G. Gonçalves, L. V. Dutra, P. M. L. A. Graça, J. B. Drake, "Vegetation profiles in tropical forests from multibaseline interferometric synthetic aperture radar, field, and lidar measurements," *J. Geophys. Res.*, vol. 114, 2009.

[40] R. N. Treuhaft, F. Goncalves, J. R. dos Santos, M. Palace, M. Keller, S. Madsen, F. Sullivan, and P. Graca, "Exploring Vegetation Profiles from TanDEM-X Phase, Lidar, and Field Measurements in Tropical Forests," *EUSAR 2014 - 10th European Conference on Synthetic Aperture Radar*, Berlin, Germany, 2014, pp. 1-3.

[41] A. Pulella, P. C. Bispo, M. Pardini, F. Kugler, V. Cazcarra, M. Tello, K Papathanassiou, H. Balzter, I. Rizaev, M. N. Santos, J. R. dos Santos, L. S. Araujo, and K. Tansey, "Tropical forest structure observation with TanDEM-X data," *IGARSS 2017 - 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Fort Worth, TX, 2017, pp. 918-921

[42] R. Guliaev, V. Cazcarra-Bes, M. Pardini, and K. Papathanassiou, "Forest Height Estimation by Means of TanDEM-X InSAR and Waveform Lidar Data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, Vol. 14, pp. 3084-3094, 2021.

[43] C. Choi, M. Pardini and K. Papathanassiou, "Quantification of horizontal forest structure from high resolution TanDEM-X interferometric coherences," *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, pp. 376-379, 2018.

3 FOREST BIOMASS MAPPING USING CONTINUOUS INSAR AND DISCRETE WAVEFORM LIDAR MEASUREMENTS: A TANDEM-X / GEDI TEST STUDY

C. Choi, M. Pardini, J. Armston, and K. P. Papathanassiou

IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

Published in vol.16, pp. 7675 – 7689, August 2023. DOI: 10.1109/JSTARS.2023.3302026

This Chapter is a post-print, differing from the published paper only in terms of layout and formatting

The author's contributions:

- Interferometric processing and analysis of the TanDEM-X data.
- Suggestion, implementation, and assessment of a tropical forest inversion concept.
- Writing of the manuscript.

The co-authors' contributions:

- M. Pardini and K. P. Papathanassiou provided guidance throughout the research.
- M. Pardini and K. P. Papathanassiou contributed to the main ideas, the discussion of the results, and reviewed the manuscript.
- J. Armston contributed to the discussion of the forest height to AGB from GEDI data and the advice on GEDI data.

FOREST BIOMASS MAPPING USING CONTINUOUS INSAR AND DISCRETE WAVEFORM LIDAR MEASUREMENTS: A TANDEM-X / GEDI TEST STUDY

Changhyun Choi[1,2], Matteo Pardini[1], John Armston[3], and Konstantinos P. Papathanassiou[1]
[1] German Aerospace Center, Microwaves and Radar Institute, Wessling, Germany
[2] ETH Zurich, Institute of Environmental Engineering, Zurich, Switzerland
[3] Department of Geographical Sciences, University of Maryland, College Park, MD, USA

Abstract

This article addresses the implementation of an above ground biomass (AGB) estimation scheme relying on the height-to-biomass allometry at stand level in the context of the synergistic use of continuous TanDEM-X (bistatic) interferometric synthetic aperture radar acquisitions and spatial discrete GEDI waveform lidar measurements. The estimation of forest height and horizontal forest structure from TanDEM-X data in the absence of a digital terrain model (DTM) is discussed. The possibility of estimating (top) canopy height variations independent of topographic height variations is discussed using wavelet-based scale analysis. This understanding is then exploited to define a structure index expressing the (top) canopy-only height variations in the absence of a DTM. The potential of using the derived structure information to account for the spatial variability of height-to-biomass allometry derived from the GEDI measurements is addressed. The performance of the conventional height-to-biomass allometry and the one achieved by the locally adapted implementation are compared against reference lidar measurements and discussed. The analysis is carried out using GEDI and TanDEM-X interferometric measurements and validated by using LVIS lidar measurements over the Lopé National Park, a diverse tropical forest test site in Gabon.

3.1 INTRODUCTION

The potential of spaceborne LiDAR or interferometric synthetic aperture radar (InSAR) configurations to measure forest height at spatial scales of about or below 1 hectare (ha) motivates the use of the so-called forest height-to-biomass allometry at stand level. Accordingly, the above ground biomass (AGB) B of a stand is expressed in terms of an exponential allometric relationship as a function of its top canopy height H [1]-[5]:

$$B = \alpha_0 \cdot H^{\beta_0} \tag{3.1}$$

where α_0 is the allometric level and β_0 is the allometric exponent. The allometric exponent β_0 defines the underlying allometric relationship defined by species composition, growth conditions,

and development stage. The allometric level α_0 accounts for anthropogenic or natural variations in stand density resulting from differences in basal area, age composition, thinning operations, and/or different disturbance effects.

Obviously, any practical application of (3.1) requires knowledge of the two allometric parameters α_0 and β_0 and of their spatial variability in addition to accurate measurements of top canopy height. And while the allometric exponent can remain constant over larger scales, the allometric level may vary locally at much smaller scales.

This article addresses the implementation of an AGB estimation scheme relying on a stand-level height-to-biomass allometry as defined in (3.1), in the context of a synergistic combination of data provided by two different Earth observation missions: the DLR's InSAR TanDEM-X mission [6] and the NASA's waveform Lidar GEDI mission [7]. GEDI samples forest structure by means of lidar waveforms in a more or less dense grid and provides forest height measurements and a set of waveform metrics that allow to estimate AGB [7], [8]. Complementarily, TanDEM-X provides a continuous high spatial resolution InSAR data set with inherent sensitivity to (vertical) forest structure. In the context of (1), spatially continuous forest height estimates derived from the interferometric TanDEM-X data can be used to obtain spatially continuous AGB estimates as long as (α_0, β_0) are known. While the GEDI forest height and biomass measurements can be used to estimate the allometric exponent β_0 and to define the general height-to-biomass allometry at regional or even finer scales, they may be not able to derive the faster varying allometric level α_0. The question is therefore if, and if so, how accurate the spatial variability of α_0 can be derived (or tracked) from TanDEM-X InSAR data.

Indeed, the partial or even complete reconstruction of the 3D radar reflectivity from InSAR or tomographic SAR data and the derivation of a number of (more or less physical) structure indices related to the horizontal and/or vertical forest structure have been demonstrated in several studies [9]-[18]. In [18], a horizontal forest structure index HS derived from InSAR TanDEM-X data has been successfully used to account for the spatial variability of the allometric level in heterogeneous forests:

$$B = \alpha(HS) \cdot H^{\beta_0}. \tag{3.2}$$

and to improve biomass estimation performance. HS quantifies the height variability of the top canopy “surface.” An increase in HS indicates a more heterogeneous canopy surface in the horizontal direction and is interpreted as a sparser forest stand. There are two arguments in favor of using HS: its close correlation with the well-established stand density index [19] and thus with basal area [16] and the fact that it can be derived from InSAR TanDEM-X data. However, the estimation of HS in [18] requires the availability of a digital terrain model (DTM). The lack of appropriate DTMs for most of the forested regions limits the application of (2) for biomass estimation.

In this article, the relationship in (3.2) is applied to a very general case where only a single polarisation TanDEM-X interferogram and a set of spatially discrete GEDI waveform measurements are available, but no DTM is provided. The intention is to develop a methodologic

concept under the perspective of forest biomass inversion on a large scale rather than discussing the optimization of performance on a local scale.

Forest height is derived from TanDEM-X data. Many different ways to invert forest height from TanDEM-X data have been discussed in the literature [20-24]. In the case of TanDEM-X data acquired in the global digital elevation model (DEM) mode [6], the availability of a single polarisation interferogram allows only a highly simplified inversion implementation. This can be only done on the basis of very simplified inversion models, which often need to be supported by additional information, e.g., by using parameters derived from lidar measurements and/or an external DTM [24], [25]. Here, the methodology proposed in [24] and [25] which inverts height from TanDEM-X interferometric coherence using the available GEDI waveforms has been used. The advantage of the proposed approach is that it allows unbiased height estimates over large scales in the absence of a DTM. The price for this is a high(er) variance when compared to other approaches.

The allometric level α_0 and exponent β_0 are derived from the GEDI footprint measurements. The dependency of the allometric level on the forest structure index α(HS) is no longer possible to be established in the absence of a DTM in the context of [18]. Instead of using a DTM for removing the topographic variation in the estimation of HS, the low-pass filtered TanDEM-X DEM is used. Even though this approach is well established in interferometric data processing, here it has to be evaluated to what extent and on which spatial scales the phase variations induced by the terrain can be separated from the phase variations induced by the vegetation and to what extent it affects the performance of the HS estimation.

The rest of this article is organized as follows. Section 3.2 describes the selected forest site for the experiments, i.e., the Lopé National Park in Gabon, the experimental data (TanDEM-X and GEDI), and the reference height and biomass measurements available. Section 3.3 addresses the estimation of a horizontal structure index from TanDEM-X data in the absence of a DTM. In Section 3.4, the forest height estimation from TanDEM-X data is reviewed. The use of the derived horizontal structure information to improve forest height estimation performance is proposed. Section 3.5 addresses the derivation of the forest height-to-biomass allometry from the GEDI measurements. In Section 3.6, the biomass estimation using the height, structure, and GEDI derived height-to-biomass relations is performed and assessed. Finally, Section 3.7, concludes this article.

3.2 TEST SITES AND DATA SETS

The experiments in this study focus on an area within the Lopé National Park in Gabon covered during the AfriSAR campaign in 2016 [26], [27]. The site consists of a variety of forest structure types ranging from open savannas to undisturbed tall (sometimes exceeding 50 m) and dense forest stands. Colonizing forest (sparse forest stands mixed up with savanna) or monodominant Okoume (dense, mono-layered, tall, and dense forest stands) are two particular cases [28-32]. Biomass ranges between around 10 t/ha in savanna areas and ~600 t/ha in the dense forest areas. The terrain is hilly with many local slopes steeper than 20°. The available lidar and

radar data are described in the following and summarized with their resolution in Tables 3.1 and 3.2.

TABLE 3.1
SUMMARY OF LIDAR DATA AND TANDEM-X PRODUCTS

Data	**Product**	**Symbol / Acronym**	**Resolution**	**Grid sampling**
Small-footprint lidar	Digital terrain model / digital canopy model	DTM / CHM	1 m × 1 m	1 m × 1 m
LVIS	RH100 at footprint level	H_{LVIS}	22 m (footprint diameter)	20 m × 20 m
	Mean RH100	H_{LVIS50}	50 m × 50 m	
		$H_{LVIS100}$	100 m × 100 m	
	Above ground biomass	B_{LVIS50}	50 m × 50 m	
		$B_{LVIS100}$	100 m × 100 m	
GEDI	Level 2A RH100	H_{GEDI}	25 m (footprint diameter)	
	Level 4A AGB	B_{GEDI}		
TanDEM-X	"Few-looks" phase center heights		5 m × 5 m	5 m × 5 m
	Digital elevation model (DEM)	DEM	120 m × 120 m	
	Canopy height profiles (CHP)	CHP	25 m × 25 m	20 m × 20 m
	Top peak height	Z_{top}	25 m × 25 m	
	Horizontal structure index	σ_{top}	100 m × 100 m	
	Forest height	H_{TX}	25 m × 25 m	
		H_{TX100}	100 m × 100 m	
	Above ground biomass	AGB	100 m × 100 m	

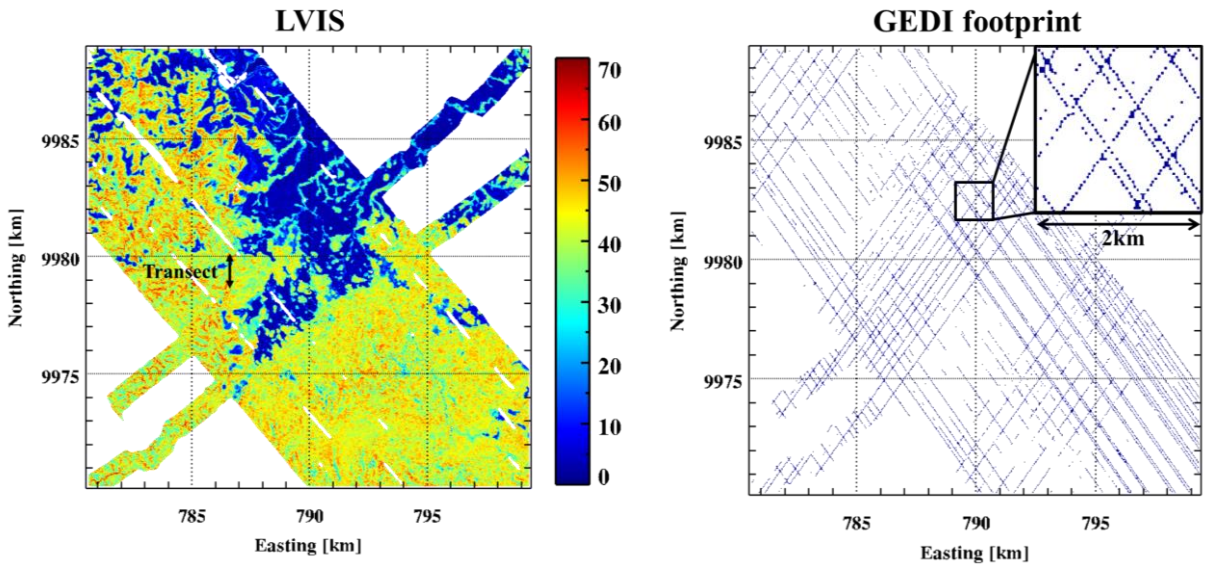
TABLE 3.2
Summary of TanDEM-X Acquisitions Parameters

Test Site	Lopé
Acquisition Date	Jan. 25, 2016
Frequency	X-band
Polarization	HH
Ground resolution (Range / Azimuth)	1.95 m / 1.99 m
Vert. wavenumber	~ 0.10 m^{-1}
HoA	~ 62.8 m
Incidence angle	~ 44.5°

Lidar full-waveform data were collected by NASA's Land and Vegetation and Ice Sensor (LVIS) in February 2016 [27]. LVIS footprints have a mean diameter of about 22 m and overlap partially on ground [27]. At each footprint, the RH100 (the height above ground at which 100% of the full-waveform energy is cumulated) is estimated from the waveform. The RH100 heights have been projected in geographic UTM coordinates and resampled at a 20 m grid. These resampled RH100 heights, denoted in the following as $\mathrm{H_{LVIS}}$, are shown in Fig. 3.1(a) and are used as a reference for the validation of the TanDEM-X height estimates. The reference LVIS heights H_{LVIS} are further averaged to a 50 m × 50 m and a 100 m × 100 m resolution. The obtained heights are referred as $\mathrm{H_{LVIS50}}$ and $\mathrm{H_{LVIS100}}$, respectively. Furthermore, two AGB maps at 50 m × 50 m (referred as $\mathrm{B_{LVIS50}}$) and 100 m × 100 m (referred as $\mathrm{B_{LVIS100}}$) resolution, both estimated from the LVIS waveforms, are used as biomass reference. These AGB maps have been obtained by means of a relationship formally equivalent to the one in (2). The RH98 is used to calculate the top forest height, while RH90, canopy cover, and regional values of wood specific gravity are used to calculate the allometric level. The relationships are further parameterized by using the field inventory plots [28], [29]. The DTM and the canopy height model (CHM) resampled at 1 m × 1 m resolution have been used as well. Both of them were derived from small-footprint (10 cm) discrete-return lidar data acquired in July 2015 covering part of the Lopé site [32].

The TanDEM-X dataset was selected to be acquired close in time to the LVIS flights. The relevant acquisition parameters are summarized in Table 3.2. The mean height of ambiguity (HoA) of about 65 m allows optimum forest height estimates in the range between 15 and 45 m [25], [33]. For shorter and taller heights, the forest height estimation performance is expected to be more or less compromised [33].

The GEDI data over Lopé available for this study have been acquired in the first 18 months of the mission between April 2018 and October 2019. For each International Space Station pass, the three GEDI lasers collected data along eight tracks separated by about 600 m in the across-track direction. Along each track, waveforms with a footprint of approximately 25 m diameter are measured every 60 m. In Lopé, this results into 12 000 footprints distributed as shown in Fig. 3.1(b) corresponding to only 0.6% of the LVIS coverage. The GEDI waveforms, the Level 2A RH100 values (referred as $\mathrm{H_{GEDI}}$) [34], as well as the derived Level 4A AGB values (referred as $\mathrm{B_{GEDI}}$) [35] have been used to initialize the height estimation from TanDEM-X coherences and to define the height-to-biomass relationship(s).



(a) (b)

Fig. 3.1. Lopé site: (a) LVIS RH100 H_{LVIS} map in meters (m), (b) footprint positions of the available GEDI measurements. Both maps are in UTM coordinates, with spacing (20 m × 20 m) in easting and northing direction, and they cover around 19 km by 19 km. The black line in (a) indicates a representative transect used for the wavelet analysis in Fig. 3.2.

3.3 HORIZONTAL STRUCTURE INDEX ESTIMATION

Originally, the estimation of the horizontal forest structure index was proposed either from tomographically reconstructed radar reflectivity profiles at L-band or from lidar waveforms [16], [17]. In [18], a similar—yet simplified—horizontal structure index has been derived from CHPs, i.e., the histograms of the InSAR "few-look" phase center heights [23], [36], [37] within a certain resolution cell. In particular, the variation of top canopy height reflected by the CHPs after the compensation of the terrain-induced (height) variations by using an available DTM was used to derive the horizontal structure index. However, in the absence of a DTM, this approach is not possible for sites with relevant topographic variations. To circumvent this rather serious limitation, the question of the existence of spatial scales at which top canopy height variations are independent or at least less affected by topographic height variations becomes important. This question is addressed in this Section.

3.3.1 WAVELET VARIANCE ANALYSIS

To evaluate the effects of both top canopy and topographic height variations on the TanDEM-X InSAR few-look phase at different spatial scales, a wavelet decomposition analysis is employed, similar to the one proposed and performed in [10] and [38]. For this, the wavelet spectrum of the TanDEM-X few-look phase center heights (as obtained by dividing the unwrapped InSAR phase by the local terrain-corrected vertical wavenumber κ_z [33], [39]) are

compared with the available CHM spectrum (representing the top canopy height variations) and the DTM spectrum (representing the topographic height variations), both derived from the small-footprint lidar data.

For the sake of simplicity, all the height maps were projected in geographic UTM coordinates and resampled on the same 1 m grid in both easting (x) and northing (y) directions. Before transforming to UTM, the TanDEM-X few-look phase center heights have been obtained from a multilooking operation with resolution 5 m × 5 m (corresponding to six independent looks, two in range, and three in azimuth) in order to reduce the phase variance induced by the interferometric decorrelation.

For each height map f(x,y), the wavelet spectrum has been calculated as a function of the (horizontal) scale parameter s as [10], [38], [40]

$$\mathrm{WS_s} = \langle c_{x,s}^2 + c_{y,s}^2 \rangle. \tag{3.3}$$

where $\langle \cdot \rangle$ indicates a moving average operator within 100 × 100 m cells introduced to reduce local fluctuations. As indicated by their subscripts, all quantities in (3) are 2D in (x,y) and depend on s, $c_{x,s}$ and $c_{y,s}$ are the coefficients along the x and y directions associated with the chosen 1D mother wavelet, respectively. Each value of s corresponds to a dilation of the mother wavelet, which is used to generate the impulse response of a filter. Its application to the input height maps along x and y provides $c_{s,x}$ and $c_{y,s}$ [38]. If the mother wavelet is a symmetric and odd function of the spatial variable, the corresponding filter resembles a differential operator producing a (mean) height difference between points at a distance corresponding to the scale [38], [40], i.e.,

$$\begin{aligned} \langle {c_{x,s}}^2 \rangle &\cong \langle [f(x + \Delta x, y) - f(x, y)]^2 \rangle \\ \langle {c_{y,s}}^2 \rangle &\cong \langle [f(x + \Delta y, y) - f(x, y)]^2 \rangle \end{aligned} \tag{3.4}$$

with Δx and Δy proportional to s. In this way, $\mathrm{WS_s}$ reflects directly 2D variations of the input heights as a function of s [10], [31]. For the wavelet decomposition the PyWavelets Python package [34] has been used with the biorthogonal 1.3 function as mother wavelet for a reliable approximation of (3.4).

Fig. 3.2 illustrates the behavior of the three (TanDEM-X few-look phase center heights, and lidar CHM and DTM) WSs as a function of the scale parameter s averaged along a representative 7-km long north–south transect in Lopé [see Fig. 3.1(a)] covered by dense forest stands and with a significant topographic variation. The relative effect of top canopy and topographic height variations on the TanDEM-X few-look phase center height at the different spatial scales becomes clearly visible: while at smaller scales (up to 10 m), the TanDEM-X phase center height (green line) is highly correlated with the CHM (blue line) and widely independent of the DTM (red line), at larger scales (>30 m) it correlates with the DTM while its dependency on the CHM decreases fast with increasing scale.

The plot makes clear that the estimation of top canopy height variations (and consequently the estimation of the horizontal structure index) by means of the CHPs without compensating for the

terrain-induced height variations is problematic at scales larger than 30 m as both top canopy and terrain height variations are relevant. In order to reflect only the top canopy height variations, the TanDEM-X few-look phase center height variation should be estimated at a 10 m scale (or finer). This becomes difficult due to the phase center height (i.e., phase) variance induced by interferometric decorrelation.

To compensate for the effect of terrain variations, the TanDEM-X few-look phase center heights are corrected once by using the DTM and once with a low-pass filtered (up to a spatial resolution of 120 m) version of its own. The use of a low-pass filtered DEM for removing the topographic variation is an established technique in interferometric SAR processing [23], [37], [42], [43]. However, the question here is to find at what scale this is best possible and how much it compromises the HS estimation performance. The value of 120 m has been chosen with reference to the behavior of TanDEM-X phase center heights in Fig. 3.2: at this scale, the top canopy height variations are attenuated (with respect to its maximum) while the topographic ones are maximized. As expected, the DTM corrected TanDEM-X few-look phase center heights (orange line) in Fig. 3.2 follow closely the CHM behavior at all scales. The self-corrected TanDEM-X few-look phase center heights (cyan line) behave similarly and follow the CHM heights across the whole range of scales as well.

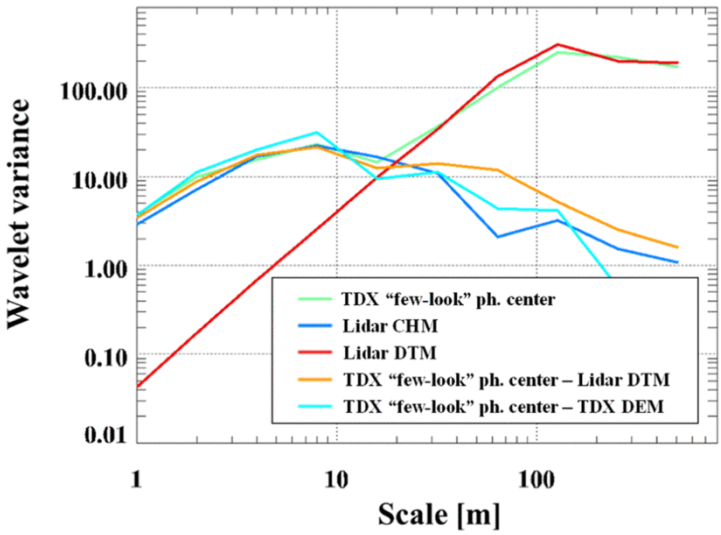


Fig. 3.2. Lopé site: wavelet variance WS as a function of the scale s for the small-footprint lidar heights (DTM and CHM) and TanDEM-X "few-look" phase center heights averaged along the representative transect shown in Fig. 3.1(a).

3.3.2 Structure Index Definition

Based on the analysis above, the horizontal structure estimation framework defined in [18] is now adapted to the self-corrected TanDEM-X few-look phase center heights. The CHPs are calculated as the histograms of the self-corrected heights, followed by smoothing with a Gaussian

window of 3 m width. The zero height of each CHP is now a local reference height provided by the low-pass filtered TanDEM-X phase center heights. As the real ground height is unknown, the determination of the top layer in which the significant canopy height variations occur is not straightforward [18]. This requires the definition of a structure index by means of a relative height. For this, only the peak at the highest height with a value above threshold, called "top" peak in the following, of each CHP in a structure cell is retained in a set $\mathrm{Z_{top}}$. The threshold is set at 10% of the absolute maximum value reached by each CHP. This thresholding operation, together with the histogram smoothing, aims at reducing the impact of using a small number of looks in the calculation of the phase center heights and to avoid the creation of insignificant peaks which may bias the structure quantification. The horizontal structure index employed here is defined as [44]

$$\sigma_{\mathrm{top}} = \sqrt{\mathrm{var}\{\mathrm{Z_{top}}\}} \tag{3.5}$$

where var{·} indicates the variance of the set. A high σtop indicates large top canopy height variations (i.e., large canopy roughness) as in the case of a sparse forest, while a low σtop indicates smaller variations (i.e., low canopy roughness) as in the case of a dense(r) forest. The use of the CHPs allows to maximize the sensitivity to the top canopy variations as it accounts only for the behavior of the "top" peak. Examples of CHPs for two relevant transects in dense and sparse forests are shown in Fig. 3.3, and the "top" peak is marked for each of them. It is apparent that in the sparse forest case, the "top" peak heights vary in a larger height interval than in the dense forest case, leading to a larger σ_{top}.

With reference to the spatial grid samplings and resolutions in Table 3.1, σ_{top} is derived by means of (3.5) at 100 m resolution (i.e., for 100 my× 100 m structure cells) as follows.

1. The single look complex TanDEM-X phase center height is multilooked to 5 m × 5 m cells (6 looks).
2. The terrain-induced height variations are compensated by subtracting a low-pass filtered (to a spatial resolution of 120 m) version of its own.
3. The CHPs are calculated at 25 m resolution (i.e., 25 m × 25 m cells) from 25 phase center height samples and for each CHP $\mathrm{Z_{top}}$ is derived.
4. The variance of $\mathrm{Z_{top}}$ within a 100 m × 100 m structure cells, i.e., across 25 $\mathrm{Z_{top}}$ samples (equivalent to 16 independent samples), is estimated and used in (3.5) to calculate σ_{top}.

The obtained map of Fig. 3.4 shows sparse forest areas (σ_{top} > 5 m) surrounded by denser ones (σ_{top} < 5 m) in the southeastern part of the site, while they are distributed along the slopes in the northwestern one.

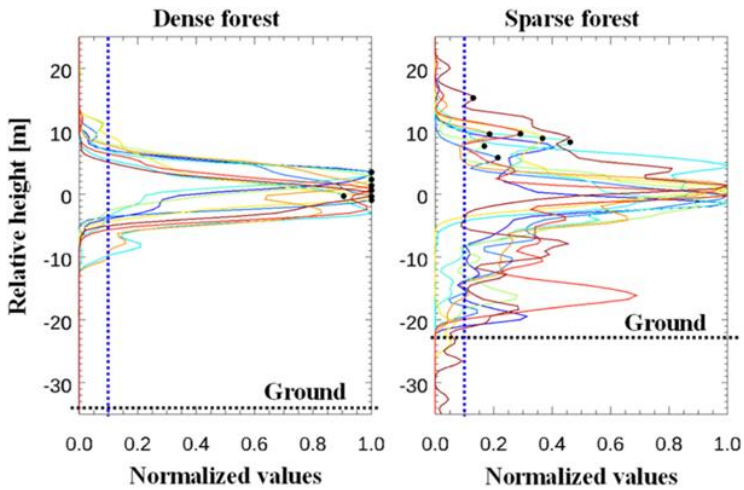


Fig. 3.3. Examples of consecutive CHPs extracted along two transects in dense and sparse forest stands (areas A and B in the following). Each CHP corresponds to a 25 m × 25 m cell on ground. The zero of the vertical axis represents the low-pass filtered TanDEM-X DEM height in the location of each CHP. The average ground height is reported (horizontal black dashed lines). For each CHP, the peak at the maximum height with a value above the threshold indicated by the vertical blue dashed line is retained in Z_{top}. The extracted peaks are denoted with a black dot.

Although the proposed horizontal structural index is able to distinguish dense from sparse forest stands, even to some extent, σ_{top} becomes ambiguous at forest nonforest transitions, misinterpreting the step-like height change as increased top canopy variations, and often classifying the transition zone as sparse forest. Such border areas can be identified by using a forest/non-forest mask [45]. When forest and non-forest attributed samples are present within a 100 m estimation cell, the cell is set as a border area. In this way, the 3% of the forested area in the site is classified as a border area and excluded. In order to exclude an additional error factor from the interferometric phase, hilly areas (i.e., slopes larger than 15°) corresponding to ∼10% of the forested area were masked as well, resulting in the exclusion of low coherence areas characterized by higher phase noise.

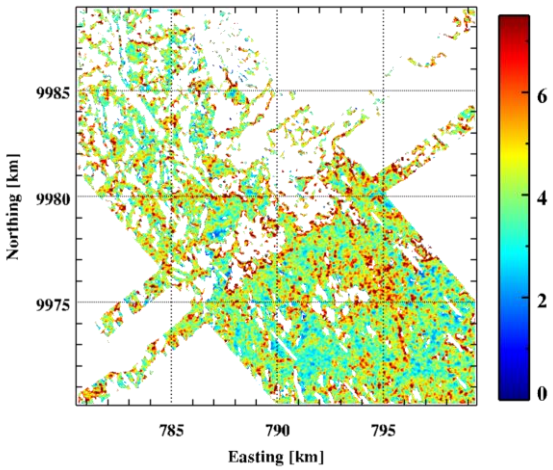


Fig. 3.4. Lopé site, same area as in Fig. 3.1. Map of σ_{top} at 100 m × 100 m resolution.

3.4 FOREST HEIGHT ESTIMATION FROM TANDEM-X COHERENCE MAGNITUDE

3.4.1 METHODOLOGY

For a bistatic single-polarization TanDEM-X acquisition, the InSAR complex coherence can be factorized as [39], [46]:

$$\tilde{\gamma}(\kappa_z) = \gamma_{SNR} \cdot \gamma_{rg}(\kappa_z) \cdot \gamma_Q \cdot \tilde{\gamma}_V(\kappa_z) \tag{3.6}$$

where γ_{SNR} is the additive noise (SNR) decorrelation, $\gamma_{rg}(\kappa_z)$ is the range spectral decorrelation, and γ_Q is the quantization decorrelation. Forest height (H_V) is estimated from the volume decorrelation contribution $\gamma_V(\kappa_z)$:

$$\tilde{\gamma}_V(\kappa_z) = \exp(i\kappa_z z_0) \frac{\int_0^{H_V} F(z) \exp(i\kappa_z z) dz}{\int_0^{H_V} F(z,\vec{w})\, dz} \tag{3.7}$$

where $F(z)$ is the vertical distribution of scatterers, H_V is the forest (top canopy) height [that defines the upper boundary of $F(z)$) and z_0 a reference height corresponding to the lower boundary of $F(z)$. κ_z is the vertical (interferometric) wavenumber defined for bistatic interferometers as:

$$\kappa_z = \frac{2\pi}{\lambda} \frac{\Delta\theta}{\sin(\theta_0 - a)} \tag{3.8}$$

where λ is the wavelength, $\Delta\theta$ is the change of the incidence angle induced by the spatial baseline, θ_0 is the nominal incidence angle, and a is the range terrain slope. The terrain slopes can be obtained from an available DEM, and for this, the TanDEM-X DEM has been used.

Following the approach proposed in [25] the available GEDI waveforms are used to derive a "mean" vertical reflectivity profile over a whole TanDEM-X scene. For this, first, the so-called profile matrix [P] is formed with columns of the GEDI waveforms $P_i(z)$ in the scene normalized to unit height (and resampled to a common number of height samples). Accordingly, the number of rows of [P] is given by the number of height samples and the number of columns by the number of available GEDI waveforms. From the profile matrix, a covariance matrix [R] is formed:

$$[R] = [P]\,[P]^T \tag{3.9}$$

where $[\cdot]^T$ indicates the transpose operation, and then diagonalised:

$$[R] = [U]\,[\Lambda]\,[U]^T \tag{3.10}$$

where $[\Lambda]$ contains the (positive) eigenvalues a_i and $[U]$ the eigenvectors $\overline{P}_i(z)$ of $[R]$. The eigenvectors $\overline{P}_i(z)$ of $[R]$ are then used to compose the mean reflectivity profile:

$$P_{mean}(z) = \sum_{i=1}^{M} a_i \overline{P}_i(z) \tag{3.11}$$

where M represents the number of eigenvectors used to compose $P_{mean}(z)$. In the following, only the first eigenfunction $P_{mean}(z) = \overline{P}_1(z)$ has been used for defining the mean reflectivity profile. The low-frequency profile component given by the first eigenvector is more appropriate for describing forest reflectivity over larger spatial scales, since the higher order profile components may locally mismatch with the actual reflectivity due to the natural spatial heterogeneity of the forest structure. In terms of the achieved estimation performance, the low-frequency profile component is sufficient because the effect of the vertical reflectivity on the volume decorrelation is smaller compared to the effect of the forest height.

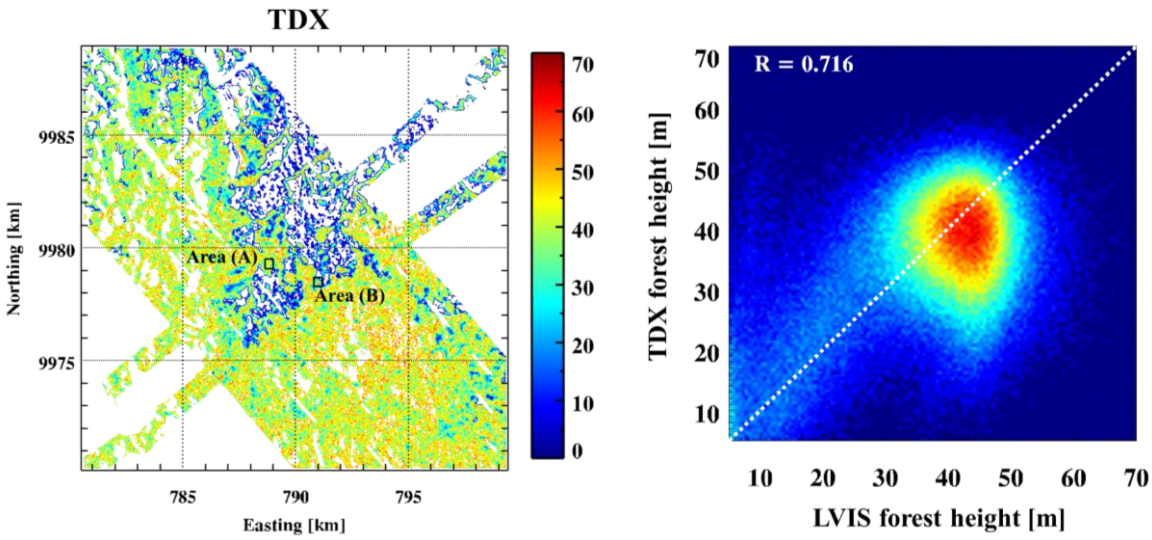
The use of the the mean reflectivity profile $P_{mean}(z)$ instead of the vertical distribution of scatterers $F(z)$ in (3.7) leads to a determined inversion problem of two unknowns, i.e. H_V and z_0 that can be inverted by a single complex observation $\tilde{\gamma}_V(\kappa_z)$. It can be further simplified to a single dimensional inversion problem (with a single unknown H_V) by accounting only the absolute values of (3.7).

The forest height inversion was performed using a coherence estimation window of 25 m × 25 m (corresponding to about 150 independent looks), compensating for non-volumetric decorrelation contributions (γ_{SNR}, γ_{rg} and γ_Q [46]), deriving the mean profile $P_{mean}(z)$ using all available GEDI waveforms in the scene by means of (3.11) and using both the volume coherence and the mean profile in (7). Samples with $|\gamma_V(\kappa_z)| < 0.25$ and heights higher than 52 m (i.e., the expected maximum top height for the actual vertical wavenumber) were discarded. In a final step, the available GEDI RH100 heights, H_{GEDI}, were used, as proposed in [25], to compensate any residual global bias affecting the vertical wavenumber – height product $\kappa_z H_V$. The obtained heights were projected in geographic UTM coordinates and resampled in a 20 m grid, and are indicated with H_{TX} in the following.

Fig. 3.5 (a)-(b) shows the TanDEM-X height estimates and their validation against the LVIS heights (H_{LVIS}) by means of a 2D histogram. The performance is consistent with what was already reported in [25], confirming the robustness of the inversion with respect to the number of available waveforms. The considerable underestimation of a number of stands with heights above 40 m is partly due to the large vertical wavenumber and partly due to the limited X-band penetration in the dense(r) stands. The underestimation introduced by a too large vertical wavenumber is well known and has been discussed for the actual TanDEM- X / GEDI case in [25].

It is clear that the adapted forest height inversion approach relies on a number of critical assumptions and compromises. The use of a single “mean” vertical reflectivity profile over a whole TanDEM-X scene makes an adaptation to the spatial forest structure heterogeneity. At the same time, there are inherent differences between the nadir lidar waveforms and the side-looking X-band reflectivity. However, such assumptions and compromises are necessary for obtaining

forest height from single-pol single-baseline TanDEM-X data. A detailed performance analysis of the implemented forest height inversion can be found in [25].



(a) (b)

Fig. 3.5. Lopé site, same area as in Fig. 3.1: (a) TanDEM-X forest height H_{TX} in meters (m); (b) comparison (2D histogram) between H_{TX} and H_{LVIS}. The height estimation performance in the areas (A) and (B) indicated in (a) is further investigated in Fig. 3.5.

3.4.2 Structure Dependency of the Height Estimation Bias

While the height underestimation due to too large vertical wavenumbers can be avoided by using a smaller spatial baseline, the underestimation due to limited penetration is a fundamental limitation that cannot be easily overcome, if not at all. Nevertheless, a first-order correction is attempted, based on the assumption that even in dense stands there may be points where the X-band pulses penetrate to the ground and allow undistorted height inversion.

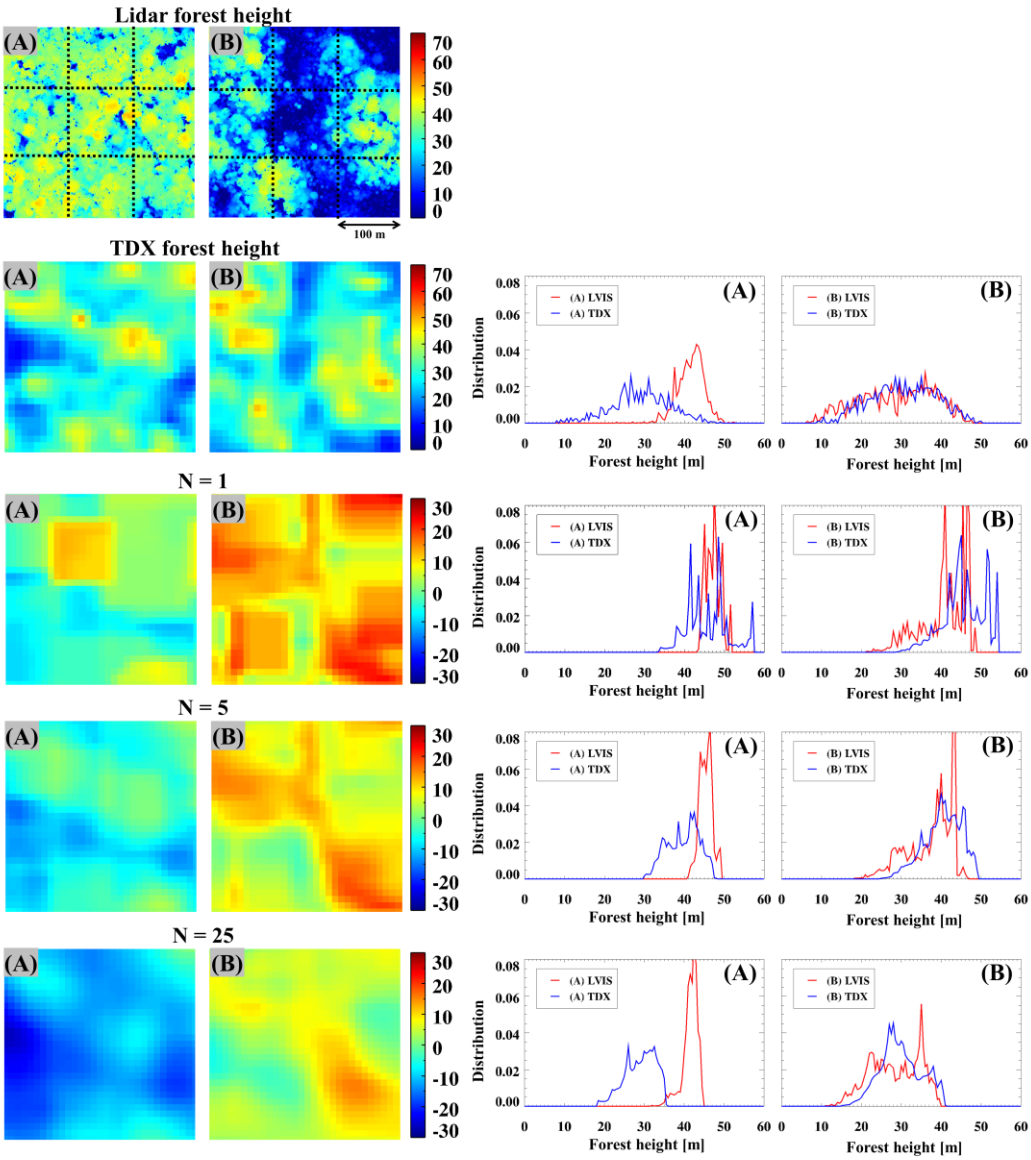
To explore this, 25 H_{TX} samples, equivalent to 16 independent estimates with 25 m resolution (see Table 3.1), are aggregated to a height estimate H_{TX100} at 100 m resolution

$$H_{TX100} = \frac{1}{N}\sum_{n=1}^{N} H_{TX}. \tag{3.12}$$

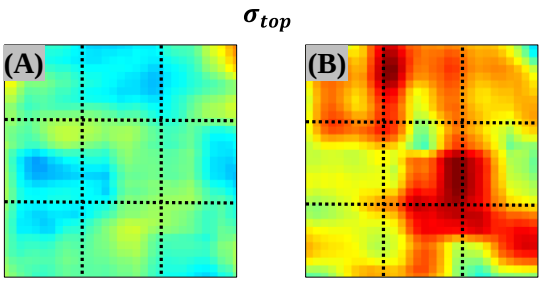
For sparse stands, the mean of the N=25 H_{TX} height samples are taken, while for dense stands only the mean of the tallest N=5 H_{TX} samples are taken. To discriminate between sparse and dense stands, the horizontal structure index σ_{top} is used.

The effect of this approach is demonstrated using two representative forest areas shown in Fig. 3.6, a dense one (A) and a sparse(r) one (B) [see Fig. 3.6 (b)], each about 300 m × 300 m. In area (A), the CHM, shown in the top row, varies about 10 m around a mean height of 45 m, indicating a rather dense forest. The estimated H_{TX} heights at 25 m resolution are shown in the

second row. From the comparison between the histogram of the estimated heights and that of the RH100 LVIS heights (H_{LVIS}) of the same area an underestimation of about 20 m is evident. Differently, in area (B), the CHM indicates a sparse forest. Here the histograms of H_{TX} and H_{LVIS} are very similar, except for heights lower than 15 m which appear overestimated in H_{TX}. For the rest, the histograms indicate an unbiased inversion performance, supported by a qualitative comparison between the CHM heights and the TanDEM-X heights in the first and second rows, respectively. In the third, fourth, and fifth rows, the aggregated forest height maps [in the sense of (3.12)] of the two areas and the associated height histograms corresponding to a 100 m resolution (e.g., 100 m × 100 m) are shown. For the maps shown in the third row, only the tallest of the 25 available height estimates H_{TX} is used, in the fourth row, the mean of the tallest 5 estimates, while in the fifth row, the mean of all 25 height estimates. With respect to $H_{LVIS100}$, comparing the two areas becomes obvious that while using the tallest height estimate compensates for the underestimation in the dense area (A), the sparse area (B) is heavily overestimated. The opposite is the case when using the mean of all estimated heights: in the dense area (A), the 100 m resolution heights still underestimate the reference heights, while in the sparse area (B), the 100 m resolution heights appear unbiased.



(a)

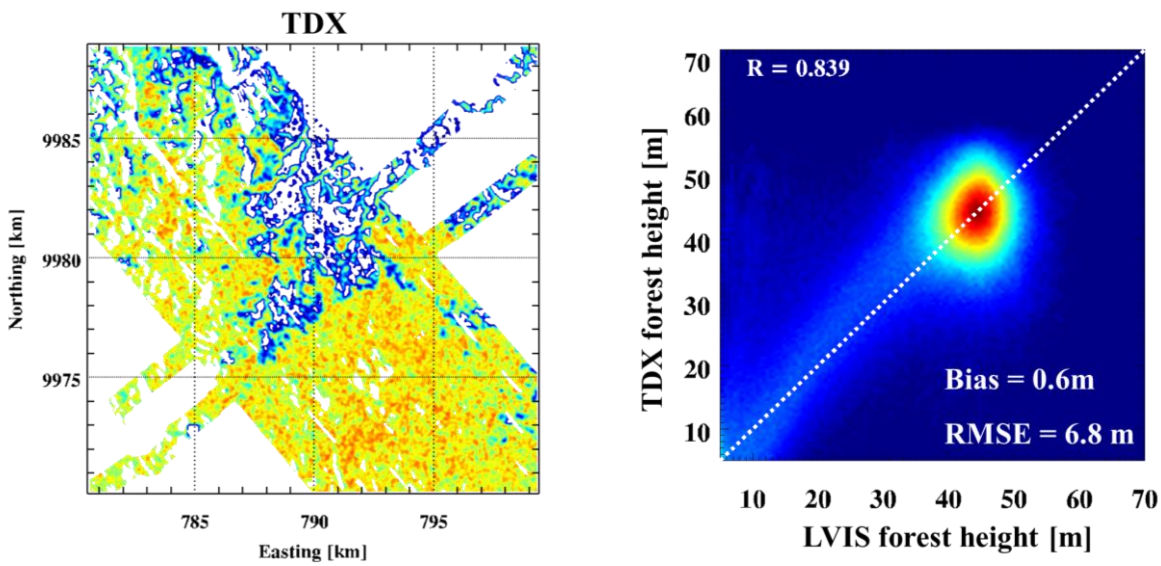


(b)

Fig. 3.6. Lopé site. (a) Height estimates and related errors for representative dense (area A) and sparse (area B) areas. The locations of areas A and B are illustrated in Fig. 3.5(a). First row: small-footprint lidar CHM. Second row: TanDEM-X H_{TX} (25 m resolution). Third, fourth, and fifth rows: height errors $H_{TX100} - H_{LVIS100}$ (100 m resolution) for N=1, 5, 25, respectively, as in (3.12) on the left; and histograms of H_{TX} and H_{LVIS} (25 m resolution) on the right. (b) Maps of σ_{top} over areas A and B. In the maps in (a) and (b) the black dashed lines separate nonoverlapping height resolution cells measuring 100 m × 100 m.

The final performance is shown in Fig. 3.7 in terms of a 2D histogram comparing H_{TX100} and $H_{LVIS100}$. As expected, N=5 corrects the height estimation bias in the tall/dense stands between 40 and 50 m. The residual (positive) bias of around 5 m in this height interval occurs in correspondence with the sparse stands, for which N=5 is not optimal. In contrast, N=25 minimizes the bias in the short(er)/sparse(r) stands. Notice that, in this case, a residual bias between 5 and 10 m persists especially for mean heights between 10 and 20 m and cannot be corrected further by (3.12). The result does not change significantly when taking the 3 or 7 tallest heights: 5 are selected because they correspond to the H100 metric that refers to the 100 tallest trees in a hectare. As single trees are not seen, the corresponding 20% of height estimates are used. Regarding the σ_{top} threshold for separating between sparse and dense stands, its selection is rather uncritical. Here, a value of 6 m was used. In principle, the threshold has to differentiate between very open and closed canopy forest. For a closed canopy forest, the H100 metric applied by using the average of the tallest 5 heights of the 25 samples does not change when the lower heights are underestimated. For open canopy forests, the top canopy height is becoming less and less appropriate with decreasing tree density. In this case, the mean of all heights is more representative.

As final remark, it was verified that if an optimal N is employed for each cell in the scene, the same conclusion on the residual bias would apply and the RMSE would improve from 6.8 to only 6.3 m. This confirms that the suboptimality of the processing has only minor effects on the final performance.



(a) (b)

Fig. 3.7. Lopé site, same area as in Fig. 3.1: (a) TanDEM-X forest height H_{TX100} in meters (m); (b) comparison (2D histogram) between H_{TX100} and $H_{LVIS100}$.

3.5 DERIVATION OF FOOTPRINT-LEVEL LIDAR ALLOMETRY

3.5.1 SAMPLING EFFECT

According to (3.1), the AGB is an exponential function of top canopy height H with constant allometric exponent β_0 over larger scales. For the case where also the allometric factor α is assumed to be constant over the whole site, the resulting allometric parameters $\alpha = \alpha_0$ and β_0 can be directly derived from the available forest height and AGB measurements, collected in the vectors $\mathbf{h}_{\mathrm{lid}}$ and $\mathbf{b}_{\mathrm{lid}}$, through a least-squares regression as

$$\min_{\alpha_0,\beta_0} \left\| \mathbf{b}_{\mathrm{lid}} - \alpha_0 \mathbf{h}_{\mathrm{lid}}^{\beta_0} \right\|^2. \tag{3.13}$$

The dependence of forest height-biomass allometry on local stand conditions [3]-[5], particularly their density, can be accounted for by the dependence of the allometric level $\alpha = \alpha(\sigma_{\mathrm{top}})$ on the horizontal structure index. In this sense, (3.13) can be modified to consider N_α uniformly distributed and non-overlapping structure intervals centered at $\left\{\sigma_{\mathrm{top}_i}\right\}_{i=1}^{N_\alpha}$ in the range of values of σ_{top}. For each of these intervals, the allometric levels $\left\{\alpha\left(\sigma_{\mathrm{top}_i}\right)\right\}_{i=1}^{N_\alpha}$ and a reference allometric exponent β_0 are jointly estimated as:

$$\min_{\left\{\alpha\left(s_{\mathrm{top}_i}\right)\right\}_{i=1}^{N_\alpha},\beta_0} \left\| \begin{bmatrix} \mathbf{b}_{\mathrm{lid},1} \\ \vdots \\ \mathbf{b}_{\mathrm{lid},i} \\ \vdots \\ \mathbf{b}_{\mathrm{lid},N_\alpha} \end{bmatrix} - \begin{bmatrix} \alpha\left(s_{\mathrm{top}_i}\right) \mathbf{h}_{\mathrm{lid},1}^{\beta_0} \\ \vdots \\ \alpha\left(s_{\mathrm{top}_i}\right) \mathbf{h}_{\mathrm{lid},i}^{\beta_0} \\ \vdots \\ \alpha\left(s_{\mathrm{top}_{N_\alpha}}\right) \mathbf{h}_{\mathrm{lid},N_\alpha}^{\beta_0} \end{bmatrix} \right\|^2 \tag{3.14}$$

where $\mathbf{b}_{\mathrm{lid,i}}$ and $\mathbf{h}_{\mathrm{lid,i}}$ are the vectors containing the lidar AGB and height values, respectively, for the generic i-th s_{top} interval centered at σ_{top_i}.

Accordingly, the height-to-biomass relationship and the dependence of the allometric level $\alpha = \alpha(\sigma_{\mathrm{top}})$ on the horizontal structure index are derived from the available set of forest height and AGB measurements. In all the considered cases, the optimization (3.14) was carried out using N_α =50 for σtop varying between 0 and 10 m. In the case of GEDI, a more or less sparse (depending on latitude and cloud cover) sampling of forest height and AGB measurements at footprint level is available. In order to evaluate how much the available measurements are sufficient to derive allometric relations, four different scenarios are considered: the full data set scenario using all available LVIS forest height (H_{LVIS100}) and AGB (B_{LVIS100}) measurements and three thinned scenarios derived by randomly reducing the full LVIS data set along simulated GEDI ground tracks according to three different cloud cover rates (50%, 75%, and 90%). Each of

the three thinned scenarios has been generated 200 times (trials), each time by removing a different set of measurements. The number of lidar height and biomass measurements available in each σ_{top} interval varies between 50 and 400 in the full data set, but it can be reduced to around 10 in the most challenging sampling scenario.

The top row of Fig. 3.8 shows the obtained allometries derived according to (3.13) for the case of constant $\alpha = \alpha_0$. In the reference full data set case, the 2D histogram shows the distribution of B_{LVIS100} as a function of H_{LVIS100}, while the white dashed lines represent the allometry for the fitted α_0 (central line) and for the cases in which α_0 is increased/decreased by 30%. In the other three plots, the blue dashed lines represent the allometries obtained in each of the 200 trials. The allometric relationship results are very stable with respect to the number of available samples, even in the case where only 10% of the samples are available (e.g., 90% cloud cover).

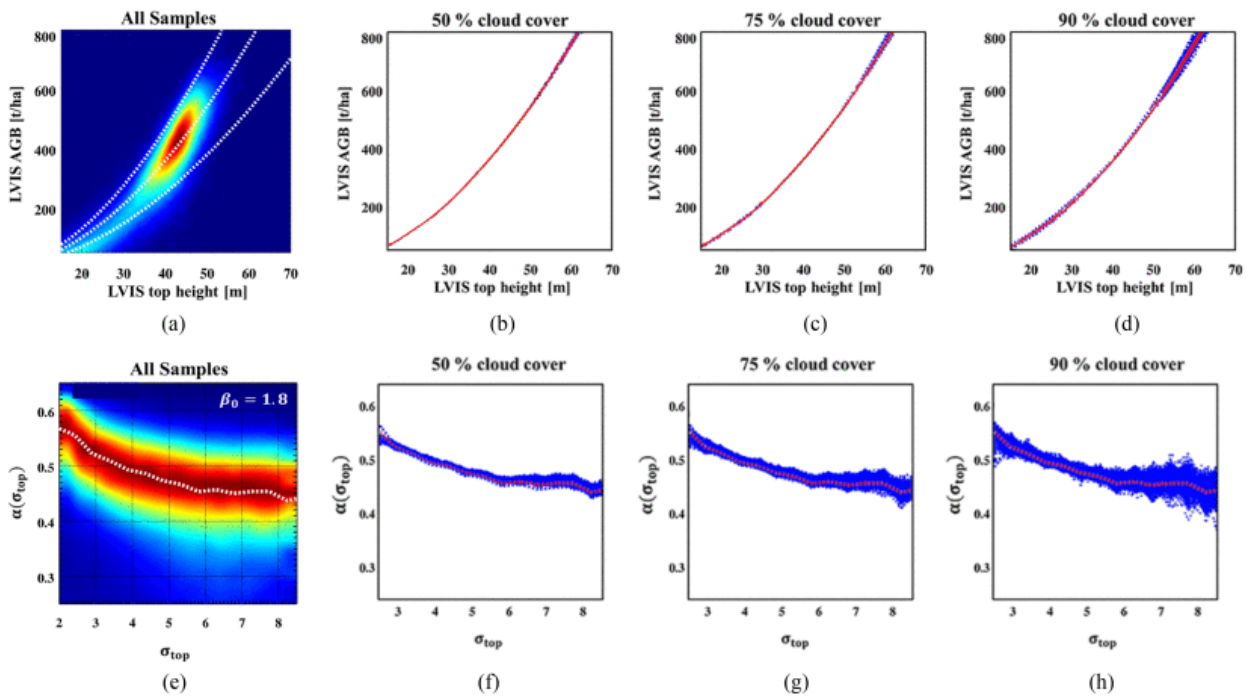


Fig. 3.8. Lopé site, same area as in Fig. 3.1. Top row: (a) 2D histogram relating HLVIS100 and BLVIS100 in correspondence of all the GEDI sampling positions in Fig. 3.1(b); the central white dashed line represents the height-to-biomass relationship with constant $\alpha = \alpha_0$ obtained using (3.13), while the upper/lower dashed lines represent the cases in which the fitted α_0 is increased/decreased by the 30%. Panels (b)–(d): the blue dashed lines represent the allometric relationships for $\alpha = \alpha_0$ obtained using (3.13) in each trial of the simulated 50%, 75%, and 90% cloud covers, respectively. The same relationship obtained in the full-sampling case [as in panel (a)] is reported in red for reference. Bottom row: (e) 2D histogram relating the allometric factor α and σtop using H_{LVIS100}, B_{LVIS100} and β_0=1.8 in correspondence with all the GEDI sampling positions in Fig. 3.1(b); the white dashed line represents $\alpha(\sigma_{\text{top}})$ obtained using (3.14). Panels (f)–(h): the blue dashed lines represent the relationships represents $\alpha(\sigma_{\text{top}})$ obtained using (3.14) in each trial of the simulated 50%, 75%, and 90% cloud covers. The same relationship obtained in the full-sampling case [as in panel (e)] is reported in red for reference.

The bottom row of Fig. 3.8 shows the obtained allometries derived according to (3.14) for the case of a variable allometric level $\alpha = \alpha(\sigma_{\text{top}})$. In the reference full data set case, the 2D histograms show the distribution of $\alpha = B_{\text{LVIS100}}/H_{\text{LVIS100}}^{\beta_0}$ at the sampling locations with $\beta_0 = 1.8$. It is apparent that the obtained allometry is less stable than with $\alpha = \alpha_0$ when the number of available samples decreases because (3.14) demands more samples to define every allometric parameter in each interval. This is especially true with increasing forest heterogeneity. While for the cases of 50% and 75% of cloud cover $\alpha = \alpha(\sigma_{\text{top}})$ can be reconstructed, in the case of 90% cloud cover the $\alpha = \alpha(\sigma_{\text{top}})$ relationship cannot be established any longer.

In the case of limited samples, one possible tradeoff is to increase the intervals used to fit the $\alpha = \alpha(\sigma_{\text{top}})$ relationship [see (3.14)] and so doing to reduce their number N_α. In this case, a more robust allometric relationship could be obtained at the cost of a low(er) structure resolution. However, the (real) GEDI samples available over Lopé correspond to almost 50% cloud cover case making the reconstruction of both allometries possible.

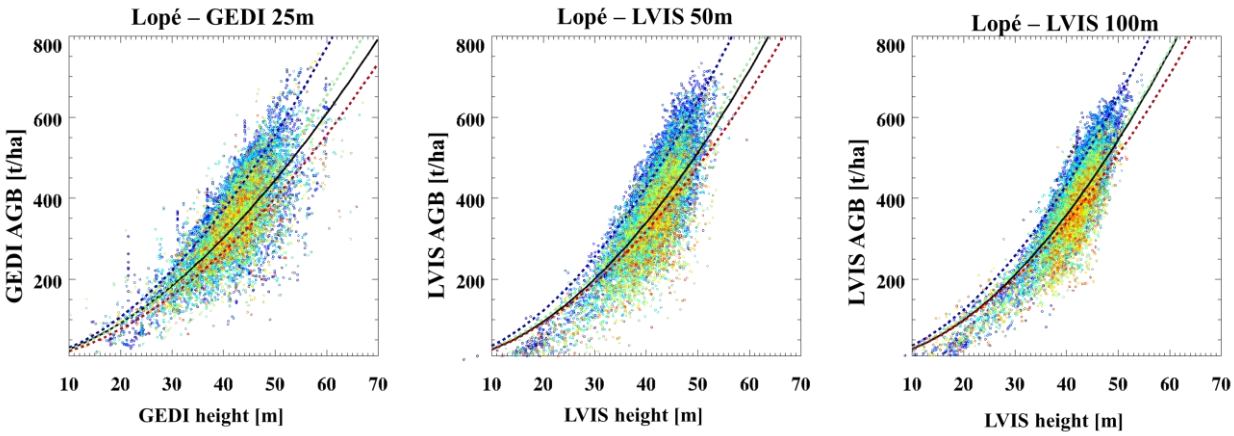
3.5.2 Scale Effect

After investigating the effect of available samples on the reconstructed allometry, the next question to face concerns the spatial scales on which the allometry is addressed. While GEDI provides measurements (forest height and AGB) at the footprint level of approximately 25 m diameter, the structure index σ_{top} used to refine the allometry is estimated at a 100 m resolution. A lower resolution of σ_{top} would not include a statistically relevant number of CHPs, thus not providing a significant structure description.

In order to investigate the effect of this scale discrepancy on the parameterization of the height-to-biomass relationship, the allometric exponent and the (structure-dependent) allometric level are derived at two different spatial resolutions, 50 and 100 m using the reference LVIS biomass estimates and compared with the ones obtained by using the 25 m GEDI estimates H_{GEDI} and B_{GEDI}. The obtained allometries are shown in Fig. 3.9. For each resolution, the regression has been performed individually. For the case of a constant allometric coefficient in all three resolutions, a very similar parameterization has been obtained: $\alpha_0 = 0.454$ and $\beta_0 = 1.76$ at 25 m, $\alpha_0 = 0.392$ and $\beta_0 = 1.83$ at 50 m, and $\alpha_0 = 0.383$ and $\beta_0 = 1.85$ at 100 m, respectively.

The color of each point in the scatterplot corresponds to a different value of σ_{top}. As expected, for the same height, the allometric factor decreases with increasing σ_{top} as a consequence of a decrease in (forest) density. This demonstrates the potential of using σ_{top} to adapt the height-to-biomass allometry to the local forest conditions. The adaptation is more effective toward the extremes of the σtop range where the correlation between σ_{top} and the allometric factor is higher. At the 25 m scale, the adaptation appears less effective as σ_{top} appears less correlated to the allometric factor. This can be a result of the large-scale difference between the GEDI samples

and σ_{top}. At the same time, the rather low geolocation accuracy of the GEDI footprints of about 10 m at 1 sigma [47] is not supportive.



(a) (b) (c)

Fig. 3.9. Lopé site, same area as in Fig. 3.1. Scatterplots between (a) H_{GEDI} and B_{GEDI}, (b) H_{LVIS50} and B_{LVIS50}, and (c) $H_{LVIS100}$ and $B_{LVIS100}$. The points in the scatterplot correspond to a spatially homogenous subset of the acquired GEDI footprint coverage. The color of each point corresponds to a different value of σ_{top} (expressed in meters) with the same color map as in Fig. 3.4. The continuous black line represents the height-to-biomass allometry parameterized by $\alpha = \alpha_0$. The dashed colored lines represent the case $\alpha = \alpha(\sigma_{\text{top}})$ for three value of σ_{top} (blue: 2 m, green: 5 m, red: 9 m).

3.6 FOREST BIOMASS ESTIMATION AND VALIDATION

The performance of the height-to-biomass allometry at 100 m resolution is now addressed by exploring the established dependency of the allometric factor σ on the horizontal structure index σtop.

In order to assess the improvement obtained by using the adaptive allometric relation assuming no uncertainty in the forest height, first the LVIS heights $H_{LVIS100}$ are used in the constant (with $\alpha = \alpha_0$) and in the adaptive [with $\alpha = \alpha(\sigma_{\text{top}})$] allometric relations derived using H_{GEDI} and B_{GEDI}. In the latter case, the relationship between the allometric factor and σ_{top} is shown in Fig. 3.10 (white dashed line). The optimization (3.14) was carried out using $N_\alpha = 50$ for σtop varying between 0 and 10 m. The obtained AGB maps for both cases are shown on the top row (left and middle respectively) of Fig. 3.11. At the right, the reference BLVIS100 derived at 100 m resolution is shown. The validation plots of the obtained AGB maps against the reference AGB are shown on the top row of Fig. 3.12. The constant $\alpha = \alpha_0$ allometry already provides sensitive results, but the high AGB levels are consistently underestimated and low AGB levels tend to be overestimated. Both effects are compensated when applying the adaptive $\alpha = \alpha(\sigma_{\text{top}})$ allometry that provides almost unbiased estimates.

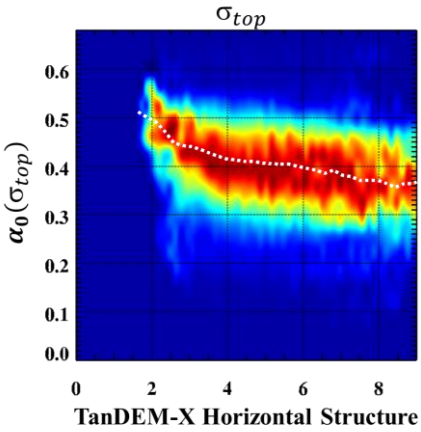
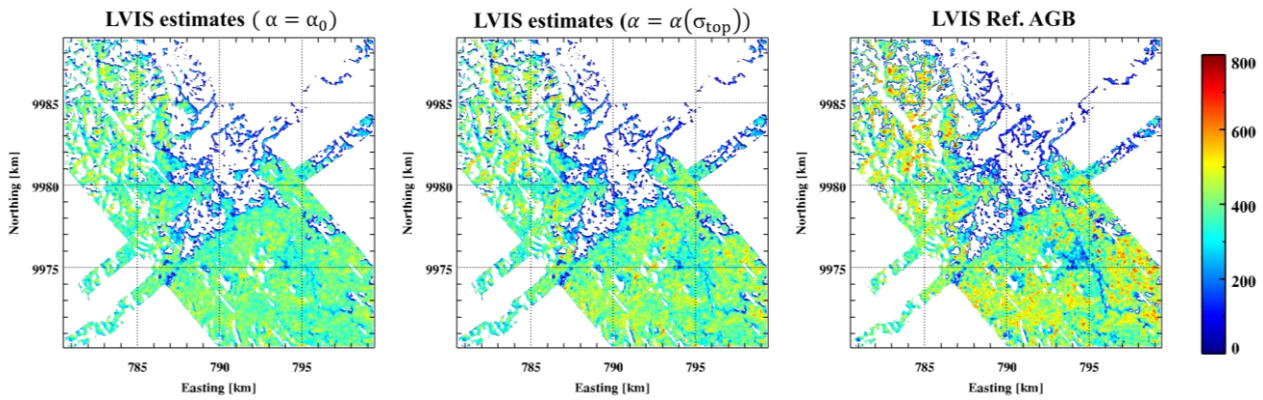
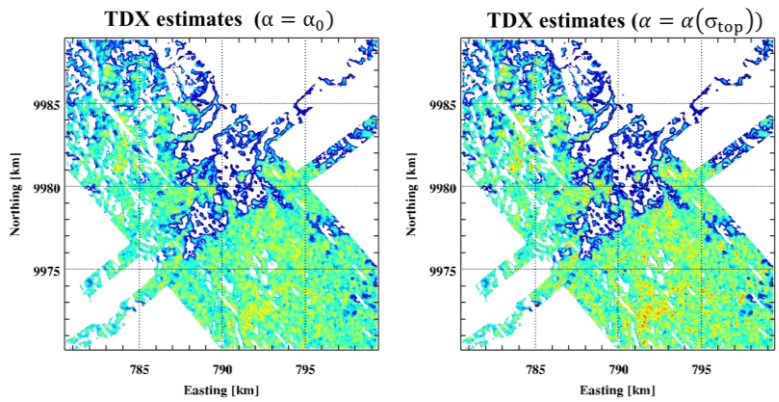


Fig. 3.10. Lopé site, same area as in Fig. 3.1. 2D histograms relating α_0 and σ_{top} using H_{GEDI} and B_{GEDI}. The white dashed lines represent the relationship $\alpha(\sigma_{\text{top}})$ obtaine using (3.14).



(a)



(b)

Fig. 3.11. Lopé site, same area as in Fig. 3.1. AGB maps obtained from allometric relationships derived using H_{GEDI} and B_{GEDI}, and using (a) $H_{LVIS100}$ with $\alpha = \alpha_0$ (right), $H_{LVIS100}$ with $\alpha = \alpha(\sigma_{\text{top}})$ (middle), and reference $B_{LVIS100}$ (left); and (b) H_{TX100} with fixed $\alpha = \alpha_0$ (left), and with $\alpha = \alpha(\sigma_{\text{top}})$.

The same procedure is now repeated using instead of the LVIS heights $H_{LVIS100}$ the TanDEM-X heights H_{TX100}. The obtained AGM maps are shown on the bottom row of Fig. 3.11, while the corresponding validation plots on the bottom row of Fig. 3.12. The forest height uncertainty dominates the obtained performance. Nevertheless, the adaptive allometry successfully compensates for the overestimation of the lower AGB range as well as the underestimation of the upper AGB range of the constant allometry, allowing for practically unbiased estimates.

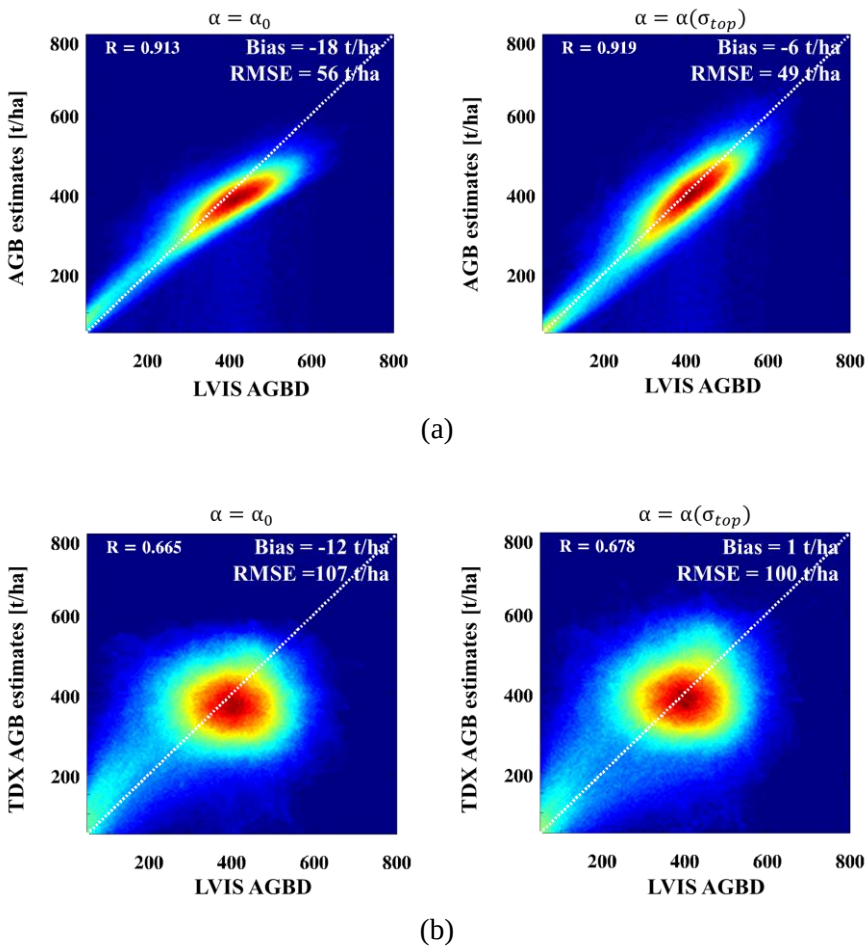


Fig. 3.12. Comparison (2D histograms) between the reference $B_{LVIS100}$ and the estimated AGB values using (a) $H_{LVIS100}$ and (b) H_{TX100}. In both (a) and (b), $\alpha = \alpha_0$ is used in the left panel, and a variable allometric factor $\alpha = \alpha(\sigma_{top})$ is used in the right panel.

3.7 CONCLUSION

The use of the forest height-to-biomass allometry, as addressed in (3.1), in the context of continuous TanDEM-X InSAR measurements and discrete height and biomass GEDI measurements over a diverse tropical test site, the Lopé National Park in Gabon, is discussed. Important points are the importance of forest structure and the possibility of deriving and using a forest structure index from TanDEM-X data in the absence of a DTM.

Two points are particularly critical to the success of such an approach: 1) the knowledge of the allometric parameters α_0 and β_0 that define the height-to-biomass allometry and their spatial variation in heterogeneous forests; and 2) the ability of TanDEM-X interferometric measurements to provide unbiased forest height estimates in dense forest conditions. In both cases, forest structure and its spatial variability play a decisive role. The proposed methodology illustrates how the sensitivity of the TanDEM-X interferometric measurements of horizontal forest structure, given by the high spatial resolution and the high attenuation at X-band, can be used to account for the forest heterogeneity and support this way both points.

For this purpose, a horizontal structure index was proposed, similar to the one in [18], but modified to be derived from relative height variations and thus not requiring a DTM. The height variations induced by the topography were compensated by using a low-pass filtered version of the interferogram to access the spatial variations of the top canopy layer. However, as the proposed structural index only takes into account relative height variations, it cannot reliably recognise whether the height changes occur in the upper tree canopy or within the volume. To maximise the sensitivity to the top canopy variations, CHPs are used. The horizontal structure index is then given by the spatial variance of the CHP "top" peaks within a structure cell. Accordingly, a large variance resulting from large canopy height variations is associated with a large canopy roughness and interpreted as "sparse forest." On the other hand, a low variance resulting from low canopy height variations is associated with low canopy roughness and is interpreted as a close canopy, e.g., a "denser forest."

The derived horizontal structure index is used to improve both the forest height estimation and the forest height-to-biomass allometry performance.

The main limitation in forest height estimation is the underestimation of dense stands caused by the limited penetration at X-band that leads to biased (underestimated) heights. Assuming that even in dense stands there may be points where the X-band pulses penetrate to the ground, an unbiased estimation can be attempted at the expense of spatial resolution. For this, the derived horizontal structure index is used to distinguish between "sparse" and "dense" forests: for the "sparse" forest stands a 100 m height estimate is obtained by averaging 25 height estimates at 25 m, for the "dense" forest stands only the mean of the highest 5 estimates is considered.

Looking on the forest height-to-biomass allometry now, the underlying allometry can be derived from the GEDI footprint measurements, i.e., the RH100 heights and the associated AGB values. The simulation of scenarios with different sampling densities shows that the derived underlying allometry is robust to the number of available footprint measurements used to define it, as long as the samples remain representative of forest conditions. This does not appear to be a critical limitation as the underlying allometry, which is primarily dependent on large scale forest attributes (as species composition and the site growth conditions), remains at larger scales. However, this underlying allometry in Lopé underestimates high AGB levels as the spatial forest heterogeneity cannot be represented by a single allometric relation.

The horizontal structure index is here used to adapt the underlying height-to-biomass allometry to the spatial (stand density) heterogeneity for improving biomass estimation performance. This has been attempted by expressing the allometric level as a function of the

derived horizontal structure index and using the GEDI footprint measurements, i.e., the RH100 heights, and the associated AGB values to reconstruct this dependency. The obtained monotonous decreasing dependence of the allometric level on the horizontal structural index points to the presence of a real(istic) correlation. This reflects on the achieved biomass estimation performance improvement, when compared to the use of a single allometry. This improvement has been obtained even if the horizontal structure index cannot distinguish vertical variations of density along height. The largest remaining uncertainty contribution in terms of bias and/or variance has been seen to be attributable to the propagation of the height estimation uncertainty.

Even if a single test site has been investigated, the intention of this article is to develop a methodological concept under the perspective of forest biomass inversion on a large scale rather than discussing the optimization of performance on local scales. Nevertheless, the same concept applied within the same test site, but with different acquisition configurations in terms of incidence angles and vertical wavenumber, leads to the same conclusions. Surely, in the future some local optimization may improve the final performance. The extension of this analysis to other test sites is, however, complicated by the availability of large-scale continuous data for validation, especially in terms of biomass.

The results characterize the potential and the limitations of TanDEM-X interferometry for characterizing forest conditions. On the one hand, the high attenuation rates at X-band and the resulting limited penetration into the forest volume maximize the interferometric sensitivity to the spatial variations of the top canopy layer and make it especially appropriate for the characterization of the horizontal forest structure. At the same time, the high spatial resolution of the TanDEM-X interferograms and its continuous measurement nature allows the estimation of forest structure variations at spatial scales relevant for the characterization of the horizontal forest structure. On the other hand, the same limited penetration into the forest volume, which favors the horizontal forest structure characterisation, limits the height estimation performance and makes the characterisation of the vertical forest structure at reasonable scales almost impossible.

Finally, the proposed methodology and the obtained results demonstrate the synergetic potential of the continuous TanDEM-X and the discrete GEDI measurements. This is because both measurements 1) are at the same time similar enough due to the high sensitivity to the geometrical architecture of the canopy and the high spatial resolution common to both configurations, facilitating a common interpretation and) are different enough because of the different acquisition geometries and measurement approaches to carry independent information.

3.8 ACKNOWLEDGMENT

The lidar AfriSAR data sets were provided by the Land, Vegetation and Ice Sensor (LVIS) team in Code 61A at NASA Goddard Space Flight Center with support from the University of Maryland, College Park, Maryland, USA.

3.9 REFERENCES

[1] T. Mette, K. Papathanassiou, I. Hajnsek, H. Pretzsch and P. Biber, “Applying a common allometric equation to convert forest height from Pol-InSAR data to forest biomass,” IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, pp. 269–272, 2004.

[2] G. P. Asner, J. Mascaro, H. C. Muller-Landau, G. Vieilledent, R. Vaudry, M. Rasamoelina, J. S. Hall, and M. van Breugel, “A universal airborne LiDAR approach for tropical forest carbon mapping,” Oecologia, vol.168, no. 4, pp. 1147–1160, Apr. 2012.

[3] T. Mette, K. Papathanassiou and I. Hajnsek, “Biomass estimation from polarimetric SAR interferometry over heterogeneous forest terrain,” IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, 2004, pp. 511–514.

[4] A. Torano Caicoya, F. Kugler, I. Hajnsek, and K. Papathanassiou, “LargeScale Biomass Classification in Boreal Forests with TanDEM-X Data,” IEEE Trans. Geosci. Remote Sens., vol. 54, no. 10, pp. 5935–5951, Oct. 2016.

[5] N. Knapp, R. Fischer, and A. Huth, “Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states,” Remote Sens. Environ., vol. 205, Feb. 2020, 111597.

[6] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, M. Zink, "TanDEM-X: A satellite formation for high-resolution SAR interferometry", IEEE Trans. Geosci. Remote Sens., vol. 45, no. II, pp. 3317-3341, 2007.

[7] R. Dubayah, J. B. Blair, S. Goetz, L. Fatoyinbo, M. Hansen, S. Healey, M. Hofton, G. Hurtt, J. Kellner, S. Luthcke, and J. Armston, "The global ecosystem dynamics investigation: High-resolution laser ranging of the earth's forests and topography", Sci. Remote Sens., vol. 1, pp. 100002. 2020.

[8] L. Duncanson, J. R. Kellner, J. Armston, R. Dubayah, D. M. Minor, S. Hancock, and C. Zgraggen, “Aboveground biomass density models for NASA’s Global Ecosystem Dynamics Investigation (GEDI) lidar mission.” Remote Sens. Environ., vol. 270, pp. 112845, Mar. 2022.

[9] M. Brolly, M. Simard, H. Tang, R. O. Dubayah, and J. P. Fisk, “A lidar-radar framework to assess the impact of vertical forest structure on interferometric coherence,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 9, no. 12, pp. 5830–5841, Dec. 2016.

[10] E. C. De Grandi, E. Mitchard, and D. Hoekman, “Wavelet based analysis of TanDEM-X and LiDAR DEMs across a tropical vegetation heterogeneity gradient driven by fire disturbance in Indonesia,” Remote Sens., vol. 8, no. 8, pp. 641–667, Aug. 2016.

[11] P. Da Conceição Bispo, M. Pardini, K. P. Papathanassiou, F. Kugler, H. Balzter, D. Rains, J. R. dos Santos, I. G. Rizaev, K. Tansey, M. N. dos Santos, and L. Spinelli Araujoi, “Mapping forest successional stages in the Brazilian Amazon using forest heights derived from TanDEM-X SAR interferometry,” Remote Sens. Environ., vol. 232, pp. 111194, Jul. 2019.

[12] S. Erasmi, M. Semmler, P. Schall, and M. Schlund, “Sensitivity of bistatic TanDEM-X data to stand structural parameters in temperate forests,” Remote Sens., vol. 11, pp. 2966, Dec. 2019.

[13] J. I. H. Askne, H. J. Persson, and L.M.H. Ulander, “On the sensitivity of TanDEM-X observations to boreal forest structure,” Remote Sens., vol. 11, pp. 1644, Dec. 2019.

[14] M. J. Soja, H. J. Persson, and L. M. H. Ulander, “Estimation of Forest Height and Canopy Density from a Single InSAR Correlation Coefficient,” IEEE Geosci. and Remote Sens. Letters, vol. 12, no. 3, pp. 646-650, 2014.

[15] S. Solberg, E. H. Hansen, T. Gobakken, E. Naessset, and E. Zahabu, “Biomass and InSAR height relationship in a dense tropical forest,” Remote Sens. Environ., vol. 192, pp. 166-175, Apr. 2017.

[16] M. Tello Alonso, V. Cazcarra Bes, M. Pardini, and K. Papathanassiou, “Forest structure characterization from SAR tomography at L-band,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 11, no. 10, pp. 3402–3414, Oct. 2018.

[17] V. Cazcarra Bes, M. Tello Alonso, R. Fischer, M. Heym, and K. Papathanassiou, “Monitoring of forest structure dynamics by means of Lband SAR tomography,” Remote Sens., vol. 9, no. 12, pp. 1229–1250, Nov. 2017.

[18] C. Choi, M. Pardini, M. Heym, K. Papathanassiou, “Improving Forest Height-to-Biomass Allometry with Structure Information: A TanDEM-X Study,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 14, 2021.

[19] L. Reineke, “Perfecting a stand-density index for even aged forests,” J. Agricultural Research, vol. 46, no. 7, pp. 627–638, 1933.

[20] H. Chen, S. R. Cloude, D. G. Goodenough, D.A. Hill, & A. Nesdoly, “Radar forest height estimation in mountainous terrain using Tandem-X coherence data.” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 11, no. 10, pp. 3443-3452, 2018.

[21] M. Schlund, A. Wenzel, N. Camarretta, C. Stiegler, and S. Erasmi. "Vegetation canopy height estimation in dynamic tropical landscapes with TanDEM-X supported by GEDI data," Methods in Ecology and Evolution, 2022.

[22] W. Qi, S. K. Lee, S. Hancock, S. Luthcke, H. Tang, J. Armston, and R. Dubayah. "Improved forest height estimation by fusion of simulated GEDI Lidar data and TanDEM-X InSAR data." Remote Sens. Environ., vol. 221, pp. 621-634, 2019.

[23] Y. Lei, R. Treuhaft, and F. Gonçalves. "Automated estimation of forest height and underlying topography over a Brazilian tropical forest with single-baseline single-polarization TanDEM-X SAR interferometry." Remote Sens. Environ., vol. 252, pp. 112132, Jan. 2021.

[24] R. Guliaev, V. Cazcarra-Bes, M. Pardini, and K. Papathanassiou, “Forest Height Estimation by Means of TanDEM-X InSAR and Waveform Lidar Data,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., Vol. 14, pp. 3084-3094, Feb. 2021.

[25] C. Choi, V. Cazcarra-Bes, R. Guliaev, W. Qi, M. Pardini, J. Armston, R. Dubayah, and K. Papathanassiou, “Large-Scale Forest Height Mapping by Combining TanDEM-X and GEDI Data,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 16, pp. 2374-2385, Feb. 2023.

[26] I. Hajnsek, M. Pardini, M. Jäger, R. Horn, J. S. Kim, H. Jörg, and K. Papathanassiou, Technical assistance for the development of airborne SAR and geophysical measurements during the AfriSAR campaign, Final technical report, ESA contract no. 4000114293/15/NL/CT. Available at: https://earth.esa.int/documents/10174/134665/AfriSAR-Final-Report.

[27] L. Fatoyinbo, J. Armston, M. Simard, S. Saatchi, M. Denbina, M. Lavalle, M. Hofton, N. Pinto, S. Hancock, H. Tang, S. Marselis, B. Hawkins, L. Duncanson, B. Blair, C. Hansen, Y. Lou, R. Dubayah, S. Hensley, C. Silva, J. Poulsen, N. Labriere, N. Barbier, K. Jeffery, D. Kenfack, A. Alonso, G. Moussavou, S. Lewis, and K Hibbard, “The NASA AfriSAR Campaign: Airborne SAR and Lidar Measurements of Tropical Forest Structure and Biomass in Support of Future Space Missions,” Remote Sens. Environ., vol. 264, pp. 112533, Oct. 2021.

[28] J. Armston, H. Tang, S. Hancock, S. Marselis, L. Duncanson, J. Kellner, M. Hofton, J.B. Blair, T. Fatoyinbo, and R.O. Dubayah, AfriSAR: Gridded Forest Biomass and Canopy Metrics Derived from LVIS, Gabon, 2016. ORNL DAAC, Oak Ridge, Tennessee, USA. doi: https://doi.org/10.3334/ORNLDAAC/1775.

[29] N. Labriere, S. Tao, J. Chave, K. Scipal, T. L. Toan, K. Abernethy, A. Alonso, N. Barbier, P. Bissiengou, T. Casal, S. J. Davies, A. Ferraz, B. Hérault, G. Jaouen, K. J. Jeffery, D. Kenfack, L. Korte, S. L. Lewis, Y. Malhi, H. R. Memiaghe, J. R. Poulsen, M. Réjou-Méchain, L. Villard, G. Vincent, L. J. T. White, and S. S. Saatchi, “In Situ Reference Datasets from the TropiSAR and AfriSAR Campaigns in Support of Upcoming Spaceborne

Biomass Missions," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 11, no. 10, pp. 3617–3627, Oct. 2018.

[30] S. M. Marselis, H. Tang, J. D. Armston, K. Calders, N. Labrière, and R. Dubayah, "Distinguishing vegetation types with airborne waveform lidar data in a tropical forest-savanna mosaic: A case study in Lopé National Park, Gabon," Remote Sens. Environ., vol. 216, pp. 626–634, Oct. 2018.

[31] E. T. A. Mitchard, S. S. Saatchi, L. J. T. White, K. A. Abernethy, K. J. Jeffery, S. L. Lewis, M. Collins, M. A. Lefsky, M. E. Leal, I. H. Woodhouse, and P. Meir, "Mapping tropical forest biomass with radar and spaceborne LiDAR in Lopé National Park, Gabon: overcoming problems of high biomass and persistent cloud," Biogeosci., vol. 9, no. 1, pp. 179–191, Jan. 2012.

[32] C. A. Silva, S. S. Saatchi, M. Garcia, N. Labriere, C. Klauberg, A. Ferraz, V. Meyer, K. J. Jeffery, K. Abernethy, L. White, K. Zhao, S. L. Lewis, and A. T. Hudak, "Comparison of Small- and Large-Footprint Lidar Characterization of Tropical Forest Aboveground Structure and Biomass: A Case Study from Central Gabon," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 11, no. 10, pp. 3512-3526, Oct. 2018.

[33] F. Kugler, S. K. Lee, I. Hajnsek, and K. P. Papathanassiou, "Forest Height Estimation by Means of Pol-InSAR Data Inversion: The Role of the Vertical Wavenumber," IEEE Trans. Geosci. Remote Sens., vol. 53, no. 10, pp. 5294–5311, Oct. 2015.

[34] R. O. Dubayah, M. Hofton, J. B. Blair, J. Armston, H. Tang, and S. Luthcke, GEDI L2A elevation and height metrics data global footprint level V001, 2020b. NASA EOSDIS Land Processes DAAC, Sioux Falls, South Dakota, USA. doi: https://doi.org/10.5067/GEDI/GEDI02A.001.

[35] R. O. Dubayah, J. Armston, J. R. Kellner, L. Duncanson, S. P. Healey, P. L. Patterson, S. Hancock, H. Tang, M.A. Hofton, J. B. Blair, and S. B. Luthcke, GEDI L4A Footprint Level Aboveground Biomass Density, Version 1, 2021. ORNL DAAC, Oak Ridge, Tennessee, USA. doi: https://doi.org/10.3334/ORNLDAAC/1907.

[36] R. Treuhaft, B. D. Chapman, J. Roberto Dos Sãntos, F. G. Gonçalves, L. V. Dutra, P. MLA de Graça, and J. B. Drake. "Vegetation profiles in tropical forests from multibaseline interferometric synthetic aperture radar, field, and lidar measurements," Journal of Geophysical Research: Atmospheres, vol. 114, no. D23, Dec. 2009.

[37] R. Treuhaft, F. Gonçalves, J. Roberto dos Santos, M. Keller, M. Palace, S. N. Madsen, F. Sullivan, and P. MLA de Graça. "Tropical-forest biomass estimation at X-band from the spaceborne TanDEM-X interferometer," IEEE Geosci. and Remote Sens. Letters, vol. 12, no. 2 pp. 239-243, Aug. 2014.

[38] E. C. De Grandi, E. Mitchard, I.H. Woodhouse and G. D. De Grandi, "Spatial wavelet statistics of SAR backscatter for characterizing degraded forest: A case study from Cameroon", IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens., vol. 8, no. 7, pp. 3572-3584, May. 2015.

[39] F. Kugler, D. Schulze, I. Hajnsek, H. Pretzsch, and K. Papathanassiou, "TanDEM-X Pol-InSAR Performance for Forest Height Estimation," IEEE Trans. Geosci. Remote Sens., vol. 52, no. 10, pp. 6404-6422, Oct. 2014.

[40] G. D. De Grandi and E. C. De Grandi, "Spatial Analysis for Radar Remote Sensing of Tropical Forests", Boca Raton, CRC Press, 2021, pp. 104-110.

[41] G. R. Lee, R. Gommers, F. Wasilewski, K. Wohlfahrt, and A. O'Leary, "PyWavelets: A Python package for wavelet analysis," Journal of Open Source Software, vol. 4, no. 36, pp. 1237, Apr. 2019.

[42] R. Lanari, G. Fornaro, D. Riccio, M. Migliaccio, K. P. Papathanassiou, J.R. Moreira, M. Schwabisch, L. Dutra, G. Puglisi, G. Franceschetti, amd M. Coltelli, "Generation of digital elevation models by using SIR-C/X-SAR multifrequency two-pass interferometry: the Etna case study," IEEE Trans. Geosci. Remote Sens., vol. 34, no. 5, pp. 1097-1114, Sep. 1996.

[43] G. H. X. Shiroma, and M. Lavalle. "Digital terrain, surface, and canopy height models from InSAR backscatter-height histograms." IEEE Trans. Geosci. Remote Sens., vol. 58, no. 6, pp. 3754-3777, Jan. 2020.

[44] C. Choi, M. Pardini and K. Papathanassiou, “Quantification of horizontal forest structure from high resolution TanDEM-X interferometric coherences,” IEEE International Geoscience and Remote Sensing Symposium, Valencia, pp. 376-379, 2018.

[45] M. Martone, P. Rizzoli, C. Wecklich, C. González, J. L.Bueso-Bello, P. Valdo, A. Moreira, “The Global Forest/Non-Forest Map from TanDEM-X Interferometric SAR data”, Remote Sens. Environ., vol. 205, pp. 352-373, Feb. 2018.

[46] M. Martone, B. Bräutigam, P. Rizzoli, C. Gonzalez, M. Bachmann, and G. Krieger, “Coherence evaluation of TanDEM-X interferometric data,” ISPRS J. Photogramm. Remote Sens., vol. 73, pp. 21–29, Sep. 2012.

[47] R. O. Dubayah, J. Beck, B. Wirt, J. Armston, M. Hofton, S. Luthcke, and H. Tang, Global Ecosystem Dynamics Investigation (GEDI) Level 02 User Guide: version 2.0, USGS contract no. G15PC00012. Available at: https://lpdaac.usgs.gov/documents/986/GEDI02_UserGuide_V2.pdf

4 LARGE SCALE FOREST HEIGHT MAPPING BY COMBINING TANDEM-X AND GEDI DATA

C. Choi, V. Cazcarra-Bes, R. Guliaev, M. Pardini, K. P. Papathanassiou, W. Qi, J. Armston, and R. Dubayah

IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

Published in vol.16, pp. 2374 – 2385, February 2023. DOI: 10.1109/JSTARS.2023.3244866

This Chapter is a post-print, differing from the published paper only in terms of layout and formatting

The author's contributions:

- Interferometric processing and analysis of the TanDEM-X data.
- Suggestion, implementation, and assessment of forest height inversion and mosaicking procedures.
- Writing of the manuscript.

The co-authors' contributions:

- M. Pardini and K. P. Papathanassiou provided guidance throughout the research.
- M. Pardini and K. P. Papathanassiou contributed to the main ideas, the discussion of the results, and reviewed the manuscript.
- V. Cazcarra-Bes and R. Guliaev contributed to the implementation of forest height inversion and performance analysis.
- J. Armston, W. Qi, and R. Dubayah contributed to the discussion of the parameterization of GEDI data for forest height inversion.

LARGE SCALE FOREST HEIGHT MAPPING BY COMBINING TANDEM-X AND GEDI DATA

Changhyun Choi[1,2], Victor Cazcarra-Bes[3], Roman Guliaev[1], Matteo Pardini[1], Konstantinos P. Papathanassiou[1], Wenlu Qi[4], John Armson[4], and Ralph Dubayah[4]

[1] German Aerospace Center, Microwaves and Radar Institute, Wessling, Germany

[2] ETH Zurich, Institute of Environmental Engineering, Zurich, Switzerland

[3] Capella Space, San Francisco, CA, USA

[4] Department of Geographical Sciences, University of Maryland, College Park, MD, USA

Abstract

The present study addresses the development, implementation, and validation of a forest height mapping scheme based on the combination of TanDEM-X interferometric coherence and GEDI waveform measurements. The very general case where only a single polarisation TanDEM-X interferogram, a set of spatially discrete GEDI waveform measurements, and no DTM are available is assumed. The use of GEDI waveforms to invert the TanDEM-X interferometric measurements is described together with a set of performance criteria implemented to ensure a certain performance quality. The emphasis is set on developing a methodology able to invert forest height at large scales. Combining 595 TanDEM-X scenes and about 15 million GEDI waveforms, a spatially continuous 25-m resolution forest height map covering the whole of Tasmania Island is achieved. The derived forest height map is validated against an airborne lidar-derived canopy height map available across the whole island.

4.1 INTRODUCTION

Accurate forest height measurements at subhectare scales are critical for characterizing the successional state of forests and/or their disturbance regime. At the same time, forest height can be related to forest biomass through allometric models, so used to initialize (or constrain) model estimates of above-ground biomass [1-4]. Despite their importance for forest inventory and modeling, forest height measurements on the ground remain difficult, so the generation of accurate high spatial resolution forest height maps over large areas remains a remote sensing challenge.

The introduction of polarimetric synthetic aperture radar (SAR) interferometry (Pol-InSAR) at the end of the 90s was a decisive step toward measuring forest height accurately at large scales [5-10]. Relying on the inherent sensitivity of the interferometric coherence to the vertical structure of volume scatterers, Pol-InSAR techniques have been established alongside lidar measurement

techniques for accurate estimation of forest height in the context of airborne and spaceborne applications [11-14]. From the proposed Pol-InSAR estimation algorithms, model-based ones have in general proved more robust and with a better performance. Such estimation algorithms model the vertical reflectivity in the forest by means of a two-layer model (accounting for vegetation and a ground scattering contribution) [5], [15], and forest height is then obtained by inverting the established model using interferometric measurements at different polarizations or vertical wavenumbers (e.g., spatial baselines) [10], [16].

Launched in 2010, TerraSAR-X add-on for Digital Elevation Measurement (TanDEM-X) introduced a new era in spaceborne radar remote sensing allowing single-pass interferometric measurements from space in a bistatic configuration [17-18]. The sensitivity of the interferometric TanDEM-X coherence to the vertical forest structure and especially to forest height initiated a large number of studies on forest height estimation from TanDEM-X SAR interferometry (InSAR) data across all possible forest types and conditions [15],[19-24]. However, since the conventional TanDEM-X global digital elevation model (DEM) observation mode is limited to a single polarization interferometric acquisition, only a highly simplified implementation of a single-layer model is possible, which often requires additional external information, such as an external digital terrain model (DTM), for its inversion [16], [25]. The achieved performance is, in general, remarkably good as long as the forest conditions allow sufficient penetration to ensure the “visibility” of the whole (vertical) forest extent, and the forest (and terrain) heterogeneity can be matched by the simplified single-layer parameterization of the vertical reflectivity. While the limited penetration is physical and must be accepted as long as the underlying topography is unknown, the limited ability of the inversion model to adapt to the local forest and terrain conditions can be addressed if a larger observation space is available [16].

The wide availability of airborne lidar data triggered several attempts to use lidar measurements to compensate for the underdetermination of the forest height inversion problem when addressed in terms of single-polarimetric TanDEM-X acquisitions. Besides using the lidar-derived DTM to directly enable the forest height inversion [22-23], [25-27], lidar data have also been used to constrain individual model parameters reducing, in this way, the dimensionality of the forest height inversion problem [28-33]. More recently, the direct use of lidar waveforms to define the full *X*-band vertical reflectivity profile has been proposed [34]. This allows, depending on the spatial density of the available waveform measurements, adaptation to local forest and terrain conditions and improved performance, especially in spatially heterogeneous forest conditions.

The launch in late 2018 and subsequent operation of NASA's Global Ecosystem Dynamics Investigation (GEDI) Mission, has been a critical development for large-area forest height estimation [35-36]. GEDI provides dense and well-distributed lidar waveform measurements across Earth's tropical and temperate forests and defines an ideal framework to explore the synergetic use of waveform lidar and interferometric *X*-band SAR measurements at local to global scales. Indeed, the synergies between TanDEM-X and waveform lidar data [34] have been confirmed by several studies using the TanDEM-X/GEDI framework [30-31], [37-38].

This study investigates the potential of this synergy for large-scale forest height mapping for the case where only a single polarisation TanDEM-X interferogram, a set of spatially discrete GEDI waveform measurements, and no DTM are available. The emphasis is set on developing a methodology able to invert forest height at large scales. Accordingly, decisions and tradeoffs are taken in favor of optimizing the data handling and processing flow and the overall inversion performance accounting for minimizing the effect of systematic errors. This is done accepting that the performance achieved may be locally inferior compared to more sophisticated, but also more complex, inversion approaches. Section 4.2 describes the selected forest site for the experiments, i.e., Tasmania in Australia, the experimental data (TanDEM-X and GEDI), and the reference height measurements available. Section 4.3 addresses the estimation of forest height based on the combination of TanDEM-X and GEDI data. In Section 4.4, the forest height estimation performance is addressed. A performance model is introduced and used to define the optimum height inversion range. Section 4.5 describes the actual processing and main results over the test site. Finally, in Section 4.6, the key findings are summarized and the conclusions are drawn.

4.2 TEST SITES AND DATASETS

4.2.1 TASMANIA

The Tasmanian island, located in the south of Australia, is about 68401 km^2 large with an extent of about 350 km in longitude and 330 km in latitude. The island is predominantly forested by temperate rain forests and has mountainous as well as flat regions. The forested area is very dynamic with fire events, logging activities and re-growth, which makes the test site atractive for evaluating different forest conditions.

4.2.2 LIDAR CANOPY HEIGHT MODEL

From 2008 to 2019 the Government of Tasmania performed several airborne laser scanning campaigns to map a large part of the island. One of the products of these campaigns was a Canopy Height Model (CHM) of 2 m spatial resolution [39]. The CHM data, mosaicked together to a single raster dataset are shown in Fig. 4.1 (a) together with a map that indicates the acquisition year for each area in Fig. 4.1 (b). Although the CHM model results from data acquired over a period of about 10 years, the availability of such a large and continuous forest height dataset allows the evaluation of the proposed methodology not only at a local but also at larger (country-wide) scales.

For the validation of the TanDEM-X heights associated with the upper canopy height, the H100 height [16] (corresponding to the mean height of the 100 tallest trees within a hectare or equivalent to the mean of the 6 tallest trees within a 25 m × 25 m window) was derived from the CHM as the mean of the tallest 3 CHM values (accounting for their 2 m spatial resolution) within

a 25 m × 25 m window. The obtained H100 values correspond well to GEDI's RH98 heights, as shown in Fig. 4.2.

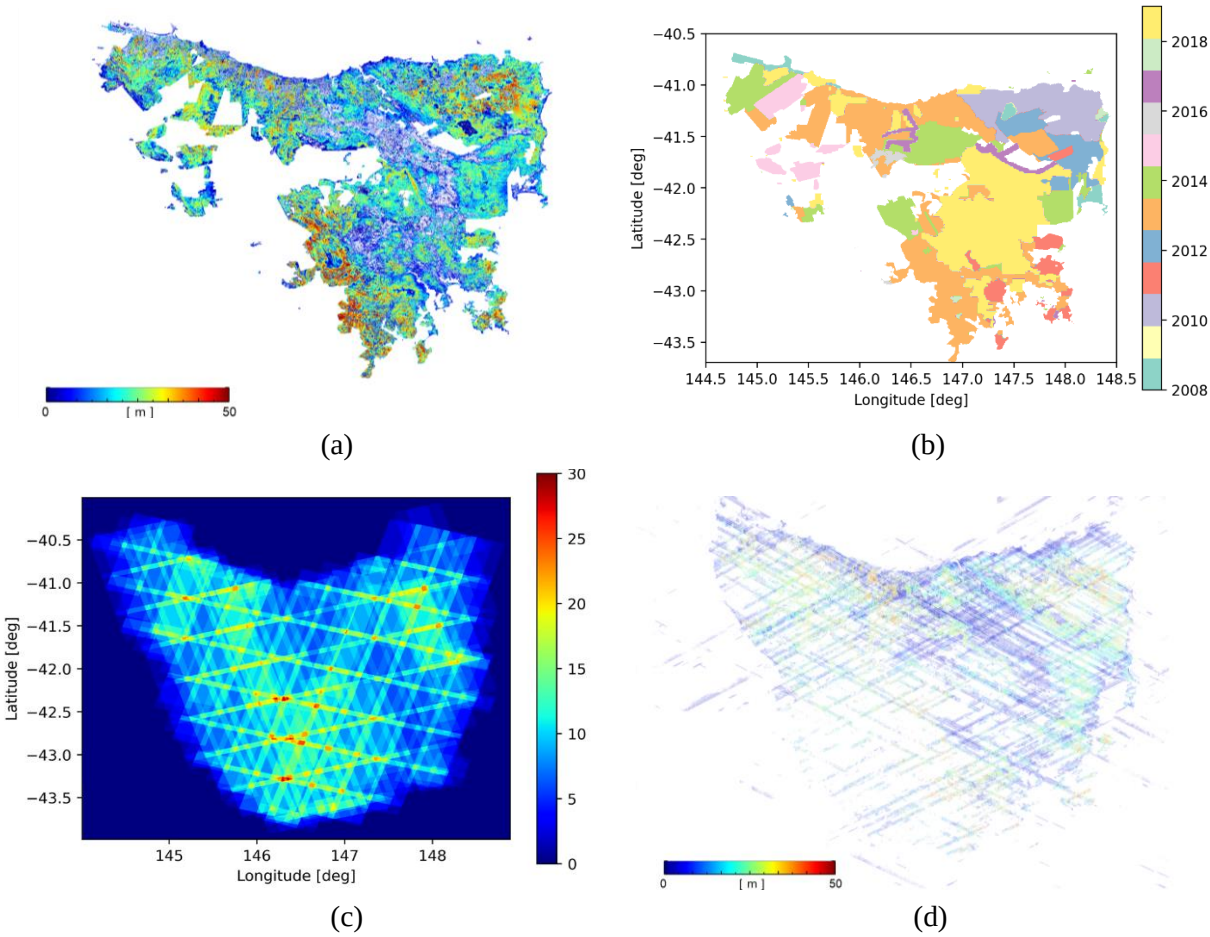


Fig. 4.1. Tasmania: (a) 25 m resolution CHM (H100) map from lidar in meters (m). (b) Date of each lidar acquisition. (c) Footprints of TanDEM-X acquisitions. The color represents the number of available images at each location. (d) GEDI RH98 map in meters over footprint positions of the available GEDI measurements. All maps are in geographic coordinates, with spacing (about 100 m × 100 m) in longitude and latitude direction.

4.2.3 Forest / Non-Forest Map

To avoid inconsistent heights over areas where the (inversion) model is not applicable, the TanDEM-X-derived forest/non-forest (FNF) map was used to mask water and urban areas [40]. For all remaining areas, forested and non-forested height estimates are computed and accounted for in the validation.

4.2.4 TANDEM-X DATASET

In this study, 595 conventional Coregistered Single look Slant range Complex (CoSSC) data products acquired and processed in an HH polarized strip map mode during the global DEM [18] and change DEM [41]. TanDEM-X mission phases between 2011 and 2019 have been used. The dataset includes ascending and descending acquisitions with a typical resolution of about 3 m in range and azimuth and with vertical wavenumbers ranging from 0.05 to 0.15 rad/m (with a median of about 0.1 rad/m). Their coverage map is shown in Fig. 4.1 (c).

4.2.5 GEDI DATASET

The GEDI instrument consists of three lasers (operating with a wavelength of 1064 nm) that are split into 4 beams and dithered along-track to produce a total of eight parallel acquisition tracks, spaced approximately 600 m apart on the Earth's surface in the cross-track direction. Along each acquisition transect, waveforms with approximately 25-m footprints are spaced every 60 m along the track. About 15 million GEDI footprints acquired in the 18 first months of the mission between April 2018 and October 2019 are used in this study. Their spatial distribution is shown in Fig. 4.1 (d), covering an area of less than 1% of the whole island.

From the available GEDI footprints, only those with a signal-to-noise ratio sufficient to penetrate through the forest with up to 95% canopy cover have been used [42]. The geolocation accuracy of the footprint data has been improved to ~10 m at 1-sigma in Version 2 [43]. The GEDI waveforms and the Level 2A RH98 values [36] have been used to initialize the height estimation from the TanDEM-X coherences.

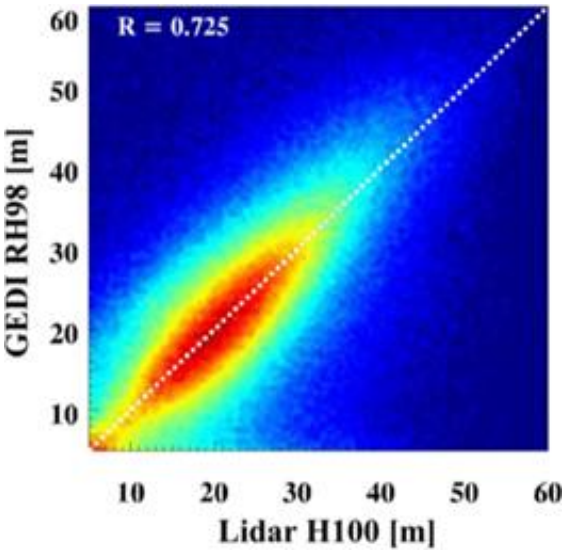


Fig. 4.2. Comparison (2-D histogram) between airborne lidar H100 measurement and GEDI RH98 measurement.

4.3 FOREST HEIGHT INVERSION

This Section introduces the forest height estimation methodology based on the combination of TanDEM-X and GEDI data.

4.3.1 TANDEM-X COHERENCE ESTIMATION

The (complex) interferometric coherence $\tilde{\gamma}_{\mathrm{Obs}}$ is obtained from the interferometric image pair s_1 and s_2 as

$$\tilde{\gamma}_{\mathrm{Obs}} = \frac{<s_1 s_2^*>}{\sqrt{<s_1 s_1^*><s_2 s_2^*>}} \tag{4.1}$$

where <·> denotes the expectation value. In the TanDEM-X global DEM and change DEM mission phases, s_1 and s_2 are acquired in a bistatic interferometric strip map mode where one of the two satellites transmit (in H polarisation) and both satellites receive the scattered signal quasi simultaneously (in H polarisation). In the absence of temporal decorrelation $\tilde{\gamma}_{\mathrm{Obs}}$ comprises two main decorrelation contributions

$$\tilde{\gamma}_{\mathrm{Obs}} = \tilde{\gamma}_{\mathrm{Sys}} \tilde{\gamma}_{\mathrm{Scat}}. \tag{4.2}$$

The first term, $\tilde{\gamma}_{\mathrm{Sys}}$, includes the decorrelation effects includes the decorrelation effects induced by the nonideal SAR system. The most prominent system decorrelation contribution is the (additive) noise decorrelation γ_{SNR}. Modelling the received signal to be composed by the scattering amplitude a and the noise amplitude n, i.e. $s = a + n$, γ_{SNR} can be written as [25]

$$\gamma_{\mathrm{SNR}} = \frac{1}{1+\mathrm{SNR}^{-1}} = \frac{A}{A+N} = \frac{A}{P} \tag{4.3}$$

where $\mathrm{SNR} = A/N$ is the signal-to-noise ratio, with $P = A + N$ the received power, $A = |a|^2$ the scattered power and $N = |n|^2$ the noise power.

A second contribution to the system decorrelation $\tilde{\gamma}_{\mathrm{Sys}}$ is γ_{Quan}, which is caused by the lossy raw data compression process [44-45]. In the case of TanDEM-X, the received backscattered signal is first digitized by an 8-b analog-to-digital converter and then further compressed by a block adaptive quantizer with a compression rate of 8:3 (or 8:4). This is associated with a coherence loss of about 3.5% (or 1%) [45-46].

The second term, $\tilde{\gamma}_{\mathrm{Scat}}$, reflects the phase stability of the scatterer under the different incidence angles induced by the interferometric baseline. After range and azimuth spectral filtering [5] $\tilde{\gamma}_{\mathrm{Scat}}$ reduces to the volume decorrelation contribution $\tilde{\gamma}_{\mathrm{Vol}}$ [5-6]

$$\tilde{\gamma}_{\mathrm{Vol}}(\kappa_z) = \frac{\int_{z_0}^{z_0+h_V} F(z)\exp(i\kappa_z z)dz}{\int_{z_0}^{z_0+h_V} F(z)\,dz} \tag{4.4}$$

with $F(z)$ is the vertical reflectivity function (also referred to as the vertical reflectivity profile) expressing the vertical distribution of scatterers seen by the interferometer. The vertical direction is expressed by z and z_0 indicates the position of the underlying ground. Accordingly, $F(z)$ depends on the frequency and polarisation of the interferometer as well as on the interferometric acquisition geometry. The upper bound of $F(z)$ is given by $z_0 + h_V$ which in the case of a forest scatterer corresponds to the (top) forest height. The lower bound of $F(z)$ is given by z_0. For the bistatic TanDEM-X acquisition the vertical (interferometric) wavenumber κ_z is defined as [47]

$$\kappa_z = \frac{2\pi}{\lambda}\frac{\Delta\theta}{\sin(\theta)} = \frac{2\pi}{\lambda}\frac{\Delta\theta}{\sin(\theta_0-\alpha)} \tag{4.5}$$

where θ is the local incidence angle given as the difference between the (nominal) incidence angle θ_0 and the terrain slope angle in the range direction α, λ the wavelength, and $\Delta\theta$ the change of the incidence angle induced by the spatial baseline (i.e., the spatial separation of the two satellites). The vertical wavenumber is often expressed by the so-called height of ambiguity HOA = $2\pi/\kappa_z$, i.e., the height that corresponds to the interferometric phase of 2π. The local incidence angle is estimated using the terrain slope angle in the range direction α derived from theTanDEM-X DEM.

Before using the interferometric coherences for inversion, γ_{SNR} and γ_{Quan} need to be compensated. The compensation of γ_{SNR} is performed using the Noise Equivalent Sigma Zero (NESZ) patterns along range provided for each of the two SLC (Single Look Complex) images (i.e., TSX and TDX) in the CoSSC data product [16]. Accordingly, the SNR for each of the two SLCs is calculated from the NESZ pattern, and the individual backscattering coefficient sigma nought σ_0

$$\mathrm{SNR}^{\mathrm{TSX}} = \frac{\sigma_0^{\mathrm{TSX}} - \mathrm{NESZ}^{\mathrm{TSX}}}{\mathrm{NESZ}^{\mathrm{TSX}}} \text{ and } \mathrm{SNR}^{\mathrm{TDX}} = \frac{\sigma_0^{\mathrm{TDX}} - \mathrm{NESZ}^{\mathrm{TDX}}}{\mathrm{NESZ}^{\mathrm{TDX}}} \tag{4.6}$$

and the SNR induced decorrelation is then obtained as [16]

$$\gamma_{SNR} = \frac{1}{\sqrt{\left(1+\frac{1}{\mathrm{SNR}^{\mathrm{TSX}}}\right)\left(1+\frac{1}{\mathrm{SNR}^{\mathrm{TDX}}}\right)}} \tag{4.7}$$

Using the γ_{SNR} values obtained from (4.7) and assuming the γ_{Quan} accounts constantly 3.5% of the decorrelation ($\gamma_{Quan} = 0.965$), the volume decorrelation contribution $\tilde{\gamma}_{Vol}$ is obtained as

$$\tilde{\gamma}_{Vol}(\kappa_z) = \tilde{\gamma}_{Obs}(\kappa_z)/(\gamma_{SNR} \cdot \gamma_{Quan}) \tag{4.8}$$

After calibration, the interferometric coherences are ready to be used for inversion. Any volume decorrelation values greater than 1 (due to the inherent standard deviation of interferometric coherence) are set to 1.

4.3.2 THE MEAN VERTICAL REFLECTIVITY PROFILE

Following the approach proposed in [34], the GEDI waveforms $P_i(z)$ are used to approximate the vertical reflectivity profile $F(z)$ in (4.4). Because GEDI's waveform measurements are spatially discrete a mean reflectivity profile, a representative for a larger area, has to be used. The generation of such a mean reflectivity profile from the available GEDI waveforms follows the approach proposed in [34]. First, the available waveforms $P_i(z)$ normalized to unit height and resampled to a common number of height samples are used as columns of the so-called profile matrix [P]. Accordingly, the number of rows of $[P]$ is given by the sampling in the height dimension and the number of columns is the number of available GEDI waveforms n.

From the profile matrix [P], a covariance matrix [R] is formed

$$[R] = [P]\,[P]^T \tag{4.9}$$

(where $[\cdot]^T$ indicates the transpose operator) and diagonalized

$$[R] = [U]\,[\Lambda]\,[U]^T \tag{4.10}$$

where $[\Lambda]$ contains the positive eigenvalues a_i (sorted in descending order) and $[U]$ the eigenvectors $\overline{P}_i(z)$ of [R] where i={1, … M} and M equals the number of the GEDI waveforms used to form the profile matrix [P]). $\overline{P}_i(z)$ are then used to compose the mean reflectivity profile

$$F_m(z) = \frac{1}{N}\sum_{i=1}^{N} a_i \overline{P}_i(z) \tag{4.11}$$

where $N < M$ represents the number of eigenvectors used to compose $F_m(z)$. In the following N=1 is used so that $F_m(z) = \overline{P}_1(z)$. Considering only the dominant profile component and omitting higher order eigenvectors corresponds to an attempt to represent all forest states within the scene by a common low-frequency structural component. This can be justified by the fact that higher order structural components play often a secondary role as the structure dependency of the volume coherence is weaker compared to its dependency on height. Finally, the use of $F_m(z)$ in (4.4) allows the estimation of forest height h_V for each $\tilde{\gamma}_{Vol}(\kappa_z)$ sample.

However, there is a number of points that have to be considered. The mean reflectivity profile is estimated for each TanDEM-X scene individually using all available GEDI waveform within the scene. Fig. 4.3 (a) shows an example of a mean reflectivity profile $F_m(z)$ representative for a whole TanDEM-X scene obtained by combining the more than 5000 GEDI waveforms located within the scene. Due to the height normalisation and the subsequent resampling of the GEDI waveforms when forming profile matrix [P], the mean reflectivity profile becomes a long “tail” i.e. an asymptotic behavior towards higher heights, especially when the scene is dominated by short(er) forest stands. This, if not accounted, biases the obtained forest height estimates later. To avoid this, the mean reflectivity profile is limited by an intensity threshold: the profile is cut at the height where the (profile) intensity decreases below a given (arbitrary) threshold (3 dB in the

following) with respect to the highest located reflectivity maximum. After this, the profile is re-normalized between 0 and 1 (i.e. from 0 to 1) to obtain the final $F_m(z)$ (see Fig. 4.3 (b)).

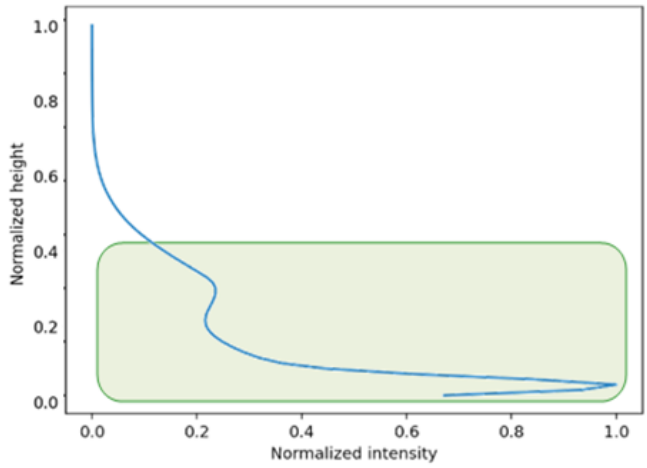
Furthermore, the use of a single $F_m(z)$ for all (forest) samples within the scene implies an intrinsic error depending on how well $F_m(z)$ represents the underlying reflectivity profile $F(z)$ and its spatial variability within the scene. One way to address this problem is to reduce the area in which $F_m(z)$ is estimated. In this sense, one can segment a TanDEM-X scene in multiple areas estimating for each a mean reflectivity profile. However, there is a trade-off between number of GEDI waveforms available and the size of the area to be represent by the mean reflectivity profile: with decreasing area size, the number of GEDI waveforms within the area decreases making the estimation of the mean profile less robust. On the other hand, the fact that the mean reflectivity profile is estimated and used on a TanDEM-X scene basis can lead to inverting the same sample in different scenes with different mean reflectivity profiles. This can be the case for partially overlapping TanDEM-X scenes.

4.3.3 Modification of Profile with Height

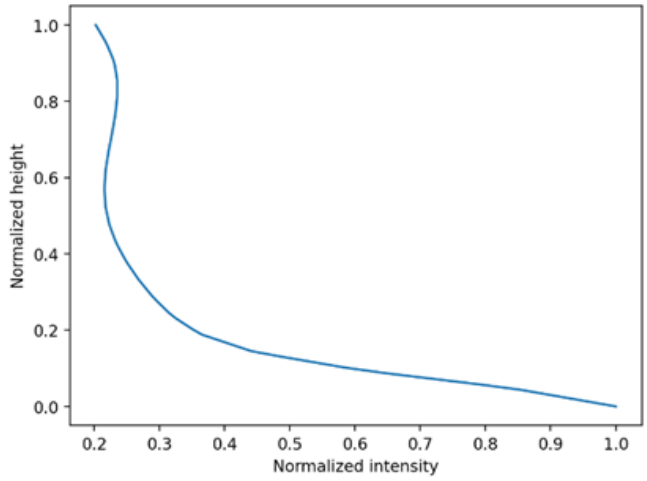
The high attenuation at lidar and X-band radar frequencies implies for both GEDI and TanDEM-X measurements a high sensitivity to the "visible" geometric architecture of the forest volume. This, together with the comparable spatial resolutions of the two measurements, makes the GEDI waveforms $P(z)$ and the X-band vertical reflectivity profiles $F(z)$ for many forest types similar. And it is this similarity that justifies the use of P(z) as an approximation for the vertical reflection profile F(z). However, it is also clear that GEDI waveforms and the X-band vertical reflectivity profiles are not equal. The differences are induced primarily by the different acquisition geometries the two configurations operate.

The nadir-looking GEDI geometry causes a stronger ground contribution - especially in open canopy forest conditions – than in the TanDEM-X reflectivity, where the side-looking geometry results in a larger path through the canopy and thus a stronger attenuation of the ground contribution. This leads to an overrepresentation of the ground contribution in the mean reflectivity profile that can bias the obtained height estimates.

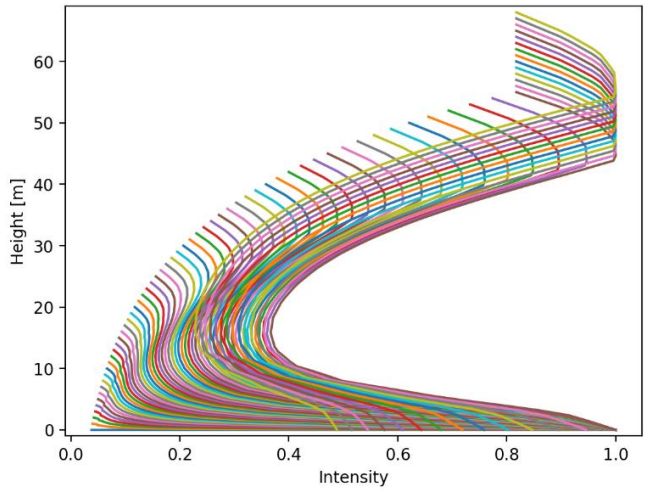
The overrepresentation of the ground contribution can be clearly seen in the Fig. 4.3 (b) where the mean reflectivity profile has a much stronger ground than canopy contribution. While in open canopy forest conditions the ground contribution is overrepresented, in denser and/or closed canopy forest conditions the ground contribution in the GEDI waveforms is weaker than the one expected at X-band (vertical) reflectivity. In this case the ground contribution in the mean reflectivity profile is underrepresented. This makes the direct use of the mean profile $F_m(z)$ as derived from the GEDI waveforms sub-optimum. Equally sub-optimum, in terms of performance, is also the use of the same profile for the whole height range, i.e. the assumption of the same distribution of scatterers for all forest heights (from 5 m to 70 m).



(a)



(b)



(c)

Fig. 4.3 (a) Example of a mean reflectivity profile obtained by using all GEDI waveforms within a single TanDEM-X scene. The mean reflectivity profile is normalized by its height and the intensity is normalized by its maximum power. (b) Final mean profile after long tail removal in (a). (c) Set of profiles was calculated from the mean reflectivity profile in (b) after the height-dependent attenuation correction of (4.12).

In order to account for both effects, the overrepresentation of the ground contribution and the variation of the distribution of scatterers with forest height, an empirical height dependent attenuation correction on $F_m(z)$ is performed

$$F_{mh}(z) = F_m(z) \cdot e^{-\varepsilon(z)/\cos(\theta)} \tag{4.12}$$

where $\varepsilon(z)$ is a height dependent attenuation factor

$$\varepsilon(z) = \varepsilon_0(h_{Ref} - z). \tag{4.13}$$

Accordingly, contributions below a certain reference height h_{Ref} are attenuated (stronger with decreasing height) while contributions above h_{Ref} are amplified. The reference height h_{Ref} is set by the mean of all GEDI RH98 heights within the scene. The constant attenuation factor ε_0 expresses the difference in the effective attenuations of the side-looking X-band and the nadir-looking lidar propagation and is universally fixed at 0.1 dB/m for all scenes. This is, of course, a strong claim that cannot be confirmed by any theoretical or experimental means. However, the main function here is to attenuate the overrepresented ground scattering contribution in the mean profile $F_m(z)$ derived from the GEDI waveforms, especially for smaller stands, rather than to provide an accurate model. In this sense, any similarly chosen attenuation value would have a very similar effect.

One should note here that the difference between the two profiles, the TanDEM-X and the GEDI one, is more a question of forest density than of forest height. However, the absence of any large-scale structure or density information makes it impossible to account for this. On the other hand, the significant errors introduced by the overrepresented ground contribution make its generic attenuation (as performed by (4.12) and (4.13)) in terms of overall performance more effective even if for a number of cases this general correction leads to worse estimates.

Once the attenuation factor ε_0 and the reference height h_{Ref} are defined, (4.12) provides for each mean reflectivity profile $F_m(z)$ a set of profiles $F_{mh}(z)$, one for each height. Fig. 4.3 (c) shows an example of a set of 70 profiles each associated to a forest height from 1 m up to 70 m).

The use of $F_{mh}(z)$ in (4.4) allows now the estimation of forest height h_V for each $\tilde{\gamma}_{Vol}(\kappa_z)$ sample. The inversion can be implemented by means of a simple look-up table (LUT) that maps the absolute value of the volume decorrelation to a forest height accounting for the actual (local) vertical wavenumber κ_z. Fig. 4.4 (a) shows an example of such a volume decorrelation – forest height LUT for different vertical wavenumbers κ_z.

4.4 HEIGHT INVERSION PERFORMANCE ANALYSIS

In order to evaluate the quality of the obtained forest height estimates on a sample basis a set of performance criteria are introduced that allow to mask sub-optimum height estimates and to compensate systematic (height) offsets.

4.4.1 Lower Coherence Level

Lower coherence values are associated to a higher (amplitude and phase) variance that propagates into the height estimates. To avoid such estimates, a coherence threshold has been introduced to mask out the heights obtained from samples with a lower (absolute) coherence. A threshold value of 0.3 has proven to be appropriate and is used throughout the following.

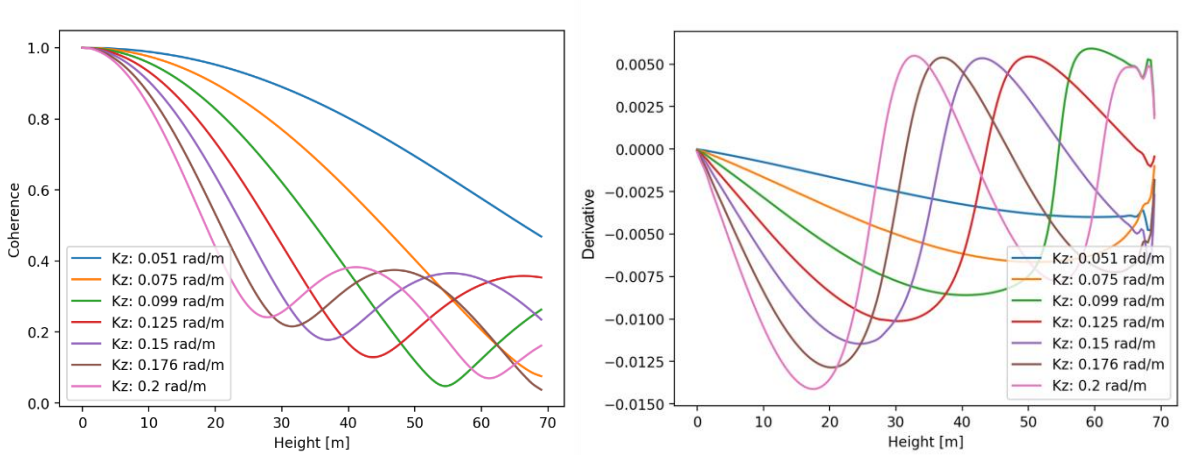


Fig. 4.4. (a) Look-up tables for the same set of profiles in Fig. 4.3 (c) for different vertical wavenumbers. (b) Derivatives of the look-up tables.

4.4.2 Vertical Wavenumber Performance

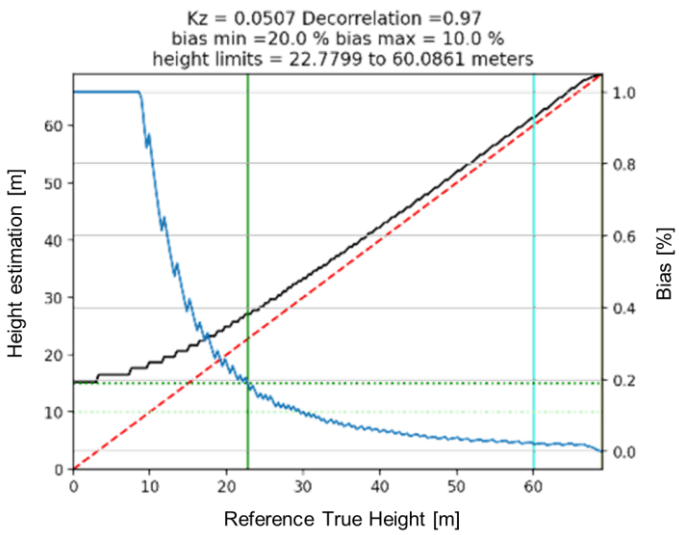
Knowledge of the reflectivity profile allows a simplified forest height inversion performance evaluation accounting for two key performance effects: 1) inversion biases induced by a residual nonvolumetric decorrelation contribution affecting primarily the lower height ranges; and 2) the reduced sensitivity (or even saturation) to higher heights induced by too large or too small vertical wavenumbers.

For a given forest height h_V and reflectivity profile $F_{mh}(z)$ the expected interferometric volume coherence $\tilde{\gamma}_{Vol}(\kappa_z)$ can be estimated using (4.4). The effect of residual (non-volumetric) decorrelation contributions can be addressed by applying a residual decorrelation contribution γ_{Res} to the volume decorrelation $\tilde{\gamma}_{Vol}(\kappa_z)$ and by inverting the distorted coherence $\gamma_{Res} \cdot \tilde{\gamma}_{Vol}(\kappa_z)$ (instead of $\tilde{\gamma}_{Vol}(\kappa_z)$). The height estimates obtained for the whole height range are then compared to the original ones to estimate the height-dependent bias induced by γ_{Res}. This estimation bias can be visualized by plotting the estimated heights against the reference heights as shown in Fig. 4.5 for three different vertical wavenumbers. The dotted red line indicates the 1:1 line, the black line estimated heights and the blue line the estimation bias as a percentage. The three plots, each for the same set of reflectivity profiles $F_{mh}(z)$ but different vertical wavenumbers ((a) κ_z =0.05 rad/m, (b) κ_z =0.1 rad/m and (c) κ_z =0.15 rad/m) clearly indicate the over estimation in the lower height range and how it decreases with increasing vertical wavenumber and/or increasing height.

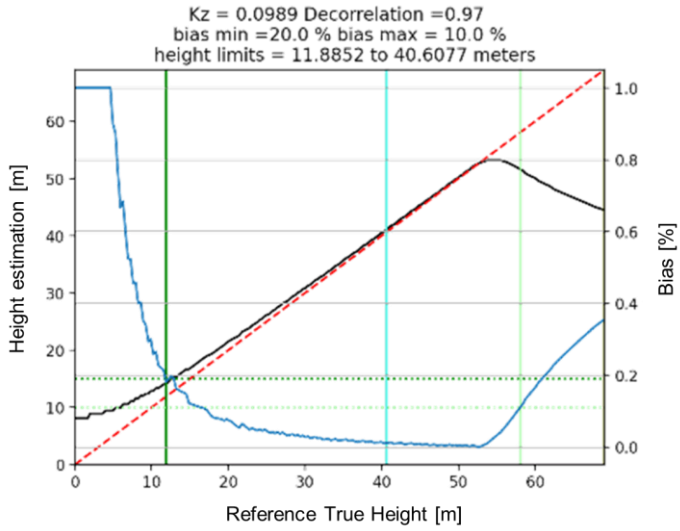
The obtained estimation bias can be used to define the optimum forest height performance range by means of an upper and lower limit for each individual height estimate using the reflectivity profiles $F_{mh}(z)$ and the associated vertical wavenumber κ_z . The upper and lower limits of the optimum forest height performance range are defined by means of a performance threshold that can be the same or different for the upper and lower case. For example, in Fig. 4.5, a 20% error threshold for the lower heights is indicated by the vertical dark green line, and a 10% error threshold for the higher heights is indicated by the vertical light green line. According to the plots for a vertical wavenumber $\kappa_z \approx 0.1$ rad/m (middle), this leads to a lower height limit of about 12m and an upper height limit of about 58m.

The reduced sensitivity at higher heights induced by too large vertical wavenumbers is accounted for means of the height derivative $\partial|\tilde{\gamma}_{Vol}(\kappa_z, h_V, F_{mh}(z))| \,/\, \partial h_V$ for a given vertical wavenumber κ_z. The height derivatives for the profile assumed in Fig. 4.3 (c) are plotted on Fig. 4.4 (b) clearly demonstrating the loss of sensitivity with increasing height and/or vertical wavenumber. The associated loss of performance can be accounted by introducing a threshold on the derivative. A sensitive (yet conservative) way to do this is to accept height estimates only up to the height that corresponds to the minimum of the $\partial|\tilde{\gamma}_{Vol}(\kappa_z, h_V, F_{mh}(z))| \,/\, \partial h_V$ derivative. This height limit is indicated by the light blue vertical line in the performance plots of Fig. 4.5 (b). For a vertical wavenumber of $\kappa_z \approx 0.1$ rad/m, this leads to an upper height limit of about 40 m. This is the height above that the inherent sensitivity of the observation configuration (expressed by the vertical wavenumber) degrades. Estimates above are rejected.

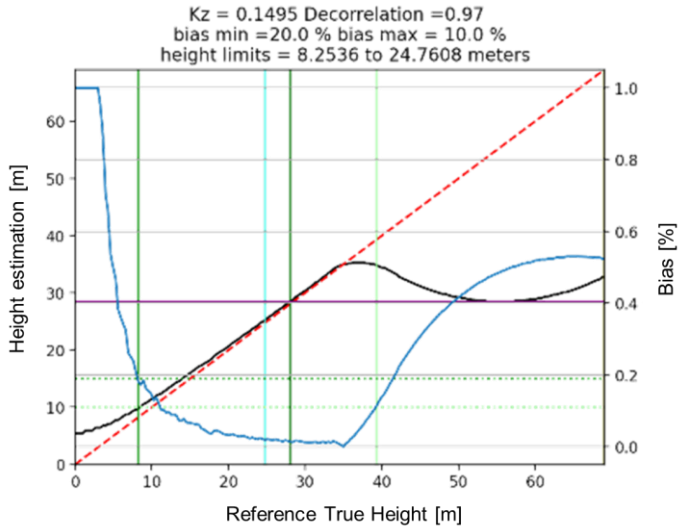
For defining the upper height limit of the validity range, the most restrictive height limit of the two performance criteria is selected. Of course, there is always a trade-off between applying less restrictive thresholds to allow more valid pixels at the price of potentially higher height errors. Accordingly, for this case only height estimates between 15 m and 40 m are accepted.



(a)



(b)



(c)

Fig. 4.5. Performance plots for the same set of profiles $F_{mh}(z)$ at different vertical wavenumbers ((a) κ_z=0.05 rad/m, (b) κ_z=0.1 rad/m, (c) κ_z=0.15 rad/m). The dotted red line indicates the 1-1 line, the black line estimated heights using a residual decorrelation of 0.97 and the blue line the estimation bias in percent.

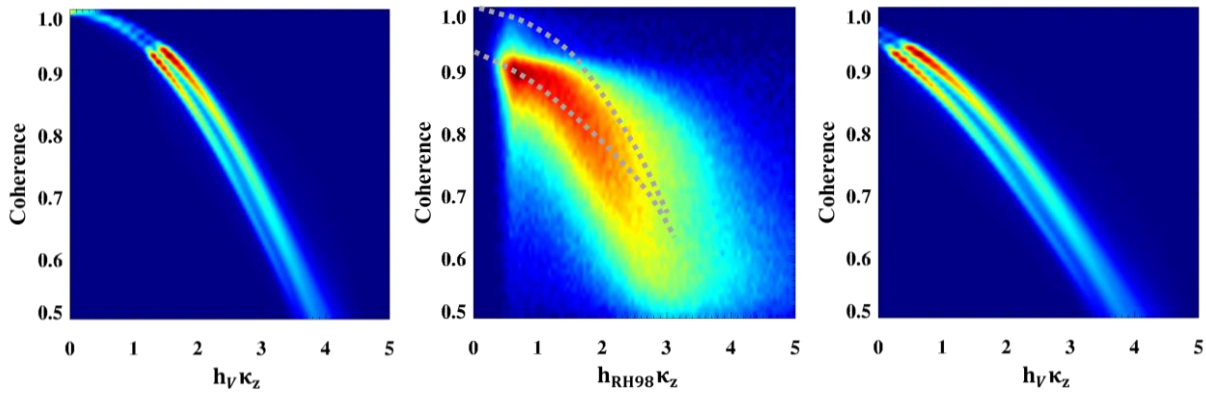
4.4.3 Global Bias Correction

The proposed performance model accounts for biases induced by residual non-volumetric decorrelation contributions and/or sub-optimal vertical wavenumbers assuming the validity of the vertical reflectivity profile(s) $F_{mh}(z)$. The bias (and the inaccuracy) induced by the profile mismatch remain unaccounted. These can be significant where the assumed vertical reflectivity profile $F_{mh}(z)$ is very different from the (actual) underlying reflectivity.

In order to compensate for the height bias arising from a systematic profile mismatch across the whole data set the GEDI RH98 heights are employed. Fig. 4.6 (a) shows on the top the interferometric volume coherence $|\tilde{\gamma}_{Vol}(\kappa_z)|$ plotted vs. the product of the estimated forest height $h_V$with the local vertical wavenumber κ_z, i.e. $h_V\kappa_z$ at the locations of the GEDI footprints. In Fig 4.6 (b), the same plot is shown, but this time the interferometric volume coherence $|\tilde{\gamma}_{Vol}(\kappa_z)|$ is plotted vs. the product of the GEDI RH98 height h_{RH98} with the local vertical wavenumber κ_z , i.e. $h_{RH98}\kappa_z$. The difference in the two plots (Fig. 4.6 (a) & (b)) reflects a systematic bias – independent of the local vertical wavenumber – primarily due to the profile mismatch and residual non-volumetric decorrelation contributions. This bias can be compensated by means of ordinary least square bisector (OLS) fitting

$$h_{RH98}\kappa_z = a_1(h_V\kappa_z) + a_0 \tag{4.14}$$

where a_0 and a_1 are the fitting coefficients. Fig. 4.6 (c) shows the interferometric volume coherence $|\tilde{\gamma}_{Vol}(\kappa_z)|$ plotted vs. the product of the corrected (by means of (4.14)) forest height estimate h_V after with the local vertical wavenumber κ_z, i.e. $h_V\kappa_z$.



(a) (b) (c)

Fig. 4.6. comparison (2D histogram) between TDX interferometric coherence magnitude and (a) $h_V\kappa_z$ before correction (b) $h_{RH98}\kappa_z$ (c) $h_V\kappa_z$ after correction. The gray dashed line in (b) indicates TDX height inversion model before and after linear regression based on $h_{RH98}\kappa_z$.

4.5 DATA PROCESSING AND RESULTS

The proposed processing flow and their main outputs, as described in Sections 4.3 and 4.4, are shown in Fig. 4.7. For each of the 595 TanDEM-X scenes (CoSSC data products) first the volume decorrelation contribution $\tilde{\gamma}_{Vol}$ and the terrain-corrected vertical wavenumber κ_z are estimated, as described in Section III.A. For each scene, the available GEDI waveforms are used to calculate

the mean profile following the procedure outlined in Section 4.3.B. The attenuation correction, discussed in Section 4.3.C, is applied to the mean profile to derive a set of reflectivity profiles that are used to estimate forest height from each volume decorrelation sample. For each obtained height estimate, the estimation bias is calculated using the set of profiles and the terrain-corrected vertical wavenumber as described in Section 4.4.B. A 20% height bias threshold (for the lower heights) and a (upper) height threshold associated to the minimum of the $\partial|\tilde{\gamma}_{Vol}(\kappa_z, h_V, F_{mh}(z))| / \partial h_V$ derivative is used to derive the validity map to mask out sub-optimum performance samples. Similarly, samples with a (absolute) coherence lower than 0.3 are discarded (Section 4.4.A). In total, about 2% of forest areas are excluded. After processing and inverting each of the TanDEM-X scenes, the forest height estimates along with an estimate of the height bias and the validity masks for the entire island are available and are georeferenced.

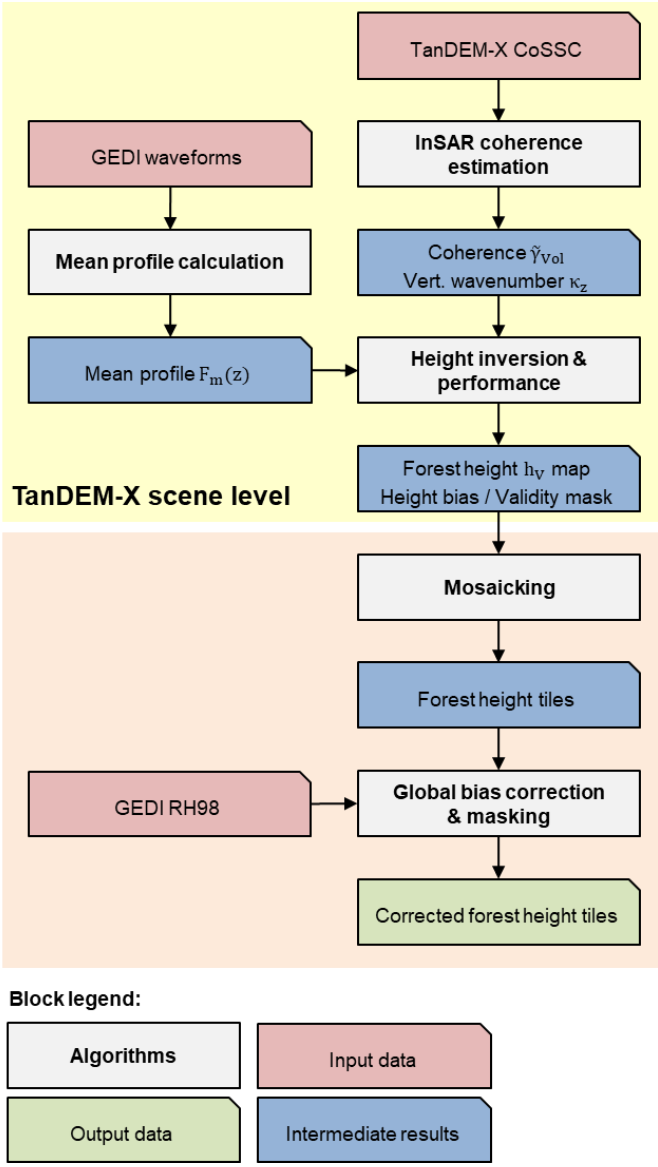
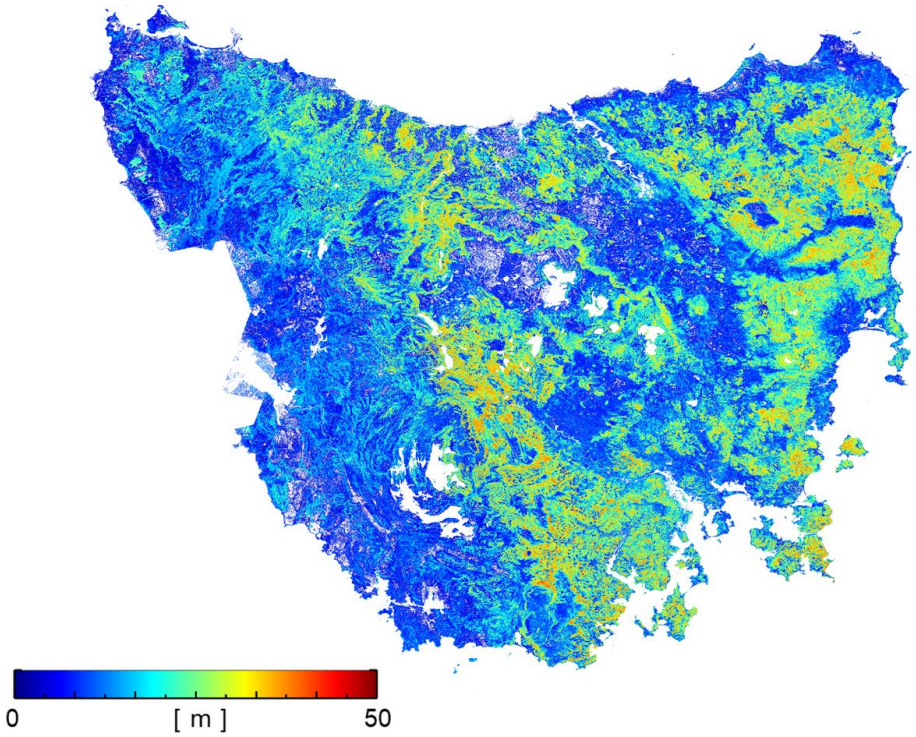


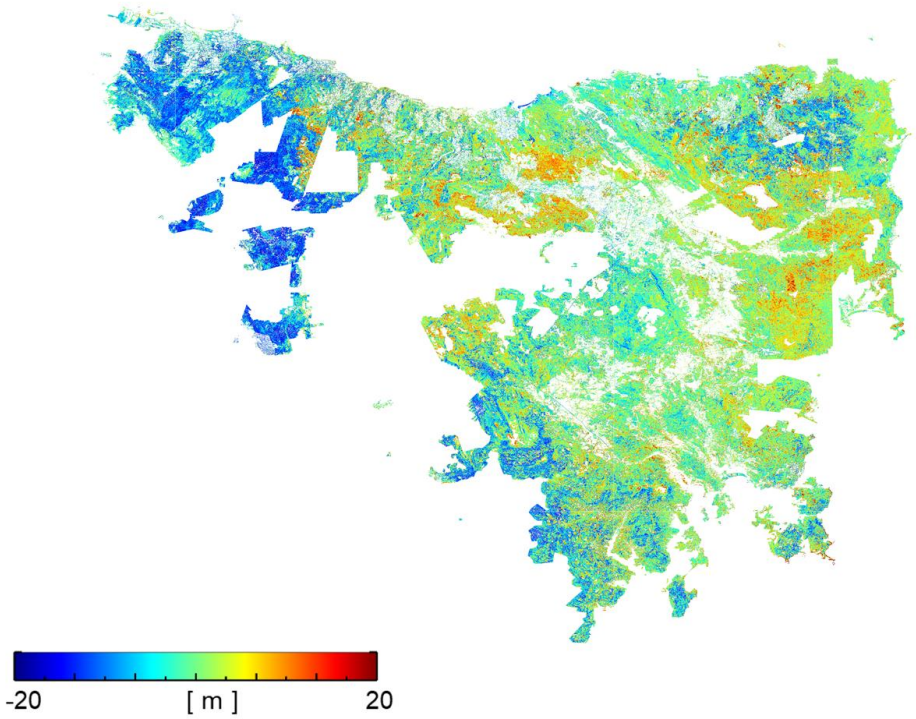
Fig. 4.7. Flow chart of the implemented forest height scheme.

After processing and inverting each of the 595 TanDEM-X scenes, the forest height estimates along with an estimate of the height bias error and the validity masks for the entire island are available and are georeferenced. The next step is to mosaic the individual maps into 1 degree × 1 degree (latitude, longitude) non-overlapping tiles. A single location (within a tile) may be covered by multiple TanDEM-X scenes and it may be associated with several forest height estimates. In such a case, and in order to obtain a single height estimate, either the available estimates can be combined (for example on the basis of their individual estimation biases or by forming their median value) or a single one is selected in order to fulfill certain vertical wavenumber, acquisition timing or performance selection criteria. Each of these approaches has its own pros and cons and can lead under certain circumstances to outliers. In the following the approach of using the "mean κ_z" (select the closest κ_z to mean κ_z value of all candidates) height value of the valid height estimates at each location is followed. The next step is the global bias correction that is performed once as described (4.14) in Section 4.4. Finally, water and settlement bodies are masked by using the forest / non- forest map.

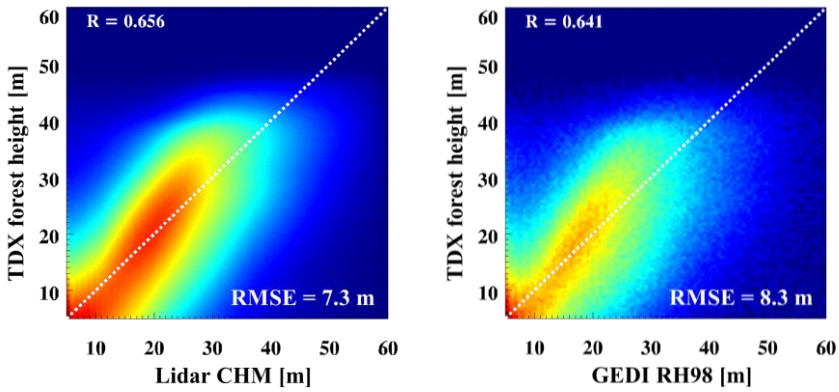
Fig. 4.8 (a) shows the final forest height map at 25 m resolution for the complete island of Tasmania. The heights are validated against the reference CHM. Fig. 4.8 (b) shows the difference between the obtained forest heights and the reference CHM across the whole island. The majority of the samples have an error of ± 5 meters with mean value of -0.08 meters. While over large parts of the island the height difference is very homogenous, locally spatial patterns of over- and under-estimation are clearly visible. The reason for these patterns is manifold. Many of these areas can be attributed to the extensive wildfires occurring in the years between the lidar and the TanDEM-X acquisition, while the underestimation on the western part may be due to the dense forest conditions there. The overestimation of the eastern part is partially due to the forest change and partially due to the mismatch of the derived reflectivity profiles and the real underlying reflectivity. Finally, Fig. 4.8 (c) shows the validation of the estimated forest heights against the CHM accounting only the valid height estimates with an RMSE of 7.3 m with a Pearson coefficient of 0.66.



(a)



(b)



(c)

Fig. 4.8. Tasmania: (a) Forest height maps obtained from TDX interferometric coherences and GEDI waveforms at 25 m resolution. (b) Forest height difference map (TDX – lidar CHM) (c) comparison between TDX forest height and lidar CHM / GEDI RH98.

4.6 PERFORMANCE DISCUSSION AND CONCLUSION

The proposed inversion scheme overcomes the limitations in large scale forest height estimation caused by the limited dimensionality of TanDEM-X observations with single polarisation by using the waveform and height measurements from GEDI.

Besides the compensation of all non-volumetric decorrelation contributions (by means of (4.8)) and the terrain correction of the vertical wavenumber (by means of (4.5)), the forest height inversion performance critically depends on the ability of the derived set of reflectivity profiles to match the real underlying reflectivity. Even if the effect of the reflectivity profile on the interferometric coherence is, when compared to the effect of the height itself, only secondary, it remains significant enough to have a decisive impact on the achieved performance. A mismatch between the assumed and the underlying reflectivity introduces a (positive or negative) bias on the estimated forest height. In contrast, a well matching reflectivity profile is able to correctly interpret (4.4) for a wide range of vertical wavenumbers and provide the "same" height estimates for different vertical wavenumbers – of course, under the performance constraints imposed by the actual vertical wavenumber (as discussed in Section 4.4). In this sense, in order to test how well the derived reflectivity profiles, match the real underlying reflectivity, forest height estimation performance and its consistency for different vertical wavenumbers are compared. Fig. 4.9 provides a representative overview of such a performance comparison on the basis of a single TanDEM-X scene. In Fig. 4.9 (a) the reference CHM is shown for an area of 50 × 50 km located the northern part of the island while in Fig. 4.9 (b) the GEDI RH 98 heights are shown. The forest height map (at 25 m resolution) obtained from a TanDEM-X scene acquired on 19th Jun 2011 with a (mean) vertical wavenumber of κ_z=0.125 (HoA=50m) along a descending orbit is in Fig. 4.9 (c). The validation plot against the CHM is shown in Fig. 4.10 (a) with a performance described by an RMSE of 6.6 m and a Pearson coefficient of 0.63. A second forest height map from a TanDEM-X scene acquired also along a descending orbit on 23rd Dec 2011 with an approx. 40%

smaller (mean) vertical wavenumber of κ_z=0.0757 (HoA=83m) is shown in Fig. 4.9 (d). The corresponding validation plot against the CHM is shown in Fig. 4.10 (b) with a very similar performance described by an RMSE of 7.2 m and a Pearson coefficient of 0.65. The correlation plot of the two derived height maps is shown in Fig. 4.11 (a) confirming their relatively good agreement.

However, even if the derived set of reflectivity profiles is up to a degree able to describe the mean underlying reflectivity, the single set of profiles used for the entire extent of a TanDEM-X scene is not able to capture the spatial variability of the underlying reflectivity within the scene. For an (spatially) uncorrelated mismatch between the assumed and the underlying reflectivity, the effect of the local height biases will appear across the whole scene in form of an increased height variance. However, when the mismatch between the assumed and the underlying reflectivity becomes spatially correlated, as for example when the underlying reflectivity changes locally because of the (local) incidence, the height estimates will be (locally) biased. This can be the case in the presence of (positive or negative) terrain slopes or when scenes acquired along ascending and descending orbits are combined. In order to obtain an impression about the magnitude of this effect forest height obtained from TanDEM-X scenes acquired along ascending and descending orbits are compared.

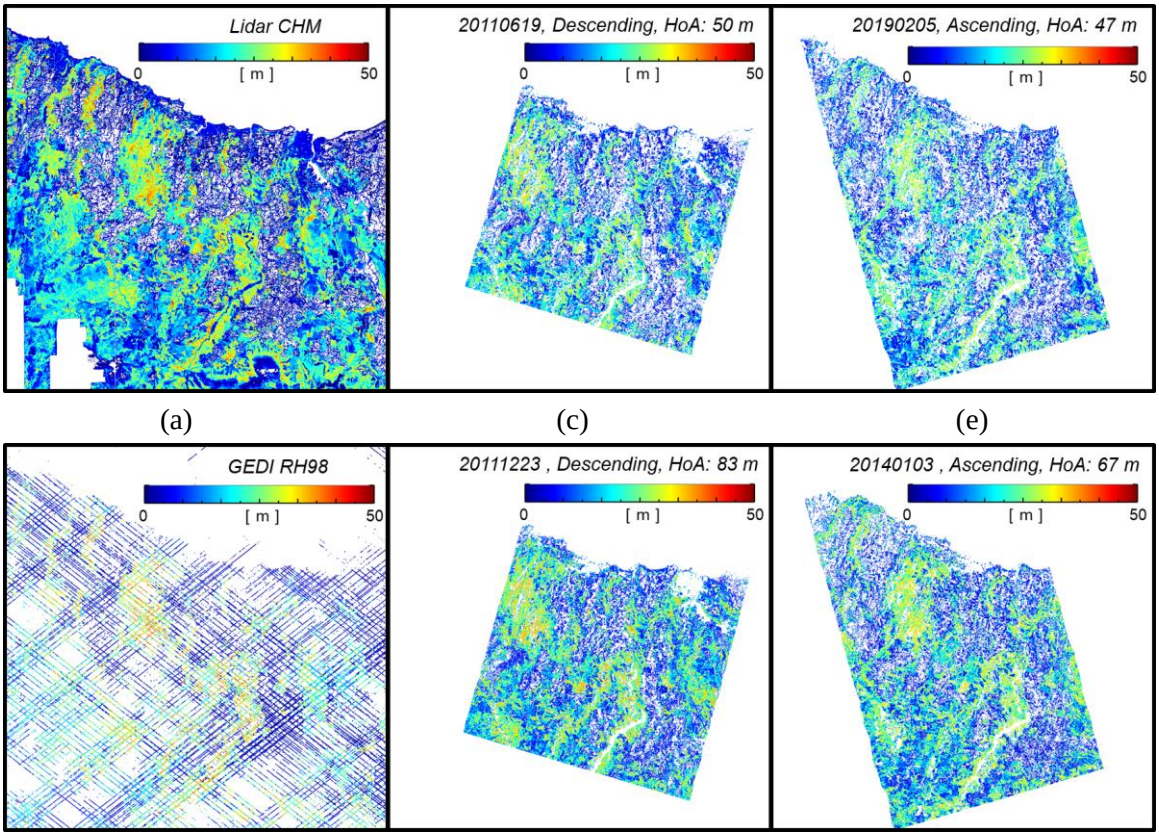
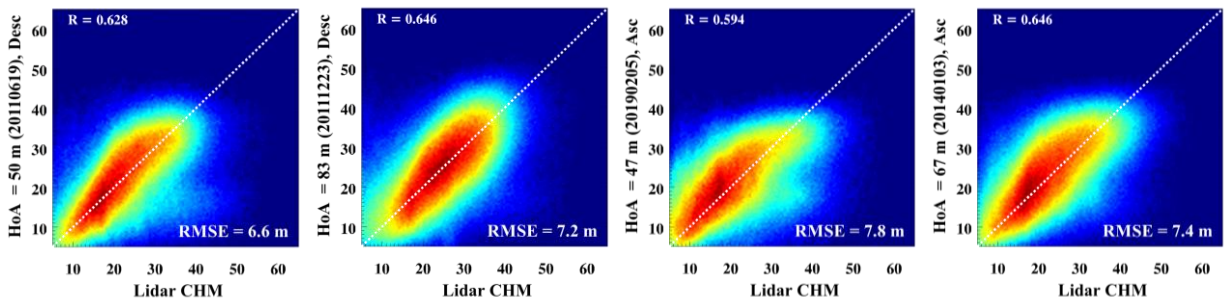
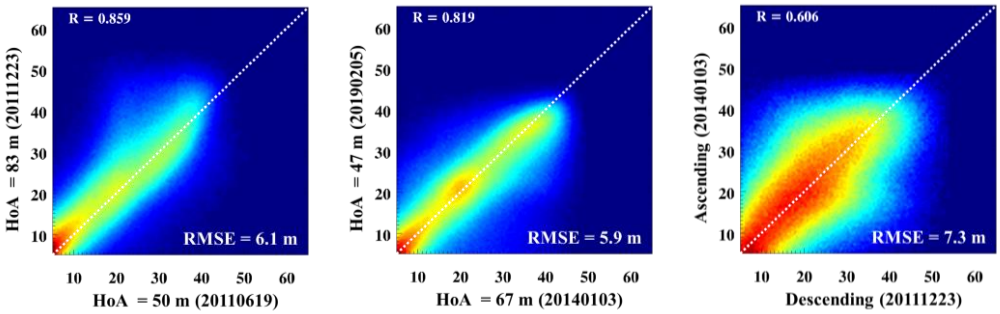


Fig. 4.9. Tasmania: forest height maps from lidar, GEDI, and four different TDX images. All maps are in geographic coordinates, with spacing about 20 m × 20 m in longitude and latitude direction, and they cover around 50 km by 50 km;



(a) (b) (c) (d)

Fig. 4.10. Comparison between lidar CHM and TDX forest heights acquired on (a) 2011/06/19, (b) 2011/12/23, (c) 2019/02/05, and (d) 2014/01/03.



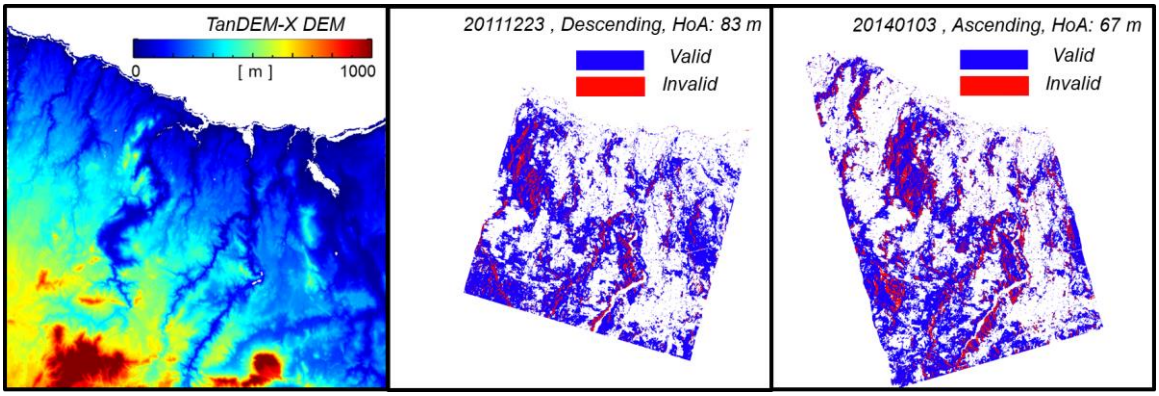
(a) (b) (c)

Fig. 4.11. Comparison between TDX forest heights. (a) With different heights of ambiguity or vertical wavenumber in descending orbit. (b) With different heights of ambiguity or vertical wavenumber in ascending orbit. (c) With different orbit directions, i.e., ascending and descending.

The forest height map obtained from a TanDEM-X scene acquired on 5th Feb 2019 with a (mean) vertical wavenumber of κ_z=0.133 (HoA = 47m) along an ascending orbit is shown in Fig. 4.9 (e). The validation plot against the CHM is shown in Fig. 4.10 (c) indicating an RMSE of 7.8 m and a correlation 0.59. Finally, in Fig. 4.9 (f), the forest height map obtained from a second TanDEM-X scene acquired along an ascending orbit on 3rd Jan 2014 with a (mean) vertical wavenumber of κ_z=0.0938 (HoA=67m) that overlaps significantly with the scene shown in Fig. 4.9 (d) (acquired about two years earlier along a descending orbit with a very similar vertical mean wavenumber). The validation plot against the CHM is shown in Fig. 4.10 (d) indicating a similar performance with an RMSE of 7.4 m and a correlation of 0.65. While the agreement between the height estimates from the ascending orbits is similar to the correlation between the height estimates from the descending orbits, the agreement between the ascending and descending orbits is (as expected) somehow lower with an RMSE of about 5.9 m and 7.3 m and correlation values of 0.82 and 0.61 (Fig 4.11 (b) & (c)). It is important to remark here that the lidar derived CHM is by itself associated to a certain uncertainty as well as those the different imaging geometries may cause an additional error when comparing the CHM with the radar derived height maps at sample level.

Finally, the performance on terrain slopes is shown in Fig. 4.12. In Fig. 4.12 (a), the TanDEM-X 90 m DEM of the site is shown while Fig 4.12 (b) & (c) the height validity maps for the two acquisitions on 23rd Dec 2011 (descending orbit) and 3rd Jan 2014 (ascending orbit). 20% and 13% of the forested areas respectively are masked out because of performance criteria respectively. For both data sets, most of the positive slope areas are excluded. Most of height estimates on slopes are on negative slope allowing to obtain valid estimates on 98% of the samples when combining both orbits. The common valid, gentle sloped areas, are characterized by a stronger height deviation due to the mismatch of the assumed and the underlying reflectivity profile.

Of course, the proposed combination of GEDI and TanDEM-X data, cannot avoid significant absolute height errors, arising from fundamental limitations, such as the limited penetration at X-band. In this case, the interferometric volume coherence no longer represents the full vertical extent of the forest, and, regardless of the assumed vertical reflectivity profile, leads to underestimated forest heights. Accordingly, underestimation of high and/or dense stands because of insufficient X-band penetration cannot be avoided and, more importantly, it is very difficult if not impossible to be detected without additional information (as for example the availability of a DTM). Nevertheless, the achieved overall performance demonstrates the potential of combining lidar and interferometric SAR measurements for large scale forest structure mapping.



(a) (b) (c)

Fig. 4.12. Tasmania: (a) TanDEM-X DEM, height validity maps for TanDEM-X images acquired in (a) 2011/12/23 and (c) 2014/01/03, respectively. All maps are in geographic coordinates, with spacing about 20 m × 20 m in longitude and latitude direction, and they cover around 50 km by 50 km.

4.7 ACKNOWLEDGMENT

This work was performed in the frame of a joint research project between DLR and NASA's Carbon Monitoring System (CMS) program in support of NASA Grant #80NSSC20K0023 to the University of Maryland. The authors would like to thank A. Sward of the Tasmanian Department of Primary Industries, Parks, Water and Environment, Australia, for providing the Lidar CHM data.

4.8 REFERENCES

[1] P. Köhler, P. and A. Huth, “Towards ground-truthing of spaceborne estimates of above-ground life biomass and leaf area index in tropical rain forests,” Biogeosciences, vol.7, pp. 2531-2543, 2010.

[2] R. O. Dubayah, S. L. Sheldon, D. B. Clark, M. A. Hofton, J. B. Blair, G. C. Hurtt, and R. L. Chazdon, “Estimation of tropical forest height and biomass dynamics using lidar remote sensing at la Selva, Costa Rica,” J. Geophys. Res., Biogeosciences, vol. 115, 2010.

[3] R. Q. Thomas, G. C. Hurtt, R. O. Dubayah, and M. H. Schilz, “Using lidar data and a height-structured ecosystem model to estimate forest carbon stocks and fluxes over mountainous terrain,” Can. J. Remote Sens., vol. 34, pp. 351–363, 2008.

[4] G. C. Hurtt, J. Fisk, R. Q. Thomas, R. Dubayah, P. R. Moorcroft, and H. H. Shugart, “Linking models and data on vegetation structure,” Geophys. Res., Biogeosciences, vol.115, 2010.

[5] S. R. Cloude and K. P. Papathanassiou, “Polarimetric SAR interferometry,” IEEE Trans. Geosci. Remote Sens., vol. 36, pp. 1551–1565, 1998.

[6] K. P. Papathanassiou and S. R. Cloude, “Single-baseline polarimetric SAR interferometry,” IEEE Trans. Geosci. Remote Sens., vol. 39, pp. 2352–2363, 2001.

[7] R. N. Treuhaft and P. R. Siqueira, “The vertical structure of vegetated and surfaces from interferometric and polarimetric radar,” Radio Sci., vol. 35, pp. 141–177, 2000.

[8] R. N. Treuhaft and S. R. Cloude, “The structure of oriented vegetation from polarimetric interferometry,” IEEE Trans. Geosci. Remote Sens., vol. 37, no. 5, pp. 2620–2624, 1999.

[9] S. R. Cloude, “Polarization coherence tomography,” Radio Sci., vol. 41, 2006.

[10] F. Kugler, S. K. Lee, and K. P. Papathanassiou, “Forest height estimation by means of Pol-InSAR data inversion: The role of the vertical wavenumber,” IEEE Trans. Geosci. Remote Sens., vol. 53, 2015.

[11] I. Hajnsek, F. Kugler, S. Lee, and K. Papathanassiou, “Tropical forest parameter estimation by means of Pol-Insar: The INDREX II campaign,” IEEE Trans. Geosci. Remote Sens., vol. 47, pp. 481–493, 2009.

[12] G. Krieger, K. P. Papathanassiou, and S. R. Cloude, “Spaceborne polarimetric SAR interferometry: Performance analysis and mission concepts,” EURASIP J. Appl. Sig. Process., vol. 20, pp. 3272–3292, 2005.

[13] A. Moreira, G. Krieger, I. Hajnsek, K. Papathanassiou, M. Younis, P. Lopez-Dekker, S. Huber, M. Villano, M. Pardini, M. Eineder, and F. De Zan, “Tandem-L: A highly innovative bistatic SAR mission for global observation of dynamic processes on the earth’s surface,” IEEE Geosci. Remote Sens. Mag., vol. 3, pp. 8–23, 2015.

[14] S. Quegan, T. Le Toan, J. Chave, J. Dall, J. F. Exbrayat, D. H. T. Minh, M. Lomas, M. M. D'alessandro, P. Paillou, K. Papathanassiou, and F. Rocca, “The European space agency BIOMASS mission: Measuring forest above-ground biomass from space,” Remote Sens. Environ., vol. 227, pp. 44–60, 2019.

[15] M. J. Soja, H. Persson, and L. M. H. Ulander, “Estimation of forest height and canopy density from a single InSAR correlation coefficient,” IEEE Geosci. Remote Sens. Lett., vol. 12, pp. 646–650, 2015.

[16] F. Kugler, D. Schulze, I. Hajnsek, H. Pretzsch, and K. Papathanassiou, “TanDEM-X Pol-InSAR performance for forest height estimation,” IEEE Trans. Geosci. Remote Sens., vol. 52, pp. 6404–6422, 2014.

[17] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, et al., "TanDEM-X: A satellite formation for high-resolution SAR interferometry", IEEE Trans. Geosci. Remote Sens., vol. 45, no. II, pp. 3317-3341, 2007.

[18] G. Krieger, M. Zink, M. Bachmann, B. Bräutigam, D. Schulze, M. Martone, P. Rizzoli, U. Steinbrecher, J. W. Antony, F. De Zan, and I. Hajnsek, “TanDEM-X: A radar interferometer with two formation flying satellites,” Acta Astron., vol. 89, pp. 83–98, 2013.
[19] H. J. Persson and J. E. S. Fransson, “Comparison between TanDEM-X and ALS based estimation of above ground biomass and tree height in boreal forests,” Scand. J. Forest Res., vol. 32, pp. 306–319, 2017.
[20] R. Treuhaft, F. Gonçalves, J. R. dos Santos, M. Keller, M. Palace, S. N. Madsen, F. Sullivan, and P. M. Graça, “Tropical-forest biomass estimation at X-Band from the spaceborne TanDEM-X interferometer,” IEEE Geosci. Remote Sens. Lett., vol. 12, pp. 239–243, 2015.
[21] S. Solberg, E. Næsset, T. Gobakken, and O. M. Bollandsås, “Forest biomass change estimated from height change in interferometric SAR height models,” Carbon Balance and Management, vol.9, pp.1-12, 2014.
[22] H. J. Persson, H. Olsson, M.J. Soja, L.M. Ulander, and J.E. Fransson, “Experiences from large-scale forest mapping of Sweden using TanDEM-X data,” Remote sens., vol.9 no.12, p.1253, 2017.
[23] H. J. Persson, M. J. Soja, J. E. S. Fransson, and L. M. H. Ulander, "National Forest Biomass Mapping Using the Two-Level Model," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol.13, pp.6391-6400, 2020.
[24] M. J. Soja, H. J. Persson, and L. M. H. Ulander, “Modelling and Detection of Deforestation and Forest Growth in Multitemporal TanDEM-X Data”, IEEE J. Sel. Top. Appl. Remote Sens. Earth Obs., 2018.
[25] C. Choi, M. Pardini, M. Heym, K. Papathanassiou, “Improving Forest Height-to-Biomass Allometry with Structure Information: A TanDEM-X Study,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 14, 2021.
[26] J. I. H. Askne, J. E. S. Fransson, M. Santoro, M. J. Soja, and L. M. H. Ulander, “Model-based biomass estimation of a hemi-boreal forest from multitemporal TanDEM-X acquisitions,” Remote Sens., vol. 5, pp. 5574–5597, 2013.
[27] M. Schlund, F. von Poncet, S. Kuntz, C. Schmullius, and D. H. Hoekman, “TanDEM-X data for aboveground biomass retrieval in a tropical peat swamp forest,” Remote Sens. Environ., vol. 158, pp. 255–266, 2015.
[28] G. Sun, K. J. Ranson, Z. Guo, Z. Zhang, P. Montesano, and D. Kimes, “Forest biomass mapping from lidar and radar synergies,” Remote Sens. Environ., vol.115, pp.2906-2916, 2011.
[29] M. Brolly, M. Simard, H. Tang, R. O. Dubayah and J. P. Fisk, "A Lidar-Radar Framework to Assess the Impact of Vertical Forest Structure on Interferometric Coherence," IEEE J. Sel. Top. Appl. Remote Sens. Earth Obs., vol.9, 2016.
[30] W. Qi and R. O. Dubayah, "Combining Tandem-X InSAR and simulated GEDI lidar observations for forest structure mapping," Remote Sens. Environ., vol.187, p.253–266, 2016.
[31] W. Qi, S. K. Lee, S. Hancock, S. Luthcke, H. Tang, J. Armston, and R. Dubayah, "Improved forest height estimation by fusion of simulated GEDI Lidar data and TanDEM-X InSAR data," Remote Sens. Environ., vol.221, pp.621–634, 2019.
[32] A. Olesk, J. Praks, O. Antropov, K. Zalite, T. Arumae, and K. Voormansik, “Interferometric SAR coherence models for characterization of hemiboreal forests using TanDEM-X data”, Remote Sens., 8, 1–23, 2016.
[33] C. Gómez, J.M. Lopez-Sanchez, N. Romero-Puig, J. Zhu, H. Fu, W. He, Y. Xie, & Q. Xie, “Canopy height estimation in mediterranean forests of Spain with TanDEM-X data,” IEEE J. Sel. Top. Appl. Remote Sens. Earth Obs., vol. 14, pp. 2956–2970, 2021.
[34] R. Guliaev, V. Cazcarra-Bes, M. Pardini, and K. Papathanassiou, “Forest Height Estimation by Means of TanDEM-X InSAR and Waveform Lidar Data,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 14, pp. 3084-3094, 2021.

[35] R. Dubayah, J. B. Blair, S. Goetz, L. E. Fatoyinbo, M. Hansen, S. Healey, M. Hofton, G. Hurtt, J. Kellner, S. Luthcke, J. Armston, H. Tang, L. Duncanson, S. Hancock, P. Jantz, S. Marselis, P. L. Patterson, W. Qi, and C. Silva, "The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography," Science of Remote Sens., vol.1, 2020.

[36] R. Dubayah, M. Hofton, J.B. Blair, J. Armston, H. Tang, and S. Luthcke, "GEDI L2A elevation and height metrics data global footprint level V002" NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/GEDI/GEDI02_A.002, 2021.

[37] H. Chen, S. R. Cloude, and J. C. White, "Using GEDI Waveforms for Improved TanDEM-X Forest Height Mapping: A Combined SINC + Legendre Approach," Remote Sens., 2021.

[38] M. Schlund, A. Wenzel, N. Camarretta, C. Stiegler, and S. Erasmi, "Vegetation canopy height estimation in dynamic tropical landscapes with TanDEM-X supported by GEDI data," Methods in Ecology and Evolution, 2022.

[39] M. Martone, P. Rizzoli, C. Wecklich, C. González, J. L. Bueso-Bello, P. Valdo, and A. Moreira, "The Global Forest/Non-Forest Map from TanDEMX Interferometric SAR data", Remote Sens. Environ., vol. 205, 2018.

[40] https://listdata.thelist.tas.gov.au/opendata/index.html#LIST_Tree_Canopy_Height

[41] M. Lachaise, M. Bachmann, T. Fritz, M. Huber, B. Schweißhelm, and B. Wessel, "The TanDEM-X Change DEM: the new temporal DEM of the TanDEM-X Mission," in Proceedings of EUSAR, pp.1-6, 2021.

[42] S. Hancock, J. Armston, M. Hofton, X. Sun, H. Tang, L. I. Duncanson, J. R. Kellner, and R. Dubayah, "The GEDI Simulator: A Large-Footprint Waveform Lidar Simulator for Calibration and Validation of Spaceborne Missions," Earth Space Sci., vol. 6, pp. 294–310, 2019.

[43] J. Beck, B. Wirt, J. Armston, M. Hofton, S. Luthcke, and H. Tang, "Global Ecosystem Dynamics Investigation (GEDI) Level 02 User Guide. Document version, 2," 2021.

[44] M. Martone, N. Gollin, P. Rizzoli, and G. Krieger, "Performance-Optimized Quantization for SAR and InSAR Applications," IEEE Trans. Geosci. Remote Sens., 2022.

[45] M. Martone, B. Bräutigam, P. Rizzoli, C. Gonzalez, M. Bachmann, and G. Krieger, "Coherence evaluation of TanDEM-X interferometric data," ISPRS J. Photogramm. Remote Sens., vol. 73, pp. 21–29, 2012.

[46] P. Rizzoli, L. Dell'Amore, J. L. Bueso-Bello, N. Gollin, D. Carcereri, and M. Martone, "On the Derivation of Volume Decorrelation From TanDEM-X Bistatic Coherence," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 15, pp.3504-3518, 2022.

[47] H. Lee and J.G. Liu, "Analysis of topographic decorrelation in SAR interferometry using ratio coherence imagery," IEEE Trans on Geosci Remote Sens., 39(2), pp.223-232, 2001.

5 CONCLUSIONS

This Section summarizes the whole work, and synthesizes the outcomes and open issues. The main findings, as well as the key points investigated for each Section, are reported in Section 5.1, and the research questions raised in Section 1.3 are answered. Moreover, the implications of this work on future investigations are discussed in Section 5.2.

5.1 SUMMARY

5.1.1 GENERAL CONCLUSIONS ON THE ROLE OF GEDI AND TANDEM-X DATA IN THE COMBINATION

The development of the algorithms within the combination framework proposed in this thesis, and the related experimental analysis, have led to a few conclusions concerning the role of GEDI lidar data and of TanDEM-X InSAR data in the combination. The developed algorithms and obtained results show an optimum combination and a great synergy potential of the discrete GEDI and the continuous TanDEM-X measurements. This is because both measurements are at the same time similar enough due to the high sensitivity to the geometrical architecture of the canopy and the high spatial resolution common to both configurations, and are different enough because of the different acquisition geometries and measurement approaches to carry independent information.

GEDI lidar data are indispensable for the definition of the forest height-to-biomass allometry. However, their role in the estimation of forest height from TanDEM-X data might seem less critical as it can be performed even in the absence of any lidar measurements. But the lack of lidar measurements affects the estimation of forest height in tall / dense / wet forest conditions due to the high attenuation at X-band.

In contrast, the horizontal structure index can be derived from TanDEM-X data alone. TanDEM-X data have been demonstrated especially appropriate for the quantitative characterization of the horizontal forest structure in terms of the spatial variability of the top canopy (phase center) height because of:

(i) the high attenuation of X-band waves into forest volume, limiting in turn the penetration, and maximizing the sensitivity of the TanDEM-X coherence to the top canopy layer. This is the same low penetration that limits the height estimation performance and makes the characterization of the vertical forest structure almost impossible (at reasonable scales);

(ii) the high interferometric accuracy of TanDEM-X as a consequence of the single-pass implementation, allowing to access the structure-induced coherence variation;

(iii) the high spatial resolution of the TanDEM-X coherence measurements and their spatial continuity (which GEDI cannot supply), allowing the estimation of structure-induced variations at spatial scales relevant for the characterization of the horizontal forest structure.

5.1.2 Answers To The Research Questions

Q1. How far can a horizontal forest structure index derived from TanDEM-X data be used to locally adapt the height-to-biomass allometry in heterogeneous forests, and improve biomass estimation performance?

The horizontal structure index considered in this thesis expresses top canopy height variations. A large index value is interpreted as a large top canopy height variation and is associated with a large canopy roughness as in the case of a sparse forest, while a low index value is interpreted as a low top canopy height variation and associated with a low canopy roughness as in a dense(r) forest. The index proposed in [1] considers vertical reflectivity profiles, and cannot be applied directly to the TanDEM-X case, as no reflectivity profiles can be reconstructed from a single InSAR coherence. The use of TanDEM-X canopy height profiles (CHP) calculated as the histograms of the InSAR phase center heights as an approximation of the reflectivity profiles at a resolution of 25 m × 25 m has been proposed instead. A critical requirement is the compensation of the terrain height variations from the phase center heights, so that they can express only canopy variations. After the calculation of the structure index, the lidar height and biomass data are used to derive a continuous relationship between the allometric factor and the structure index.

In Chapter 2, this concept is demonstrated in three forest test sites (Lopé, Mabounie and Mondah) in Gabon covered within the AfriSAR 2015-2016 campaign [2] in which airborne LVIS lidar height (DTM and RH100) and biomass data are available. In this analysis, the (continuous) lidar DTM is compensated from the TanDEM-X coherences. The horizontal structure index is derived at a resolution (here 100 m × 100 m) coarser than the CHP resolution. Within each structure resolution cell, the index is calculated as the number of peaks of all the TanDEM-X CHPs contained in a height interval extending for a fraction of the maximum height above ground. The results achieved for each test site indicate that, indeed, this index is able to adapt a more general height-to-biomass allometry (100 m × 100 m resolution) to local forest (density) conditions by changing the allometric factor. The estimation biases appearing in the conventional single height-to-biomass allometry are widely compensated improving the overall biomass RMSE up to 30%. Importantly, the ability to establish a single height-to-biomass allometry common to the three sites supported by the horizontal structure index has been demonstrated. This result indicates the potential of generalizing the height-to-biomass allometry at large(r) scale where the structure index is provided.

Chapter 3 addresses the GEDI–TanDEM-X case in which the direct compensation of a DTM is not possible. Due to the sparsity of the lidar footprint, even an interpolation of the measured terrain height may not provide the required accuracy for a reliable compensation. Then, a

methodology has been devised to compensate a “relative” terrain height variation, and a new horizontal structure index from the resulting CHP has been defined (see the answer to **Q2**). The real data analysis over the Lopé site shows a monotonically decreasing relationship between the allometric factor and the structure index, as one would reasonably expect. Accordingly, the biomass estimation performance improves when compared to the use of a single allometry. The largest remaining uncertainty contribution (in terms of bias and / or variance) is attributed to the propagation of the height estimation uncertainty.

Q2. In which way can a horizontal forest structure index be estimated from TanDEM-X data in the absence of a DTM?

This question has been addressed in Chapter 3. A wavelet analysis has been carried out to understand at which scale the canopy height variations are independent, or at least less affected by the terrain height variations. In the Lopé site, it has been found that the TanDEM-X phase center height variations are not affected by the terrain height variations at a scale of 10 m (or even finer). In order to reduce the degradation of the phase center height due to phase noise variance, coarser structure resolutions should be employed. In this case, in order to reflect only the top canopy height variations, and in absence of a DTM, the TanDEM-X phase center heights calculated with a resolution of 5 m × 5 m are corrected by a filtered version of its own at a low resolution of 120 m × 120 m. This resolution corresponds to a wavelet scale maximizing the terrain height variations and minimizing the canopy ones.

The definition of the structure index must be adapted accordingly. The CHPs are now calculated with the self-corrected phase centers, but in this case the definition of a top layer is not straightforward. Thus, the structure index is now defined only as the statistical variance of the heights of the top peaks of all CHP within the structure resolution cell, here set again to 100 m × 100 m. The larger the variance, the larger the corresponding top canopy height variation, and the sparser the forest.

The experimental analysis over the Lopé site showed that the proposed horizontal structural index is able to distinguish dense from sparse forest stands, as the index in Chapter 2 does. However, in absence of a DTM the structure index becomes ambiguous at forest / non-forest transitions, misinterpreting the step-like height change as increased top canopy variations and classifying the transition zone as dense forest. A forest/non-forest mask allows to identify these ambiguous areas and to exclude them from subsequent analyses.

Q3. What is the role of a horizontal forest structure index in compensating the forest height estimation bias in dense (tropical) forests in the absence of a DTM?

This question has been addressed in Chapter 3. The basic assumption is that even in dense stands there might be some small areas in which the X-band waves penetrate until the ground and provide unbiased height estimates. If so, the estimation bias can be corrected even in absence of a DTM at the cost of spatial resolution.

Starting from high resolution (25 m × 25 m) height estimates, the horizontal structure index is thresholded to identify dense and sparse stands. Based on this classification, heights are aggregated in different ways within each final height resolution (100 m × 100 m) cell. Following an optimization process, in dense stands the highest three estimates are averaged, while in sparse stands the average is extended to all the estimates. A noticeable improvement (bias reduction) of the estimation performance (evaluated against GEDI heights) has been found, although a full bias compensation is not possible.

Q4. Which is the ability of the GEDI waveform sampling in the parameterization of a height-to-biomass allometry as a function of sampling density and resolution?

This question has been addressed in Chapter 3. The ability of the GEDI waveform sampling in parameterizing the height-to-biomass allometry has been evaluated with respect to (i) the sampling density, and (ii) resolution effects.

The Monte Carlo simulation of scenarios with different sampling density has indicated that the derivation of the allometric coefficients for a single allometry is very robust against the decrease of the available number of waveform samples, as long as they continue to represent the forest conditions in the scene. This does not appear to be a critical limitation, as the underlying allometry, depending primarily on the species composition and the site growth conditions, remains valid at larger scales. However, the performance of this underlying allometry on Lopé has been seen to be insufficient to provide accurate biomass estimates as the heterogeneous forest stand conditions in the site cannot be represented by a single allometric relation. The obtained allometry parameterized by the structure index is less stable at the decrease of the available samples. For instance, for a very extended cloud cover affecting the 90% of the samples the structure-dependent allometric factor cannot be reconstructed anymore.

Resolution effects arise from the fact that GEDI provides height and biomass measurements at footprint level (25 m diameter), while the horizontal structure index used to refine the height-to-biomass allometry is at coarser resolution (100 m × 100 m). The effect of this discrepancy has been analyzed in this case by estimating the (structure-dependent) allometric coefficient by means of LVIS gridded height and biomass measurements at 50 and 100 m resolutions in Lopé, and by comparing the obtained allometry with the one obtained by using GEDI data at 25 m resolution. In the last case, the adaptation of the allometry to structure appears less effective than in the other two, maybe indeed due to the large difference of resolution. However, the geolocation error of the GEDI footprints might have played a role in this case.

Q5. In which way can GEDI waveforms and heights initialize and/or correct TanDEM-X forest height inversion in the absence of a DTM?

This question has been addressed in Chapter 4. In general, the initialization of TanDEM-X height inversion at fine resolution by means of GEDI waveforms is motivated by a common sensitivity to the canopy geometry induced by a similar attenuation at such short wavelengths.

But GEDI waveforms and X-band vertical reflectivity profiles are not the same due to the different acquisition geometries in which the two configurations operate, causing a likely stronger ground contribution in the GEDI waveforms than in the TanDEM-X reflectivity induced by the nadir-looking geometry.

Additionally, since only a limited number of GEDI waveforms are available for a TanDEM-X scene, the inversion could be initialized by using a "mean" GEDI waveform for all the TanDEM-X pixels. This strategy does not account for the instrinsic spatial variability of structure across the scene. One single profile may not be suitable to represent this variability, as a consequence an empirical correction of the lidar attenuation as a function of the candidate height in the inversion has been proposed and applied. In this, the mean GEDI height within a scene is used to identify the height at which no correction is performed.

After the height-dependent attenuation correction, the bias and the inaccuracy induced by a residual profile mismatch remain still unaccounted. These can be significant where the assumed TanDEM-X vertical reflectivity profile is very different from the actual situation. In order to compensate for the height bias arising from a systematic profile mismatch across the whole data set, the GEDI RH98 heights are employed. An ordinary linear square bisector (OLS) fitting of the product of the estimated vertical wavenumber – height ($\kappa_z h_V$) at the locations of the GEDI footprints successfully corrected the bias and mismatch.

Q6. Which performance can be achieved in height estimation by the implemented GEDI-TanDEM-X combination for wide areas (e.g. country-wide) applications?

This question has been addressed in Chapter 4 considering large scale height inversion results over Tasmania. It is remarkable that the proposed approach (including height inversion from the corrected "mean" GEDI profile, selection of the reliable estimates by filtering out low coherence values and inappropriate vertical wavenumber values for the expected height range, and global bias correction accounting for a residual profile mismatch) can circumvent the constrains imposed by one single-pol TanDEM-X acquisition on their ability to reconstruct forest heights on large scales by using GEDI's waveform and height measurements. The accuracy of the obtained forest height maps with a spatial resolution of 25 m appears to be enough (in an absolute and/or relative context) to be significant in forest mapping applications.

In the analysed Tasmania case, even though the height difference over the most parts of the test site is very homogenous and robust, the proposed combination of GEDI and TanDEM-X data cannot avoid significant absolute height errors (7 m of RMSE and 0.66 of Pearson coefficient against lidar CHM) at local scale. These discrepancies arise from fundamental limitations, such as a local variability of GEDI–TanDEM-X profile mismatch, and limited penetration at X-band, that can lead to relevant underestimation of forest height in tall and/or dense stands.

The forest height inversion performance critically depends on the ability of the derived set of reflectivity profiles to match the real underlying reflectivity. In contrast, with no mismatch, the assumed reflectivity profile can interpret the variation of the InSAR coherence for a wide range of vertical wavenumbers, providing consistent height estimates for different vertical

wavenumbers (obviously under the performance constraints imposed by the actual vertical wavenumber). The experimental results point to this conclusion. Spatially correlated mismatches can appear for instance in presence of slopes, or for images acquired in ascending and descending orbits combined together.

5.2 OUTLOOK

The conclusions of this Thesis summarized in Section 5.1 highlight also some open issues to be faced and additional research questions to be answered from which it is possible to identify future research directions.

A large scale (country-wide) implementation of the forest height mapping through the GEDI – TanDEM-X combination has been presented. To improve the estimation performance, strategies to derive locally-variable approximations of the X-band reflectivity profiles should be investigated further. Maybe the horizontal structure index could play a role in it. Certainly, the availability of height, structure and biomass maps at (sub-)hectar resolution is important not only for monitoring purposes, but also for initializing forest models towards the quantitative characterization of other parameters (e.g. gross primary productivity and net ecosystem exchange [3].), disturbance, carbon fluxes, or even to constrain dynamic models (like FORMIND [4]) to elaborate predictions in time. In this regard, experiments have already been started regarding the implementation of both large scale structure and biomass mapping through the calculation of the proposed horizontal structure index and height-to-biomass allometry, respectively. First results are reported in Fig. 5.1 and 5.2. Fig. 5.1 shows the global distribution of horizontal structure as measured by TanDEM-X phase variations and Fig. 5.2 shows the global distribution of forest AGB as estimated by horizontal structure index and height-to-biomass allometry at 100 m resolution. GEDI RH98 and AGB data were used to initialize the allometry.

A thorough performance evaluation of these large scale structure and biomass products is currently in progress, which is expected to end up in an analysis of still unexplored trade-offs especially when it comes to the definition of the height-to-biomass allometry and the relationship between the allometric factor and the structure index. Indeed, the experiments in Chapter 2 have shown the potential of generalizing the height-to-biomass allometry at large scale, however three aspects should be further evaluated in this context: (i) the difference of resolution between the GEDI samples and the structure index, (ii) the representativeness of the lidar sampling in the parameterization of the allometric factor by means of the structure index, and (iii) the ambiguities of the horizontal structure index in this process.

It has already been noted in Chapter 3 of this thesis that the difference of resolution between the GEDI footprint-level biomass and the TanDEM-X structure index reduces the adaptation of the allometry. In this regard, it could be worth investigating if it is possible to obtain (sparse) biomass estimates with the same (hectar) resolution of the structure index, for instance at the crossings of the GEDI orbits. However, the possible improvement brought by the recovered resolution difference is to be traded-off with the lower number of GEDI aggregated biomass estimates available for the relationship derivations, and at the same time the representativeness of

the structural heterogeneity induced by their spatial distribution. Notice that such investigation would not be possible at a test site level as it would lead to an irrelevant quantity of aggregated biomass estimates.

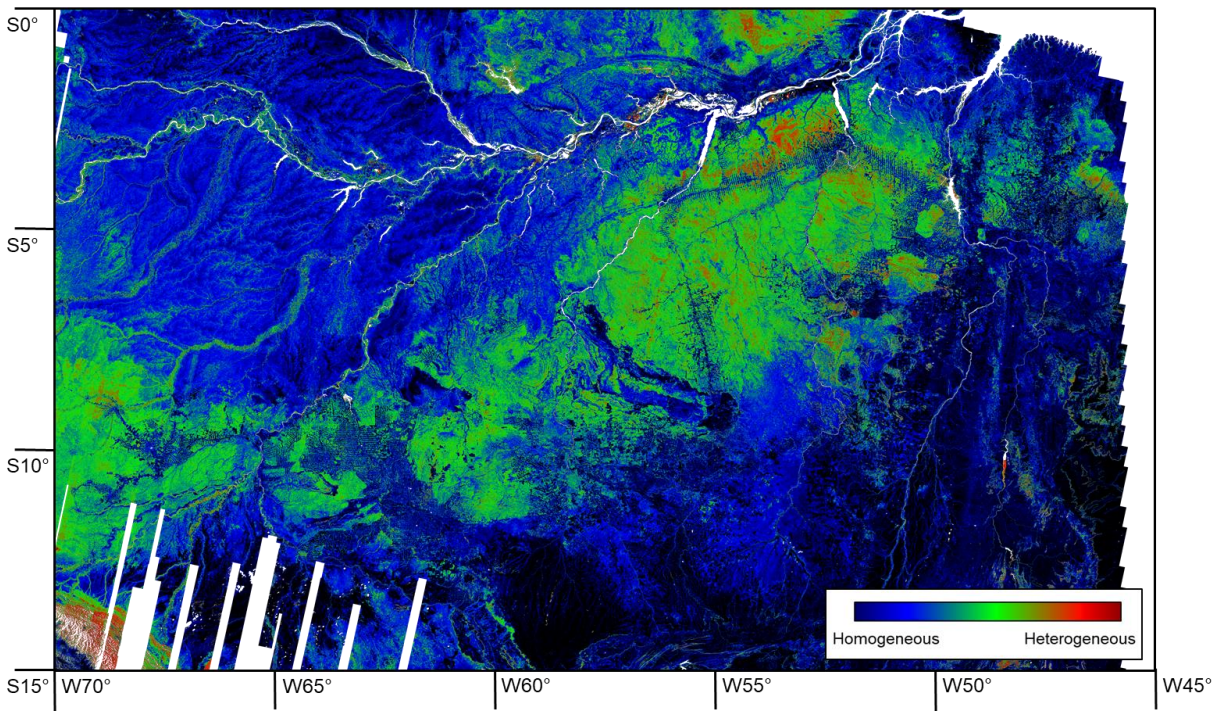


Fig. 5.1. TanDEM-X forest structure map over Brazilian Amazon forest. The map is in geographic lat-lon coordinates, with resolution of 100 m × 100 m. The acquisition period is between 2010 and 2020.

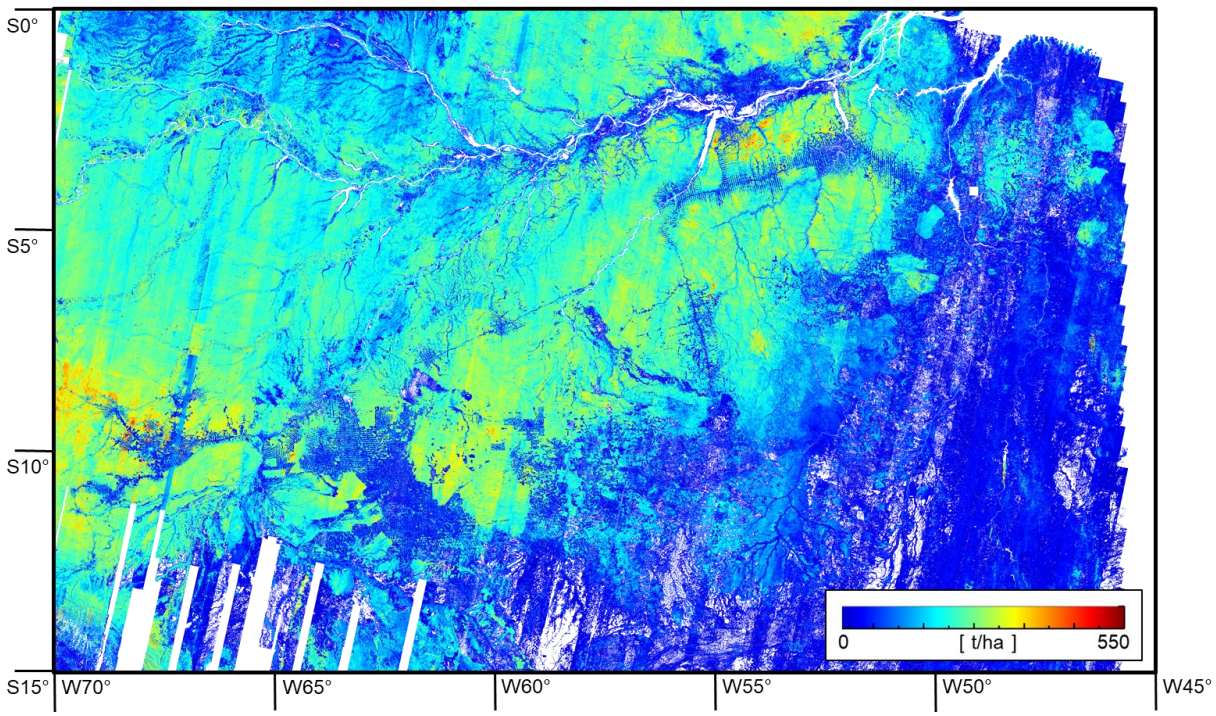


Fig. 5.2. TanDEM-X AGB map over the Brazilian Amazon forest. The map is in geographic lat-lon coordinates, with resolution of 100 m × 100 m. The acquisition period is between 2010 and 2020.

The limit in the number of lidar samples and its consequential under-representativeness of some structure levels can be counteracted by reducing the structure resolution of the allometric factor. The trade-off between the structure resolution needed to achieve a robust allometry and the possible loss in biomass estimation performance due to a less detailed allometry should be characterized.

In Chapter 2 it has also been concluded that the structure index can reconstruct a decreasing allometric factor for increasing sparsity, as it is reasonable to expect, even in absence of a DTM at a test-site scale. Although this trend seems to be confirmed at large scale by the first experiments above, some more attention should be paid to the dispersion of the allometric factor for each structure index value. Its characterization could lead to improved and more appropriate definitions of the structure index. The availability of continuous and extended (possibly small footprint) airborne lidar acquisitions with some physical understanding brought by ground measurements is a critical requirement to advance the interpretation of the structure index. It should be remembered that the structure index as defined now is a single number that actually should represent and distinguish among a wide range of structure types. In this sense, more significant definitions may also better explore the resolution of the TanDEM-X data, and evaluate structure by means of multi-scale representations. A wavelet analysis may provide an appropriate framework.

An attractive opportunity offered by the TanDEM-X mission is the possibility to retrieve forest changes. Indeed, up to now, all land surfaces were covered several times, although more or less dense and/or long time series have been realized over a few test sites. In particular, two global forest coverages with a reasonable time differences (2010-2013 and 2018-2020) and appropriate interferometric sensitivity are available [5]. The GEDI–TanDEM-X combination framework could then be extended towards the characterization of height, structure and biomass changes. However, it is understood that a change should not be evaluated by means of a mere difference of estimates, but rather by means of a differential procedure in the InSAR coherence domain to maximize accuracy and sensitivity even to the smaller changes. More importantly, the change framework should be able to differentiate between geometric structural changes (induced by growth, management, logging, mortality, disturbance etc.) and dielectric changes (induced by change of water content induced by seasonality, rainfalls, droughts, etc). Both of them define the change of the underlying vertical reflectivity function, resulting into a change of InSAR coherence. Past investigations [6] have already reported the effect of seasonal dielectric changes on the estimation of forest height from TanDEM-X data. In contrast, the horizontal structure index might be more robust to dielectric changes [1] since it focuses on spatial (phase center) height variations, rather than absolute measurements, thus being more sensitive to geometric changes. Being sensitive to the geometric canopy architecture, lidar waveform measurements could be used to constrain one temporal end of the geometric change component, and to allow the estimation of the dielectric one. The resulting characterization of the dielectric change effects would bring the combination of lidar and SAR measurements to a new scientific and application perspective. This direction of investigation is very challenging and ambitious, and has not been explored yet.

Finally, in the very near future the ESA BIOMASS mission will be launched with the objective of mapping forest height and biomass of tropical forests globally [7]. It is a SAR mission operating

at P-band (wavelength around 70 cm) that will provide tomographic, interferometric and polarimetric data. The chosen wavelength maximizes penetration until the ground even for the densest stands, however international regulations only allow a small bandwidth at P-band for signals transmitted from space, which in turn results into a low spatial resolution. For this reason, the BIOMASS mission will provide height and biomass at a resolution of 200 m. A multifrequency synergistic combination framework could then be established including GEDI, TanDEM-X and BIOMASS measurements in which the unique complementary capabilities of each sensor could be exploited. For instance, the use of BIOMASS derived forest height estimates with a TanDEM-X derived horizontal structure index (calculated using the BIOMASS ground topography) and a reference height-to-biomass allometry derived from GEDI data could be investigated. The P-band BIOMASS measurements are used to overcome the X-band penetration issues in forest height estimation, but also in the horizontal structure index calculation; the TanDEM-X measurement resolution is critical for the derivation of the index at the scales relevant for characterizing horizontal structure in contrast to the low resolution of the BIOMASS data; the GEDI measurements continue to be indispensable for the definition of the allometry. A critical condition to realize this combination is the ability to transfer the different measurements and products across scales. A successful development of such combination framework would pave the way to the definition of a new generation of physical products for monitoring forests from space.

5.3 REFERENCES

[1] M. Tello Alonso, V. Cazcarra Bes, M. Pardini, and K. Papathanassiou, “Forest structure characterization from SAR tomography at L-band,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol.11, pp.3402–3414, 2018.

[2] I. Hajnsek, M. Pardini, M. Jäger, R. Horn, J. S. Kim, H. Jörg, and K. Papathanassiou, Technical assistance for the development of airborne SAR and geophysical measurements during the AfriSAR campaign, Final technical report, ESA contract no. 4000114293/15/NL/CT. Available at: https://earth.esa.int/documents/10174/134665/AfriSAR-Final-Report.

[3] L. Bauer, N. Knapp, and R. Fischer, “Mapping Amazon Forest Productivity by Fusing GEDI Lidar Waveforms with an Individual-Based Forest Model,” *Remote Sens.*, vol.13, 2021.

[4] P. Köhler, and A. Huth, “The effects of tree species grouping in tropical rainforest modelling: simulations with the individual-based model FORMIND,” *Ecological Modelling*, vol.109, pp.301-321, 1998.

[5] M. Lachaise, M. Bachmann, T. Fritz, M. Huber, B. Schweißhelm, and B. Wessel, “The TanDEM-X Change DEM: the new temporal DEM of the TanDEM-X Mission,” *in Proceedings of EUSAR*, pp.1-6, 2021.

[6] F. Kugler, D. Schulze, I. Hajnsek, H. Pretzsch and K. Papathanassiou, "TanDEM-X Pol-InSAR performance for forest height estimation," *IEEE Trans. Geosci. Remote Sens.*, vol.52, pp.6404–6422, 2014.

[7] T. Le Toan, S. Quegan, M.W.J. Davidson, H. Balzter, P. Paillou, K. Papathanassiou, S. Plummer, F. Rocca, S. Saatchi, H. Shugart, and L. Ulander, “The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle,” *Remote Sens. Environ.*, vol.115, pp.2850-2860, 2011.

6 CURRICULUM VITAE

Personal Information

Name:	Changhyun Choi
Date of Birth:	27th January 1994 (Busan, South Korea)
Institution:	German Aerospace Center (DLR)
	Microwaves and Radar Institute (HR)
Work Address:	Münchener Str. 20, 82234 Weßling
Email:	changhyun.choi@dlr.de; cch947@gmail.com;

Education

2017 – 2022 PhD Studies
ETH Zurich, Institute of Environmental Engineering, Zurich, Switzerland.
Thesis: *"Combining Tandem-X Interferometric SAR and GEDI Lidar Measurements for Improving Forest Height, Structure and Biomass Estimates"*

2014 – 2017 Master of Science in Earth and Environmental Sciences
Seoul National University (SNU), Seoul, South Korea
Thesis: *"Determining the Optimum Conditions for Generating Fine Topography in Tidal Flats Using an In-SAR System"*

2011 – 2014 Bachelor of Science in Earth and Environmental Sciences
Bachelor of Science in Astronomy
Seoul National University (SNU), Seoul, South Korea
Thesis: *"Quantitative Area Mapping of Shinmoedake Volcanic Ash Using Time-series InSAR Coherence"*

Training Courses

DLR (German Aerospace Center) Graduate Programs

- The basics of communication
- Speed reading
- Basics of project management

- Effective scientific presentations
- Project management – follow up
- Effective scientific presentations – follow up
- How to publish in peer-reviewed journals

External Programs

- 2021, 3-day course IDEA League Summer School - “Climate Change and University Campuses”. RWTH Aachen University
- 2019, 5-day course “Advanced Course on Radar Polarimetry”. European Space Agency (ESA).
- 2017, 5-day course “SAR Principles and Applications”. Carl-Cranz-Gesellschaft e. V.

Teaching Experiences

Spring Semester 2017 Satellite Geophysics and Lab.
Spring Semester 2017 Earth –system Science Lab.
Fall Semester 2016 Earth –system Science Lab.
Spring Semester 2016 Earth –system Science Lab.
Seoul National University (SNU), Seoul, South Korea

Honors and Awards

2021	Best presentation award in VeKNI (The Korean Scientists and Engineers Association in the FRG)
2020	VeKNI (The Korean Scientists and Engineers Association in the FRG) Scholarship Award
2017	German Academic Exchange Service Scholarship (DAAD)
2016	Scholarship from Seoul National University Foundation
2015	Best presentation award in Korea Remote Sensing Society
2015	Scholarship from Seoul National University Foundation
2015	Scholarship for Honor Students
2014	Scholarship from Seoul National University Foundation
2011 – 2014	National Science and Technology Scholarship

Journal Publications

First author

1. **Choi, C.,** Pardini, M., Papathanassiou, K., "Large Scale Forest Height Mapping by Combining TanDEM-X and GEDI data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.
2. **Choi, C.,** Pardini, M., Armston, J., and Papathanassiou, K., "Forest Biomass Mapping Using Continuous InSAR and Discrete Waveform Lidar Measurements: A TanDEM-X / GEDI Test Study," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2023.
3. **Choi, C.,** Pardini, M., Heym, M., and Papathanassiou, K., "Improving Forest Height-To-Biomass Allometry with Structure Information: A Tandem-X Study," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021.
4. **Choi, C.** and Kim, D.J., "Optimum baseline of a single-pass In-SAR system to generate the best DEM in tidal flats," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018.

Co-author

1. M. Pardini, J. Armston, W. Qi, S. K. Lee, M. Tello, V. Cazcarra Bes, C. Choi, K. P. Papathanassiou, R. O. Dubayah, and L. E. Fatoyinbo, "Early lessons on combining lidar and multi-baseline SAR Measurements for forest structure characterization," *Surveys in Geophysics*, vol. 40, pp. 803–837, Jul. 2019.
2. Kang, K.M., Kim, D.J., Hwang, J.H., **Choi, C.,** Nam, S., Kim, S., Cho, Y.K., Byun, D.S. and Lee, J., "Establishment of thermal infrared observation system on Ieodo Ocean Research Station for time-series sea surface temperature extraction," *The Sea*, vol.22, pp.57-68, 2017.

Conference Publications

1. **Choi, C.,** Guliaev, R., Pardini, M. and Papathanassiou, K., 2022. "Large Scale TanDEM-X and GEDI Tropical Forest Height Mapping with Structure Information," in *Proceedings of EUSAR*, Leipzig, Germany, 2022.
2. **Choi, C.,** Guliaev, R., Pardini, M. and Papathanassiou, K.P., 2022. "Fusion of TanDEM-X and GEDI Data for Forest Height Mapping in the Brazilian Amazon," in *Proceedings of IGARSS*, Virtual, 2022.

3. Cazcarra-Bes, V., Pardini, M., **Choi, C.,** Guliaev, R. and Papathanassiou, K., "TanDEM-X and GEDI Data Fusion for a Continuous Forest Height Mapping At Large Scales," In *Proceedings of IGARSS*, Virtual, 2021.
4. **Choi, C.,** Pardini, M., Guliaev, R. and Papathanassiou, K., "An Investigation of Forest Height Estimation by using TanDEM-X Local InSAR Phase Center Variations," in *Proceedings of EUSAR*, Virtual, Germany, 2021.
5. **Choi, C.,** Guliaev, R., Cazcarra-Bes, V., Pardini, M. and Papathanassiou, K.P., "Forest Height Estimation from Tandem-X InSAR Coherence Magnitude Towards Large Scale Applications," in *Proceedings of IGARSS*, Virtual, 2020.
6. **Choi, C.,** Guliaev, R., Pardini, M. and Papathanassiou, K., "On the Role of Structure Metrics in the Combination of the GEDI Waveforms and TanDEM-X Coherences for Biomass Estimation," in *Proceedings of ESA PolInSAR Workshop*, Frascati, Italy, 2019.
7. **Choi, C.,** Pardini, M. and Papathanassiou, K., "A Structure-Based Framework for the Combination of GEDI and Tandem-X Measurements Over Forest Scenarios," in *Proceedings of IGARSS*, Yokohama, Japan, 2019.
8. **Choi, C.,** Pardini, M. and Papathanassiou, K., "Quantification of Horizontal Forest Structure from High Resolution Tandem-X Interferometric Coherences," in *Proceedings of IGARSS*, Valencia, Spain, 2018.
9. **Choi, C.,** Pardini, M., Armston, J., Papathanassiou, K.P. and Dubayah, R., "Relating TanDEM-X Local InSAR Phase Center Variations to Lidar Full Waveforms over Forest Scenarios," in *Proceedings of EUSAR*, Aachen, Germany, 2018.
10. Kim, D.J., Jung, J., **Choi, C.,** Kang, K.M., Kim, S.H. and Hwang, J.H., "A Long Baseline Airborne SAR Interferometry for Tidal Flat Mapping," in *Proceedings of URSI AP-RASC*, Seoul, Korea, 2016.
11. Kim, D.J., **Choi, C.,** Jung, J., Kang, K.M., Kim, S.H. and Hwang, J.H., "Measurements of Intertidal Flat Topography Using a Long-Baseline Airborne Interferometric SAR," in *Proceedings of IGARSS*, Beijing, China, 2016.
12. **Choi, C.** and Kim, D., "Evaluation of Long-Baseline TanDEM-X DEM in Tidal Flat," in *Proceedings of IGARSS*, Beijing, China, 2016.