REMOTE SENSING ASSESSMENT OF TROPICAL FOREST CANOPY HEIGHT, ABOVEGROUND BIOMASS, AND REGROWTH IN MAI NDOMBE PROVINCE, DEMOCRATIC REPUBLIC OF THE CONGO

By

Herve Kashongwe

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Geography – Doctor of Philosophy

2023

ABSTRACT

Tropical forests account for about half the world's terrestrial carbon, primarily stored in aboveground biomass (AGB), and so play a key role in the terrestrial carbon cycle. Central African tropical forests constitute the second largest continuous block of tropical forests in the world, and the majority are in the Democratic Republic of the Congo (DRC). Despite this, they are relatively understudied compared to other tropical regions and their contribution to the global terrestrial carbon cycle is not well unknown. The overarching research goal of this dissertation is to quantify AGB storage in Central African tropical forest regrowth following forest cover loss, to improve our understanding of the carbon sequestration potential of forest regrowth in Central Africa. The following three research hypotheses, each comprising a dissertation chapter (Chapters 2-4) and all concerning DRC Mai Ndombe province, are addressed:

#1. Using dry and wet season Landsat 8 imagery will significantly improve forest height prediction (with airborne LiDAR training data) compared to using single season images.

#2. Using a six month time period of Global Ecosystem Dynamics Investigation (GEDI) observations is sufficient for forest AGB assessment with 10% uncertainty at Reducing Emission from Deforestation and Degradation (REDD+) project scale.

#3. The impact of regrowth species differences on mature tree AGB (at 25 m GEDI footprint scale) will be less than the 10% REDD+ forest AGB reporting uncertainty.

ACKNOWLEDGEMENTS

Foremost, I would like to express my special and heartfelt thanks to my doctoral advisor, Dr David Roy. My successful completion of a doctoral degree in Remote Sensing Geography is because of Dr Roy. He spent considerable time making constructive comments, asking relevant questions, offering worthy advice, and nurtured me into an all-rounded scholar and person. Thank you! Clearly, my doctoral committee deserves my thanks for their willingness to work with me and provide their valuable advice and encouragement through this process. I wish to extend my sincere gratitude to Drs. Geoffrey Henebry, David Skole and Leo Zulu. I would also like to thank Wildlife Works Carbon LLC for providing me the opportunity to work within their REDD+ project areas in Democratic Republic of the Congo and use their experienced team of foresters during my fieldwork.

I am very grateful to my former boss Dr. Jean Robert Bwangoy for believing in me by introducing me to my advisor Dr. Roy, and for his continued support throughout my time in graduate school. I am also grateful to the MSU Department of Geography Graduate Academic Program Coordinator Joni Burns and Program Director Dr. Nathan Moore for helping me overcome any administrative difficulties and assisting me getting funding for my dissertation research.

Finally, I would also like to express my deepest gratitude to my parents, Zacharie and Josephine, for providing me with strength, words of encouragement and belief in myself to succeed, my kids — Aurore and Liam — who have been a great source of inspiration, and my wife and best friend, Sabrinette, for her sacrifices and patience throughout my time in graduate school. Without them none of this would have been possible. I love you all. Thank you.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AGB	Aboveground Biomass
ALS	Airborne Laser Scanning
DRC	Democratic Republic of the Congo
FAO	United Nations Food and Agriculture Organization
GEDI	Global Ecosystem Dynamics Investigation
IPCC	Intergovernmental Panel on Climate Change
ISS	International Space Station
LiDAR	Light Detection and Ranging
Mg ha^{-1}	Megagrams per hectare
MNRP	Mai Ndombe REDD+ project
MRV	Monitoring Reporting and Verification
NASA	National Aeronautics and Space Administration
OMGA	Overall Mean GEDI AGB
RADAR	Radio Detection and Ranging
REDD	Reducing Emission from Deforestation and Degradation
UNFCCC	United Nations Framework Convention on Climate Change
WWC	Wildlife Works Carbon LLC

CHAPTER 1. INTRODUCTION

1.1 Research Background

Tropical forests are important global ecosystems that sequester and store large quantities of carbon, primarily in above ground biomass (AGB) (Hubau et al., 2020; Harris et al., 2021) with deforestation and forest degradation comprising ~20% of global annual net carbon emissions from human activities (Friedlingstein et al., 2022). In the last several decades international forest carbon mitigation and monitoring frameworks, including the Reducing Emissions from Deforestation and Degradation and enhancement of carbon stocks program (REDD+), have been initiated that seek to compensate developing countries for reducing carbon emissions from forested lands.

Financial support for REDD+ projects requires measured, reported, and verified (MRV) estimates of forest AGB to quantify avoided carbon emissions achieved by conservation of forest biomass (UNFCCC, 2014). Conserving secondary and recovering degraded forests is important because it enables regrowth in previously deforested or degraded tropical forests. Emerging evidence suggests that the amount of CO_2 absorbed by regrowing forests is significant (Cook-Patton et al., 2020) and potentially provides a low-cost mechanism for carbon sequestration and an effective pathway to help mitigate climate change (Houghton et al., 2015; Chazdon et al., 2016).

Central African forests are relatively understudied but are thought to be more resistant to climate extremes than other humid tropical forests (Bennett et al., 2021, Saatchi et al., 2021). They constitute the second largest continuous block of tropical forests in the world, and the majority are in the Democratic Republic of the Congo (DRC) (Vancutsem et al., 2021). Despite this situation, few studies on forest AGB (Xu et al., 2017) and forest recovery (Baccini et al., 2017; Bauters et al., 2019) exist in Central Africa mainly because the turbulent political history of most Central African countries has been such that national forest inventory systems are either underdeveloped or absent (Chave et al., 2019).

Forest AGB is ideally estimated using destructive methods that involve cutting and weighing trees, a process that is time consuming and expensive. A more conventional approach to estimating forest AGB is using allometric methods, which parameterize AGB as a function of tree diameter at breast height, height, and wood density characteristics of the species (Chave et al., 2005). In the tropics, these data are difficult to collect and can vary geographically because of variable forest structure and composition and because of deforestation and forest degradation (Baccini et al., 2012; Chave et al., 2014).

It is well established that Light Detection and Ranging (LiDAR) remote sensing can measure forest canopy height, which can then be used with standard allometric equations to estimate AGB. Airborne LiDAR has been used with allometric models and Landsat data at national scale to map forest AGB in Central Africa (Xu et al., 2017) but the high acquisition cost of Airborne Laser Scanning (ALS) data remains an obstacle for repeat large area monitoring (Wulder et al., 2012; Bustamante et al., 2016). Moreover, in the tropics, cloud cover is particularly prevalent (Kovalskyy & Roy, 2013) and so several years of satellite derived reflectance are needed. This may reduce mapping accuracy because of seasonal changes in the forest reflectance associated with wet and dry seasons especially in equatorial tropical forests where stands in secondary forests have been shown to have substantial dry and wet season reflectance differences (Tyukavina et al., 2018; Zutta et al., 2023). Different strategies have been used to handle the issue of cloud obscuration in the tropics including the use of RADAR remote sensing data that is less affected by cloud cover because of the longer electromagnetic wavelengths used. Synthetic Aperture Radar (SAR) remote sensing data often acquired in the longer wavelength C-band (e.g., Sentinel-1) and L- band (e.g., Advanced Land Observing Satellite; ALOS-2) have been used to derive AGB in the tropics, but the C-band and L-band SAR backscattering signal has been reported to saturate at

AGB values ≥150 Mg ha^{-1} (Mermoz et al., 2015; Sinha et al., 2015; Migolet et al., 2022), resulting in an underestimation of AGB (Heinrich et al., 2023). Deriving forest AGB in Central Africa using SAR data alone complicates the issue as these forests have been shown to store up to 500 Mg ha^{-1} of AGB (Xu et al., 2017; Silva et al., 2018). Two new SAR satellite missions to improve forest AGB estimation in the tropics are set to start collecting data in 2024. These include the L-band SAR satellite mission NISAR (NASA-ISRO Synthetic Aperture Radar) a joint project between the National Aeronautics and Space Administration and the Indian Space Research Organization (Lisano et al., 2023) and the P-band SAR satellite mission BIOMASS from the European Space Agency (ESA) (Quegan et al., 2019).

In contrast to airborne LiDAR data that are expensive to acquire and have very limited aerial coverage, satellite LiDAR remote sensing data, also referred to as spaceborne LiDAR data, provide freely available data that can be used to generate local to global scale maps of forest canopy height and AGB for most of the world's temperate and tropical forests. The Geoscience Laser Altimeter System (GLAS) onboard the Ice, Cloud, and land Elevation Satellite's (ICESat-1) mission was the first spaceborne LiDAR instrument to acquire global height observations of the Earth (Schutz, 2005) from 2003 to 2009 that provided invaluable calibration data to generate regional to global scale forest canopy height (Simard et al., 2011; Hansen et al., 2016) and AGB (Baccini et al., 2014; Xu et al., 2021) maps.

Recently, the Advanced Topographic Laser Altimeter System (ATLAS) instrument onboard the ICESat-2 mission was launched and has been collecting LiDAR height data since 2018 with near-global coverage between ±88° latitude. It has been recently used to generate large area forest canopy height and AGB maps (Narine et al., 2019; Li et al., 2020). However, these instruments because primarily designed to quantify ice sheet dynamics, are not optimized for vegetation mapping (Montesano et al., 2015; Potapov et al., 2021). The recent NASA Global Ecosystem Dynamics Investigation (GEDI) spaceborne LiDAR onboard the International Space Station (ISS) is the first spaceborne LiDAR specifically designed for vegetation structure retrieval between ±51.6° latitude with a focus on improved quantification of forest AGB, modeling land surface carbon budgets, and understanding of the effects of vegetation structure on habitat quality and biodiversity (Dubayah et al., 2020). The GEDI mission was planned to last for two years (Dubayah et al., 2020), collected nearly four years (from 18 April 2019 to 16 March 2023) of waveform LiDAR measurements in 25 m diameter footprints. GEDI footprints are spaced 600 m apart across track and 60 m apart along track with infrequent revisit coverage due to ISS orbit changes and high frequency sensor orientation variations (Beck et al., 2021).

The first GEDI science data were released in 2020 and have been used to map forest canopy height at near-global scale (Potapov et al., 2021; Lang et al., 2022). These initial studies used six months of Version 1.0 GEDI relative height product. Recently, the GEDI data were reprocessed with improved algorithms and sensor knowledge (Beck et al., 2021) and the Version 2.1 GEDI product suite was expanded to include a GEDI footprint-level AGB product (Duncanson et al., 2022) that has been used to develop a 1 km near-global vegetation GEDI AGB map (Dubayah et al., 2022a). This ability of GEDI to provide near-global 25 m footprint estimates of forest AGB is of great importance for the REDD+ forest carbon mitigation framework especially in Central Africa where their implementation has been hampered by the lack of forest AGB inventories data (Umunay et al., 2017). Although there is high confidence that GEDI can retrieve measurements that allow estimation of AGB at scale, less is known about how well its operational deployment performs for measurement of AGB to support REDD+ projects. This includes an understanding of the appropriate time period to collect sufficient GEDI observations for reliable forest AGB assessment.

The GEDI relative height product also enables tree height chronosequences to be generated and has been used recently to estimate Amazonian secondary forest regrowth rates without considering tree species (Milenkovic et al., 2022). This has not been undertaken in Central Africa where secondary forests regrowth rates are particularly poorly documented. This is critical because nearly half the world's tropical forests are secondary forests and are a major component of human-modified tropical landscapes (FAO, 2020), and their regrowth has the potential to play a key role in biodiversity conservation (Brancalion et al., 2018; Lennox et al., 2018), landscape restoration (Strassburg et al., 2019), determining species competition and secondary succession (Paine et al., 2015; Poorter et al., 2021), and climate change mitigation (Chadzon et al., 2016; Lewis et al. 2019; Heinrich et al., 2023).

Secondary forest regrowth is traditionally quantified using repeat *in situ* measurements over the lifetime of individual trees. Collecting *in situ* measurements in the tropics is prohibitively expensive and time consuming, and compounded by difficulties in reliably dating forest cover loss. Moreover, obtaining sufficient samples to reliably quantify regrowth at landscape scale is extraordinarily challenging as absolute growth rates typically change with tree size, age, and species. Early successional, secondary tropical forests in Central Africa typically grow quickly and have low wood density and are replaced by high wood density shade tolerant species (White, 1983; Huntley, 2023). Detailed information and data on regrowth rates in Central Africa are sparse. Few studies have documented forest regrowth over tropical Africa all of which used repeated *in situ* measurements including measurements of tree diameter at breast height (Nasi, 1997; Madron et al., 2000; Fayolle., 2012; Hubau et al., 2019), tree height (Ross et al., 1954; Swaine et al., 1983), or tree rings (Madron et al., 2003; Groenendijk et al., 2014). To quantify secondary forest regrowth from satellite data in Central Africa, precise time series Landsat-based maps of forest and forest-change to date and locate forest disturbances are required. These maps can now be provided by the 30 m global forest change product of Hansen et al. (2013) for the past 20 years (2001-2021). These data, combined with the GEDI relative height product enables tree height chronosequences to be generated to quantify secondary forest regrowth and associated AGB storage in Central Africa, and improve our understanding of their carbon sequestration potential.

1.2. Research Hypotheses

The overarching research goal of this dissertation is to quantify AGB storage in Central African tropical forest regrowth following forest cover loss, to improve our understanding of the carbon sequestration potential of forest regrowth in Central Africa. The following three hypotheses are addressed.

Hypothesis #1. Satellite derived reflectance has been used to predict tree height with models trained using LiDAR data. However, reflectance saturation occurs over dense vegetation conditions and factors, including the degree of canopy cover, the leaf area index, soil background reflectance, understory vegetation, and shadows, modify reflectance. Typically, mature tree stands in Central African tropical forests have similar reflectance throughout the year whereas stands in secondary forests, that are characterized by trees with lower height, lower canopy cover and so more apparent understory vegetation, have more evident wet and dry season reflectance differences. **Therefore, using dry and wet season satellite reflectance imagery is hypothesized to significantly improve forest height prediction (with airborne LiDAR training data) over Mai Ndombe province compared to using single season images.**

A central advantage of LiDAR compared to optical wavelength reflectance data is that structural information such as tree height can be retrieved directly from LiDAR. LiDAR derived tree height data are commonly used to train classifiers applied to optical wavelength satellite data to produce large areas tree heights maps (Hansen et al., 2016; Xu et al., 2017; Matasci et al., 2018). This assumes a good relationship between optical wavelength reflectance and tree height. Taller trees are likely to have more leaves that may be exhibited in reflectance data (for example, greater absorption at visible wavelengths and so lower reflectance, and greater leaf scattering and so higher reflectance at NIR wavelengths) (Asner, 1998; Ollinger, 2011). However, factors such as the soil background, understory vegetation, canopy cover, forest structure (3D arrangement of individual trees and arrangement of trees relative to pixels), leaf area index, shadows, and species composition, can affect reflectance (Asner, 1998; Lobell et al., 2002; Fensholt et al., 2006; Yang et al., 2006; Wei et al., 2019) and so will complicate reflectance tree height relationships. Moreover, over dense canopies, reflectance saturates and becomes insensitive to changes in structural attributes (Myneni et al., 2002; Mutanga et al., 2023). In addition, equatorial tropical forests at moderate resolution may have similar reflectance throughout the year (Abernethy et al., 2018; Sun et al., 2022); whereas stands in secondary forests, that are characterized by trees with lower height, lower canopy cover and so more apparent understory vegetation, have substantial wet and dry season reflectance differences (Tyukavina et al., 2018; Zutta et al., 2023). The effect of the Landsat acquisition seasonality on tree height mapping accuracy in Central Africa has not been documented, but using wet and dry seasons satellite reflectance imagery is expected to significantly improve forest height prediction over using single season imagery as it will help capture variation in reflectance that has been observed in secondary forests between wet and dry season.

Hypothesis #2. The GEDI LiDAR onboard the International Space Station (ISS) is a sampling instrument that generated 25 m footprint relative height and aboveground biomass (AGB) products for nearly 4 years. Financial support for REDD+ projects requires measured, reported, and verified (MRV) estimates of AGB to quantify avoided emissions achieved by conservation of forest biomass in secondary forest dominated landscapes. The Intergovernmental Panel on Climate Change's (IPCC) good practice guidance recommends 10% uncertainty for REDD+ forest AGB reporting. Recent studies have used six months of GEDI data to produce wall-to-wall forest canopy height and AGB maps at national to near global scale. Although there is high confidence that GEDI can retrieve measurements that allow estimation of AGB at scale, less is known about how well its operational deployment performs for measurement of AGB to support REDD+ projects. **Using a six month time period of GEDI observations is hypothesized to be sufficient for forest AGB assessment with 10% uncertainty at REDD+ project scale in Mai Ndombe province.**

The GEDI LiDAR onboard the International Space Station (ISS) has been used to generate 25 m footprint relative height and AGB products for nearly 4 years (Dubayah et al., 2020). REDD+ projects require forest AGB inventories to quantify avoided carbon emissions achieved by conserving forest biomass in secondary forest dominated landscapes (UNFCCC, 2014) and the Intergovernmental Panel on Climate Change's (IPCC) good practice guidance recommends a maximum uncertainty of 10% (IPCC 2003).

Recently, the GEDI AGB footprint data for a 28 month were used to develop a near-global 1 km vegetation GEDI AGB map (Dubayah et al., 2022). The map was not validated in a spatially explicit manner but 169 country level AGB totals were compared to United Nations Food and Agriculture Organization (FAO) country AGB estimates with a 0.86 reported coefficient of determination (Dubayah et al., 2022). Less is known about how well GEDI performs for measurement of AGB to support REDD+ projects. This includes an understanding of the appropriate time periods required to collect sufficient GEDI observations for reliable forest AGB assessment.

GEDI is a LiDAR sampling instrument with areal acquisition coverage built up over time. Each orbit senses 25 m diameter footprints in eight tracks spaced 600 m apart across swath, and 60 m apart along track, and the footprints may not be acquired in a spatially and temporally consistent manner due to ISS orbit changes and high frequency sensor orientation variations (Beck et al., 2021; Roy et al., 2021). In addition, a number of factors reduce the availability of GEDI AGB footprint retrievals, including clouds, ephemeral surface water, and sensitivity of the GEDI retrievals under high dense canopy cover or low AGB conditions (Kellner et al., 2021; Dubayah et al., 2022; Duncanson et al., 2022). The number of GEDI orbits, and so observation time period, needed to characterize AGB at REDD+ project scale in Central Africa to within the 10% IPCC requirement is unknown, but it is expected to be comparable than the six-observation time period recently used to derive national to near-global scale tree height and AGB maps (Potapov et al., 2021; Lang et al., 2022; Liang et al., 2023).

***Hypothesis #3*.** Forests recovering from deforestation and forest degradation provide a potential natural climate mitigation strategy that is unaccounted for in REDD+ project reporting. The GEDI relative height product enables tree height chronosequences to be generated and has been used recently to estimate Amazonian secondary forest regrowth rates without considering tree species. This has not been undertaken in Central Africa where secondary forests regrowth rates are particularly poorly documented. **The impact of forest regrowth species differences on AGB accumulation in mature trees (at 25 m GEDI footprint scale) is hypothesized to be less than the 10% REDD+ forest AGB reporting uncertainty in Mai Ndombe province.**

As stated in section 1.1 tropical forest regrowth provides an important natural climate mitigation strategy and potentially an effective low-cost mechanism for carbon sequestration that is often unaccounted for in REDD+ forest AGB reporting (Chazdon et al., 2016; Lewis et al., 2019). However, secondary forest regrowth rates in Central Africa are poorly documented due to reasons including sparse forest inventory, difficulties to date forest cover loss, and to obtain enough samples of tree height and/or stem diameter to quantify forest regrowth.

The availability of the GEDI relative height product (Dubayah et al., 2021) provides new opportunities to build tree height chronosequences to quantify tropical forest regrowth at landscape scale which has been used recently to estimate secondary forest regrowth rates at two sites in the Amazon (Milenkovic et al., 2022). Notably, they estimated growth rates using a space-for-time imputation with median forest heights per stand age class without considering tree species. However, the impact of forest regrowth species differences on AGB accumulation in Central African forest is unknown. However, because forest regrowth species in Central Africa are usually dominated by light-demanding pioneer trees with low wood densities (Huntley et al., 2023), it is expected to be less than the 10% REDD+ forest AGB reporting uncertainty at 25 m GEDI footprint scale and so may not necessary be considered.

1.3. Overview of the Following Chapters

- Chapter 2 addresses hypothesis #1 and assesses the sensitivity of airborne LiDAR and Landsat-8 based dominant canopy height and AGB mapping with respect to the season of Landsat acquisition in Mai Ndombe province, DRC. This chapter has been published in the journal *Remote Sensing.*
- Chapter 3 addresses hypothesis #2 and examines the amount of Global Ecosystem Dynamics Investigation (GEDI) data (number orbits, days, and months) needed to reliably

characterize tropical forest AGB at REDD+ project scale in Mai Ndombe province, DRC. This chapter has been published in the journal *Science of Remote Sensing*.

- Chapter 4 addresses hypothesis #3 and assesses the ability of GEDI to quantify Central African tropical forest regrowth AGB storage in Mai Ndombe province post-forest cover loss areas. This chapter is in preparation for submission to the journal *Forest Ecology and Management.*
- Chapter 5 summarizes the findings of the three research hypotheses and provides recommendations for future research.

REFERENCES

Abernethy, K., Bush, E. R., Forget, P. M., Mendoza, I., & Morellato, L. P. C. 2018. Current issues in tropical phenology: a synthesis. *Biotropica*. 50, 477–482.

Asner, G. P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*. 64, 234–253.

Baccini, A. G. S. J., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., & Houghton, R. 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*. 2, 182–185.

Baccini, A., & Asner, G.P. 2014. Improving pantropical forest carbon maps with airborne LiDAR sampling. *Carbon Management*. 4, 591-600.

Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D., & Houghton, R. 2017. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science*. 358, 230–234.

Bauters, M., Vercleyen, O., Vanlauwe, B., Six, J., Bonyoma, B., Badjoko, H., & Boeckx, P. 2019. Long-term recovery of the functional community assembly and carbon pools in an African tropical forest succession. *Biotropica*. 51, 319–329.

Beck, J., Wirt, B., Luthcke, S., Hofton, M., Armston, J. 2021. Global Ecosystem dynamics investigation (GEDI) level 02 user guide. Document version 2.0, April 2021. U.S. *Geological Survey, Earth Resources Observation and Science Center, Sioux Falls, South Dakota, USA*.

Bennett, A. C., Dargie, G. C., Cuni-Sanchez, A., Tshibamba Mukendi, J., Hubau, W., Mukinzi, J. M., & Lewis, S. L. 2021. Resistance of African tropical forests to an extreme climate anomaly. *Proceedings of the National Academy of Sciences*. 118, e2003169118.

Brancalion, P. H., Niamir, A., Broadbent, E., Crouzeilles, R., Barros, F. S., Almeyda Zambrano, A. M., & Chazdon, R. L. 2019. Global restoration opportunities in tropical rainforest landscapes. *Science Advances*. 5, eaav3223.

Bullock, E.L., Healey, S.P., Yang, Z., Acosta, R., Villalba, H., Insfrán, K.P., Melo, J.B., Wilson, S., Duncanson, L., Næsset, E., Armston, J., Saarela, S., Ståhl, G., Patterson, P.L., & Dubayah, R. 2023. Estimating aboveground biomass density using hybrid statistical inference with GEDI lidar data and Paraguay's national forest inventory. *Environmental Research Letters*. 18, 085001.

Bustamante, M. M., Roitman, I., Aide, T. M., Alencar, A., Anderson, L. O., Aragão, L., & Vieira, I. C. 2016. Toward an integrated monitoring framework to assess the effects of tropical forest degradation and recovery on carbon stocks and biodiversity. *Global Change Biology*. 22, 92–109.

Castillo-Santiago, M. A., Ricker, M., & de Jong, B. H. 2010. Estimation of tropical forest structure from SPOT-5 satellite images. *International Journal of Remote Sensing*. 31, 2767–2782.

Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., & Yamakura, T. 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*. 145, 87–99.

Chave, J., Davies, S. J., Phillips, O. L., Lewis, S. L., Sist, P., Schepaschenko, D., & Saatchi, S. 2019. Ground data are essential for biomass remote sensing missions. *Surveys in Geophysics*. 40, 863–880.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B., & Vieilledent, G. 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*. 20, 3177–3190.

Chazdon, R.L., Broadbent, E.N., Rozendaal, D.M., Bongers, F., Zambrano, A., Aide, T.M., Balvanera, P., Becknell, J.M., Boukili, V., Brancalion, P.H., & Craven, D. 2016. Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. *Science Advance*. 2, 1–10.

Cook-Patton, S. C., Leavitt, S. M., Gibbs, D., Harris, N. L., Lister, K., Anderson-Teixeira, K. J., & Griscom, B. W. 2020. Mapping carbon accumulation potential from global natural forest regrowth. *Nature*. 585, 545–550.

Dubayah, R., Armston, J., Healey, S.P., Bruening, J.M., Patterson, P.L., Kellner, J.R., Duncanson, L., Saarela, S., Ståhl, G., Yang, Z., Tang, H., Blair, J.B., Fatoyinbo, L., Goetz, S., Hancock, S., Hansen, M., Hofton, M., Hurtt, G., & Luthcke, S. 2022a. GEDI launches a new era of biomass inference from space. *Environmental Research Letters*. 17, 095001.

Dubayah, R. O., Armston, J., Kellner, J. R., Duncanson, L., Healey, S. P., Patterson, P. L., & Luthcke, S. B. 2022b. GEDI L4A Footprint Level Aboveground Biomass Density, Version 2.1. *ORNL DAAC, Oak Ridge, Tennessee, US*A. (Last accessed: 18 August 2023).

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., & Silva, C. 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*. 1, 100002.

Dubayah, R., Hofton, M., Blair, J., Armston, J., Tang, H., & Luthcke, S. 2021. GEDI L2A Elevation and Height Metrics Data Global Footprint Level V002. *NASA EOSDIS Land Processes DAAC*. (Last accessed: 18 August 2023).

FAO. 2020. Global Forest Resources Assessment 2020: *Main Report, Rome, Italy*. Retrieved from http://www.fao.org/documents/card/en/c/ca9825en (Last accessed: 30 August 2023).

Fayolle, A., Engelbrecht, B., Freycon, V., Mortier, F., Swaine, M., Rejou-Mechain, M., Doucet, J.L., Fauvet, N., Cornu, G., & Gourlet-Fleury, S. 2012. Geological substrates shape tree species and trait distributions in African moist forests. *PLoS One*. 7, e42381.

Fensholt, R., Sandholt, I., & Stisen, S. 2006. Evaluating MODIS, MERIS, and VEGETATION vegetation indices using in situ measurements in a semiarid environment. *IEEE Transactions on Geoscience and Remote Sensing*, 44. 1774-1786.

Friedlingstein, P., O'sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., & Zheng, B. 2022. Global carbon budget 2022. *Earth System Science Data*. 2022, 1–159.

Groenendijk, P., Sass-Klaassen, U., Bongers, F., & Zuidema, P.A. 2014. Potential of tree-ring analysis in a wet tropical forest: A case study on 22 commercial tree species in Central Africa. *Forest Ecology and Management*. 323, 65–78.

Hansen, M. C., Potapov, P., Moore, R., Hancher, M., Turubanova, S., Tyakavina, A., Townsend, A. R. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342. Data available on-line from: http://earthenginepartners.appspot.com/science-2013-global-forest

Hansen, M. C., Potapov, P. V., Goetz, S. J., Turubanova, S., Tyukavina, A., Krylov, A., & Egorov, A. 2016. Mapping tree height distributions in Sub-Saharan Africa using Landsat 7 and 8 data. *Remote Sensing of Environment*. 185, 221–232.

Harris, N. L., Gibbs, D. A., Baccini, A., Birdsey, R. A., De Bruin, S., Farina, M., & Tyukavina, A. 2021. Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change*. 11, 234–240.

Heinrich, V. H., Vancutsem, C., Dalagnol, R., Rosan, T. M., Fawcett, D., Silva-Junior, C. H., & Aragão, L. E. 2023. The carbon sink of secondary and degraded humid tropical forests. *Nature*. 615, 436–442.

Hérault, B., Bachelot, B., Poorter, L., Rossi, V., Bongers, F., Chave, J., & Baraloto, C. 2011. Functional traits shape ontogenetic growth trajectories of rain forest tree species. *Journal of ecology*. 99, 1431–1440.

Houghton, R.A., Byers, B., & Nassikas, A.A. 2015. A role for tropical forests in stabilizing atmospheric CO2. *Nature Climate Change*. 5, 1022–1023.

Hubau, W., De Mil, T., Van den Bulcke, J., Phillips, O.L., Angoboy Ilondea, B., Van Acker, J., & Beeckman, H. 2019. The persistence of carbon in the African forest understory. *Nature Plants*. 5, 133–140.

Hubau, W., Lewis, S. L., Phillips, O. L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., ... & Zemagho, L. 2020. Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature*. 579, 80–87.

Huntley, B. J. 2023. The Guineo-Congolian Rain Forest Biome. In Ecology of Angola: Terrestrial Biomes and Ecoregions (pp279-304). Cham: *Springer International Publishing*.

IPCC. 2003. Good practice guidance for land use, land-use change and forestry. Institute for Global Environmental Strategies, Kanagawa. Available at: https://www.ipcc.ch/publication/good-practice-guidance-for-land-use-land-use-change-and-forestry/. (Last accessed: 07 November 2023).

Kellner, J.R., Armston, J., & Duncanson, L. 2021. Algorithm theoretical basis document (ATBD) for GEDI level-4A (L4A) footprint level aboveground biomass density. In: Document Version 1, August 2021. *University of Maryland, College Park, Maryland, USA*.

Kovalskyy, V., & Roy, D. P. 2013. The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30 m Landsat data product generation. *Remote Sensing of Environment*. 130, 280–293.

Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J.D. 2022. Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*. 268, 112760.

Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Del Bom Espirito-Santo, F., & de Oliveira Jr, R. 2005. Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical research letters*. 32, 22.

Lennox, G. D., Gardner, T. A., Thomson, J. R., Ferreira, J., Berenguer, E., Lees, A. C., & Barlow, J. 2018. Second rate or a second chance? Assessing biomass and biodiversity recovery in regenerating Amazonian forests. *Global Change Biology*. 24, 5680–5694.

Lewis, S. L., Wheeler, C. E., Mitchard, E. T., & Koch, A. 2019. Restoring natural forests is the best way to remove atmospheric carbon. *Nature*. 568, 25–28.

Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., & Chen, H. 2020. High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel and Landsat-8 data. *International Journal of Applied Earth Observation and Geoinformation*. 92, 102163.

Liang, M., Duncanson, L., Silva, J. A., & Sedano, F. 2023. Quantifying aboveground biomass dynamics from charcoal degradation in Mozambique using GEDI Lidar and Landsat. *Remote Sensing of Environment*. 284, 113367.

Lisano, M., Edelstein, W., Xaypraseuth, P., Hoffman, P., Guerrero, A.M., Silverman, M., Ferdowsi, B., & Krishnamurthy, A. 2023. NASA-ISRO Synthetic Aperture Radar (NISAR): Towards Observatory-level Integration and Testing. *In, 2023 IEEE Aerospace Conference*.

Lobell, D. B., & Asner, G. P. 2002. Moisture effects on soil reflectance. *Soil Science Society of America Journal*. 66, 722–727.

Madron, L.D. 2003. Accroissements Diamétriques du bété and iroko in Central African. Tropical rainforest. *CIRAD Bois et Forêts Des Tropiques*. 275, 83–87.

Madron, L.D., Nasi, R., & Detienne, P. 2000. Accroissements Diamétriques De quelque Essence en Forêt Dense Africaine. *CIRAD Bois et Forêts Des Tropiques*. 263, 63–74.

Matasci, G., Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Bolton, D.K., & Bater, C.W. 2018. Three decades of forest structural dynamics over Canada forested ecosystems using Landsat time-series and lidar plots. *Remote Sensing of Environment*. 216, 697–714.

Mermoz, S., Réjou-Méchain, M., Villard, L., Le Toan, T., Rossi, V., & Gourlet-Fleury, S. 2015. Decrease of L-band SAR backscatter with biomass of dense forests. *Remote Sensing of Environment*. 159, 307–317.

Migolet, P., Goïta, K., Pambo, A.F.K., & Mambimba, A.N. 2022. Estimation of the total dry aboveground biomass in the tropical forests of Congo Basin using optical, LiDAR, and RADAR data. *GIScience & Remote Sensing*. 59, 431–460.

Milenković, M., Reiche, J., Armston, J., Neuenschwander, A., De Keersmaecker, W., Herold, M., & Verbesselt, J. 2022. Assessing Amazon rainforest regrowth with GEDI and ICESat-2 data. *Science of Remote Sensing*. 5, 100051.

Montesano, P. M., Rosette, J., Sun, G., North, P., Nelson, R. F., Dubayah, R. O., & Kharuk, V. 2015. The uncertainty of biomass estimates from modeled ICESat-2 returns across a boreal forest gradient. *Remote Sensing of Environment*. 158, 95–109.

Mutanga, O., Masenyama, A., & Sibanda, M. 2023. Spectral saturation in the remote sensing of high-density vegetation traits: A systematic review of progress, challenges, and prospects. *ISPRS Journal of Photogrammetry and Remote Sensing*. 198, 297–309.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., & Running, S. W. 2002. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote sensing of environment*. 83, 214–231.

Narine, L. L., Popescu, C. S., & Malambo, L. 2019. Synergy of ICESat-2 and Landsat for mapping forest aboveground biomass with deep learning. *Remote Sensing*. 11,1503.

Nasi, R. 1997. Les peuplements D'Okoumés au Gabon leur dynamique et croissance en zone côtière. *CIRAD Bois et Forêts Des Tropiques*. 251,5–27.

Neumann, T. A., Martino, A. J., Markus, T., Bae, S., Bock, M. R., Brenner, A. C., & Thomas, T. C. 2019. The Ice, Cloud, and Land Elevation Satellite–2 Mission: A global geolocated photon product derived from the advanced topographic laser altimeter system. *Remote sensing of environment*. 233, 111325.

Paine, C. T., Amissah, L., Auge, H., Baraloto, C., Baruffol, M., Bourland, N., & Hector, A. 2015. Globally, functional traits are weak predictors of juvenile tree growth, and we do not know why. *Journal of Ecology*. 103, 978–989.

Poorter, L., Rozendaal, D. M., Bongers, F., Almeida, D. J. S., Álvarez, F. S., Andrade, J. L., & Westoby, M. 2021. Functional recovery of secondary tropical forests. *Proceedings of the National Academy of Sciences*. 118, e2003405118.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., & Hofton, M. 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*. 253, 112165.

Ollinger, S.V. 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist*. 189, 375–394.

Quegan, S., Le Toan, T., Chave, J., Dall, J., Exbrayat, J. F., Minh, D. H. T., & Williams, M. 2019. The European Space Agency BIOMASS mission: Measuring Forest above-ground biomass from space. *Remote Sensing of Environment*. 227, 44–60.

Ross, R. 1954. Ecological Studies on the Rain Forest of Southern Nigeria III Secondary succession in the Shashaa Forest reserve. *Journal of Ecology*. 42, 259–282.

Roy, D.P., Kashongwe, H.B., & Armston, J. 2021. The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. *Science of Remote Sensing*. 4, 100024.

Saatchi, S., Longo, M., Xu, L., Yang, Y., Abe, H., Andre, M., & Elmore, A. C. 2021. Detecting vulnerability of humid tropical forests to multiple stressors. *One Earth*. 4, 988–1003.

Schutz, B.E., Zwally, H.J., Shuman, C.A., Hancock, D., & DiMarzio, J.P. 2005. Overview of the ICESat Mission. *Geophysical Research Letters*. 32, L21S01.

Silva, C. A., Saatchi, S., Garcia, M., Labriere, N., Klauberg, C., Ferraz, A., & Hudak, A. T. 2018. Comparison of small-and large-footprint lidar characterization of tropical forest aboveground structure and biomass: a case study from Central Gabon. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 11, 3512–3526.

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. 2011. Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences*. 116 (G4).

Sinha, S., Jeganathan, C., Sharma, L.K., & Nathawat, M.S. 2015. A review of radar remote sensing for biomass estimation. *International Journal of Environmental Science and Technology*. 12, 1779–1792.

Strassburg, B. B., Beyer, H. L., Crouzeilles, R., Iribarrem, A., Barros, F., de Siqueira, M. F., & Uriarte, M. 2019. Strategic approaches to restoring ecosystems can triple conservation gains and halve costs. *Nature Ecology & Evolution*. 3, 62–70.

Sun, Y., Knyazikhin, Y., She, X., Ni, X., Chen, C., Ren, H., & Myneni, R. B. 2022. Seasonal and long-term variations in leaf area of Congolese rainforest. *Remote Sensing of Environment*. 268, 112762.

Swaine, M.D., & Hall, J.B. 1983. Early Succession on Cleared Forest Land in Ghana. *Journal of Ecology*. 71, 601–627.

Tyukavina, A., Hansen, M. C., Potapov, P., Parker, D., Okpa, C., Stehman, S. V., & Turubanova, S. 2018. Congo Basin forest loss dominated by increasing smallholder clearing. *Science Advances*. 4, eaat2993.

Umunay, P.M., Gregoire, T.G., & Ashton, M.S. 2017. Estimating biomass and carbon for Gilbertiodendron dewevrei (De Wild) Leonard, a dominant canopy tree of African tropical Rainforest: Implications for policies on carbon sequestration. *Forest Ecology and Management*. 404, 31–44.

UNFCCC – United Nations Framework Convention on Climate Change. 2014. Warsaw Framework for REDD-plus. Available at: https://unfccc.int/topics/land-use/ resources/warsaw-framework-for-redd-plus

Vancutsem, C., Achard, F., Pekel, J. F., Vieilledent, G., Carboni, S., Simonetti, D., & Nasi, R. 2021. Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Science Advances*. 7, eabe1603.

Wei, S., Fang, H., Schaaf, C. B., He, L., & Chen, J. M. 2019. Global 500 m clumping index product derived from MODIS BRDF data (2001–2017). *Remote Sensing of Environment*. 232, 111296

West, T. A., Wunder, S., Sills, E. O., Börner, J., Rifai, S. W., Neidermeier, A. N., & Kontoleon, A. 2023. Action needed to make carbon offsets from tropical forest conservation work for climate change mitigation. *Science*. 381, 873

White, F. 1983. The vegetation of Africa. *United Nations Educational, Scientific and Cultural Organization (UNESCO)*. 20, 1-354.

Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., & Gobakken, T. 2012. Lidar sampling for large-area forest characterization: A review. *Remote sensing of environment*. 121, 196–209.

Xu, L., Saatchi, S. S., Shapiro, A., Meyer, V., Ferraz, A., Yang, Y., & Ebuta, D. 2017. Spatial Distribution of Carbon Stored in Forests of the Democratic Republic of Congo. *Nature Scientific Report*. 7, 15030.

Xu, L., Saatchi, S. S., Yang, Y., Yu, Y., Pongratz, J., Bloom, A. A., & Schimel, D. 2021. Changes in global terrestrial live biomass over the 21st century. *Science Advances*. 7, eabe9829.

Yang, W., Huang, D., Tan, B., Stroeve, J. C., Shabanov, N. V., Knyazikhin, Y., & Myneni, R. B. 2006. Analysis of leaf area index and fraction of PAR absorbed by vegetation products from the terra MODIS sensor: 2000-2005. *IEEE Transactions on Geoscience and Remote Sensing*. 44, 1829–1842.

Zuideema, P.A, & Jakovac, C.C. 2023. Regrowing tropical forests are a sizeable carbon sink, *Nature*. 615, 398–399.

Zutta, B. R., Salinas, N., Cosio, E. G., Tito, R., Aragón, S., Nina-Quispe, A., & Roman-Cuesta, R. M. 2023. Satellite-derived forest canopy greenness shows differential drought vulnerability of secondary forests compared to primary forests in Peru. *Environmental Research Letters*.18, 064004.

CHAPTER 2. DEMOCRATIC REPUBLIC OF THE CONGO TROPICAL FOREST CANOPY HEIGHT AND ABOVEGROUND BIOMASS ESTIMATION WITH LANDSAT-8 OPERATIONAL LAND IMAGER (OLI) AND AIRBORNE LIDAR DATA: THE EFFECT OF SEASONAL LANDSAT IMAGE SELECTION

This chapter was previously published by *Remote Sensing*: Kashongwe, H. B., Roy, D. P. & Bwangoy, J. R.B. 2020. Democratic Republic of the Congo Tropical Forest Canopy Height and Aboveground Biomass Estimation with Landsat-8 Operational Land Imager (OLI) and Airborne LiDAR data: the effect of seasonal Landsat image selection. *Remote Sensing*. 12 (9) 1360.

2.1 Introduction

Tropical forests play a key role in the terrestrial carbon cycle with globally significant amounts of carbon stored as aboveground biomass (AGB) (Houghton et al., 2005; Meyer et al., 2019; Hubau et al., 2020). National inventories of forest AGB are incomplete and imprecise in many tropical countries for several reasons and primarily because tropical forests have highly variable structure and composition that make them difficult to survey (Clark, D.B. & Clark, D.A., 2000; Saatchi et al., 2011; Ferraz et al., 2016). Forest AGB can be measured using destructive methods, i.e., by cutting and weighing trees which is time consuming and expensive, and so more conventionally is estimated using allometric methods that typically parametrize AGB as a function of tree diameter at breast height, or tree height, and species wood density information (Chave et al., 2005). Reliable application of allometric models depends on reliable measurement of these tree biometric parameters (Lu, 2005; Skole et al., 2011; Clark et al., 2012).

Remotely sensed data have been used to estimate tropical forest tree height and AGB. Early work used statistical approaches applied to optical wavelength passive satellite data and using allometric models (Foody et al., 2003; Disney et al., 2018). Airborne and terrestrial Light Detection and Ranging (LiDAR) remote sensing provides new capabilities for estimating tree canopy structure and has the potential to improve or even replace allometric models (Disney et al., 2018; Csillik et al., 2019). Regional to national scale tree height maps have been derived using either spaceborne or airborne LiDAR data at select locations to derive tree height training data that are used to train classifiers applied to optical wavelength satellite data, typically sensed by Landsat, with AGB derived by applying allometric models to the mapped tree heights (Hansen et al., 2016; Xu et al., 2017a; Matasci et al., 2018).

In many regions cloud is prevalent at the time of overpass of Landsat, especially in the tropics (Asner, 2001; Lindquist et al., 2008; Kovalskyy & Roy, 2013), and the Democratic Republic of Congo (DRC) in western Central Africa is one of the cloudiest tropical regions (Dommo et al., 2018). Different strategies have been used to handle this cloud issue. For example, Staben et al. (2018) mapped tree heights in the Northern territory, Australia, using single date cloud-free 30 m Landsat-5 TM or Landsat-7 ETM+ images and airborne LiDAR tree height training data. Other researchers have extracted multi-temporal metrics from Landsat time series. For example, Xu et al. (2017a) mapped tree height and AGB across the DRC using the medians (i.e., 50th percentiles) of the red, near infrared (NIR), and the two shortwave Landsat-8 Operational Land Imager (OLI) reflective wavelength bands acquired over three years, with airborne LiDAR tree height training data. Thus, the predictor variables at adjacent pixel locations may have been selected from different years and seasons. Hansen et al. (2016) mapped tree heights in Sub-Saharan Africa using 30 m multi-temporal metrics extracted from two years of Landsat-7 ETM+ and Landsat-8 OLI reflectance with tree height training data derived from Geoscience Laser Altimeter System (GLAS) data. However, it is unknown how many cloud free-observations were used, the Landsat 7 ETM+ had about 22% fewer observations due to the scan line corrector issue, and Hansen et al. (2016) did not examine the effect of the Landsat acquisition seasonality on the mapping accuracy.

Using longer time periods of Landsat data to extract temporal metrics or to select cloud-free images will increase the possibility of obtaining cloud-free observations. However, this may reduce the capability to reliably map tree height and AGB because of seasonal changes in the forest reflectance associated with wet and dry seasons and because of inter-annual variations, for example, due to drought, and other factors that are subject to ongoing research (Myneni et al., 2007; Brando et al., 2010; Manoli et al., 2018).

In this paper the sensitivity of DRC tree height and AGB estimation with respect to the season of Landsat acquisition is examined for the first time. The study area, composed of approximately 10,000 km^2 of tropical forest in the western part of the DRC, is dominated by dense tropical evergreen rainforest and is often cloudy with distinct wet and dry seasons. Four airborne discrete return LiDAR 10 km × 2 km transects flown in 2014 were used to derive dominant canopy height training and test data. A single dry and a single wet season Landsat-8 Operational Land Imager (OLI) image acquired over the study area were considered. This was because they were the only images acquired in a two-year period around the LiDAR flight dates that had low (<20%) cloud cover and that were cloud-free over the locations of the airborne LiDAR transects.

The following experiments were undertaken independently three times to map and assess the study area 30 m dominant canopy height and AGB using (i) only the wet season Landsat-8 image, (ii) only the dry season Landsat-8 image, and (iii) both images. A random forest regression estimator was trained using 50% (n = 2639) of the LiDAR 30 m heights and using 30 m Landsat-8 predictor variables defined by the green, red, NIR, and shortwave reflective wavelength bands and spectral band ratios. The dominant canopy height prediction accuracy was evaluated using the remaining 50% of the LiDAR 30 m heights not used to train the model. The random forest regression estimator was used to map the dominant canopy height at 30 m for all the study area and then the corresponding 30 m AGB was mapped by application of recent allometric equations (Xu et al., 2017a). Field measurements were used to validate the AGB at the equivalent of 43 30 m Landsat pixel locations.

The paper is structured as follows. First, the study area and the data used are described (Section 2.2), followed by the processes used to map and assess the dominant canopy height and AGB results (Section 2.3), and then the results (Section 2.4). The sensitivity of the dominant canopy height and AGB results with respect to the season of the Landsat acquisition and implications of the research are discussed (Section 2.5) followed by the conclusion (Section 2.6).

2.2 Study Area and Data

2.2.1 Study Area

The study area covers 137 km × 80 km of Mai Ndombe province in the DRC (Figure 1). It was selected because it contains airborne LiDAR data and field plot data that can be used to derive AGB. In addition, it is located within a single Landsat Path (180) and Row (061) which reduces Landsat data processing complexity as overlapping images sensed from adjacent Landsat orbits (Egorov et al., 2018) do not need to be processed. The study area falls in the tropical monsoon climate zone and because it lies close to the Equator has two wet and two dry seasons (Samba et al., 2007). The Inongo weather station, operated by the Congolese national society of meteorology, is situated near the center of the study area. The annual mean temperature is 24 °C (Einzmann et al., 2012) and the annual total rainfall is 1800 mm that falls typically over about 115 days (Bwangoy et al., 2010). The main wet season is from September to December. The principal dry season extends from June to August, with a secondary dry season from January to February.

The majority of the study area is covered by dense tropical evergreen rainforest with low lying parts that can be flooded in the wet seasons and includes the northern end of Lake Mai Ndombe (Lindquist et al., 2008; Mayaux et al., 2000). People subsist on the terra firme non-forest rural complex (evident in Figure 2.1a in pink tones), primarily growing cassava, corn, sorghum, upland rice, and peanuts and practicing slash and burn agriculture (Bwangoy et al., 2012; Molinario et al., 2015). The main forest characteristics are high tree crown cover (70–100%) with mature tree heights of 35–45 m and predominantly evergreen heterogeneous shade tolerant species (Bwangoy et al., 2010; Mayaux et al., 2000). The interior forest is relatively undisturbed but at risk of deforestation and degradation due to unregulated resource exploitation and limited governance on timber harvesting, charcoal production, and mining (Zhuravleva et al., 2013; Shapiro et al., 2016).

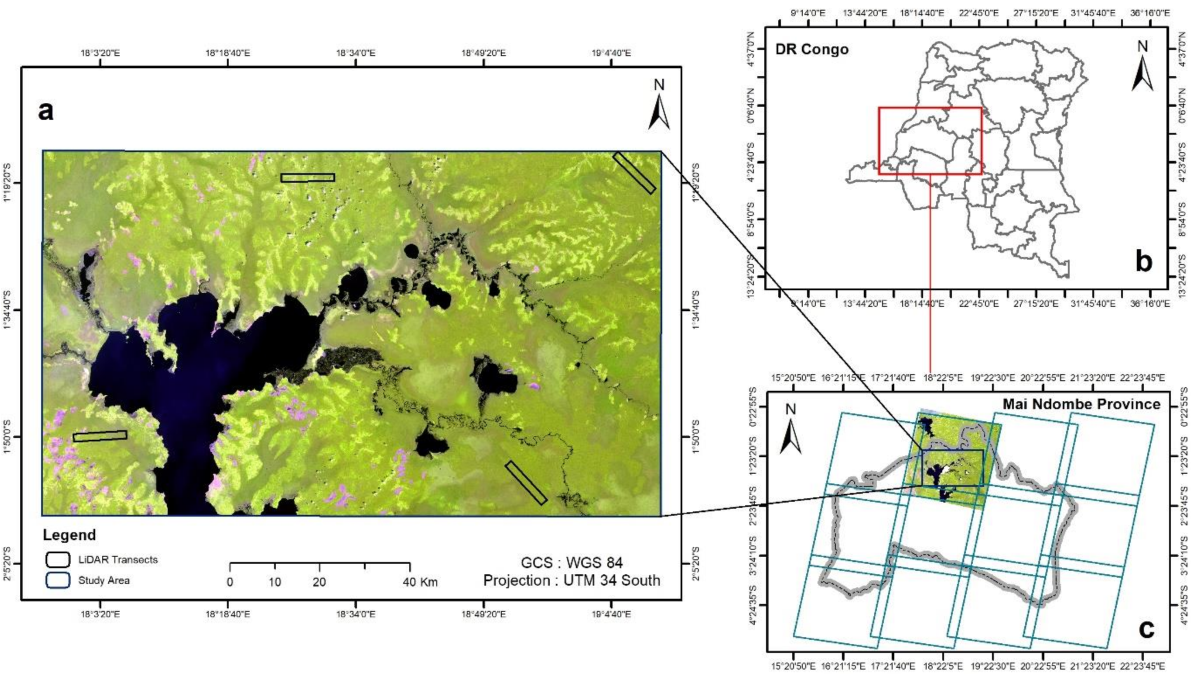

Figure 2.1 Study area showing, **(a)** the extent (black boxes) of the four airborne LiDAR transects superimposed on wet season December 8th 2014 Landat-8 false color (1610 nm, 865 nm, 655 nm) surface reflectance, **(b)** the study area location in Mai Ndombe province within the Democratic Republic of Congo, **(c)** the Landsat (WRS) 185 × 170 km image path/row coordinate map, the study area falls within Landsat Path 180 and Row 061.

2.2.2 Airborne LiDAR Transect Data

Airborne discrete-return LiDAR transects were flown across the DRC in support of a World Wildlife Fund (WWF) carbon mapping and modelling project (Xu et al., 2017a). The transects were selected based on a systematic random sampling design where a 1° × 1° grid was overlaid on the forest cover map of the country produced by Observatoire satellital des forest d'Afrique Centrale (OSFAC) (Xu et al., 2017a). An Orion M300 LiDAR was flown at 700 m aboveground level with a 150 kHZ pulse frequency and a laser beam divergence of 0.25 mRad to provide an average density of 4 returns per square meter with a nominal footprint size of 0.17 m. The airborne LiDAR data were flown to have positional spatial errors no greater than 0.05 m horizontally and 0.10 m vertically. The data are available categorized as ground or non-ground returns (Xu et al., 2017b). Four airborne transects were flown over the study area, three were flown on July 1st 2014 (the N.W., S.W., and S.E. transects illustrated in Figure 1(a)) and one on August 12th 2014 (the N.E. transect). Each transect is approximately 2 km wide and 10 km long, i.e., each covers approximately 20 km^2 (2000 hectares).

2.2.3 Landsat-8 Operational Land Imager (OLI) Data

The Landsat-8 Operational Land Imager (OLI) provides 30 m optical wavelength images with improved radiometry and geolocation compared to previous Landsat sensors (Roy et al., 2014). The most recent Collection 1 Landsat images that are defined with per-pixel cloud and quality information (Dwyer et al., 2018) were used. In order to select contemporaneous imagery, two years of Landsat-8 OLI images acquired from July 2013 to August 2015, i.e., from one year before to one year after the airborne LiDAR data acquisitions, were considered. Only those OLI images that had <20% cloud cover (defined by the Collection 1 "land cloud cover" metadata) and that were cloud-free over the four LiDAR transects were selected. In total, out of 48 OLI images acquired over the two years, only two images, sensed July 14th 2013 (i.e., dry season) and December 8th 2014 (i.e., wet season) met these selection criteria. This is not surprising given the prevalence of cloud at the time of Landsat overpass that is evident in Figure 2.2.

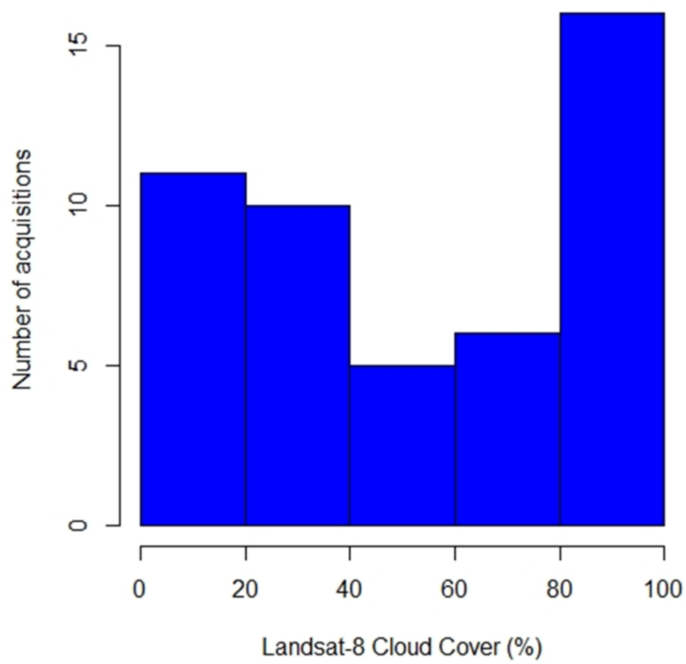

Figure 2.2 Histogram showing the percentage cloud cover in the 48 Landsat-8 Operational Land Imager (OLI) images acquired over the study area (Landsat Path 180 Row 061, Figure 1) for two years from July 2013 to August 2015. In total, out of 48 images, 11 had a cloud cover < 20%, 23 had cloud cover ≥50%, and 16 had cloud cover ≥80%.

The Landsat-8 OLI has nine 30 m reflective wavelength (435 nm to 2200 nm) bands. In this study OLI bands 3 (Green, ~560 nm), 4 (Red, ~655 nm), 5 (NIR, ~865 nm), 6 (Shortwave Infrared, ~1610 nm), and band 7 (Shortwave Infrared, ~2200 nm) were used. The two OLI blue bands were not used because of their sensitivity to atmospheric scattering (Dwyer et al., 2018; Vermote et al., 2016). The surface reflectance rather than top of atmosphere reflectance was used to minimize the effects of atmospheric contamination that can be particularly significant over the tropics due to high water vapor content and biogenic and pyrogenic aerosols. The surface reflectance imagery were obtained from the United State Geographical Survey (USGS) website (https://earthexplorer.usgs.gov) and are derived using the Land Surface Reflectance Code (LaSRC) (Vermote et al., 2016).

2.2.4 Aboveground Biomass Field Plot Validation Data

Field plot data used to validate AGB were collected in support of the Mai Ndombe Emission Reductions Program, a World Bank coordinated program (under the Forest Carbon Partnership Facility Carbon Fund) that aims to provide benefits for the local population while reducing greenhouse gas emissions from deforestation and forest degradation (World-Bank et al., 2016). The field plots were in undisturbed primary forest (Xu et al., 2017b). The field plot data were collected November 2015 where the airborne LiDAR data had been flown the previous year (recall the LiDAR data were flown July and August 2014). The field plot data were collected in two of the study area LiDAR transects (the NE and SW transects illustrated in Figure 2.1a). In each transect, two field plots of one hectare situated <10 km apart were surveyed on the ground. Thus, there were four field plots in the study area.

Each field plot (1 ha) was divided into sixteen 25 m × 25 m (0.0625 ha) parcels. In each parcel, the diameter at breast height (1.3 m) for all trees with diameter >10 cm was measured. The species of each tree was identified by ecologists following the procedure described in Xu et al. (2017b) and the tree wood density was assigned using the Global Wood Density Database for tropical trees (Chave et al., 2005). In cases where tree identification was not possible, the mean wood density of the plot was assigned. The tree AGB in each parcel was derived using a standard allometry model (Chave et al., 2014) as was undertaken by (Xu et al. 2016b):

$$\mathrm{AGB} = \frac{10^{-3}}{a}\sum_{i=1}^{n} exp\,(-1.803 - 0.976\,e + 0.976\,ln(\rho_i) + 2.673\,ln(d_i) - 0.0299(ln(d_i))^2) \tag{1}$$

where *AGB* is the estimated tree aboveground biomass in the parcel (Mg ha^{-1}), *a* is the parcel area (ha), *d* is the diameter at breast height of each tree in the parcel (cm), ρ is the wood density of each tree in the parcel (g cm^{-3}), n is the number of trees in the parcel, and e is an environmental stress parameter that depends on the seasonality of temperature, precipitation, and the climatic water deficit (Chave et al., 2014). There were 16 parcels and so 16 AGB estimates were derived as (1) for each one hectare field plot, except for one field plot (located over the study area NE LiDAR transect) where there were 15 AGB estimates as no data was collected in one of its parcels. This provided a total of 63 AGB 25 m × 25 m parcel estimates.

No field measurements of the non-forest vegetation, i.e., grasses and shrubs, were made. The AGB of grasses and shrubs in tropical forests is not well documented, and although a minor fraction of tree AGB, post tropical forest disturbance (e.g., due to deforestation, degradation, or fire) grass, shrubs, and tree saplings grow rapidly (Souza et al., 2005). However, in this study, only tree AGB was considered.

2.3. Methods

2.3.1 Forest Mask Classification

A 30 m forest mask was derived so that only the forest parts of the Landsat images would be considered. This is needed as about a quarter of the study area is composed of water and seasonally inundated soil (Figure 2.1.). The Landsat images were acquired in the dry and wet seasons and so the lake level and the soil and cloud conditions were different between the images. In addition, the Landsat cloud and shadow mask was not always reliable, which has been observed by (Qiu et al., 2019; Egorov et al., 2019). Therefore, in order to provide a reliable forest mask, a supervised random forest classification applied to both images was undertaken, which is a state of the practice land cover classification approach (Wulder et al., 2018), and used to define forest, water, permanent wet soil (wet in both Landsat images), dry soil, cloud, and shadow classes. Training data were derived by visual interpretation of both Landsat images. Care was taken to ensure that the forest training pixels did not include mixed forest and non-forest pixels (e.g., over forest edges and small forest clearings) as they likely contain shrubs and grasses whose AGB is unknown. Care was also taken to select training samples across the study area and to ensure that the proportion selected among the different classes reflected the visually estimated study area class proportions in order to provide approximately similar class training portions as found by random sampling (Zhang et al., 2017).

A total of 7280 training pixels of 30 m were collected composed of forest (67% of the pixels), water (25%), permanent wet soil (3%), dry soil (3%), cloud (1%), and shadow (1%) classes. The classification predictor variables were defined by the Landsat-8 OLI surface reflectance for bands 3, 4, 5, 6, and 7. In addition, normalized difference band ratios, defined like the normalized difference vegetation index (NDVI), for every possible two band combination of these bands were derived. This provided a total of 11 predictor variables. These bands and ratios have been used before for Landsat land cover classification (Zhang et al., 2017; Hansen et al., 2016; Yan & Roy, 2015).

The training data were used to develop a random forest classification tree using the default parameter settings, i.e., 500 trees were grown with each tree built using 63.2% of the training data selected randomly with replacement and three predictor variables (the square root of the number of predictor variables) randomly selected (Breiman, 2001). The random forest classification was applied to the 11 predictor variables at every 30 m study area pixel. The land cover classification was checked by visual comparison with the Landsat-8 OLI images and with Google Earth high resolution images. No formal quantitative per-pixel assessment of the classification accuracy was undertaken as the objective here was to develop a conservative forest mask used to discard non-forest pixels from the tree height and AGB mapping.

2.3.2 LiDAR Dominant Canopy Height Quantification

The data were processed using FUSION, a public software designed by the U.S. Forest Service to analyze LiDAR data (McGaughey et al., 2016). The Landsat 30 m pixel grid was used to define a coordinate system.

First, the number of LiDAR ground returns in different sized grid cells (1 m, 2 m, 2.5 m, and 3 m side dimensions) aligned with the Landsat coordinate system were examined to determine an appropriate grid cell dimension for the subsequent processing. In the DRC national carbon mapping study undertaken by Xu et al. (2016b) a 2 m grid cell dimension was used. However, for the four study area transects, we found that a 2.5 m grid cell dimension was more appropriate as with smaller grid cells there were usually no ground returns in each grid cell. As an example, Figure 2.3 illustrates in detail the number of ground returns for the four different grid cell dimensions considered. The percentage of illustrated grid cells with no ground return data were 85%, 60.4%, 49.3%, and 39.7% for 1 m, 2 m, 2.5 m, and 3 m, grid cell dimensions, respectively. In Figure 2.3, the greatest ground return density is in the north east and occurs where there are no trees.

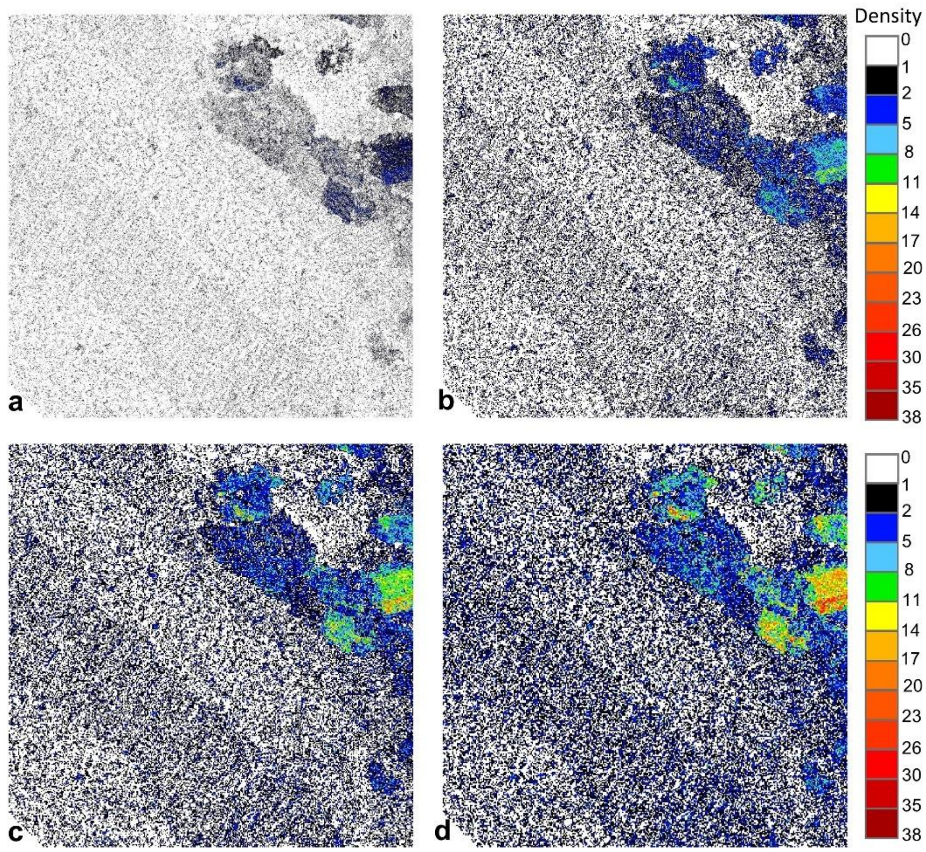

Figure 2.3 Illustration of the sensitivity of the airborne LiDAR ground returns density to grid cell size, showing the number of ground returns in **(a)** 1 m × 1 m grid cells (85% contain no ground return data), **(b)** 2 m × 2 m grid cells (60.4% contain no data), **(c)** 2.5 m × 2.5 m grid cell (49.3% contain no data), and **(d)** 3 m × 3 m grid cells (39.7% contain no data). Example results for a 1 km^2 portion of the NE study area LiDAR transect (Figure 2.1a).

The discrete-return airborne LiDAR transect data categorized as ground returns were used to generate a 2.5 m ground height digital terrain model (DTM) by averaging the heights of the ground returns falling in each 2.5 m grid cell. Some DTM grid cells had no data (e.g., white in Figure 2.3) and the DTM gaps were interpolated from neighboring DTM grid cell values by natural neighbor interpolation that has been recommended for LiDAR processing (Bater & Coops, 2009) and has elegant interpolation properties, i.e., no parameters are used, the interpolated values are guaranteed to be within the range of the samples used and to pass through the input samples, and are smooth everywhere except at the locations of the input samples (Ju et al., 2012).

A canopy height model (CHM) was estimated by extracting the DTM height from the maximum first return height in each 2.5 m grid cell and only considering first returns with heights >1 m. This is a common approach in tropical forests if the LiDAR first returns are not particularly noisy (Asner et al., 2014; Leitold et al., 2015; Réjou-Méchain et al., 2015; Xu et al., 2016).

The dominant canopy heights in 30 m grid cells aligned with the Landsat 30 m pixel grid were derived by taking the mean of the 2.5 m CHM values falling in each 30 m grid cell. The mean rather than another metric, such as the maximum or the median, was used as it provides a reliable representation of forest structure and has been used in other LiDAR based tropical forest studies (Xu et al., 2016; Mascaro et al., 2011; Asner et al., 2011). The dominant canopy height was derived only for 30 m grid cells ≥75% covered by LiDAR data (i.e., containing ≥108 2.5 m canopy height values). This resulted in a proportion of the 30 m grid cells along the LiDAR transect edges being discarded from the analysis.

The above processing was also repeated independently for the LiDAR data falling over the 25 m × 25 m AGB field parcels (Section 2.4) and using the field plot corner locations to define a coordinate system. A 2.5 m DTM was generated, then, as above, any DTM gaps were filled by natural neighbor interpolation, canopy heights were estimated for each 2.5 m grid cell, and then the dominant canopy heights in the 25 m × 25 m grid cells falling over each 25 m × 25 m field plot parcel were derived.

2.3.3 Dominant Forest Canopy Height Prediction and Accuracy Assessment

The dominant forest canopy height was predicted at each Landsat 30 m pixel using the established non-parametric supervised random forest regression estimator (Breiman, 2001). Other researchers have also used this approach (Xu et al., 2017a; Staben et al., 2018; Xu et al., 2016). Only the 30 m pixels classified as forest (Section 2.3.1) were considered.

The response variable was defined by the 30 m dominant canopy height data (Section 3.2) sampled systematically every four pixels (120 m) north and south across each LiDAR transect. A four pixel sampling interval was used to reduce spatial autocorrelation effects that can introduce biases into the forest height prediction (Miller et al., 2007). The four pixel sampling interval was selected beacause it is >100 m which is the distance that canopy heights in Mai Ndombe province were found to be significantly different from forest edge canopy heights (Shapiro et al., 2016).

A total of 5278 pixels with 30 m dominant canopy height response variables and 11 associated predictor variables were extracted. The predictor variables were defined by the Landsat-8 OLI surface reflectance for bands 3, 4, 5, 6, and 7. In addition, normalized difference band ratios, defined like the NDVI, for every possible two band combination of these bands were derived. The 5278 response and predictor values were divided into two equally sized portions, one portion was used to train the random forest regression and the other to test it. To ensure that a full range of forest canopy heights were used in both the training and testing, the following sampling procedure was used. The 5278 30 m dominant canopy height values were ranked into ascending canopy height order. Every second sample in the ranked list was selected as training data (n = 2639) and the remainder were used to define the test data (n = 2639). The dominant canopy heights for the training data ranged from 2.71 m to 43.99 m and for the test data ranged from 2.65 m to 42.71 m.

The random forest regression estimator was trained using the 2639 30 m dominant canopy height pixel training values and the 11 corresponding predictor values. The default random forest regression parameter settings were used, i.e., 500 trees were grown with each tree built using 63.2% of the training data selected at random with replacement and 3 predictor variables (one third the number of predictor variables) randomly selected (Beiman, 2001). The resulting random forest regression tree was applied to the 11 predictor variables at every forest mask Landsat-8 OLI pixel location to generate a 30 m dominant forest canopy height map.

The random forest regression prediction accuracy was assessed by application of the random forest regression tree to the 2639 test predictor variables. The resulting 2639 random forest regression predicted canopy heights were compared with the test 30 m dominant forest canopy heights and the Root Mean Square Error (RMSE) between them derived. In addition, scatterplots comparing the predicted and test 30 m dominant canopy heights were generated and Ordinary Least Squares (OLS) regressions between the data and the goodness of fit (R^2) and regression confidence (p value) statistics were derived.

The above process was undertaken independently three times, using the Landsat-8 OLI predictor variables extracted from (a) only the dry season (July 14th 2013) image, (b) only the wet season (December 8th 2014) image, and (c) both images. This resulted in three dominant forest canopy height maps and three accuracy assessments.

2.3.4 Aboveground Biomass Mapping

The aboveground biomass (AGB) was derived at each 30 m pixel location with a dominant forest canopy height estimate as:

$$AGB = 1.88\, h^{1.55} \tag{2}$$

where AGB is the predicted aboveground biomass (Mg ha^{-1}) and h is the 30 m dominant forest canopy height predicted by the random forest regression tree (Section 3.3). This allometric equation was defined by Xu et al. (2017a) by statistically fitting 92 pairs of dominant canopy heights (derived using the same airborne LiDAR data as this study but extracted from more (33) LiDAR transects flown across the main forest types of the DRC) with field AGB estimates (derived as described in Section 2.2.4).

Three AGB maps were generated by application of Equation (2) to the 30 m dominant forest canopy height maps generated using predictor variables derived from the Landsat-8 OLI (a) dry season, (b) wet season, and (c) both images.

2.3.5 Aboveground Biomass Map Accuracy Assessment

The study area 30 m AGB maps were validated by comparing them with the field plot AGB data that were defined in 25 × 25 m parcels (Section 2.4). The Landsat pixels and field parcels have different sizes and are not aligned. Consequently, the 25 × 25 m parcel AGB estimates falling under each 30 × 30 m Landsat pixel location were weighted to derive an equivalent 30 m field AGB estimate as:

$$\mathrm{AGB}^{30} = \frac{\sum_{i=1}^{n} \mathrm{AGB}_i^{25} f_i}{\sum_{i=1}^{n} f_i} \quad (3)$$

where AGB^{30} is the 30 × 30 m AGB field estimate derived from the *n* (typically 4 but sometimes 2 or 1) parcel AGB estimates (AGB^{25}) that fall under the 30 m Landsat pixel location, and f_i is the fraction of the 30 × 30 m Landsat pixel area occupied by parcel *i*. As some Landsat 30 m pixels fall along the forest plot edges, and so include areas with no field AGB estimate information, i.e., , $\sum_{i=1}^{n} f_i \leq 1$, only 30 m pixel locations with $\sum_{i=1}^{n} f_i \geq 0.5$ were considered. Thus, if a 30 m pixel was less than 50% covered by field plot parcels it was not considered.

The RMSE between AGB^{30} and the corresponding mapped 30 m AGB values (Section 3.4) were derived. Scatterplots comparing these data were generated and OLS regressions between the data and the goodness of fit (R^2) and regression confidence (p value) statistics were derived to quantify the correspondence of the data.

2.4. Results

2.4.1 Dominant Forest Canopy Height Maps

Figure 4 shows the predicted 30 m dominant forest canopy heights for the study area derived using the Landsat-8 OLI predictor variables generated using (a) only the dry season, (b) only the wet season, and (c) both images. White shows the pixels that were classified as either water, permanent wet soil, dry soil, cloud, or shadow, and that were masked off from the subsequent AGB analysis. The masked off pixels include Lake Mai Ndombe evident in the wet season Landsat-8 OLI image (Figure 2.1a) and also capture most of the small rivers including streams with small axis dimensions greater than about half a 30 m pixel. Clouds and shadows located mostly in the North West that occurred in the dry season Landsat-8 OLI image are also apparent. Typically, forest edges and small forest clearings were classified as one of the non-forest classes (usually as wet soil or water) which is not a problem as these masked off pixels likely contain shrubs and grassed whose AGB is unknown.

The 30 m dominant forest canopy height maps derived using the predictor variables collected from the dry season (Figure 2.4a) and wet season (Figure 2.4b) Landsat-8 OLI images have a different evident spatial distribution. The differences are most evident around Lake Mai Ndombe and in the vicinity of several of the rivers. The mapped results derived using the predictor variables collected from both images (Figure 2.4c) tend to have intermediate or lowe nopy heights. Despite r ca these geographic differences, the mean 30 m dominant forest canopy height for the study area was similar between the three maps and was 20.6 m (Figure 2.4a), 20.8 m (Figure 2.4b), and 20.4 m (Figure 2.4c). The maximum 30 m dominant forest canopy height was 36.30 m, 37.22 m, and 37.23 m, respectively. The minimum 30 m dominant forest canopy height found in the three maps was very similar, within 0.01 m, and was approximately 4.0 m.

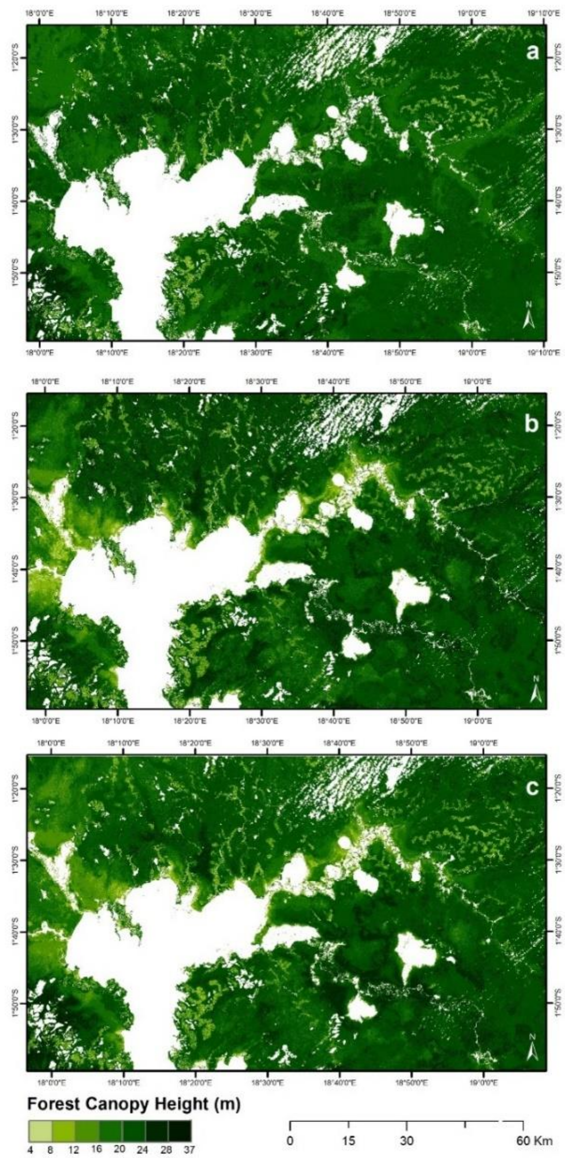

Figure 2.4 Predicted 30 m dominant forest canopy height derived by the random forest regression tree using Landsat-8 OLI predictor variables collected from **(a)** the dry season July 14th 2013 Landsat-8 OLI image, **(b)** the wet season December 8th 2014 Landsat-8 OLI image, and **(c)** both images.

2.4.2 Dominant Forest Canopy Height Prediction Accuracy Assessment

The dominant forest canopy height prediction accuracy was assessed, as described in Section 3.3, by application of the random forest regression tree to the 2639 test pixels that were not used to train the tree. This was undertaken three times using the trees derived with Landsat-8 OLI predictor variables generated from (a) the dry season, (b) the wet season, and (c) both images. Figure 5 shows scatterplots comparing the test and the predicted 30 m dominant canopy height values. There are two clouds of dots evident in the scatterplots, the larger cloud corresponds to tall trees > 20 m present in the mature tropical evergreen forest parts of the four LiDAR transects, and the other corresponds to shorter forest canopies about 18 m high that occur predominantly around the Lake Mai Ndombe and often in the S.E. and S.W. LiDAR transects.

The OLS regressions of the plotted data are shown in red in Figure 5. In all three cases the regressions are significant (p values < 0.05) with slopes less than unity and intercepts > 12 m. The random forest regression underestimates and overestimates the heights for pixels dominated by tall and short trees, respectively. The predicted 30 m dominant canopy heights are similar to the test height values only for dominant canopy heights around 22 m. The wet season results (Figure 2.5b) have the lowest OLS regression R^2 (0.28) and the greatest RMSE (4.43 m). The dry season results (Figure 2.5a) are slightly improved with an 0.36 R^2 and a 4.17 m RMSE. The prediction accuracy is best when both the Landsat-8 OLI images were used (Figure 2.5c) with an 0.47 R^2, a 3.84 m RMSE, and the OLS regression slope is closer to unity (0.42) and the intercept is closer to zero (12.41 m). This 3.84 m RMSE value corresponds to about 17% of the mean of the 2639 test pixel canopy height values (22 m). These results indicate that using both the dry and wet season Landsat-8 OLI images provides more accurate dominant canopy height prediction.

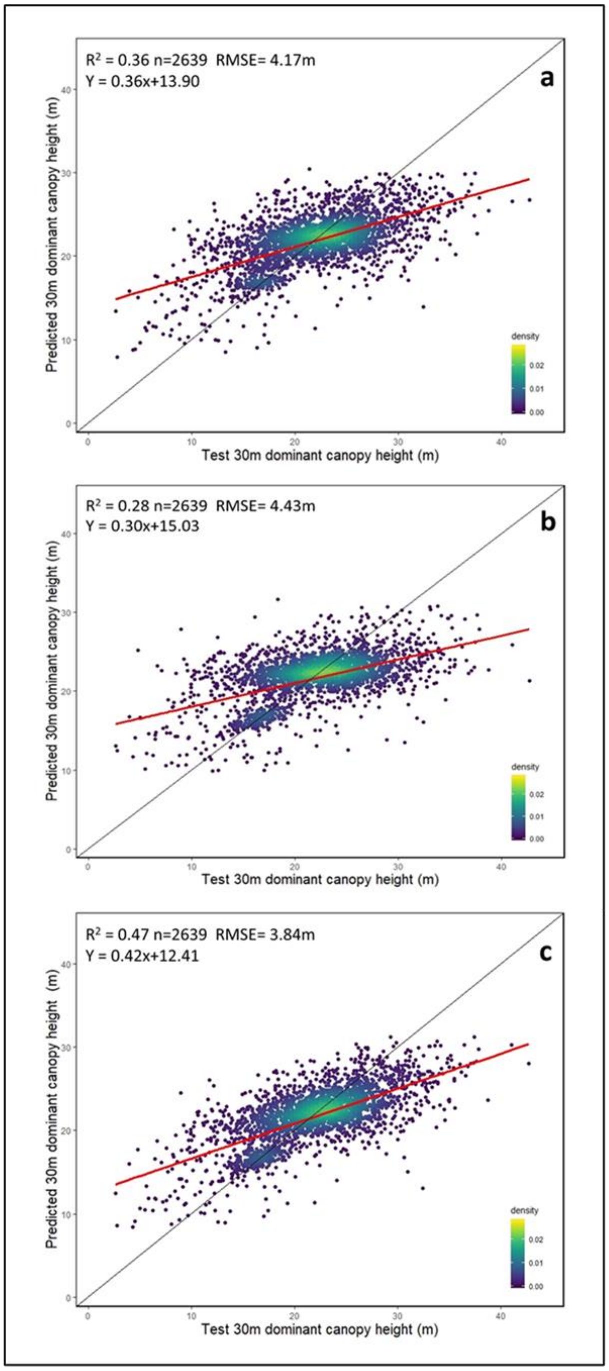

Figure 2.5 Scatterplots comparing the 30 m dominant canopy heights of the 2639 test pixels and the random forest regression tree predicted values. Results shown for the regression trees derived using predictor variables from **(a)** the dry season July 14th 2013 Landsat-8 OLI image, **(b)** the wet season December 8th 2014 Landsat-8 OLI image, and **(c)** both images. The point densities, calculated using a 100 × 100 quantization of the plot axes, are displayed with a rainbow color scale.

2.4.3 Aboveground Biomass Maps

Figure 2.6 shows the 30 m AGB biomass maps derived from the 30 m dominant forest canopy height maps (Figure 2.4) using Equation (2). The same broad patterns as the dominant forest canopy height maps are observed, which is expected given that the AGB is proportional to the dominant canopy height.

The mean study area AGB was 206 Mg ha^{-1}, 211 Mg ha^{-1}, and 204 Mg ha^{-1} for the AGB maps generated using the dry season, wet season and both images dominant forest canopy height maps, respectively. The maximum AGB was found at the 30 m pixels with greatest dominant canopy height and was 493.5 Mg ha^{-1} (tree height 36.3 m), 511.8 Mg ha^{-1} (tree height 37.22 m), and 511.8 Mg ha^{-1} (tree height 37.23 m) for the dry season, wet season, and both image maps, respectively. The minimum 30 m AGB among the three maps was very similar, within 0.01 Mg ha^{-1}, and was approximately 16 Mg ha^{-1}.

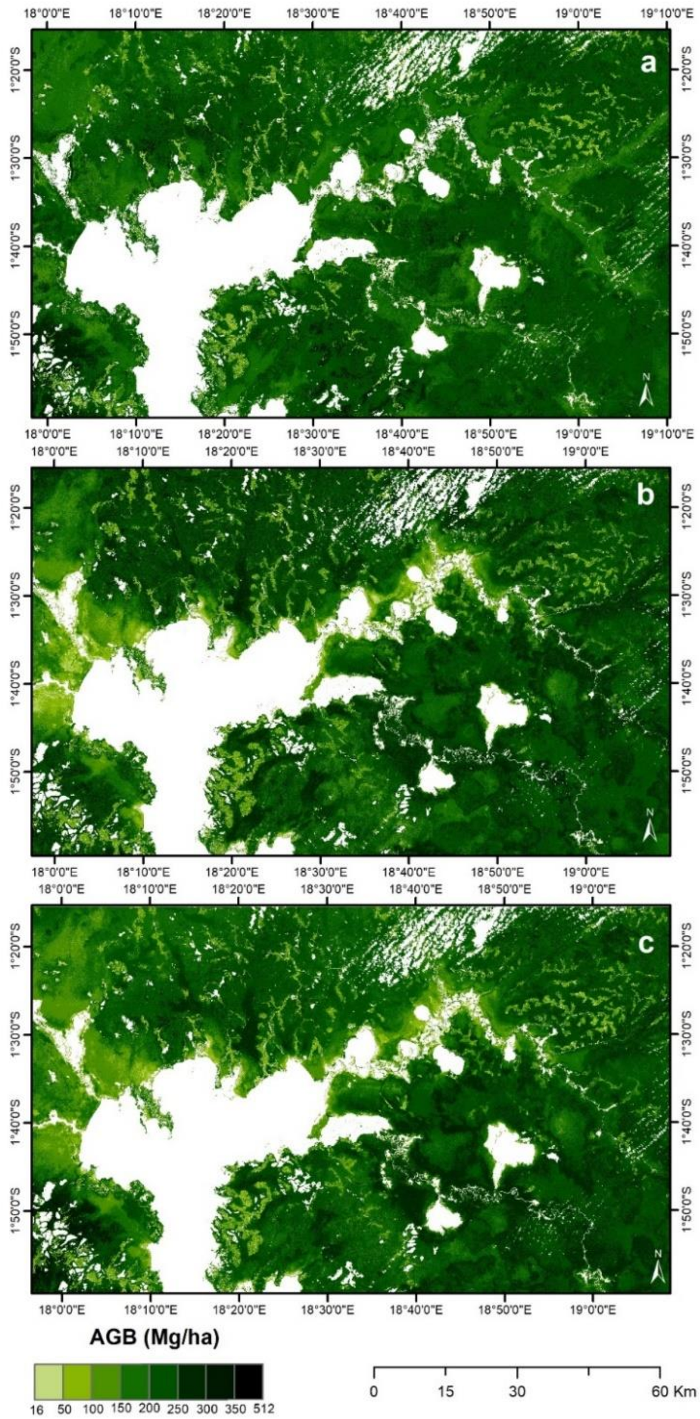

Figure 2.6 Estimated 30 m aboveground biomass (AGB) derived from the 30 m predicted dominant forest canopy height maps (Figure 2.4) generated using **(a)** the dry season July 14th 2013 Landsat-8 OLI image, **(b)** the wet season December 8th 2014 Landsat-8 OLI image, and **(c)** both images.

2.4.4 Aboveground Biomass Validation

Figure 2.7 shows scatterplots comparing the 30 m AGB derived from the remotely sensed data (Figure 2.6) and the 30 m area weighted AGB field estimates (AGB^{30}) over the four one-hectare field plots. The three scatterplots compare the same AGB^{30} with AGB predicted using forest canopy heights generated from the dry season (Figure 2.7a), wet season (Figure 2.7b), and both (Figure 2.7c) Landsat-8 OLI images. There were 63 AGB 25 m × 25 m parcel estimates but after the area weighting to 30 m (Equation 3) and application of the constraint that 50% of the 25 m × 25 m parcels with AGB estimates must fall under a 30 m pixel (Section 2.3.5), there were 43 pairs of values. The 43 plotted values are color coded to designate which of the four field plots the AGB^{30} were derived from. They illustrate that two of the field plots (purple and green) had higher AGB^{30} and that there was a wide range of values from about 96 Mg ha^{-1} to 503 Mg ha^{-1}. However, this range is smaller than present in the estimated 30 m AGB study area maps shown in Figure 6 that had AGB that varied from approximately 16 to 512 Mg ha^{-1}.

The OLS regressions of the plotted data are shown in red and were insignificant for the dry (Figure 2.7a) and wet (Figure 2.7b) season derived image results (p values > 0.05) with small R^2 values, of 0.05 and 0.07, respectively. Conversely, the OLS regression results for the AGB estimated using the dominant forest canopy heights derived from both Landsat images (Figure 2.7c) was more significant ($p = 0.03$) with a 0.11 R^2 value, and a slope closer to unity (0.135) and an intercept closer to zero (193 Mg ha^{-1}). The RMSE values for the wet and dry season results were 92.43 Mg ha^{-1} and 87.76 Mg ha^{-1}, respectively, and smaller, 83.77 Mg ha^{-1}, for the combined image results. The 83.77 Mg ha^{-1} RMSE value corresponds to about 41% of the mean study area mapped AGB (204 Mg ha^{-1}) (Figure 2.6c). However, clearly, the mapped AGB is over-estimated below about 225 Mg ha^{-1} and under-estimated above this value (Figure 2.7c).

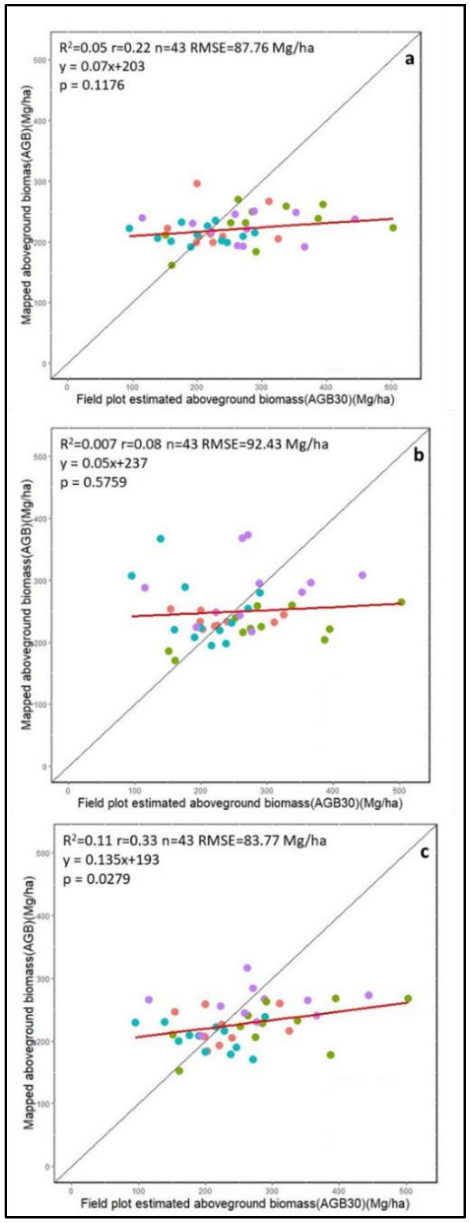

Figure 2.7 Scatterplots comparing the area weighted field plot above ground biomass (AGB^{30}) (Equation 3) with the corresponding 30 m aboveground biomass (Figure 6) derived from the 30 m predicted dominant forest canopy height maps generated using **(a)** the dry season July 14th 2013 Landsat-8 OLI image, **(b)** the wet season December 8th 2014 Landsat-8 OLI image, and **(c)** both images. The dots are color coded by which one-hectare field plot the AGB^{30} were derived from.

2.5 Discussion

The AGB of the Congo Basin Forest has been poorly documented due to a lack of inventory data and research (Baccini et al., 2008). Recently, airborne LiDAR and Landsat-8 OLI data have been used to map tree height and AGB across all of the Congo Basin (Xu et al., 2017). Cloud cover at the time of Landsat overpass can be high (e.g., Figure 2.2), reducing the ability to obtain cloud-free imagery needed to undertake the mapping. Consequently, Xu et al. (2017) defined Landsat predictor variables by the medians (i.e., 50th percentiles) of the red, NIR, and the two shortwave Landsat-8 OLI reflective wavelength bands acquired over three years. Thus, the predictor variables at adjacent pixel locations may have been selected from different years and seasons which is an issue if the forest and the Landsat Forest reflectance changed between years and seasons. However, the detailed processing and results of this study show, on average at the study area level, no great difference between the dry and wet season dominant canopy height and AGB results, i.e., little sensitivity to the seasonality of the Landsat imagery used. This is discussed below.

The dominant forest canopy heights and AGB estimated from the dry season Landsat-8 OLI image were on average marginally lower than those estimated from the wet season image. The RMSE between the mapped and 2639 independent test 30 m dominant canopy heights was 4.17 m (dry season) and 4.43 m (wet season) and corresponds to 20.2% and 21.2% of the mean study area mapped dominant canopy heights that were 20.6 m (dry season) and 20.8 m (wet season). The RMSE between the mapped and 43 independent field based 30 m AGB estimates was 87.76 Mg ha^{-1} (dry season) and 92.43 Mg ha^{-1} (wet season) and corresponds to 42.6% and 43.8% of the mean study area mapped AGB that was 206 Mg ha^{-1} (dry season) and 211 Mg ha^{-1} (wet season). There were seasonal geographic differences between the mapped dominant canopy height (Figure 4) and the AGB (Figure 2.6) results, in particular around Lake Mai Ndombe and in the vicinity of several of the rivers. The reasons for this are complex but may be due to seasonal vegetation condition and surface differences. For example, although the Landsat 30 m forest mask was derived conservatively, sub-pixel disturbed forest patches, and degraded forest areas with reduced live tree cover, may include shrubs and saplings that exhibit greater seasonal reflectance differences than elsewhere in the forest. In addition, in these regions, wet season flood water may have been observable at Landsat resolution through the forest canopy.

The pre-processing applied to the LiDAR and Landsat-8 OLI data were state of the practice. However, the difference between the seasonal results may have been affected by Landsat bi-directional reflectance distribution (BRDF) effects. Landsat BRDF variations are smaller than in wider field of view satellite optical wavelength data and occur due to changes in the view geometry across the image swath and to temporal changes in the solar geometry (Roy et al., 2016; Roy et al., 2020). The two Landsat-8 OLI images were not corrected for BRDF effects because they were both sensed from the same orbit and so have similar view geometry. The solar zenith angle at the center of each image was 35.85° and 32.30° for the dry and wet season images, respectively. This 3.55° solar zenith difference is small compared to the 15° Landsat field of view although variations of this magnitude can cause small reflectance variations (Roy et al., 2020). Use of both the dry and wet season Landsat-8 OLI images provided the lowest RMSE value (3.84 m) between the predicted and test dominant forest canopy heights, corresponding to 18.8% of the mean study area mapped tree height (20.4 m) derived using both images. This 18.8% RMSE is quite small and was not particularly expected as many of the study area forest Landsat pixels have high NDVI (~0.8) which is indicative of "saturated vegetation" reflectance conditions. Typically, the NDVI saturates with increased Leaf Area Index (LAI) above about LAI ~3.0 (Huete et al., 2002) and Congo basin tropical evergreen rainforest has LAI > 5 (de Wasseige et al., 2003; Favier et al., 2004). Presumably, despite this potential saturation issue, other factors related to the dominant forest canopy height could be discriminated by the random forest regression model using the wet and dry season Landsat-8 OLI images. Notably, the R^2 value was not particularly high (0.47) but the regression was significant ($p < 0.001$), indicating that the regression fit model illustrated in Fig. 5c is better than not having a model. Moreover, the R^2 is comparable to other recent study results, for example, Staben et al. (2018) reported a 0.49 R^2 between mapped and predicted forest canopy height in the Northern territory, Australia.

Using both Landsat-8 OLI images provided the most accurate AGB prediction and the lowest RMSE value (83.77 Mg ha^{-1}) between predicted and field estimated AGB. This was expected because the AGB was derived as an allometric function of the dominant canopy height which was most accurately predicted using both Landsat-8 OLI images. Although the R^2 was low (0.11) the regression was significant ($p < 0.03$), indicating that the regression fit model illustrated in Fig. 2.7c is better than not having a model. As noted earlier, the range of the field plot derived AGB (~96 Mg ha^{-1} to 503 Mg ha^{-1}) is smaller than the range of the study area estimated 30 m AGB (~16 to 512 Mg ha^{-1}), and the field plot data were available at only four 1 ha sites, which may reduce the representativeness of the validation results. However, the 83.77 Mg ha^{-1} AGB RMSE is comparable to the 89.83 Mg ha^{-1} RMSE value reported by Xu et al. (2017a) and corresponds to about ~40% of the reported mean Congo Forest AGB. This is quite a high error but is not surprising as we found that the dominant canopy heights were under-estimated for trees > ~25 m and over-estimated for trees < ~18 m (Figure 2.5). Similarly, the AGB was over-estimated below about 225 Mg ha^{-1} and under-estimated above this value (Figure 2.7).

The magnitude of the estimated mean study area AGB derived using both Landsat-8 OLI images (204 Mg ha^{-1}) is similar to that reported in other Congo Basin forest studies. For example, Baccini et al. (2008) reported 216 Mg ha^{-1} for the evergreen rainforest of Central Africa, Silva et al. (2018) reported 223 Mg ha^{-1} for Lope National Park in Central Gabon, and Xu et al. (2017a) reported 231 Mg ha^{-1} for all the DRC. Similarly, the magnitude of the maximum study area AGB derived using both Landsat-8 OLI images (511.8 Mg ha^{-1}) was comparable to the maximum Congo Basin forest AGB reported by Silva et al. (2018) and Xu et al. (2017). Thus, the maximum AGB that a hectare of Congo basin forest can store is about 500 Mg, which demonstrates the importance of these forests for carbon storage.

The predicted dominant canopy height results were derived using LiDAR transect training data acquired in July and August 2014, i.e., up to 12 and 13 months, respectively, after the dry season Landsat image acquisition, and up to 6 and 5 months, respectively, before the wet season Landsat image acquisition. The reliability of the dominant canopy height prediction will be reduced if the forest within the transects was disturbed in these periods. However, given the paucity of cloud-free satellite data we have no way to check this. The field plot data used to derive the AGB validation data were collected November 2015, i.e., 11 and 28 months after the wet and dry season Landsat images, respectively. The field plot data were collected in undisturbed primary forest and so are unlikely to have been subsequently disturbed but, again, we have not way to check this definitively. These issues underscore the difficulty in mapping and validating DRC dominant forest canopy height and AGB.

The degree to which AGB estimation can be improved using optical wavelength data is unknown. Further research to examine the effects of using additional satellite data, such as the Landsat-like Sentinel-2 data (Drusch et al., 2012), to see if the height estimation can be improved, is warranted. Using improved allometry (Skole et al., 2011) and recent spaceborne LiDAR data (Dubayah et al., 2020) is also warranted.

2.6 Conclusion

The sensitivity of airborne LiDAR and Landsat-8 OLI based dominant canopy height and AGB 30 m mapping was assessed with respect to the season of Landsat acquisition for a ~10,000 km^2 Congo Basin tropical forest study area. Experiments were undertaken independently three times to map and assess the 30 m dominant canopy height and AGB using (i) only a wet season Landsat-8 image, (ii) only a dry season Landsat-8 image, and (iii) both images. The images were predominantly cloud-free. A random forest regression estimator was used to predict and assess the 30 m dominant canopy height using LiDAR derived test and training data. The AGB was mapped using an allometric model parameterized with the dominant canopy height and was assessed by comparison with 43 field based 30 m AGB estimates.

The most accurate results were obtained using both the dry and wet season Landsat-8 OLI images together. The RMSE between the mapped and test 30 m dominant canopy heights was 3.84 m, and the RMSE between the mapped and field based AGB estimates was 83.77 Mg ha^{-1}. These RMSE values correspond to 18.8% of the mean study area mapped tree height (20.4 m) and to 41% of the mean study area mapped AGB (204 Mg ha^{-1}). The mean study area mapped AGB is similar to that reported in other Congo Basin Forest studies (Xu et al., 2017a; Baccini et al., 2008; Silva et al., 2018).

At the study area level there was little sensitivity to the seasonality of the Landsat imagery used. The study area mean dominant canopy height and AGB values were similar between seasons, within 0.19 m and 5 Mg ha^{-1}, respectively, and the RMSE between the mapped and test 30 m dominant canopy heights was 4.17 m (dry season) and 4.43 m (wet season), and the RMSE between the mapped and field based AGB estimates was 87.76 Mg ha^{-1} (dry season) and 92.43 Mg ha^{-1} (wet season). The degree to which AGB estimation can be improved using temporally richer optical wavelength data is unknown due to difficulties in obtaining cloud-free imagery. These results suggest that (i) using a single cloud-free Landsat-8 OLI image may be sufficient for airborne LiDAR and Landsat-8 OLI based dominant canopy height and AGB 30 m mapping in the Congo Basin tropical forest, but (ii) using Landsat imagery from different seasons is preferred to improve tropical forest inventories in the Congo Basin Forest.

REFERENCES

Asner, G.P. 2001. Cloud cover in Landsat observations of the Brazilian Amazon. *International Journal of Remote Sensing*. 22, 3855–3862.

Asner, G.P., Hughes, R.F., Mascaro, J., Uowolo, A.L., Knapp, D.E., Jacobson, J., Kennedy-Bowdoin, T., & Clark, J.K. 2011. High-resolution carbon mapping on the million-hectare Island of Hawaii. *Frontiers in Ecology and the Environment*. 9, 434–439.

Asner, G.P. & Mascaro, J. 2014. Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sensing of Environment*. 140, 614–624.

Baccini, A., Laporte, N., Goetz, S.J., Sun, M., & Dong, H. 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters*. 3, 45011–45019.

Bater, C.W. & Coops, N.C. 2009. Evaluating error associated with lidar-derived DEM interpolation. *Computers & Geoscience*. 35, 289–300.

Brando, P.M., Goetz, S.J., Baccini, A., Nepstad, D.C., Beck, P.S., & Christman, M.C. 2010. Seasonal and interannual variability of climate and vegetation indices across the Amazon. *Proceedings of the National Academy of Sciences of the United States of America*. 107, 14685–14690.

Breiman, L. 2001. Random Forests. *Machine Learning*. 45, 5–32.

Bwangoy, J.-R.B., Hansen, M.C., Potapov, P., Turubanova, S., & Lumbuenamo, R.S. 2012. Identifying nascent wetland forest conversion in the Democratic Republic of the Congo. *Wetland Ecology Management*. 21, 29–43.

Bwangoy, J.-R.B., Hansen, M.C., Roy, D.P., Grandi, G.D., & Justice, C.O. 2010. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. *Remote Sensing of Environment*. 114, 73–86.

Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., & Yamakura, T. 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, 145, 87-99.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B., & Vieilledent, G. 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology, 20, 3177-3190.

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

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

Csillik, O., Kumar, P., Mascaro, J., O'Shea, T., & Asner, G.P. 2019. Monitoring tropical forest carbon stocks and emissions using Planet satellite data. *Nature Scientific Reports*. 9, 17831.

de Wasseige, C., Bastin, D., Defourny, P., 2003. Seasonal variation of tropical forest LAI based on field measurements in Central African Republic. *Agricultural and Forest Meteorology*. 119, 181–194.

Disney, M.I., Boni Vicari, M., Burt, A., Calders, K., Lewis, S.L., Raumonen, P., & Wilkes, P. 2018. Weighing trees with lasers: Advances, challenges and opportunities. *Interface Focus*. 8, 20170048.

Dommo, A., Philippon, N., Vondou, D.A., Sèze, G., & Eastman, R. 2018. The June–September Low Cloud Cover in Western Central Africa: Mean Spatial Distribution and Diurnal Evolution and Associated Atmospheric Dynamics. *Journal of Climate*. 31, 9585-9603

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., & Bargellini, P. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of Environment*, 120, 25-36.

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., & Silva, C. 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*. 1, 100002.

Dwyer, J.L., Roy, D.P., Sauer, B., Jenkerson, C.B., & Zhang, H.K. 2018. Analysis Ready Data: Enabling Analysis of the Landsat Archive. *Remote Sensing*. 10, 19.

Egorov, A., Roy, D., Zhang, H., Hansen, & M., Kommareddy, A. 2018. Demonstration of Percent Tree Cover Mapping Using Landsat Analysis Ready Data (ARD) and Sensitivity with Respect to Landsat ARD Processing Level. *Remote Sensing*. 10, 209.

Egorov, A., Roy, D., Zhang, H., Li, Z., Yan, L., & Huang, H. 2019. Landsat 4, 5 and 7 (1982 to 2017) Analysis Ready Data (ARD) Observation Coverage over the Conterminous United States and Implications for Terrestrial Monitoring. *Remote Sensing*. 11, 447.

Einzmann, K., Haarpaintner, J., & Larsen, Y. 2012. Forest monitoring in Congo Basin with combined use of SAR C- & L-band. In Proceedings of the *IEEE International Geoscience and Remote Sensing Symposium*, Munich, Germany, 22–27 July 2012, pp. 6573–6576.

Favier, C., de Namur, C., & Dubois, M. 2004. Forest progression modes in littoral Congo, Central Atlantic Africa. *Journal of Biogeography*. 31, 1445–1461.

Ferraz, A., Saatchi, S., Mallet, C., & Meyer, V. 2016. Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment*. 183, 318–333.

Foody, G.M., Boyd, D.S., & Cutler, M.E.J. 2003.Predictive relations of tropical forest biomass from Landsat TM data and their transferability between regions. *Remote Sensing of Environment*. 85, 463–474.

Hansen, M.C., Egorov, A., Roy, D.P., Potapov, P., Ju, J., Turubanova, S., Kommareddy, I., & Loveland, T.R. 2011. Continuous fields of land cover for the conterminous United States using Landsat data: First results from the Web-Enabled Landsat Data (WELD) project. *Remote Sensing Letters*. 2, 279–288.

Hansen, M.C., Krylov, A., Tyukavina, A., Potapov, P.V., Turubanova, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., & Moore, R. 2016. Humid tropical forest disturbance alerts using Landsat data. *Environmental Research Letters*. 11, 034008

Houghton, R.A., 2005. Aboveground Forest Biomass and the Global Carbon Balance. *Global Change Biology*. 11, 945–958.

Hubau, W., Lewis, S. L., Phillips, O. L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., & Zemagho, L. 2020. Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature*. 579, 80-87.

Huete, A.R., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., & Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*. 83, 195–213.

Ju, J., Roy, D.P., Vermote, E., Masek, J., & Kovalskyy, V., 2012. Continental-scale validation of MODIS-based and LEDAPS Landsat ETM+ atmospheric correction methods. *Remote Sensing of Environment*. 122, 175–184.

Kovalskyy, V. & Roy, D.P., 2013. The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30m Landsat data product generation. *Remote Sensing of Environment*. 130, 280–293.

Leitold, V., Keller, M., Morton, D.C., Cook, B.D., & Shimabukuro, Y.E. 2015. Airborne lidar-based estimates of tropical forest structure in complex terrain: Opportunities and trade-offs for REDD+. *Carbon Balance Management*. 10, 1–12.

Lindquist, E.J., Hansen, M.C., Roy, D.P., & Justice, C.O. 2008. The suitability of decadal image data sets for mapping tropical forest cover change in the Democratic Republic of Congo: Implications for the global land survey. *International Journal of Remote Sensing*. 29, 7269–7275.

Lu, D. 2005. Aboveground biomass estimation using Landsat TM data in the Brazilian Amazon. *International Journal of Remote Sensing*.26, 2509–2525.

Manoli, G., Ivanov, V.Y., & Fatichi, S. 2018. Dry-Season Greening and Water Stress in Amazonia: The Role of Modeling Leaf Phenology. *Journal of Geophysical Research: Biogeosciences*. 123, 1909–1926.

Mascaro, J., Detto, M., Asner, G.P., Muller-Landau, H.C., 2011. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment*. 115, 3770–3774.

Matasci, G., Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Zald, H.S.J., 2018. Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots. *Remote Sensing of Environment*. 209, 90–106.

Mayaux, P., De Grandi, G., & Malingreau, J.P. 2000. Central African Forest Cover Revisited. *Remote Sensing of Environment*. 71, 183–196.

McGaughey, R.J. 2016. FUSION LDV: Software for LiDAR Data Analysis and Visualization, *U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station*, University of Washington, Seattle, WA,98115: pp. 1–212.

Meyer, V., Saatchi, S., Ferraz, A., Xu, L., Duque, A., Garcia, M., & Chave, J. 2019. Forest degradation and biomass loss along the Choco region of Colombia. *Carbon Balance Management*. 14, 2.

Miller, J. & Franklin, J. 2007. Aspinall, R. Incorporating spatial dependence in predictive vegetation models. *Ecological Modelling*. 202, 225–242.

Molinario, G., Hansen, M.C., & Potapov, P.V. 2015. Forest cover dynamics of shifting cultivation in the Democratic Republic of Congo: A remote sensing-based assessment for 2000–2010. *Environmental Research Letters*.10, 094009

Myneni, R. B., Yang, W., Nemani, R. R., Huete, A. R., Dickinson, R. E., Knyazikhin, Y., & Salomonson, V. V. (2007). Large seasonal swings in leaf area of Amazon rainforests. *Proceedings of the National Academy of Sciences*, 104, 4820-4823.

Qiu, S., Lin, Y., Shang, R., Zhang, J., Ma, L., & Zhu, Z. 2019. Making Landsat Time Series Consistent: Evaluating and Improving Landsat Analysis Ready Data. *Remote Sensing*. 11, 51.

Réjou-Méchain, M., Tymen, B., Blanc, L., Fauset, S., Feldpausch, T.R., Monteagudo, A., Phillips, O.L., Richard, H., & Chave, J. 2015. Using repeated small-footprint LiDAR acquisitions to infer spatial and temporal variations of a high-biomass Neotropical Forest. *Remote Sensing of Environment*. 169, 93–101.

Roy, D.P., Li, Z., Zhang, H.K., & Huang, H. A. 2020. conterminous United States analysis of the impact of Landsat 5 orbit drift on the temporal consistency of Landsat 5 Thematic Mapper data. *Remote Sensing of Environment*. 240, 111701.

Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C., & Zhu, Z. 2014. Landsat-8: Science and product vision for terrestrial global change research. *Remote sensing of Environment*, 145, 154-172.

Roy, D.P., Zhang, H.K., Ju, J., Gomez-Dans, J.L., Lewis, P.E., Schaaf, C.B., Sun, Q., Li, J., Huang, H., & Kovalskyy, V. A. 2016. General method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Remote Sensing of Environment*. 176, 255–271.

Saatchi, S., Marlier, M., Chazdon, R.L., Clark, D.B., & Russell, A.E. 2011. Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. *Remote Sensing of Environment*. 115, 2836–2849.

Samba, G., Nganga, D., &NMpounza, M., 2007. Rainfall and temperature variations over Congo-Brazzaville between 1950 and 1998. *Theoretical and Applied Climatology*. 91, 85–97.

Shapiro, A.C., Aguilar-Amuchastegui, N., Hostert, P., & Bastin, J.F. 2016. Using fragmentation to assess degradation of forest edges in Democratic Republic of Congo. *Carbon Balance Management*. 11, 11.

Silva, C. A., Saatchi, S., Garcia, M., Labriere, N., Klauberg, C., Ferraz, A., & Hudak, A. T. 2018. Comparison of small-and large-footprint lidar characterization of tropical forest aboveground structure and biomass: a case study from Central Gabon. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11, 3512-3526.

Skole, D.L., Samek, J.H., & Smalligan, M.J. 2011. Implications of allometry. *Proceedings of the National Academy of Sciences of the United States of America*.108, E12.

Souza, C.M. & Roberts, D.A. 2005. Monteiro, A.L. Multitemporal Analysis of Degraded Forest in the Southern Brazilian Amazon. *Earth Interactions*. 9, 1–25.

Staben, G., Lucieer, A., & Scarth, P. 2018. Modelling LiDAR derived tree canopy height from Landsat TM, ETM+ and OLI satellite imagery A machine learning approach. *International Journal of Applied Earth Observation and Geoinformation*. 73, 666–681.

Vermote, E., Justice, C., Claverie, M., & Franch, B. 2016. Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*. 185, 46–56.

World-Bank., 2016. Emission Reduction Program. *World Bank/Forest Carbon Partnership Facility*. 1-297.

Wulder, M.A., Coops, N.C., Roy, D.P., White, J.C., & Hermosilla, T. 2018. Land cover 2.0. *International Journal of Remote Sensing*. 39, 4254–4284.

Xu, L., Saatchi, S.S., Yang, Y., Yu, Y., & White, L. 2016. Performance of non-parametric algorithms for spatial mapping of tropical forest structure. *Carbon Balance Management*. 1, 18.

Xu, L., Saatchi, S. S., Shapiro, A., Meyer, V., Ferraz, A., Yang, Y., & Ebuta, D. 2017a. Spatial Distribution of Carbon Stored in Forests of the Democratic Republic of Congo. *Nature Scientific Reports*. 7, 15030.

Xu, L., Saatchi, S. S., Shapiro, A., Meyer, V., Ferraz, A., Yang, Y., & Ebuta, D. 2017b. Spatial Distribution of Carbon Stored in Forests of the Democratic Republic of Congo – Supplemental Information. *Nature Scientific Reports*. 7, 15030.

Yan, L. & Roy, D.P. 2015. Improved time series land cover classification by missing-observation-adaptive nonlinear dimensionality reduction. *Remote Sensing of Environment*. 158, 478–491.

Yang, J., Tian, H., Pan, S., Chen, G., Zhang, B., & Dangal, S., 2018. Amazon drought and forest response: Largely reduced forest photosynthesis but slightly increased canopy greenness during the extreme drought of 2015/2016. *Global Change Biology*. 24, 1919–1934.

Zhang, H.K. & Roy, D.P., 2017. Using the 500 m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification. *Remote Sensing of Environment*. 197, 15–34.

Zhuravleva, I., Turubanova, S., Potapov, P., Hansen, M., Tyukavina, A., Minnemeyer, S., Laporte, N., Goetz, S., Verbelen, F., & Thies, C. 2013. Satellite-based primary forest degradation assessment in the Democratic Republic of the Congo, 2000–2010. *Environmental Research Letters*. 8, 024034

CHAPTER 3. EXAMINATION OF THE AMOUNT OF GEDI DATA REQUIRED TO CHARACTERIZE CENTRAL AFRICA TROPICAL FOREST ABOVEGROUND BIOMASS AT REDD+ PROJECT SCALE IN MAI NDOMBE PROVINCE

This chapter was previously published by *Science of Remote Sensing*: Kashongwe, H.B., Roy, D.P., Skole, D.L. 2023. Examination of the amount of GEDI data required to characterize central Africa tropical forest aboveground biomass at REDD+ project scale in Mai Ndombe province. *Science of Remote Sensing*. 7, 100091.

3.1 Introduction

The Global Ecosystem Dynamics Investigation (GEDI) LiDAR onboard the International Space Station (ISS) is the first spaceborne LiDAR specifically designed for vegetation structure retrieval with a primary focus on improved forest aboveground biomass (AGB, units: Mg ha^{-1}) quantification in support of carbon cycle and climate studies (Dubayah et al., 2020). The amount of carbon in vegetation is constant, equivalent to about 47% of its dry weight of biomass (Ma et al., 2018), and so satellite AGB data provide a way to map the amount of carbon stored in vegetation biomass. The GEDI mission acquires waveform LiDAR measurements in 25m diameter footprints from 51.6° South to 51.6° North including the world's tropical forests. Tropical forests are thought to account for about half the world's terrestrial carbon and more than a third of the current terrestrial carbon sink (Hubau et al., 2020; Harris et al., 2021) with deforestation and forest degradation comprising ~20% of global annual net carbon emissions from human activities (Friedlingstein et al., 2020). In last several decades international forest carbon mitigation and monitoring frameworks, including the Reducing Emissions from Deforestation and Degradation and enhancement of carbon stocks program (REDD+), have been initiated that seek to compensate developing countries for reducing carbon emissions from forested lands. Financial support for REDD+ projects requires measured, reported, and verified (MRV) estimates of forest AGB to quantify avoided emissions achieved by conservation of forest biomass (UNFCCC, 2014). A 10% uncertainty for forest AGB estimates is recommended by the Intergovernmental Panel on Climate Change (IPCC) good practice guidance (IPCC, 2003). Robust and accurate forest carbon MRV systems are critical, and finance mechanisms are increasingly emphasizing a principle of "payments for performance" (Herold et al., 2019; Sandker et al., 2021). Currently, most REDD+ projects and national programs develop their own assessments of carbon stocks (also known as emission factors), which can vary considerably depending on how they allocate plot sampling, select measurement procedures, utilize allometric scaling models and equations, and elect to use remotely sensed data. With the recognition of the large number of samples required and the difficult task of field measurement in a forest inventory, particularly for developing countries and projects with limited human and financial resources, there is increasing interest in remote sensing-based AGB estimation and monitoring (Chave et al., 2019).

The first GEDI science data were released in 2020 and have been used to map forest canopy height including at near-global scale (Potapov et al., 2021; Lang et al., 2022). These initial studies used limited (less than six months) of Version 1.0 GEDI footprint canopy height data. Recently, the GEDI data were reprocessed with improved algorithms and sensor knowledge (Beck et al., 2021) and the Version 2.1 GEDI product suite was expanded to include a GEDI footprint-level AGB product (Duncanson et al., 2022). The good quality AGB footprint data generated over a 28 month were used to develop a near-global 1 km vegetation GEDI AGB map (Dubayah et al., 2022). The map could not be validated in a detailed spatially explicit manner but 169 country level AGB totals were compared to United Nations Food and Agriculture Organization (FAO) country estimates with a 0.86 reported coefficient of determination (Dubayah et al., 2022). Validation is challenging because the GEDI 25 m diameter footprints rarely coincide spatially with the locations of ground-based tree height or AGB measurements. Recently, information derived from simulated GEDI data derived from airborne laser scanning (ALS) data were used as independent reference data to validate the GEDI relative canopy height product (Li et al. 2023). However, in central Africa, ground-based AGB and coincident ALS data needed for GEDI AGB product validation are scarce and no validation has been undertaken (Duncanson et al., 2022).

This study examines the amount of GEDI data required to assess AGB in African tropical forests at REDD+ project scale. This is of great interest because although there is high confidence that GEDI can map AGB over large areas (Dubayah et al., 2022), less is known about how well its operational deployment performs for measurement of AGB to support REDD+ and other forest carbon MRV requirements. Notably, GEDI is a LiDAR sampling instrument with areal acquisition coverage built up over time. Each orbit senses 25 m diameter footprints in eight tracks spaced 600 m apart across swath, and 60 m apart along track, and the footprints may not be acquired in a spatially and temporally consistent manner due to ISS orbit changes and high frequency sensor orientation variations (Beck et al., 2021; Roy et al., 2021). In addition, a number of factors reduce the availability of GEDI AGB footprint retrievals, including clouds, ephemeral water, and sensitivity of the GEDI retrievals under high dense canopy cover or low AGB conditions (Kellner et al., 2021; Dubayah et al., 2022; Duncanson et al., 2022). The number of GEDI orbits, and so observation time period, needed to characterize AGB at REDD+ project scale is unknown. For the purposes of tractability, African tropical forest REDD+ projects are considered here as covering 50 × 50 km (250,000 ha) predominantly forested areas. This is based on a recent survey undertaken by the Center for International Forestry Research that reported the area of 34 African REDD+ projects (Atmadja, 2021) with a 286,876 ha mean project area that is approximately equivalent to a 50 × 50 km area. In the absence of no definitive spatially explicit AGB data available at fine spatial resolution, the mean AGB derived from a long time series of GEDI Version 2.1 AGB data were considered as “truth” for the reported analysis. The use of the mean AGB as the target parameter is consistent with measurement requirements for REDD+ projects, which use emission factors based on mean AGB per forest strata from inventories of sample plots that have similar dimensions as the GEDI footprints. Thus, the analysis reported in this study is self-consistent with respect to the Version 2.1 GEDI AGB footprint data. A total of 31 months of GEDI footprint-level AGB data acquired over tropical lowland forest in Mai Ndombe province, in the Democratic Republic of the Congo (DRC), were examined. Fifteen sites were selected in a stratified manner from five AGB stratum defined from the Version 2.1 AGB GEDI data, and with three sites selected per stratum that had low, medium and high semivariogram sill values that reflect increasing within-site AGB spatial variation. In this way the sites are expected to capture a range of forest AGB conditions. A 30m binary forest/non-forest map was derived from publicly available Landsat-derived products and only sites with >80% forest cover and good quality forest AGB footprints occurring across the site were considered. The expected minimum number of GEDI orbits required to characterize the site mean AGB to within ±5%, ±10%, and ±20% was derived by considering different combinations of GEDI orbits randomly selected from the 31 months of GEDI data. These values were converted into time periods by multiplying with a scalar coefficient that is equivalent to the average number of days required to obtain a GEDI orbit containing good quality forest AGB data at 50 × 50 km scale. Given the range in factors that affect the availability of good quality GEDI AGB data, scalar coefficients for more than 750 sites across Mai Ndombe province were considered. The results of this detailed study are illustrated and the implications for tropical forest AGB characterization are discussed.

3.2 Study Area and Data

3.2.1 Study Area

The Mai Ndombe province study area covers 123,000 km^2 of the DRC (Figure 3.1) of which about 98,000 km^2 is forested (World-bank, 2016). The study area falls close to the equator in the tropical monsoon climate zone with two wet seasons (March-May and September-December) and two dry seasons, 1800 mm mean annual rainfall, and 24 °C mean annual temperature (Samba et al., 2007; Bwangoy et al., 2010). The main forest characteristics are evergreen lowland forests with high tree crown cover (typically 70% or greater), 35–45 m mature tree heights, and heterogeneous shade tolerant tree species (Bwangoy et al., 2010; Fayolle et al., 2016; Mayaux et al., 2000). Fig. 1a shows a top of atmosphere (TOA) Landsat reflectance mosaic derived from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Landsat 8 Operational Land Imager (OLI) time series acquired in 2021 (Global Forest Change, 2021). The longer wavelength Landsat bands (1610, 865, 655 nm) are shown because the shorter wavelength green and blue Landsat bands are highly sensitive to atmospheric contamination (Roy et al., 2014). The forested areas are apparent in Fig. 1a (darker green tones) and are surrounded by successional secondary forest (lighter green tones). The interior and edges of the forest are disturbed (pink tones) due to human pressure associated with unregulated timber harvesting, charcoal production and slash-and-burn agriculture (predominantly cassava, maize, sorghum, peanuts, and upland rice) (Molinario et al., 2015; Shapiro et al., 2016). The study area includes the Mai Ndombe REDD+ project that is run by a US-based non governmental agency, Wildlife Works Carbon LLC, that has a 30-year contract (2011 to 2041) with the DRC government and receives annually carbon offset revenue through the voluntary carbon market (WWC 2012).

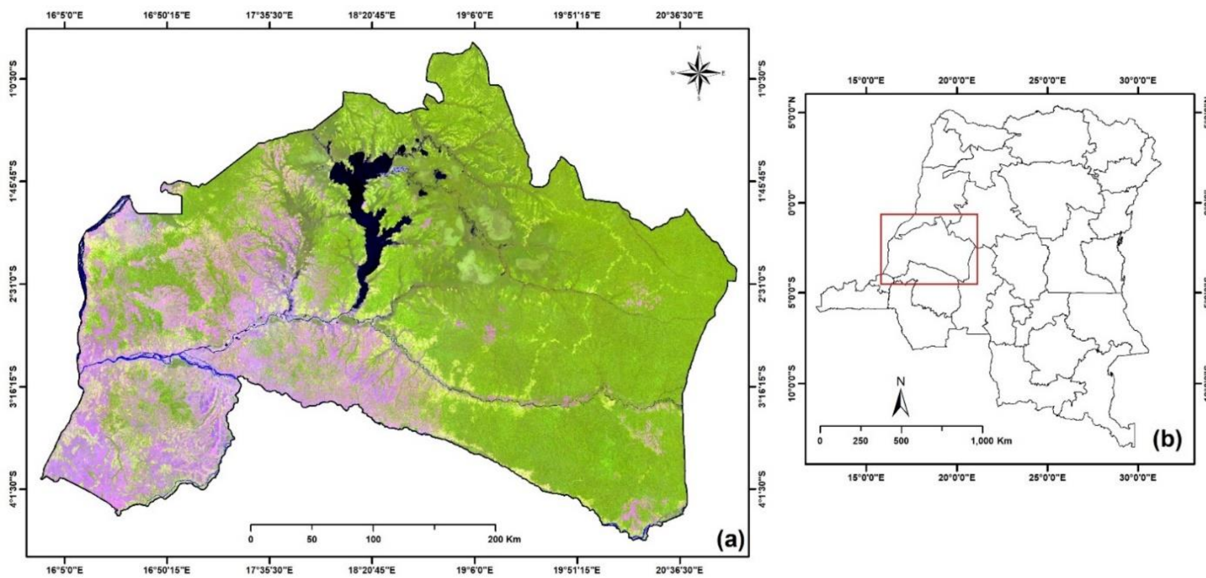

Figure 3.1 Mai Ndombe province study area **(a)** superimposed on median Landsat 30 m top of atmosphere reflectance (1610 nm, 865 nm, 655 nm) derived from Landsat 7 ETM+ and Landsat 8 OLI time series images acquired in 2021 (Global Forest Change, 2021), **(b)** location of Mai Ndombe province within the Democratic Republic of the Congo (DRC).

3.2.2 Data

3.2.2.1 Study area forest map

To undertake the research, a binary forest/non-forest study area map was derived from publicly available Landsat-derived Forest products (Hansen et al., 2013a; Global Forest Change, 2021). Previously we generated a 30 m forest map using year 2014 Landsat-8 OLI surface reflectance data (Kashongwe et al., 2020). However, as the GEDI data were first acquired in 2019, and given likely ongoing forest disturbance, a forest map was derived from the recently updated 2000-2021 global forest cover loss product and the most recent global 2010 percent tree cover product (Hansen et al., 2013a; Global Forest Change, 2021). These products are defined in geographic (latitude, longitude) coordinates with 0.00025° resolution (about 30 m at the equator) hereafter referred to for brevity as being defined with respect to 30 m pixels. The global 2010 percent tree cover product defines the maximum 30 m tree canopy cover (expressed as a percentage

0-100%) for year 2010 (Hansen et al., 2013a). The global forest cover loss product defines the first year in the period 2000 to 2021 when a 30 m pixel experienced significant canopy cover loss (Hansen et al., 2016). The global 2010 percent tree cover product has not been validated, and the global forest cover loss product was validated for a ten year period considering 1,500 locations with reported forest loss producer's and user's accuracies of 87.8% and 87.0%, respectively (Hansen et al., 2013b).

The binary forest/non-forest study area map was derived by thresholding the global 2010 percent tree cover product and then removing pixels where forest canopy loss was reported. A 30% percent tree cover threshold was applied to the 2010 percent tree cover product. This is consistent with the forest definition used by the DRC ministry of Environment and Sustainable Development in its Forest Reference Emission Level reporting, that was approved by the United Nations Framework Convention on Climate Change (DRC FREL, 2018). This canopy cover criterion is often used in national REDD+ programs and projects along with criteria for minimum area and tree height by which forest is defined in the humid tropics as "a minimum area of land of 0.5 hectare with tree canopy cover >30% with trees with the potential to reach a minimum height of 2-5 m at maturity *in situ*" (Kim et al., 2015; Keenan et al., 2015). Note that we do not use the 0.5 hectare criterion because GEDI footprints are an order of magnitude smaller (0.0491 hectare) and are sensed 60 m and 600 m apart along and across track. In addition, we do not use the minimum 2-5 m tree height at maturity criterion because GEDI height data are expected to be less reliable for low tree heights (Li et al. 2023), although this has not been quantified in tropical evergreen lowland forests, and we implicitly assume that all the trees in the selected forested sites would meet this criterion. After thresholding, all 30 m pixel locations where forest loss was detected in the global 2000-2021 forest cover loss product were removed. Thus, the forest study area map represents the intact forest cover mapped at 30 m resolution up to 2021. The forest map is shown in Figure 3.2 and illustrates that Mai Ndombe province is dominated by forest, with 81% of the 30 m pixels classified as forest.

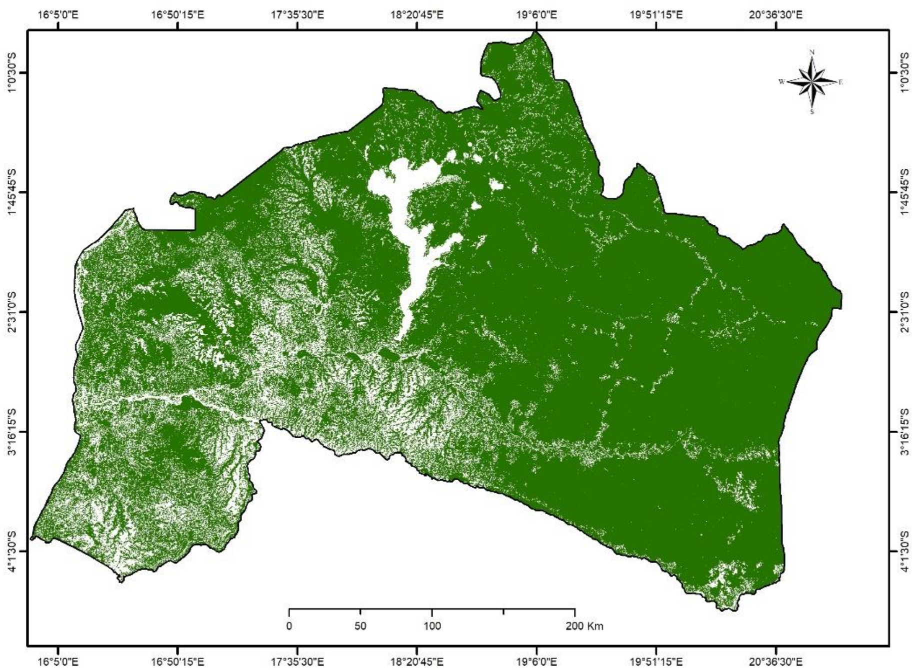

Figure 3.2 Mai Ndombe province study area 30 m forest map (forest shown dark-green, white non-forest) derived from publicly available global 2010 tree cover and 2000-2021 forest cover loss products (Global Forest Change, 2021). The black line shows the province boundary.

3.2.2.2 GEDI footprint-level aboveground biomass data

The GEDI instrument is onboard the International Space Station (ISS) and collects day and night time 1064 nm laser waveforms in 25 m diameter footprints (Dubayah et al., 2020). GEDI is composed of three lasers that sense eight ground tracks spaced 600 m apart across track and 60 m apart along track. One of the laser beams is split into two weaker beams (termed coverage beams), the two others remain at full power (termed power beams), and all are optically dithered to provide the eight ground tracks. The location of each GEDI 25 m footprint center is provided as a latitude and longitude coordinate defined in the WGS84 datum. The ISS is in a low earth orbit with a nominal 51.6° inclination and overpasses the surface at varying times of day with an approximately 3 to 4 day equatorial revisit depending on the ISS altitude (Gebelein and Eppler, 2006; Dubayah et al., 2022). Unlike repeat coverage imagers, such as Landsat where spatial and temporal coverage is limited chiefly by cloud cover (Kovalskyy and Roy, 2013), the GEDI instrument irregularly acquires surface observations in space and time, and the sampling density of the AGB footprint data are dependent on a range of factors, including cloud, that are discussed below.

The recently released GEDI Level 4A Version 2.1 footprint-level aboveground biomass (AGB) product available from the Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center (DAAC) (https://daac.ornl.gov) was used in this study. The AGB (units: Mg ha^{-1}) is derived by applying an allometric model to relative canopy height estimates (units: meters) retrieved at each GEDI footprint waveform (Duncanson et al., 2022). Seven AGB estimates are provided for each 25 m footprint as there are seven sets of relative canopy height retrievals, each generated using different algorithm setting groups to control the GEDI waveform interpretation (Beck et al., 2021). The recommended optimal AGB estimate is denoted in the GEDI Level 4A Version 2.1 footprint-level AGB product; in this study only the optimal AGB values were used. The allometric models were pre-defined with respect to continent and plant functional type by statistical comparison of AGB ground-based estimates with GEDI waveform relative heights simulated from airborne laser scanner (ALS) data (Duncanson et al., 2022). Specifically, the GEDI AGB data available over the Mai Ndombe province study area were derived using an African Evergreen Broadleaf Tree (EBT) allometric model applied to the GEDI 50th and 98th relative canopy height retrievals (i.e., the relative height to the ground below which 50% and 98% of the GEDI cumulative waveform energy was returned). The GEDI EBT model was not validated by comparison with independent AGB data due to the scarcity of ground-based AGB and coincident

ALS data. The EBT model had a 66.9% reported RMSE fit uncertainty, with no reported bias term, and the GEDI footprint-level AGB product is expected to have a comparable or greater error in African equatorial evergreen broadleaf forests (Duncanson et al., 2022).

The GEDI footprint-level AGB product has not yet been validated. As stated in the introduction, validation of GEDI data by comparison with ground-based data is challenging because GEDI footprints rarely coincide with ground measurements. This is illustrated in Figure 3.3 that shows the GEDI footprint center locations (black dots) of 31 months of GEDI good quality AGB data and the center locations of all REDD+ forest inventory field plots (red dots) where inventory data were collected in 30 m diameter plots over a five year period that included the GEDI acquisition period. The locations are shown superimposed on Landsat 30 m TOA reflectance for (a) 20 × 20 km (left) and (b) 1.1 × 1.8 km areas. For illustrative clarity, Figure 3(b) shows only the GEDI footprint center locations falling within 120 m of the forest inventory field plot center locations. The Figure 3b results are typical and illustrate that even when GEDI footprints are close to forest inventory field plot locations, the vegetation structure and so AGB may not be similar as indicated by the different 30 m pixel Landsat reflectance values at the plotted locations.

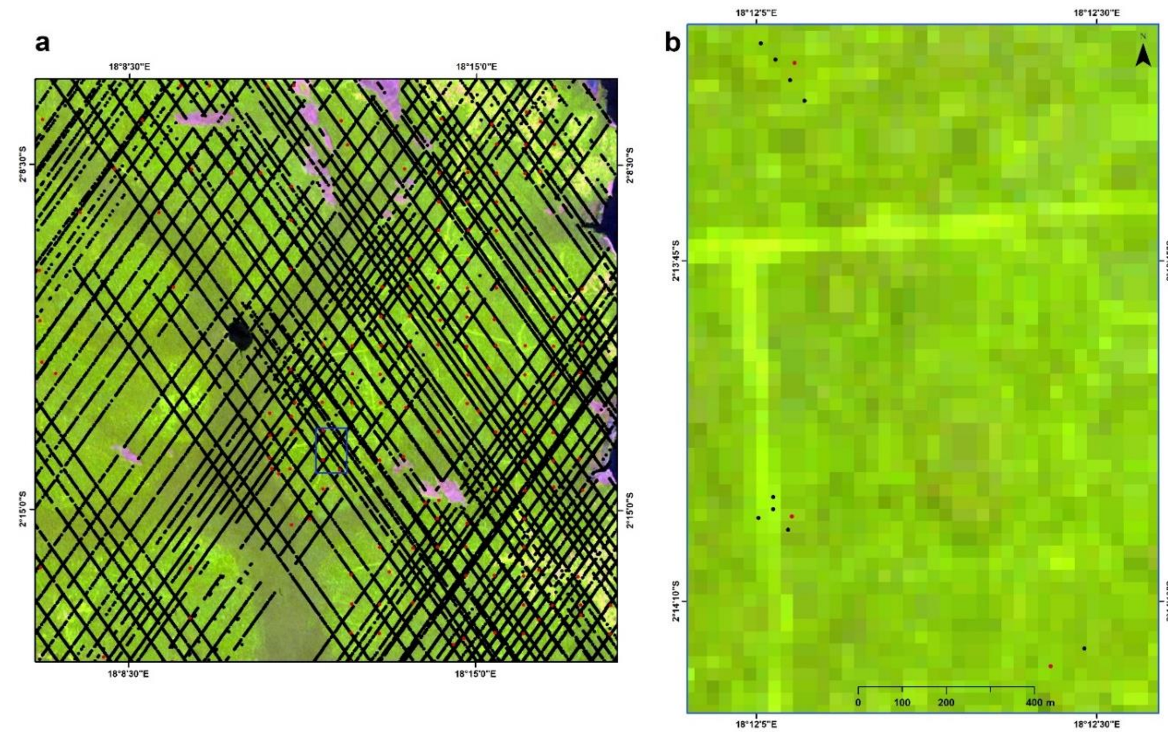

Figure 3.3 Illustration of the typical lack of coincidence between GEDI 25 m diameter footprints (center locations shown by black dots) and 30 m diameter forest inventory field plots (center locations shown by red dots) for **(a)** 20 × 20 km and **(b)** 1.1 × 1.8 km areas in the study area. The locations of all the available good quality GEDI AGB footprint data acquired over 31 months (18th April 2019 to 13th November 2021) are illustrated. The forest inventory field plots are where ground-based data were collected in the period 2017 and 2021 by Wildlife Works Carbon LLC REDD+ staff. The points are superimposed on year 2021 30 m Landsat TOA reflectance (see Figure 3.1(a) caption).

In this study all the GEDI Level 4A Version 2.1 optimal AGB footprint day and night data available over the study area at the time of writing were used. Data for more than 14.9 million GEDI 25m footprints sensed by 426 GEDI orbits acquired from 18th April 2019 to 13th November 2021 (31 months) were available. The data were first quality filtered to retain only the AGB footprint estimates labelled as good quality (L4_quality_flag = 1) as illustrated in Figure 3.3. The data were then filtered so that only GEDI footprint center coordinates that fell within a 30 m study area forest map pixel (Figure 2) were retained, which resulted in 3.8 million footprints and 381 orbits. Figure 3.4 illustrates the spatial sampling of the 31 months of good quality forest GEDI footprint AGB data for a 50 × 50 km forested site in the central eastern part of the province. Figure 4a shows the locations of the good quality forest GEDI footprints and Figs. 4b and 4c show, for geographic context, Landsat 30 m TOA reflectance and the forest map, respectively. Three different GEDI orbits are colored for illustrative purposes in Figure 4a. The eight GEDI tracks per orbit and the ascending and descending ISS orbit tracks are evident. Spatially variable surface sampling is evident with missing footprints within orbits, and areas where there are fewer orbit tracks. The GEDI footprints rarely spatially overlap between different orbits, which is related to a number of factors including variations in the ISS orbit geometry and GEDI footprint geolocation errors (Roy et al., 2021). The 31 months of good quality forest GEDI footprint AGB data illustrated in Figure 4 were acquired by 64 unique orbits. Figure 5 shows their temporal availability and there were either no, one, or at most, two GEDI orbits per day with good quality forest AGB footprint data.

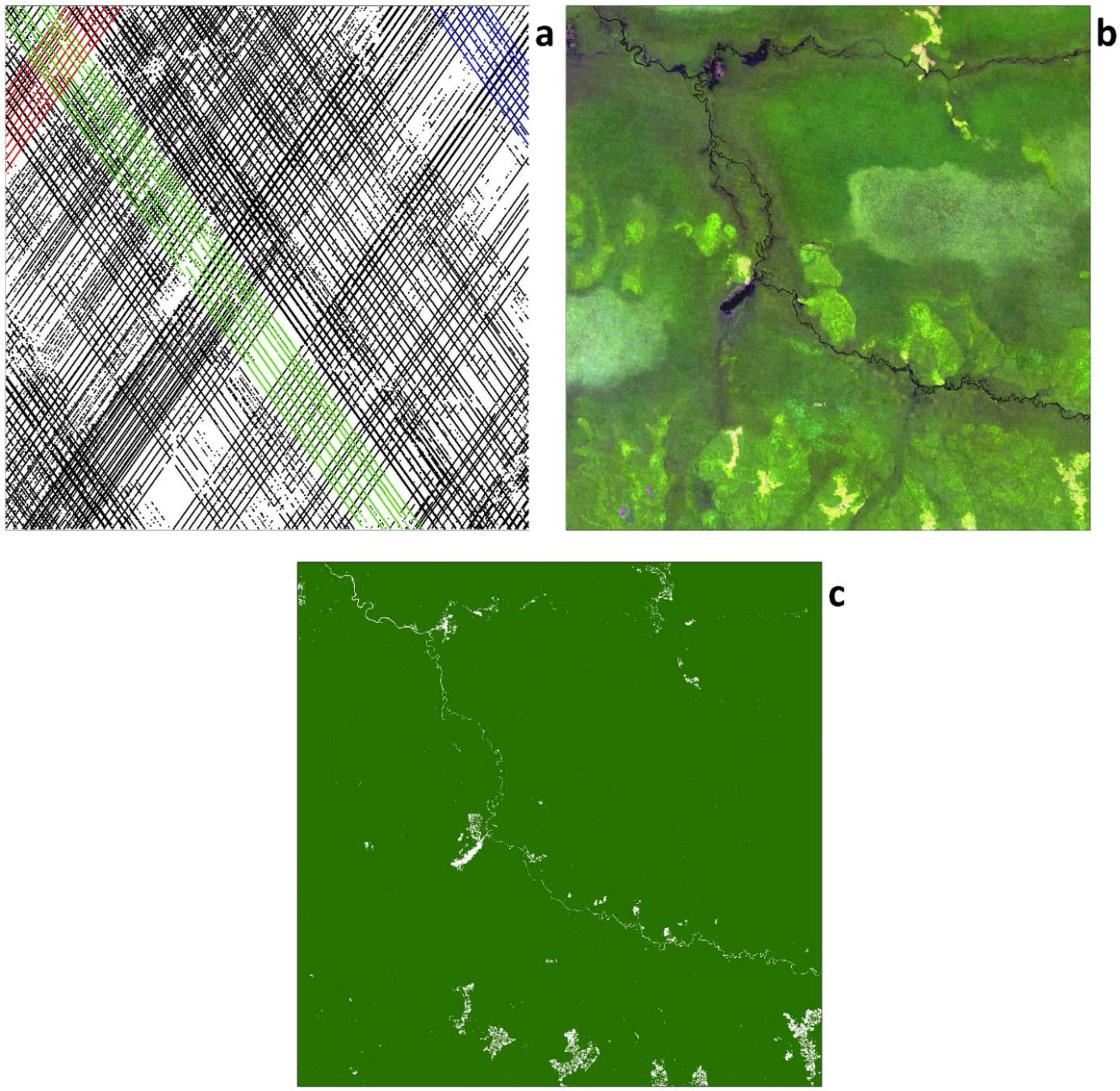

Figure 3.4 Detailed illustration for a 50 × 50 km area of (a) the center locations of each good quality AGB forest GEDI footprint available day and night over the 31 months of study data (a total of 142,277 good quality AGB forest GEDI footprints), colors show three example orbits, (b) median Landsat 30 m top of atmosphere reflectance (1610 nm, 865 nm, 655 nm) derived from Landsat 7 ETM+ and Landsat 8 OLI time series images acquired in 2021 (Global Forest Change, 2021), (c) the 30 m forest study area map (forest shown dark-green, white non-forest) subset from Figure 1. The center of the illustrated area is at 19.385°E, 2.415°S (site #1, Figure 3.7b).

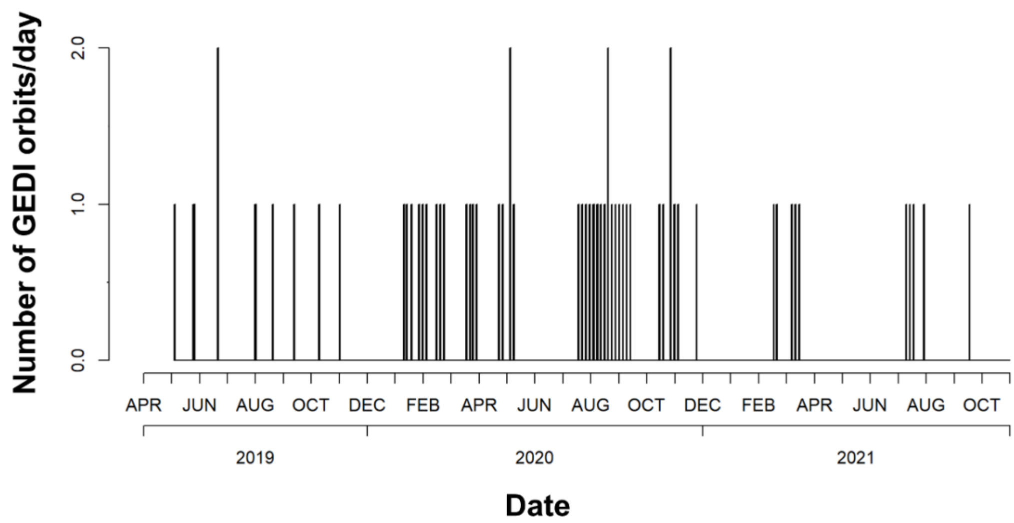

Figure 3.5 Daily number of GEDI day and night orbits with good quality AGB forest footprint data over the 31 months of study data over the 50 × 50 km area shown in Figure 4. The first and last orbits with good quality data were acquired for these example data on 18th April 2019 and 13th November 2021, respectively.

The spatial and temporal absence of GEDI AGB data evident in Figures 3.4 and 3.5 was due to several factors. First, Mai Ndombe province is often cloudy (Dommo et al., 2018) and the GEDI 1064 nm lasers do not penetrate clouds (Dubayah et al., 2020). Second, the GEDI lasers are typically absorbed by water (Beck et al., 2021) which is apparent in Figure 3a with no good quality GEDI observations over the water bodies evident in the Landsat 30 m TOA reflectance, and in addition, Mai Ndombe province forests can be seasonally flooded (Bwangoy et al., 2013). Other factors may also have prevented in undocumented ways the processing of the GEDI AGB data. For example, GEDI relative canopy heights are less reliable and cannot be retrieved when only a small proportion of the canopy is penetrated to the ground by the waveform energy (Hofton et al., 2019; Beck et al., 2021). Conversely, the GEDI AGB data may not be reliably defined for low AGB conditions, typically outside of the range of the forest ground-based AGB data used to derive the EBT model (Duncanson et al., 2022) and this is also evident in Figure 3a where there are no good quality AGB data GEDI over deforested areas (pink tones). In addition, intermittent low quality ISS and GEDI specific exterior orientation (attitude and position) measurements, combined with ISS orbit maneuvers and structural vibrations, can preclude reliable GEDI geolocation and good quality AGB retrieval (Klein and Axelrad, 2020; Roy et al., 2021).

3.3 Methods

3.3.1 Overview

Fifteen 50 × 50 km forested sites were selected that captured a representative range of GEDI AGB across the province and included sites with different within-site forest AGB spatial variation. The sites are expected to be representative of central African tropical forest AGB at REDD+ project scale. The site selection (Section 3.3.2) was complicated because only forested sites with good quality forest GEDI AGB footprint values distributed across them could be used to capture the spatial variation in forest AGB. Sites with more heterogenous AGB are expected to require more GEDI data to characterize. The sites were selected in a stratified manner from five AGB stratum defined from the GEDI data, and with three sites selected per stratum that had low, medium and high semivariogram sill values that reflect increasing within-site AGB spatial variation. For each site, the expected minimum number of GEDI orbits (n_{orbits}^{p}) required to characterize the site mean AGB to within p=±5%, ±10%, and ±20% was derived by considering different combinations of GEDI orbits randomly selected from the 31 months of GEDI data (Section 3.3.1). These three p values were considered as they bound the 10% Intergovernmental Panel on Climate Change's (IPCC) recommended accuracy requirement for designing REDD+ Measured, Reported and Verified (MRV) forest AGB sampling (IPCC 2003). The expected minimum number of days (n_{days}^{p}) required to characterize the AGB over each site was derived by multiplying the site n_{orbits}^{p} values with a scalar coefficient. The scalar coefficient was derived statistically considering the 31 months of GEDI data and is equivalent to the average number of days required to obtain a GEDI orbit containing good quality forest AGB data at 50 × 50 km scale (Section 3.3.2). As the availability of good quality GEDI AGB data is subject to the different factors described in the previous section, the scalar coefficients for a range of sites across Mai Ndombe province were considered.

3.3.2 Site Selection

Fifteen sites with 50 × 50 km area were selected across Mai Ndombe province. They were not selected randomly because although REDD+ sites are predominantly forested, randomly selected sites across the province may not be (Figure 3.2), and because GEDI footprints with good quality forest AGB may be unevenly spatially distributed (Figure 3.4a). Therefore, the sites were selected in a stratified manner. To ensure that sites were selected with good quality forest AGB GEDI values occurring across them, the site was partitioned into $n \times n$ grid cells, and only sites with good quality forest AGB GEDI values falling in each grid cell were selected. The site selection analysis was undertaken in the geographic (latitude, longitude) projection i.e., in the same projection used by the global forest cover loss and percent tree cover products and by our derived forest map. The GEDI footprint data were associated with the 30m forest map pixels by projection of the footprint center point coordinates into the forest map, and consequently without resampling errors that occur when raster data are reprojected (Roy et al., 2016).

The number of good quality forest AGB footprint values in different sized grid cells cast systematically across the province was examined to determine an appropriate grid cell dimension for the site selection. As expected, the number per grid cell increased with the grid cell size. Figure 3.6 shows a map of the number of good quality forest AGB GEDI values with respect to 0.05° grid cells. Grid cells with smaller side dimensions often had no GEDI good quality forest AGB footprint data over the 31 study months. Across the study area the number of good quality forest AGB footprint values in each 0.05° grid cell varied from zero, typically for cells with no or few 30 m forest pixels, to a maximum of 3229, with a mean of 877 and a median of 777. As this median value corresponds to a relatively high spatial sampling i.e., equivalent to 25 GEDI footprints/km^2, a 0.05° grid cell dimension was used for the site selection. In addition, over Mai Ndombe province 9 × 9 0.05° grid cells cover 50.06 × 50.06 km (calculated specifically for the 2°S province center latitude, Figure 1) which is close to 50 × 50 km.

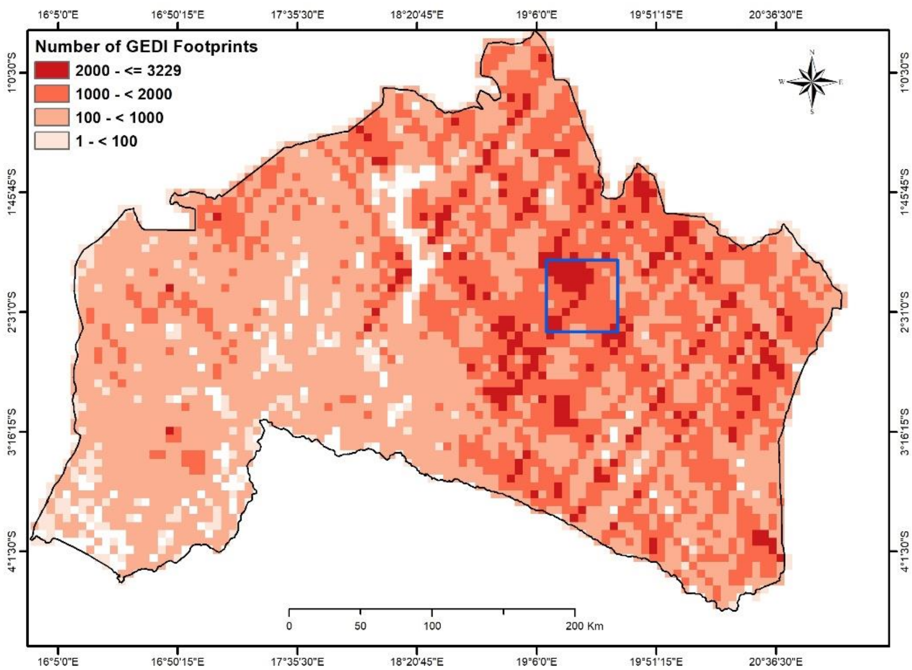

Figure 3.6 Number of good quality AGB forest GEDI footprint values in 0.05° × 0.05° grid cells acquired over 31 months (18th April 2019 to 13th November 2021) across the Mai Ndombe province study area. The blue box shows the 50 × 50 km boundary of the data illustrated in Figures 3.4 and 3.5 (the box center is at 19.385°E, 2.415°S).

Square areas defined by 9 × 9 0.05° grid cells cast systematically across the province (translating the region in 0.05° grid cell increments) were examined. Only areas that were >80% forest (as defined by the 30 m forest map, Figure 1) and that had ≥100 GEDI good quality AGB forest footprint values in each 0.05° grid cell were retained for the remainder of the site selection analysis. A histogram of the average good quality AGB forest footprint values in each of the 9 × 9 0.05° grid cell areas was used to derive four AGB thresholds to partition the province into five AGB forest strata. Three sites were selected from each AGB stratum to capture a range of within-strata forest AGB spatial variation. This was undertaken by examination of the empirical semivariogram (Cressie, 1993) of the good quality AGB forest GEDI footprint values over all the different 9 × 9 0.05° grid cell areas that fell completely within each stratum. Previously we used semivariogram analysis to assess forest height spatial variation to contextualize the impact of geolocation uncertainty on GEDI forest canopy height retrievals in Mai Ndombe province (Roy et al., 2021). Following a similar approach, the empirical semivariogram was derived in four compass directions, buffering each direction with a 22.5° tolerance and using a 23.55 km cutoff (equivalent to one-third of the diagonal length of a 50 × 50 km site) to ensure that sufficient paired GEDI good quality forest 25 m footprint AGB values samples were considered. A spherical model was used to derive in each direction the semivariogram sill (the degree of semi-variance when the points are uncorrelated) and range (the distance at which points are effectively uncorrelated, i.e., have no spatial dependence) (Barnes, 1991; Cressie, 1993; Curran and Atkinson, 1998). Only sites where the semivariogram sill values could be derived in all four directions were considered. The sill and not the range values were used because the former quantifies the spatial variability across the 50 × 50 km area and the latter determines the distance beyond which points across the area are not correlated. In each AGB stratum, three sites that had representative low, medium and high mean directional sill values were selected. Care was taken to ensure that the 15 selected sites did not spatially overlap. Each site had typically >100,000 good quality forest GEDI AGB footprint values acquired over the 31 study months. Notably, for all 15 sites, >3% of the 30 m forest pixels had good quality AGB footprint values falling within them over the 31 months.

3.3.3 Estimation of The Amount of GEDI Data Required To Characterize Redd+ Project Scale AGB

3.3.3.1 Expected minimum number of GEDI orbits required to characterize REDD+ project scale AGB

The expected minimum number of GEDI orbits required to characterize the site mean AGB was derived for each of the 15 selected sites. The overall mean GEDI AGB (OMGA, units Mg ha^{-1}) derived from all the good quality forest GEDI footprint AGB values in each site over 31 months was assumed to be representative of the actual mean site AGB. This is a reasonable assumption given that: (i) the site selection criteria ensured that only forested sites with GEDI observations distributed across them were used, and (ii) for all 15 sites, >3% of the 30 m forest pixels had good quality AGB footprint values falling within them over the 31 months which provides a greater sample than used by conventional field based sampling of tree structural attributes used to derive AGB. The expected minimum number of GEDI orbits (n_{orbits}^{p}) required to characterize the site mean AGB to within $\pm p$= 5%, 10%, and 20% was derived for each site. These three p values were considered as they bound the 10% Intergovernmental Panel on Climate Change's (IPCC) recommended accuracy requirement for designing REDD+ Measured, Reported and Verified (MRV) forest AGB sampling (IPCC, 2003).

The OMGA at each site was sequentially compared with the sample mean AGB derived using different numbers of randomly selected GEDI orbits, each containing only good quality AGB forest footprint values. A comparison of the OMGA with the sample mean AGBs derived in order of date of acquisition was not undertaken because the spatial coverage of individual orbits varies considerably (e.g., the three colored orbits in Figure 4a). Rather, the OMGA was compared with the mean AGB derived considering *different* combinations of $k = 1, 2, 3 \ldots n$ GEDI orbits, where n is the total number of GEDI orbits with good quality forest GEDI AGB data for the 31 study months at the site. The total number of unique combinations that k different orbits can be selected without consideration of their order and without replacement from n orbits is:

$$C(n,k) = \frac{n!}{(n-k)!k!} \qquad (1)$$

In this study the 15 selected sites had from 53 to 65 orbits containing good quality AGB footprint values. Thus, the total number of unique orbit combinations can be very high. For example, considering a study site with $n = 53$ there is 1 unique combination of 53 orbits $C(n = 53, k = 53)$, 53 unique combinations of 1 orbit $C(n = 53, k = 1)$, 1378 unique combinations of 2 orbits $C(n = 53, k = 2)$, 23426 unique combinations of 3 orbits $C(n = 53, k = 3)$, and about 973×10^{12} combinations of 27 orbits $C(n = 53, k = 27)$. Therefore, the OMGA was compared with the mean AGB derived considering no more than 700 different combinations of randomly selected orbits for each $k = 1, 2, 3 \ldots n$. A maximum of 700 combinations was sufficient to ensure that each orbit was likely to be included. For example, the probability of an orbit not being considered in a particular combination is greatest when $k = 1$ and the probability of a single orbit not being considered in 700 combinations is $\left(\frac{n-1}{n}\right)^{700}$ which is small, i.e., < 0.001 for a site with 53 orbits, and < 0.01 for a site with 65 orbits.

Figure 3.7 shows results for the 50 × 50 km site data illustrated in Figs. 4 and 5 that had $n = 64$ GEDI orbits with good quality forest GEDI AGB values over the 31 study months. The orange horizontal line shows the OMGA derived considering all 64 orbits and is 117.4 Mg ha^{-1} which is at the lower end of African tropical forest AGB estimated without GEDI data (Baccini et al., 2008; Lewis et al., 2013). For a given k (shown on the x-axis) each plotted dot corresponds to the mean

AGB derived from $C(n = 64, k)$ (up to a maximum of 700) different combinations of k orbits. Thus, each dot on the left hand side of the plot (k=1) is the mean AGB derived from just one orbit over the 50 × 50 km site, and there are 64 dots, and the considerable variability in the k=1 mean AGB values is due to the different locations of single GEDI orbits and their forest footprints across the site (this is evident in Figure 3.4). As expected, as combinations composed of a larger number of different orbits are considered, the mean AGB values converge to the OMGA.

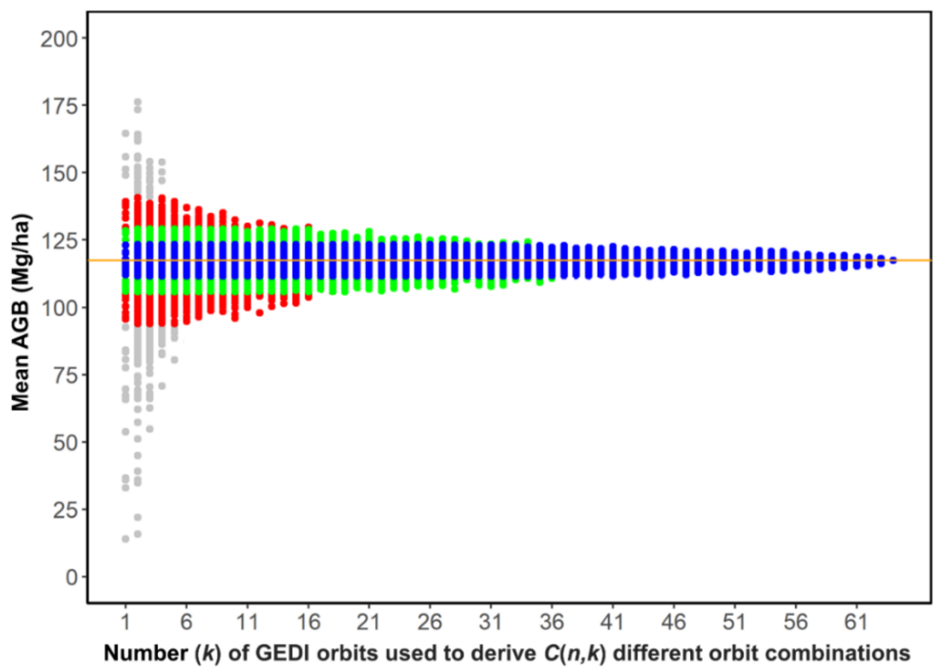

Figure 3.7 Mean AGB values (dots) derived considering $C(n = 64, k)$ (maximum 700) different combinations of k GEDI orbits. Only orbits with good quality forest GEDI AGB data over the 31 month study period are considered. Results for the 50 × 50 km site illustrated in Figure 3.4 and 3.5. The orange line shows the overall mean GEDI AGB (OMGA) i.e., the mean of all the good quality forest GEDI footprints derived from all 64 orbits containing good quality forest GEDI AGB data available over the site. The individual mean AGB values (dots) are colored as red, green, and blue if they fall within ±20%, ±10% and ±5% of the OMGA, respectively, otherwise colored grey.

The expected minimum number of GEDI orbits (n^p_{orbits}) required to characterize the site mean AGB was defined by examination of the proportion of the mean AGB values (Figure 7 dots) falling within ±p% of the OMGA for the different combinations of k = 1, 2, 3 … n GEDI orbits, where n is the total number of GEDI orbits with good quality forest GEDI AGB data for the 31 study months. As evident in Figure 7 the proportion will increase with greater k. The first k value when *all* the mean AGB estimates fall within ±p% of the OMGA was used to define the n^{p}_{orbits} value. For example, in Figure 7 the mean AGB values within ±5% of the OMGA are colored blue and all values fall within ±5% of the OMGA when $k \geq 37$ and so, for this example n^{5}_{orbits}= 37, i.e., at least 37 GEDI orbits are required to characterize the site mean AGB to within ±5%.

3.3.4 Expected minimum number of days and years required to characterize REDD+ project scale AGB

The expected minimum number of days required to characterize the site mean AGB was derived as:

$$n^{p}_{days} = s \times n^{p}_{orbits} \qquad (2)$$

where n^{p}_{days} is the expected minimum number of days required to characterize the site mean AGB to within ±p%, s is a scalar coefficient that is equivalent to the average number of days required to obtain a GEDI orbit containing good quality forest AGB data at 50 × 50 km scale, and n^{p}_{orbits} is the expected minimum number of GEDI orbits required to characterize the site mean AGB to within ±p% (Section 3.3.1).

The temporal availability of GEDI orbits containing good quality forest GEDI footprint AGB data is not uniform (e.g., Fig 5) due to factors, summarized at the end of Section 2.2, including cloud cover, ephemeral surface water presence, GEDI AGB retrieval sensitivity to forest conditions, and the GEDI orbit. The scalar coefficient s was derived taking this into account. Specifically, the temporal intervals between successive GEDI orbits containing footprint-level good quality forest AGB data needed to obtain m orbits were counted for each site. This was undertaken by searching through the site GEDI orbit time series in chronological order and counting successive days through the time series until *m* GEDI orbits with good quality forest GEDI footprint AGB values were acquired. This was undertaken many times, first counting from the first day containing a good quality AGB footprint value, and then counting again but starting from the second day, and then starting from the third day, etc., until all the days in the 31 month site time series were considered. Counts for *m* values 2, 3, 4, … *n* – 5 were derived for each site where *n* is the total number of GEDI orbits containing good quality forest AGB footprint data over the 31 months. Thus, there were always ≥5 sets of counted days per site and *m* value. For each site an ordinary least squares regression (OLS) between the counts and *m* was derived. The OLS regression was fixed through the origin, as the number of days required to obtain zero orbits should be zero, and the regression slope term was used to define the scalar coefficient *s*.

As the temporal availability of GEDI orbits with good quality AGB footprint data may vary across Mai Ndombe province, the conversion coefficient *s* was derived as above for the 15 sites and for all the other possible 50 × 50 km sites with >80% forest cover and ≥100 GEDI good quality AGB forest footprint values per 0.05° grid cell. A histogram of the 778 resulting site *s* values was examined and the minimum, mean, and maximum *s* values were used, as Equation 2, to derive the equivalent n_{days}^{p} values for the 15 sites. The expected minimum number of years (n_{years}^{p}) was also defined by dividing the n_{days}^{p} values by 365.

3.4 Results

3.4.1 Site Selection

Figure 8 shows the study area AGB (Mg ha^{-1}) derived from the 31 months of good quality AGB forest GEDI footprint data, with respect to (a) 0.05° grid cells, and (b) the 9 × 9 0.05° grid cell areas that satisfied the site selection criteria (i.e., >80% forest covered and all 81 grid cells had

≥100 GEDI good quality AGB values). Across the province lower mean 0.05° grid cell AGB values (minimum 1.2 Mg ha^{-1}) were over the sparsely forested regions in the west and greater values were over the forested regions (maximum 414.9 Mg ha^{-1}) (see Figure 3.2). There are fewer grid cells with mean AGB results in Figure 8(b) compared to Figure 3.8(a) because only those 9 × 9 0.05° grid cell areas that met the site selection criteria are shown. The minimum, mean, median, and maximum mean AGB of the Figure 3.8(b) results was 89.7, 155.5, 165.2, and 197.1 Mg ha^{-1}, respectively, which are lower than reported in other Congo Basin forest studies that did not use GEDI data. For example, Baccini et al. (2008) reported a mean and maximum AGB of 216 Mg ha^{-1} and 454 Mg ha^{-1} for the evergreen rainforest of Central Africa, and Xu et al. (2017) reported a mean and maximum AGB of 231 Mg ha^{-1} and 500 Mg ha^{-1} for all of the Democratic Republic of Congo.

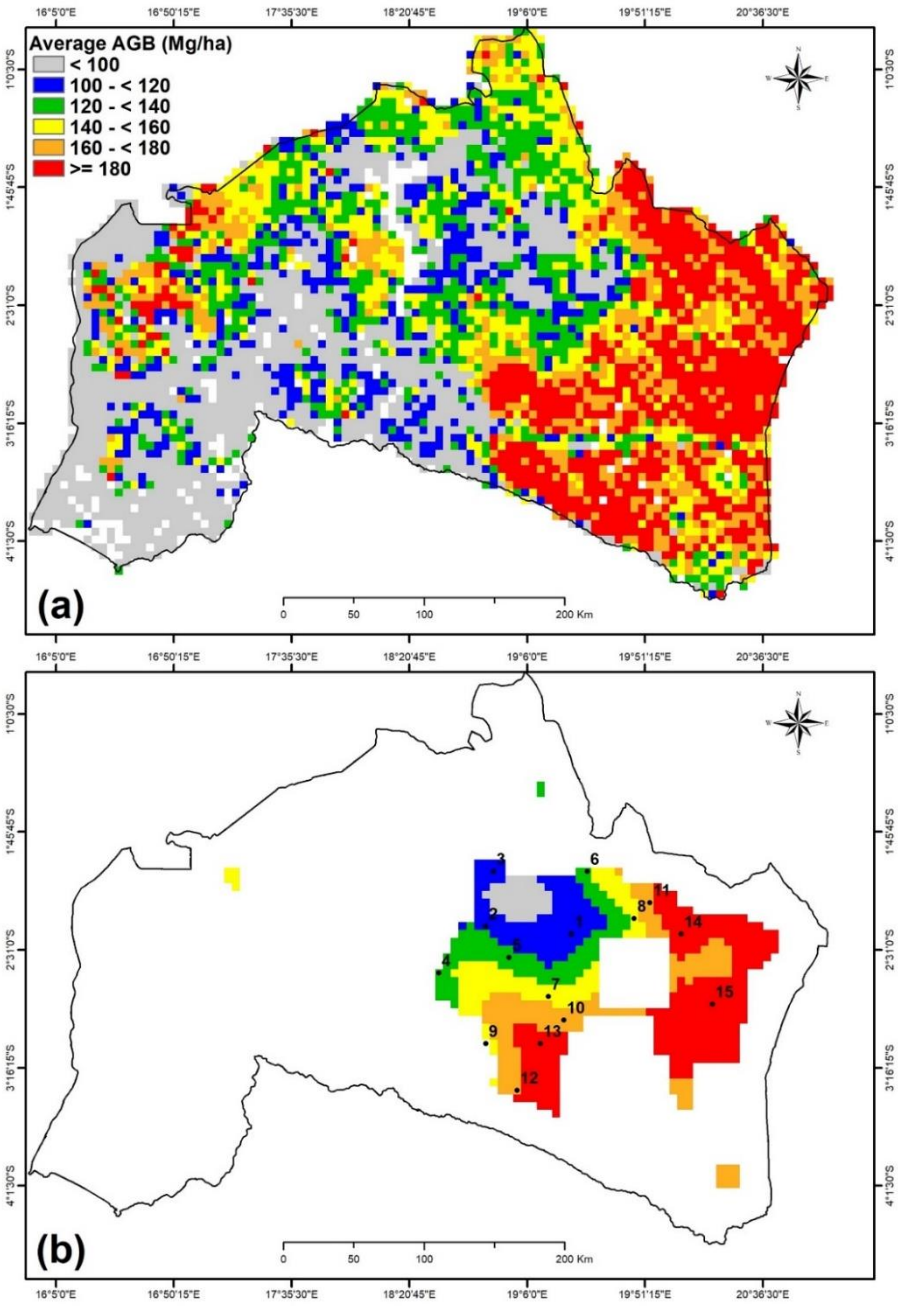

Figure 3.8 Mai Ndombe province mean AGB derived from the good quality forest GEDI AGB observations acquired over 31 months derived over **(a)** 0.05° grid cells (white areas had no good quality GEDI AGB values), **(b)** 9 × 9 0.05° grid cells areas that satisfied the site selection criteria, i.e., have >80% forest cover (as defined by the 30 m forest map) and where all 81 0.05° grid cells had ≥100 GEDI good quality forest AGB values. The colors show the five AGB strata. The black dots show the center locations of the 15 selected sites.

Figure 3.9 shows a histogram of the mean AGB values illustrated in Figure 8b. This histogram was used to select four evenly spaced mean AGB thresholds that were used to define the five AGB strata shown by the colors in Figure 3.8b. Given the fairly uniform frequency of AGB values <180 Mg ha^{-1} the thresholds were spaced every 20 Mg ha^{-1} with the highest threshold at 180 Mg ha^{-1}.

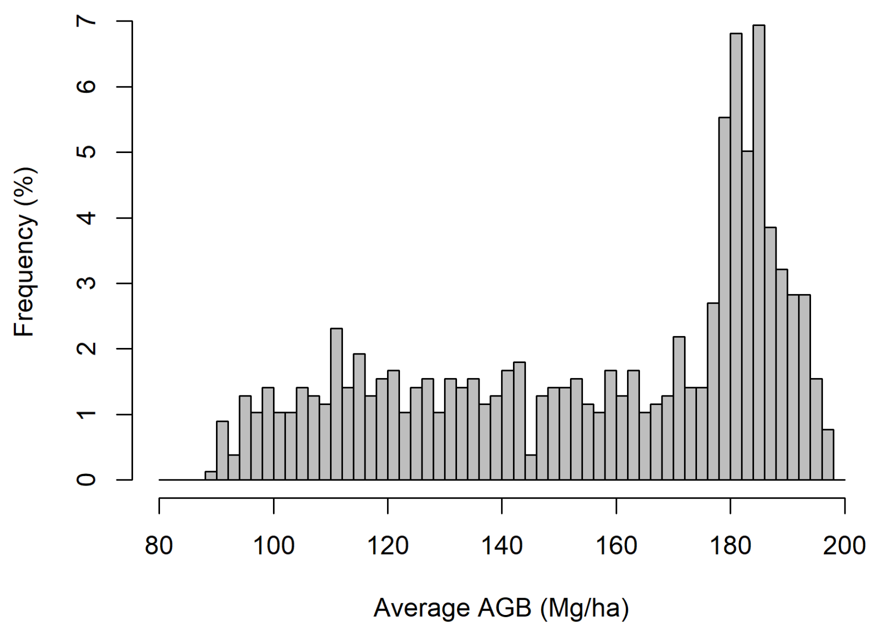

Figure 3.9 Histogram of the mean AGB values illustrated in Figure 3.8b, shown with 2 Mg ha^{-1} bin widths.

From each of the five AGB stratum, three 50 × 50 km sites were selected by the semivariogram sill analysis. The 15 selected sites are referenced by a number and their center locations are illustrated by black dots in Figure 8b. Table 1 summarizes key GEDI, forest cover, and mean directional sill value information for each site. The site OMGA values within each stratum were similar with <7.5 Mg ha^{-1} difference, and ranged among the sites from 117.39 Mg ha^{-1} (site #1) to 189.56 Mg ha^{-1} (site #15). The 15 sites had from 53 to 65 orbits containing good quality AGB footprint values, and the site forest cover varied from 81.9% (site #9) to 98.7% (site

#11). For all sites, >3% of the 30 m forest pixels had GEDI footprint center coordinates that fell within them, which may appear to be a low sample rate but compared to most field-based inventories is rather high and given that the sites are stratified by mean AGB is typically much higher than field based inventories.

Table 3.1 Summary site information listed grouped by the five AGB strata that the 15 sites were selected from (Figure 3.8b). The OMGA is the overall mean GEDI AGB derived from all the good quality forest GEDI footprint AGB values over 31 months at the site and is assumed to be representative of the actual mean site AGB. The mean directional sill value is the mean of the four sill values derived independently in four directions from the good quality forest GEDI footprint AGB values (Figure 3.10).

Site #	AGB stratum range (Mg ha^{-1}) that the site falls within	OMGA (Mg ha^{-1})	Total Number of GEDI orbits with good quality forest AGB values over 31 months (*n*)	Percentage (%) of site 30 m pixels that are forest	Percentage (%) of 30 m forest pixels with GEDI good quality AGB footprint values	Mean directional sill value
1	100 - < 120	117.39	64	98.6	4.28	4112.2
2		118.35	61	98.5	4.04	4311.1
3		116.54	64	95.7	3.91	4545.4
4	120 - < 140	130.54	53	95.2	3.3	4834.5
5		131.95	61	96.8	4.02	5103.8
6		133.41	65	96.2	3.73	5489.3
7	140 - < 160	151.90	59	97.3	3.75	6022.9
8		150.80	55	99.4	3.19	6282.2
9		151.84	54	81.9	3.87	7761.2
10	160 - < 180	172.01	55	97.3	3.58	7296.3
11		178.36	55	98.7	3.03	8345.2
12		179.19	53	88	3.33	8923.5
13	≥ 180	184.49	54	94.6	4.03	7552.5
14		184.29	59	96.8	3.36	8478.2
15		189.56	55	98.4	3.27	9058.9

Recall that sites with relatively low, medium, and high mean directional sill values were selected from each AGB stratum (shown in Table 1 last column). The mean directional sill values quantify the spatial variability across the 50 × 50 km site area. For example, Figure 3.10 illustrates the four direction empirical semivariograms for site #1 (for context the site #1 GEDI and Landsat data are illustrated in Figure 3.3). The sill and range values are indicated by the horizontal and vertical lines, respectively, and are colored by direction. The minimum directional sill value was in the north-east to south-west direction (135° orange) that is predominantly undisturbed dense forest (darker green tones in Figure 3.4b) and so has less spatial variation. The maximum directional sill value occurred in the north-south direction (90° green) that includes deforested areas in the north and south and areas of lower tree cover and likely degraded/disturbed forest cover in the south (lighter green tones in Figure 3.4b).

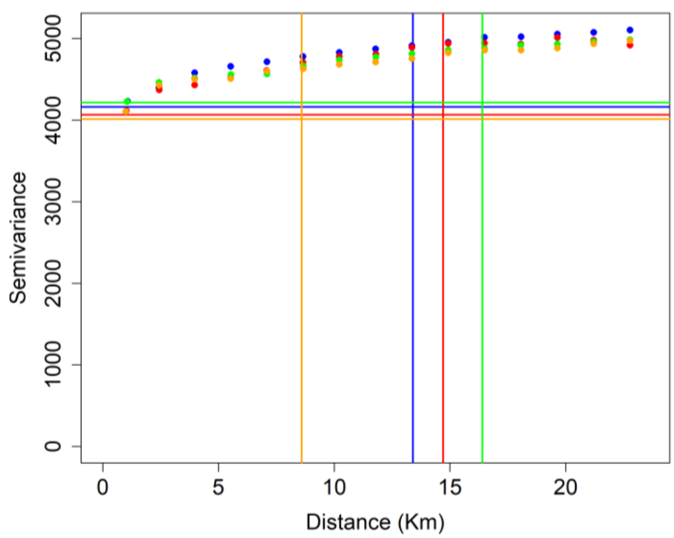

Figure 3.10 Directional empirical variograms of the 31 months of GEDI footprint-level good quality forest AGB data acquired over site #1 (illustrated in Figure 3.4-3.6, center location shown in Figure 3.8b). The colors show the directional results where 0° (blue) is east-west, 45° (red) is north-west to south-east, 90° (green) is north-south, and 135° (orange) is north-east to south-west direction. The directional sill values are shown by the horizontal lines, the directional range values are shown by the vertical lines.

Semivariogram sill values are derived from the squared differences between pairs of values within a given distance apart (Figure 3.10 x-axis) and are not normalized by the data mean (Barnes, 1991). Consequently, the sill values in Table 1 can only be meaningfully compared among the three sites in the same AGB stratum.

3.4.2 The Amount of GEDI Data Required to Characterize REDD+ Project Scale AGB

3.4.2.1 Expected minimum number of GEDI orbits required to characterize REDD+ project scale AGB

Figures 3.11 and 3.12 show the results of the analysis to derive the expected minimum number of GEDI orbits (n_{orbits}^{p}) needed to characterize the site forest AGB to within ±*p*% of the site overall mean GEDI AGB (OMGA). Figure 3.11 shows the results for sites #1-9, and Figure 3.12 shows the results for sites #10-15. Each dot illustrates the proportion of the mean AGB values falling within ±*p*% of the OMGA considering *k* different combinations of randomly selected GEDI orbits. Results for *k* = 1, 2, 3 … *n* orbits are illustrated where *n* is the number of site GEDI orbits with good quality forest GEDI AGB data over the 31 month study period. The results are colored for *p* = 20% (red), *p* = 10% (green), and *p* = 5% (blue). The proportion of the mean AGB values falling within ±*p*% of the OMGA generally increases with *k* (as discussed in Section 3.3.1 with respect to Figure 7). The n_{orbits}^{20}, n_{orbits}^{10} and n_{orbits}^{5} values are shown by the red, green and blue colored numbers respectively (also summarized in Table 2) and occur when the proportion of the mean AGB values falling within ±20%, ±10%, and ±5% of the site OMGA first reach 100%. As expected, more orbits are required to characterize the site mean forest AGB more precisely i.e., to within smaller ±*p*% values. Specifically, among the 15 sites, n_{orbits}^{5} varied from 30 – 57 orbits, n_{orbits}^{10} varied from 11 – 41 orbits, and n_{orbits}^{20} varied from 5 – 17 orbits.

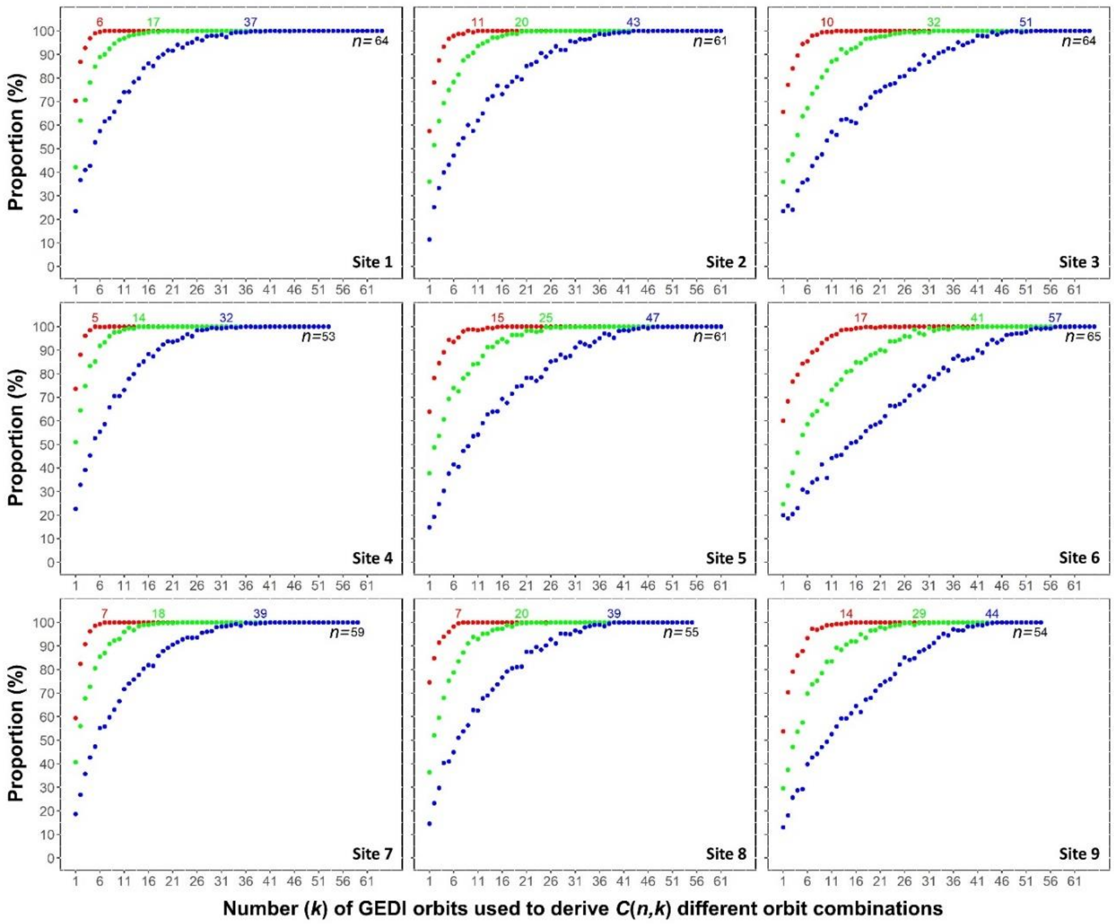

Figure 3.11 The proportion of the mean AGB values falling within ±20% (red dots), ±10% (green dots), and ±5% (blue dots) of the site OMGA derived considering $C(n = 64, k)$ different (up to a maximum of 700) combinations of randomly selected GEDI orbits. Only orbits with good quality forest GEDI AGB data over the 31 month study period are considered. The n^{20}_{orbits}, n^{10}_{orbits} and n^{5}_{orbits} values are shown by the red, green and blue colored numbers, respectively. Results for the lower mean AGB stratum sites #1 - #9 (site center locations illustrated in Figure 3.8b).

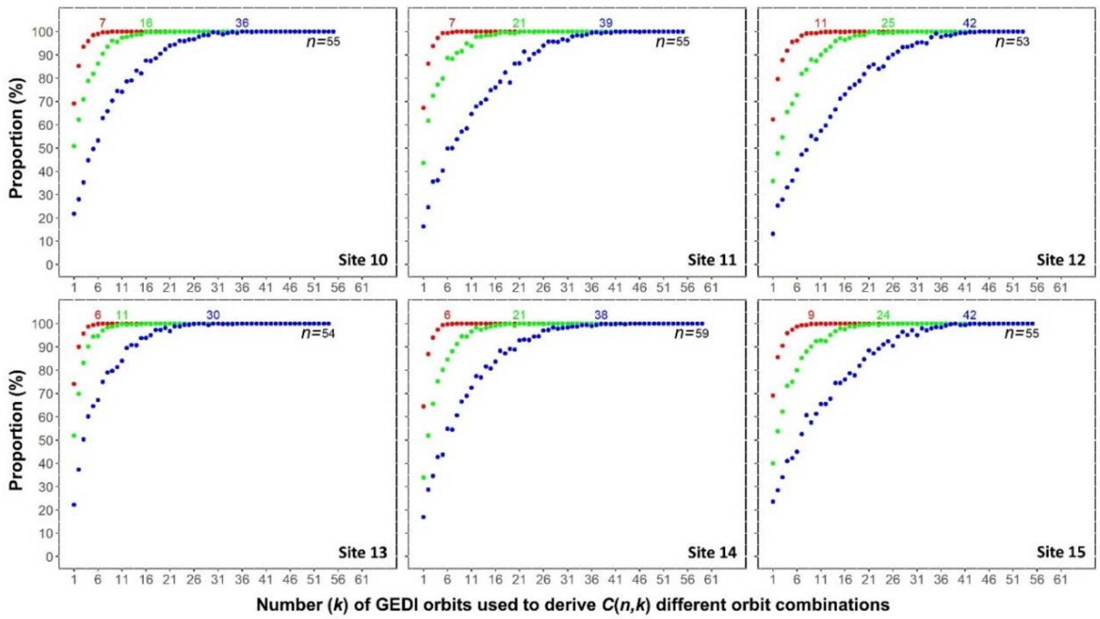

Figure 3.12 As Figure 3.11 but for the higher mean AGB stratum sites #10-15.

Table 3.2 The expected minimum number of GEDI orbits (n^{p}_{orbits}) required to characterize the site forest AGB to within ±p% of the site overall mean GEDI AGB (OMGA) (also shown in Figure 3.11 and 3.12). Site results grouped by the five AGB strata that the sites were selected from (Table 3.1).

site #	p = ±5%	p = ±10%	p = ±20%
1	37	17	6
2	43	20	11
3	51	32	10
4	32	14	5
5	47	25	15
6	57	41	17
7	39	18	7
8	39	20	7
9	44	29	14
10	36	16	7
11	39	21	7
12	42	25	11
13	30	11	6
14	38	21	6
15	42	24	9

3.4.2.2 Expected minimum number of days and years required to characterize REDD+ project scale AGB

Figure 3.13 shows a histogram of the scalar coefficient values derived considering 778 50 × 50 km sites across Mai Ndombe province. Recall that the scalar coefficient is equivalent to the average number of days required to obtain a GEDI orbit containing good quality forest AGB data at 50 × 50 km scale. Sites with more orbits containing quality forest GEDI footprint AGB values and sites with a more regular temporal occurrence of orbits tended to have smaller coefficient values. The minimum and maximum values were 9.09 and 21.50 days. The histogram is fairly symmetrical with similar mean and median values of 13.03 days and 13.19 days, respectively. Thus, 13.03 days is required on average to obtain one GEDI orbit containing good quality forest AGB data over a 50 × 50 km site in the Mai Ndombe province study area, but this period could be as low as 9.09 days and as high as 21.50 days.

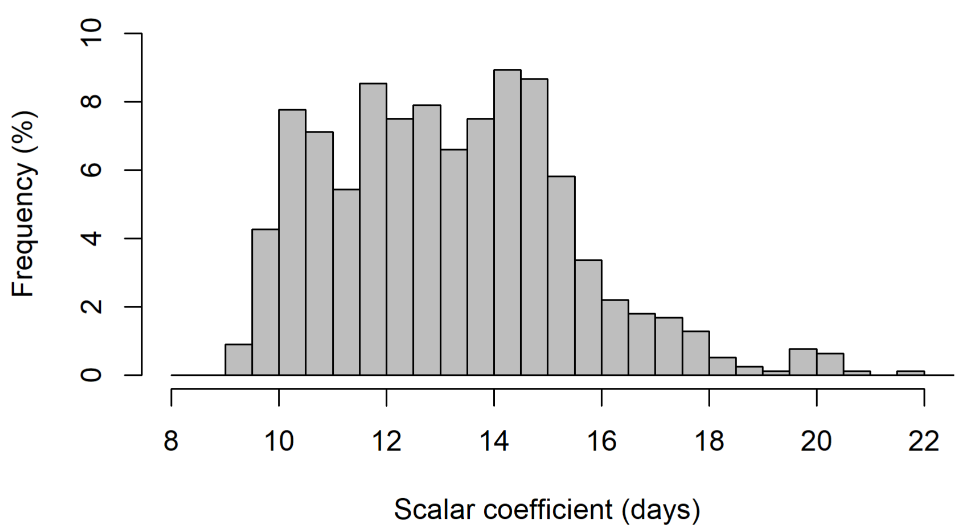

Figure 3.13 Histogram of the scalar coefficient values derived (Section 3.3.2) considering the 31 months of good quality forest GEDI AGB data falling in 778 50 × 50 km sites across Mai Ndombe province. The minimum, mean, median, and maximum scalar coefficient values are 9.09, 13.03, 13.19, and 21.50 days, respectively.

The mean scalar coefficient value (i.e., s = 13.03 days) was used in Equation 2 to convert the n^{p}_{orbits} values (Table 2) to n^{p}_{days} values (Table 3.3). In addition, the n^{p}_{years} values, defined by dividing n^{p}_{days} by 365, were derived.

Table 3.3 The expected minimum number of days (n^{p}_{days}) and years (n^{p}_{years}) required to characterize the site forest AGB to within ±p% of the site overall mean GEDI AGB (OMGA), derived with the mean 13.03 day scalar coefficient value (Figure 3.13).

	n^{p}_{days} (n^{p}_{years})		
Site #	p = ±20%	p = ±10%	p = ±5%
1	78 (0.21)	221 (0.61)	482 (1.32)
2	143 (0.39)	260 (0.71)	560 (1.54)
3	130 (0.36)	416 (1.14)	664 (1.82)
4	65 (0.18)	182 (0.5)	416 (1.14)
5	195 (0.54)	325 (0.89)	612 (1.68)
6	221 (0.61)	534 (1.46)	742 (2.03)
7	91(0.25)	234 (0.64)	508 (1.39)
8	91 (0.25)	260 (0.71)	508 (1.39)
9	182 (0.50)	377 (1.04)	573 (1.57)
10	91 (0.25)	208 (0.57)	469 (1.29)
11	91 (0.25)	273 (0.75)	508 (1.39)
12	143 (0.39)	325 (0.89)	547 (1.50)
13	78 (0.21)	143 (0.39)	390 (1.07)
14	78 (0.21)	273 (0.75)	495 (1.36)
15	117(0.32)	312 (0.86)	547 (1.50)

A longer observation time period is required to characterize the site mean forest AGB more precisely, i.e., to within smaller ±p% values. Specifically, among the 15 sites, n_{days}^{5} varied from 390 – 742 days (i.e., 1.07 – 2.03 years), n_{days}^{10} varied from 143 – 534 days (i.e., 0.39 – 1.46 years), and n_{days}^{20} varied from 65 – 221 days (i.e., 0.18 – 0.61 years).

Table 3.4 shows the range in the days n_{days}^{p} and n_{years}^{p} values derived using the minimum and maximum scalar coefficient values of 9.09 and 21.50 days. The tabulated values bound the Table 3 results. Among the 15 sites, n_{days}^{5} varied from 272 – 1225 days (i.e., 0.75 – 3.36 years), n_{days}^{10} varied from 99 – 881 days (i.e., 0.27 – 2.42 years), and n_{days}^{20} varied from 45 – 365 days (i.e., 0.12 – 1.00 years).

Table 3.4 The range in expected minimum number of days (n_{days}^{p}) and years (n_{years}^{p}) required to characterize the site forest AGB to within ±*p*% of the site overall mean GEDI AGB (OMGA), derived with minimum and maximum scalar coefficient values of 9.09 and 21.50 days (Figure 3.13).

	n_{days}^{p} (n_{years}^{p})		
Site #	*p* = ±20%	*p* = ±10%	*p* = ±5%
1	54 – 129 (0.15 - 0.35)	154 – 365 (0.42 – 1.00)	336 – 795 (0.92 - 2.18)
2	99 – 236 (0.27 - 0.65)	181 – 430 (0.50 - 1.18)	390 – 924 (1.07 - 2.53)
3	90 – 215 (0.25 - 0.59)	290 – 688 (0.8 - 1.88)	463 – 1096 (1.27 – 3.00)
4	45 – 107 (0.12 - 0.29)	127 – 301 (0.35 - 0.82)	290 – 688 (0.80 - 1.88)
5	136 – 322 (0.37 - 0.88)	227 – 537 (0.62 - 1.47)	427 – 1010 (1.17 - 2.77)
6	154 – 365 (0.42 – 1.00)	372 – 881 (1.02 - 2.42)	518 – 1225 (1.42 - 3.36)
7	63 – 150 (0.17 - 0.41)	163 – 387 (0.45 - 1.06)	354 – 838 (0.97 - 2.30)
8	63 – 150 (0.17 - 0.41)	181 – 430 (0.5 - 1.18)	354 – 838 (0.97 - 2.30)
9	127 – 301 (0.35 - 0.82)	263 – 623 (0.72 - 1.71)	399 – 946 (1.10 - 2.59)
10	63 – 150 (0.17 - 0.41)	145 – 344 (0.40 - 0.94)	327 – 774 (0.90 - 2.12)
11	63 – 150 (0.17 - 0.41)	190 – 451 (0.52 - 1.24)	354 – 838 (0.97 - 2.3)
12	99 – 236 (0.27 - 0.65)	227 – 537 (0.62 - 1.47)	381 – 903 (1.05 - 2.47)
13	54 – 129 (0.15 - 0.35)	99 – 236 (0.27 - 0.65)	272 – 645 (0.75 - 1.77)
14	54 – 129 (0.15 - 0.35)	190 – 451 (0.52 - 1.24)	345 – 817 (0.95 - 2.24)
15	81 – 193 (0.22 - 0.53)	218 – 516 (0.6 - 1.41)	381 – 903 (1.05 - 2.47)

3.5 Discussion

Careful analysis was undertaken to select 15 50 × 50 km sites that captured a representative range of Mai Ndombe province forest AGB with different within-site forest AGB spatial variation. Only sites that had >80% forest cover and where GEDI good quality AGB forest 25m footprint values occurred across each site were selected. The most recently processed GEDI footprint-level AGB (Level-4A) product data acquired from April 2019 to November 2021 (31 months) were considered. To provide a forest map contemporaneous with the GEDI data, a global 2010 30 m percent tree cover product was processed by removing locations where forest loss was detected in a global 2000-2021 30 m forest cover loss product (Global Forest Change, 2021). The resulting data set was thresholded and 30 m pixels with percent tree cover >30% were labelled as forest. A different forest definition could impact the site selection, but as this definition is the one conventionally used by REDD+ projects in the DRC, being recommended by the DRC ministry of Environment and Sustainable Development and approved by the United Nations Framework Convention on Climate Change (DRC FREL, 2018), it is appropriate and is expected to provide sites that are representative of central African forest REDD+ projects.

For each of the 15 selected sites, the expected minimum number of GEDI orbits (n^{p}_{orbits}) and the equivalent temporal observation period (n^{p}_{days}) needed to characterize the AGB using GEDI data to within ±p% of the site overall mean GEDI AGB (OMGA) was derived. The OMGA is assumed to be representative of the actual mean site AGB. This is reasonable given that 31 months of GEDI data over sites with >80% forest cover and GEDI observations distributed across them were used. In addition, notably, in each selected site >3% of the 30 m forest pixels had good quality AGB footprint values which provides a greater sample than typically used by conventional field-based tree structure sampling. We also note that 31 months is a longer period than the 28 months of GEDI data used recently to make a global vegetation GEDI AGB map (Dubayah et al., 2022). The irregular spatial and temporal coverage of GEDI orbits containing good quality forest GEDI AGB data over 50 × 50 km areas (e.g., Figures 3.4 and 3.5) meant that comparison of the OMGA with mean AGB values derived from increasing numbers of GEDI orbits ordered by their acquisition date was not meaningful. Instead, at each selected site, the OMGA was compared with the mean AGB derived considering *different* unique combinations of $k = 1, 2, 3 \ldots n$ GEDI orbits, where n is the total number of GEDI orbits with good quality forest GEDI AGB data over the 31 study months.

The results of this study indicate, as expected, that more GEDI orbits are needed to characterize the mean site AGB more precisely i.e., with smaller $\pm p$% values (Table 3.2). This is consistent with statistical theory: the mean of results obtained from a number of trials will converge to the expected value as more trials are performed (Dekking et al., 2005). The purpose of this analysis is not to confirm this but to quantify the number of GEDI orbits and so time periods need to characterize the site AGB. Among the 15 sites n^{20}_{orbits} varied from 5 – 17 orbits, n^{10}_{orbits} varied from 11 – 41 orbits, and n^{5}_{orbits} varied from 30 – 57 orbits. Thus, given that all 15 selected sites are representative of Mai Ndombe province at REDD+ project scale, then our results indicate that at least 17, 41 and 57 GEDI orbits are required to characterize REDD+ project AGB to within ±20%, ±10%, and ±5% of the mean site AGB, respectively.

Further analysis was undertaken to understand the reported n^{p}_{orbits} values. The values were compared with the site OMGA to check that they were not correlated. For all 15 sites the n^{p}_{orbits} had low and insignificant ($p > 0.1$) correlations with site OMGA of -0.40, -0.25, 0.29 for n^{5}_{orbits}, n^{10}_{orbits} and n^{20}_{orbits}, respectively. This is also apparent in Figure 14 that shows the n^{10}_{orbits} values for the 15 sites colored by the AGB stratum that each site fell within. The n^{10}_{orbits} values are plotted in Figure 3.14 against the directional semivariogram sill values. Recall that three sites with low, medium, and high mean directional sill values were selected within each stratum to capture different within-site forest AGB spatial variation. As noted in Section 3.4.1, sill values can only be meaningfully compared among sites with similar mean AGB, i.e., among the three sites in the same AGB stratum. However, there is a clear pattern of systematically greater n_{orbits}^{10} values with greater mean directional sill values within each stratum. The same pattern was also observed for the n_{orbits}^{5} and n_{orbits}^{20} results. This is expected, and supports the veracity of our analysis, as sites that are more spatially heterogenous (i.e., have greater within-site AGB spatial variation) will need more n_{orbits}^{p} to characterize the site forest AGB.

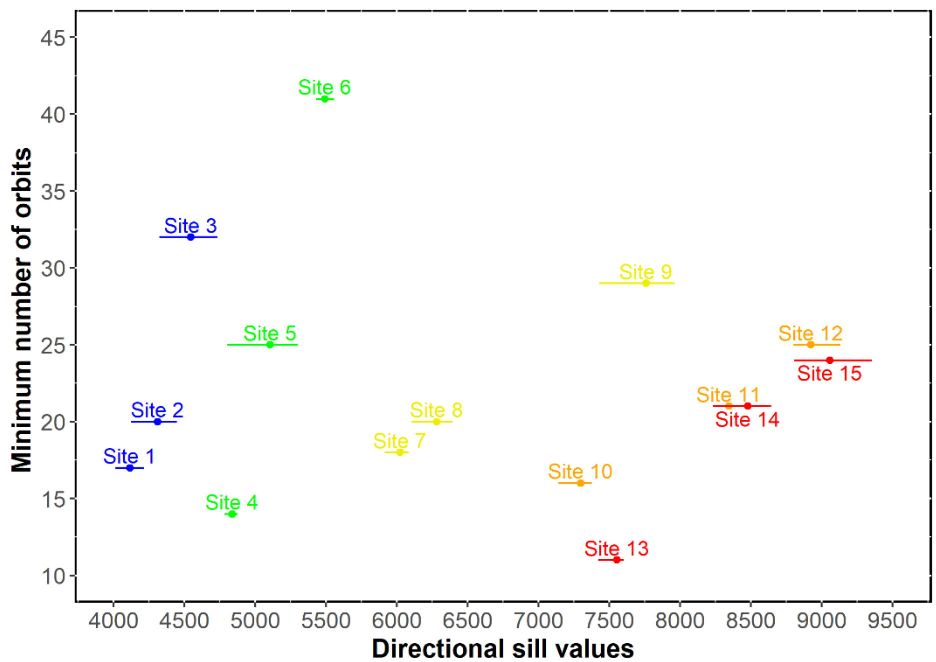

Figure 3.14 The minimum number of GEDI orbits (n_{orbits}^{p}) required to characterize the site forest AGB to within ± 10% of the OMGA plotted against site directional semivariogram sill values. The dots show the mean site directional semivariogram sill values (Table 1), the lines delimit the range of the four directional semivariogram sill values. Results are colored by the five AGB strata that each of the 15 sites were selected from.

The GEDI temporal observation period (n_{days}^{p}) needed to characterize the AGB over each site was defined as the product of the site n_{orbits}^{p} value and a scalar conversion coefficient (Equation 2). The scalar conversion coefficient was derived considering 778 50 × 50 km sites across Mai Ndombe province. The mean coefficient value was 13.03 days, indicating that on average this number of days is required to obtain one GEDI orbit containing good quality forest AGB data over a 50 × 50 km site in Mai Ndombe province. This value is somewhat expected as, although the approximate overpass time of the International Space Station (ISS) is between 3 to 4 days (Gebelein and Eppler, 2006), frequent tropical cloud cover will reduce the availability of GEDI orbit data. Notably, the GEDI 1064 nm lasers do not penetrate clouds (Dubayah et al., 2020), there is no GEDI footprint cloud flag, and the inclined ISS orbit means that the GEDI observations occur at different overpass times which complicates GEDI-based cloud analyses, particularly in the tropics where clouds can have seasonally and regionally different diurnal variation (Philippon et al., 2016; Dommo et al., 2018).

The availability of good quality GEDI AGB data may also be reduced by ephemeral surface water presence, GEDI AGB product retrieval sensitivity to forest conditions, and changes in the ISS orbit. The relative influence of these factors is unknown but among the 778 sites the minimum and maximum coefficient values were 9.09 and 21.50 days, respectively. Sites with more orbits containing quality forest GEDI footprint AGB values and sites with a more regular temporal occurrence of orbits tended to have smaller coefficient values. Given these factors it is unknown if the mean 13.03 day period is typical of all central Africa tropical forest locations. A longer observation time period was needed to characterize the site mean forest AGB more precisely (Tables 3.3 and 3.4). Among the 15 sites, using the mean scalar conversion coefficient of 13.03 days (Table 3), n_{days}^{20} varied from 65 – 221 days (i.e., 0.18 – 0.61 years), n_{days}^{10} varied from 143

– 534 days (i.e., 0.39 – 1.46 years), and n_{days}^{5} varied from 390 – 742 days (i.e., 1.07 – 2.03 years). Thus, given that all 15 selected sites are representative of the Mai Ndombe province at REDD+ project scale, then at least 212, 534 and 742 days (0.61, 1.46, and 2.03 years) are required to characterize the AGB to within ±20%, ±10%, and ±5%, of the mean site AGB, respectively. Under worst- and best-case scenarios, i.e., using 21.50 day and the 9.09 day scalar conversion coefficients, the n_{days} values scale proportionally by factors of 1.650 and 0.698 respectively. Under the worst-case scenario at least 365, 881, and 1225 days (1.00, 2.42, and 3.36 years) are required to characterize the AGB to within ±20%, ±10%, and ±5%, of the mean site AGB, respectively (Table 3.4).

The analysis undertaken in this study is internally self-consistent. However, although the data sets used are state of the practice and publicly available, they are not error free. The 2010 global percent tree cover product has not been validated, and the forest cover loss product was validated for a ten year period with reported forest loss producer's and user's accuracies of 87.8% and 87.0%, respectively (Hansen et al., 2013b). In the DRC frequent cloud cover means that the temporal reporting accuracy of the forest cover loss product, that is based on examination of Landsat-7 and -8 time series, may not report the year of forest loss correctly. Thus, the binary forest map (Figure 2) may contain forest pixels that were not forested in the GEDI 31 month time period because forest cover loss failed to be detected. The GEDI footprint-level AGB product also has error. The AGB is derived by applying an allometric model to relative canopy height estimates retrieved from each GEDI footprint waveform (Duncanson et al., 2022). The GEDI relative canopy height retrieval accuracy depends on several factors. For example, in this study both day and night GEDI data were used but nighttime data can be more reliable (Beck et al., 2021). For example, GEDI observations with beam sensitivity > 0.9 (i.e., when > 90% of the canopy cover is penetrated to the ground by the waveform energy) are more reliable than lower beam sensitivity (Duncanson et al., 2020). Only good quality GEDI observations were used which nominally have beam sensitivity > 0.95 (Kellner et al., 2021). However, the beam sensitivity retrieval is dependent on the algorithm setting group used to control the GEDI waveform interpretation and although we used only optimal setting group data this may not always be correctly defined. In addition, GEDI has geolocation errors that introduce relative canopy height biases, and so AGB biases, over spatially heterogeneous canopies and along forest edges (Roy et al., 2021). Most notably, the GEDI AGB product available over the study area were derived using an African Evergreen Broadleaf Tree (EBT) allometric model applied to the GEDI 50th and 98th relative canopy height retrievals that has a reported 66.9% RMSE fit uncertainty with no reported bias term (Duncanson et al., 2022). In addition, the GEDI AGB data may not be reliably defined for low AGB conditions, typically outside of the range of the forest ground-based AGB data used to derive the EBT model (Duncanson et al., 2022). These GEDI errors together may explain why the site GEDI AGB values reported in this study are lower than those reported in the literature that were derived without GEDI data. The Intergovernmental Panel on Climate Change (IPCC) recommended accuracy requirement for forest AGB estimates is 10% (IPCC, 2003). Thus, to meet this accuracy requirement our results suggest that at least 534 days (1.46 years) would be required for REDD+ site monitoring using the GEDI AGB product. However, if the site was located where the temporal availability of GEDI orbits containing good quality GEDI data is particularly low, for example, in this study one site had a 21.50 day scalar conversion coefficient, then our results indicate that 881 (2.42 years) are required. Notably, these time periods may encompass forest change. For example, forest disturbance and forest cover loss due to logging or forest clearing for charcoal production and subsistence agriculture can occur rapidly, and tropical forest degradation rates, although typically not documented, can also occur although at up to decadal time scales (Matricardi et al., 2020). Further, although studies of tropical forest regrowth are limited, tree canopy height increases of several meters per year have been reported (Muller-Landau et al., 2006; Cole et al., 2014). Thus, care should be taken in using an appropriate time period of GEDI AGB data for tropical forest assessment at REDD+ project scale. Despite these limitations, as tropical forest AGB inventory data are expensive and hard to collect *in situ* (Chave et al., 2019; Kohl et al., 2020) the reported results are encouraging given the relatively short time periods indicated by this study.

3.6 Conclusions

The amount of GEDI data (number of orbits and amount to time) required to characterize tropical forest AGB over 50 × 50 km (the typical size of REDD+ projects in tropical Africa) was examined considering 31 months of GEDI footprint-level AGB data across Mai Ndombe province in the Democratic Republic of the Congo. Careful analysis of the AGB product and 30 m Landsat based percent tree cover and forest cover loss products was undertaken to select 15 sites that captured a representative range of forest AGB from five GEDI-derived AGB strata, with a range of within-site forest AGB spatial variation. In addition, summary statistics of the average number of days required to obtain a GEDI orbit containing good quality forest AGB data were derived at 778 50 × 50 km sites across the province. The results of this study found that (i) more GEDI orbits were needed to characterize the mean site AGB more precisely and the number increased for sites with greater within-site AGB variation, (ii) an average of 13.03 days (minimum 9.09 days, maximum 21.0 days) was required to obtain one GEDI orbit containing good quality forest AGB data at REDD+ project scale, (iii) longer observation periods were needed to characterize the site mean forest AGB more precisely, (iv) on average among the 15 sites, periods from 65 – 221 days (0.18 – 0.61 years), 143 – 534 days (0.39 – 1.46 years), and 390 – 742 days (1.07 – 2.03 years) were required to characterize the AGB to within ±20%, ±10%, and ±5% of the mean site AGB, respectively, (v) under the worst-case scenario, where the average number of days required to obtain a GEDI orbit containing good quality forest AGB data at a site was 21.5 days, at least 365, 881, and 1225 days (1.00, 2.42, and 3.36 years) were required to characterize the AGB to within ±20%, ±10%, and ±5%, of the mean site AGB, respectively. These results imply that in Mai Ndombe province at least 0.61, 1.46, and 2.03 years are required to characterize the AGB at REDD+ project scale to within ±20%, ±10%, and ±5%, of the mean site AGB, respectively. In other REDD+ sites with markedly different forest AGB and spatial variation, for example, projects with significant areas of afforestation, these observations periods may be different. Moreover, in other central African tropical forest localities these observations periods may be different depending not just on the forest AGB and spatial variation but on the cloud cover, ephemeral surface water presence, and GEDI AGB retrieval sensitivity to the forest conditions.

REFERENCES

Atmadja, S. 2021. Summary analysis of REDD+ projects, 2018-2020, Version 4.1. available at: URL:http://www.reddprojectsdatabase.org (Last accessed : 7 December 2022).

Baccini, A., Laporte, N., Goetz, S.J., Sun, M., & Dong, H. 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters*. 3, 045011.

Barnes, R.J., 1991. The variogram sill and the sample variance. *Mathematical Geology*. 23, 673-678.

Beck, J., Wirt, B., Luthcke, S., Hofton, M., Armston, J. 2021. Global Ecosystem dynamics investigation (GEDI) Level 02 user guide. Document version 2.0, April 2021. *U.S. Geological Survey, Earth Resources Observation and Science Center (Sioux Falls, South Dakota, USA)*.

Bwangoy, J.-R.B., Hansen, M.C., Roy, D.P., Grandi, G.D., & Justice, C.O. 2010. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. *Remote Sensing of Environment*. 114, 73-86.

Bwangoy, J.-R.B., Hansen, M.C., Potapov, P., Turubanova, S., & Lumbuenamo, R.S. 2013. Identifying nascent wetland forest conversion in the Democratic Republic of the Congo. *Wetlands Ecology and Management. 21*, 29-43

Chave, J., Davies, S. J., Phillips, O. L., Lewis, S. L., Sist, P., Schepaschenko, D., & Saatchi, S. 2019. Ground Data are Essential for Biomass Remote Sensing Missions. *Surveys in Geophysics*. 40, 863-880.

Cole, L.E., Bhagwat, S.A., Willis, K.J. 2014. Recovery and resilience of tropical forests after disturbance. *Nature Communication*. 5, 1–7.

Cressie, N. 1993. *Statistics for Spatial Data*. Wiley Interscience, New York.

Curran, P.J. & Atkinson, P.M. 1998. Geostatistics and remote sensing. *Progress in Physical Geography*. 22 (1), 61–78.

Dekking, F. M., Kraaikamp, C., Lopuhaä, H. P., & Meester, L. E. 2005. *A Modern Introduction to Probability and Statistics: Understanding why and how (Vol. 488).* Springer, London.

Dommo, A., Philippon, N., Vondou, D.A., Sèze, G., & Eastman, R. 2018. The June–September Low Cloud Cover in Western Central Africa: Mean Spatial Distribution and Diurnal Evolution and Associated Atmospheric Dynamics. *Journal of Climate*. 31, 9585-9603.

DRC Forest Emission Reference Level (DRC FREL) approved by the United Nations Framework Convention on Climate Change –UNFCC. 2018. *UNFCC REDD+ web platform* available at: https://redd.unfccc.int/files/rdc_documentnerf_soumissionfinale_29112018.pdf (Last accessed : 7 Dec 2022).

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., & Silva, C. 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing.*1, 100002.

Dubayah, R., Armston, J., Healey, S., Bruening, J., Patterson, P., Kellner, J., Duncanson, L., Saarela, S., Ståhl, G., & Yang, Z. 2022. GEDI launches a new era of biomass inference from space. *Environmental Research Letters. 17*(9), 095001.

Duncanson, L., Kellner, J. R., Armston, J., Dubayah, R., Minor, D. M., Hancock, S., & Zgraggen, C. 2022. Aboveground biomass density models for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar mission. *Remote Sensing of Environment.* 270, 112845.

Duncanson, L., Neuenschwander, A., Hancock, S., Thomas, N., Fatoyinbo, T., Simard, M., Silva, C.A., Armston, J., Luthcke, S.B., Hofton, M., Kellner, J.R., & Dubayah, R. 2020. Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. *Remote Sensing of Environment.* 242, 111779.

Fayolle, A., Loubota Panzou, G.J., Drouet, T., Swaine, M.D., Bauwens, S., Vleminckx, J., Biwole, A., Lejeune, P., & Doucet, J.-L. 2016. Taller trees, denser stands and greater biomass in semi-deciduous than in evergreen lowland central African forests. *Forest Ecology and Management*. 374, 42-50.

Friedlingstein, P., O'sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., & Zaehle, S. 2020. Global Carbon Budget 2020. *Earth System Science Data*. 12, 3269– 3340.

Gebelein, J., & Eppler, D., 2006. How Earth remote sensing from the International Space Station complements current satellite-based sensors. *International Journal of Remote Sensing*. 27(13), 2613-2629.

Global Forest Change, 2021, Global Forest Change 2000–2021 Data Description document, University of Maryland, College Park, MD, Last accessed 7 Dec 2022 from https://storage.googleapis.com/earthenginepartners-hansen/GFC-2021-v1.9/download.html

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., & Townshend, J. 2013a. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*. 342: 850–53.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., & Townshend, J. 2013b. Supplementary material for High-Resolution Global Maps of 21st-Century Forest Cover Change. Science. 342: 850–53.

Hansen, M.C., Krylov, A., Tyukavina, A., Potapov, P.V., Turubanova, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., & Moore, R., 2016. Humid tropical forest disturbance alerts using Landsat data. *Environmental Research Letters.* 11, 034008.

Harris, N. L., Gibbs, D. A., Baccini, A., Birdsey, R. A., De Bruin, S., Farina, M., et.al., 2021. Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change*. 11, 234-240.

Herold, M., Carter, S., Avitabile, V., Espejo, A. B., Jonckheere, I., Lucas, R., & De Sy, V. 2019. The Role and Need for Space-Based Forest Biomass-Related Measurements in Environmental Management and Policy. *Surveys in Geophysics.* 40, 757–778.

Hubau, W., Lewis, S. L., Phillips, O. L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., & Zemagho, L. 2020. Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature.* 579, 80-87.

Hofton, M., Blair, B., Story, S., & Yi, D. 2019. Algorithm Theoretical Basis Document (ATBD) for GEDI Transmit and Receive Waveform Processing for L1 and L2 Products. *Goddard Space Flight Center, Greenbelt, MD, USA.*

IPCC. 2003. Good practice guidance for land use, land-use change and forestry. Institute for Global Environmental Strategies, Kanagawa. Available at: https://www.ipcc.ch/publication/good-practice-guidance-for-land-use-land-use-change-and-forestry/. (Last accessed: 7 Dec 2022).

Kovalskyy, V., & Roy, D. P. 2013. The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30 m Landsat data product generation. *Remote Sensing of Environment*. *130*, 280-293.

Kashongwe, H.B., Roy, D.P., & Bwangoy, J.-R.B. 2020. Democratic Republic of the Congo tropical Forest Canopy Height and Aboveground Biomass Estimation with Landsat-8 Operational Land Imager (OLI) and Airborne LiDAR Data: The Effect of Seasonal Landsat Image Selection. *Remote Sensing.* 12, 21.

Keenan, R.J., Reams, G.A., Achard, F., de Freitas, J.V., Grainger, A., & Lindquist, E. 2015. Dynamics of global forest area: Results from the FAO Global Forest Resources Assessment 2015. *Forest Ecology and Management*. 352, 9-20.

Kellner, J.R., Armston, J., & Duncanson, L. 2021. Algorithm Theoretical Basis Document (ATBD) for GEDI Level-4A (L4A) Footprint Level Aboveground Biomass Density. Document Vesion 1, August 2021. *University of Maryland*. (College Park, Maryland, USA).

Kim, D.H., Sexton, J.O., & Townshend, J.R. 2015. Accelerated deforestation in the humid tropics from the 1990s to the 2000s. *Geophysical Research Letters*. 42, 3495-3501.

Klein, S.V., & Axelrad, P. 2020. Simulation and analysis of GPS multipath for the GEDI experiment onboard the international space station. In: *Proceedings of the 33rd International Technical Meeting of the Satellite Division of the Institute of Navigation* (ION GNSS+ 2020). pp. 1338–1371.

Köhl, M., Neupane, P.R., & Mundhenk, P. 2020. REDD+ measurement, reporting and verification – A cost trap? Implications for financing REDD+MRV costs by result-based payments. *Ecological Economics.* 168, 106513.

Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. 2022. Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*. 268, 112760.

Lewis, S. L., Sonké, B., Sunderland, T., Begne, S. K., Lopez-Gonzalez, G., Van Der Heijden, G. M., & Zemagho, L. 2013. Above-ground biomass and structure of 260 African tropical forests. *Philosophical Transactions of the Royal Society London B Biological Science.* 368, 20120295.

Li, X., Wessels, K., Armston, J., Hancock, S., Mathieu, R., Main, R., Naidoo, L., Erasmus, B., & Scholes, R. 2023. First validation of GEDI canopy heights in African savannas. *Remote Sensing of Environment*. *285*, 113402.

Ma, S., He, F., Tian, D., Zou, D., Yan, Z., Yang, Y., & Fang, J. 2018. Variations and determinants of carbon content in plants: a global synthesis. *Biogeosciences*. 15, 693-702.

Matricardi, E. A. T., Skole, D. L., Costa, O. B., Pedlowski, M. A., Samek, J. H., & Miguel, E. P. 2020. Long-term forest degradation surpasses deforestation in the Brazilian Amazon. *Science*, 369, 1378-1382.

Mayaux, P., De Grandi, G., & Malingreau, J.P. 2000. Central African Forest Cover Revisited. *Remote Sensing of Environment. 71*, 183-196.

Molinario, G., Hansen, M.C., & Potapov, P.V. 2015. Forest cover dynamics of shifting cultivation in the Democratic Republic of Congo: a remote sensing-based assessment for 2000–2010. *Environmental Research Letters.* 10, 094009.

Muller-Landau, H. C., Condit, R. S., Chave, J., Thomas, S. C., Bohlman, S. A., Bunyavejchewin, S., & Ashton, P.2006. Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. *Ecology Letters.* 9, 575–588.

Philippon, N., de Lapparent, B., Gond, V., Sèze, G., Martiny, N., Camberlin, P., Cornu, G., Morel, B., Moron, V., Bigot, S., Brou, T., & Dubreuil, V. 2016. Analysis of the diurnal cycles for a better understanding of the mean annual cycle of forests greenness in Central Africa. *Agricultural and Forest Meteorology*. 223, 81-94.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A.., Blair, J.B., & Hofton, M. 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment.* 253, 112165.

Roy, D. P., Qin, Y., Kovalskyy, V., Vermote, E. F., Ju, J., Egorov, A., & Yan, L. 2014. Conterminous United States demonstration and characterization of MODIS-based Landsat ETM+ atmospheric correction. *Remote Sensing of Environment*. 140, 433-449.

Roy, D. P., Li, J., Zhang, H. K., & Yan, L. 2016. Best practices for the reprojection and resampling of Sentinel-2 Multi Spectral Instrument Level 1C data. *Remote Sensing Letters*. 7, 1023-1032.

Roy, D.P., Kashongwe, H.B., & Armston, J. 2021. The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. *Science of Remote Sensing*. 4, 100024.

Sandker, M., Carrillo, O., Leng, C., Lee, D., d'Annunzio, R., & Fox, J. 2021. The importance of high–quality data for REDD+ monitoring and reporting. *Forests*. 12, 99.

Samba, G., Nganga, D., & Mpounza, M. 2007. Rainfall and temperature variations over Congo-Brazzaville between 1950 and 1998. *Theoretical and Applied Climatology*. 91, 85-97.

Shapiro, A.C., Aguilar-Amuchastegui, N., Hostert, P., & Bastin, J.F. 2016. Using fragmentation to assess degradation of forest edges in Democratic Republic of Congo. *Carbon Balance Management*. 11, 1-15.

UNFCCC – United Nations Framework Convention on Climate Change. 2014. Warsaw Framework for REDD-plus. Available at: https://unfccc.int/topics/land-use/resources/warsaw-framework-for-redd-plus. (Last accessed on 7 Dec 2022).

World-Bank. 2016. Emission Reduction Program. *World Bank/Forest Carbon Partnership Facility*. 1-297.

WWC- Wildlife Works Carbon. 2012. The Mai Ndombe REDD+ Project (MNRP) project description document. Verified Carbon Standard-Verra program web platform available at: https://registry.verra.org/app/projectDetail/VCS/934 (Last accessed: 26 January 2023).

Xu, L., Saatchi, S. S., Shapiro, A., Meyer, V., Ferraz, A., Yang, Y., & Ebuta, D. 2017. Spatial distribution of carbon stored in forests of the Democratic Republic of Congo. *Nature Scientific Reports*, 7, 15030.

CHAPTER 4. ASSESSING CENTRAL AFRICAN POST-FOREST DISTURBANCE VEGETATION REGROWTH CARBON STORAGE WITH GEDI DATA: A FOCUSED STUDY IN THE MAI NDOMBE REDD+ PROJECT, DRC

4.1 Introduction

The Global Ecosystem Dynamics Investigation (GEDI) LiDAR onboard the International Space Station (ISS) is the first spaceborne LiDAR specifically designed for vegetation structure retrieval with a focus on improved forest structure and aboveground biomass quantification in support of carbon cycle and biodiversity studies (Dubayah et al., 2020). The GEDI mission acquires waveform LiDAR measurements in 25 m diameter footprints from 51.6° South to 51.6° North including the world's tropical forests. Tropical forests account for about half the world's terrestrial carbon (Hubau et al., 2020; Harris et al., 2021) with deforestation and degradation comprising about 20% of global annual net carbon emissions from human activities (Friedlingstein et al., 2022).

Tropical forests can grow back as 'secondary' forests after natural or human-induced disturbances that cause deforestation. Regrowing secondary forests absorb CO_2 from the atmosphere, which partially compensates for the emissions generated in their destruction. For example, Heinrich et al. (2023) estimated that regrowing tropical forests in the Amazon, Borneo and Central Africa, mitigated 26% of the emissions that originally occurred as a result of forest loss and degradation. However, this, and similar research is far from definitive, and validation involving integration of satellite- and ground-based data is required before the outcomes can be used by policymakers (Zuideema and Jakovac, 2023). This is very important as conserving secondary and recovering degraded forests, by allowing them to regrow, provides a low-cost mechanism for carbon sequestration and an effective pathway to help mitigate climate change, and may play a key role in biodiversity conservation (Chadzon et al., 2016; Lennox et al., 2018; Lewis et al., 2019). However, recent findings in Borneo tropical forests suggest that regrowing forests on previously degraded forest land was not a carbon sink but a net carbon source due to persistent carbon losses from soil organic matter and deadwood (Mills et al., 2023).

Central African tropical forests are relatively understudied compared to other tropical regions such as the Amazon (Verbeeck et al., 2011; White et al., 2021). They constitute the second largest continuous block of tropical forests in the world, and the majority are in the Democratic Republic of the Congo (DRC) (Vancutsem et al., 2021). Despite this, forest conversion in Central African tropical forest is substantial and mainly driven by shifting cultivation (Tyukavina et al., 2018), resulting in relatively large areas of regrowing secondary forests. Tropical forest regrowth rates are poorly understood. Regrowth rates are difficult to quantify for several reasons including forest accessibility issues, prohibitively expensive and time consuming repeated *in situ* measurements over the lifetime of individual trees, and difficulties associated with dating forest cover loss reliably. Moreover, obtaining sufficient samples to reliably quantify regrowth at landscape scale is extraordinarily challenging as absolute growth rates typically change with tree size, age, and species. Early successional secondary tropical forests in Central Africa typically grow quickly, have low wood density, and are replaced by high wood density shade tolerant species (White, 1983; Huntley, 2023).

Detailed information and data on regrowth rates in Central Africa are sparse and limited by sparse forest inventory (Gourlet-Fleury et al., 2013; Baccini et al., 2017; Bauters et al., 2019). The largest body of work on forest regrowth in Central Africa is based on *in situ* measurement of tree diameter at breast height (dbh) that have often been collected at places with good forest access, typically associated with commercial forestry. For example, Madron et al. (2000) collated annual dbh data for 17 commonly logged tree timber species (>10 cm dbh) in Ivory Coast, Cameroon, Central African Republic, and Gabon, for 14 years and reported mean dbh growth rates of 2-5 mm/year for primary forest tree species and ~10 mm/year for secondary forest pioneer species (*Nesogordonia papaverifera*, *Terminalia superba*, and *Aucoumea klaineana*). Fayolle et al. (2012) measured annual dbh for 31 tropical forest tree species (>10 cm dbh) in 3 Central African countries for 4 years and reported dbh increases of 5.09-11.14 mm/year for primary forest tree species and 6.37-38.2 mm/year for secondary forest pioneer species (*Musanga cecropioides and Lophira Alata* species).

Deriving tropical forest regrowth rates from *in situ* measurement of tree height is challenging because direct measurement of tree heights (for example with a clinometer or a Terrestrial Laser Scanner) is not usually practical in tropical forests because the canopy morphology (characterized by wide tree crowns, adjacent trees with different heights, high canopy cover, and high leaf area index) occludes lines of sight to the canopy top. Despite these challenges, Swaine (1983) measured tree heights annually for five years post clearance of upland evergreen forest in Ghana and reported tree height growth rates up to 4 m/year for the light dependent pioneer species *Musanga cecropioides* and lower rates of 2.4 m/year for the shade tolerant *Harungana madagascariensis*. Other researchers have inferred tree height growth rates by comparing the heights of forest stands with known ages. For example, Ross (1954) measured *Musanga cecropioides* tree heights in secondary forest stands in southern Nigeria known to be 5 and 14 years old and reported heights of 10 m and 23 m, respectively, and inferred a growth rate of 2 m/year.
Dendrochronological methods also perform poorly in the tropics because tree rings are typically not clearly exhibited as there is no single growing season to demark annual growth cycles (Worbes, 2003, Brienen and Zuidema, 2005; Ligot et al., 2019). Unsurprisingly, only a small number of tree ring studies have attempted to quantify tree growth in Central Africa. For example, Madron (2003) reported tree diameter increases measured in the Central African Republic for light dependent

secondary forest non-pioneer tree species *Mansonia altissima* and *Milicia excelsa* of 8.5 and 9.3 mm/year for juvenile (dbh <10 cm) and 4.7 and 4.5 mm/year for older trees (dbh > 110 cm), respectively. Groenendijk et al. (2014) found significantly different increases among commercial forest species in Cameroon that varied from 4.3 mm/year for primary forest tree species to 9.6 mm/year for the secondary forest tree species *Terminalia ivorensis.* It is well established that Light Detection and Ranging (LiDAR) remote sensing can accurately measure forest canopy height and that repeated LiDAR measurement provides estimates of forest regrowth. For example, Boehm et al (2023) used bi-temporal ALS data acquired in 2007 and 2011 to quantify tropical forest regrowth in Central Kalimantan, Indonesia. In another example, Riofrio et al (2023) quantified temperate forest height regrowth in South Ontario, Canada using multi-temporal ALS data acquired in 2005, 2012, and 2018. However, this approach to measure forest regrowth has not yet been implemented in Central African tropical forest because of the high cost in acquiring ALS data, which remains an obstacle for repeat large area monitoring (Wulder et al., 2012; Bustamante et al., 2016).

The availability of the GEDI relative height product (Dubayah et al., 2021) provides new opportunities to build tree height chronosequences to quantify tropical forest regrowth at landscape scale. Recently, Milenkovic et al. (2022) estimated regrowth rates at two sites in the Amazon using annually-updated Landsat-derived 30 m stand age maps and GEDI and ICESat-2 forest height retrievals without considering tree species. Although there is high confidence that GEDI can retrieve measurements that allow for the estimation of forest canopy height in the tropics, its ability to quantify Central African tropical forest regrowth rates has not yet been investigated. Certainly, in Central Africa, as discussed above, different tree species have different regrowth rates, and the species compositions within stands may change temporally. Furthermore, different species can have very different dry wood densities and so have markedly different carbon sequestration potential. To quantify forest regrowth with satellite data, precise time-series maps of forest and forest-change are needed, which are provided by the 30 m global forest cover loss (FCL) product of Hansen et al. (2013) for the past 21 years (2001-2021). These data, combined with the GEDI relative height (RH) product provides a means to quantify forest height regrowth rates in the tropics. This study investigates the ability of the GEDI relative height (RH) product to quantify forest regrowth in the Mai Ndombe forests of central Africa. The global forest cover loss (FCL) product was used to locate and date forest disturbance from 2001 to 2018 (i.e., up to the year before the first GEDI acquisitions). GEDI RH footprints encompassing mapped FCL and capturing a range of regrowing tree heights were identified. Field work was undertaken in each identified GEDI footprint to confirm the presence of regrowing forests and collect tree species inventory data. The tree species inventory data were collected to determine the study area tree species composition and examine the sensitivity of AGB in forest regrowth with respect to variation in species wood density. The GEDI footprint acquisition year and global FCL reported year were used to determine the forest stand age of each identified GEDI RH footprint inventoried. GEDI footprint height values were aggregated by forest stand age and median height for each age was derived. A GEDI tree height chronosequence was derived using median heights plotted as a function of forest stand age and an exponential regression model was fitted to the plotted data to quantify forest height regrowth. The modeled forest height regrowth was converted into AGB using a DRC tree height AGB allometric equation derived for all the DRC as a function of LiDAR heights and species wood density. The tree species inventory data were further used to determine the study area tree species composition and to investigate the study area species diversity using the Shannon diversity index.

4.2 Study Area and Data

4.2.1 Study Area

The research was located in the largest REDD+ project in the DRC, covering 299,654 ha, located near the center of Mai Ndombe province (Figure 4.1). The Mai Ndombe REDD+ project is run by the US-based organization Wildlife Works Carbon (WWC) LLC, which pursues the emerging marketplace for carbon offsets as a sustainable and scalable funding mechanism for biodiverse forest protection and community development (https://www.wildlifeworks.com/).

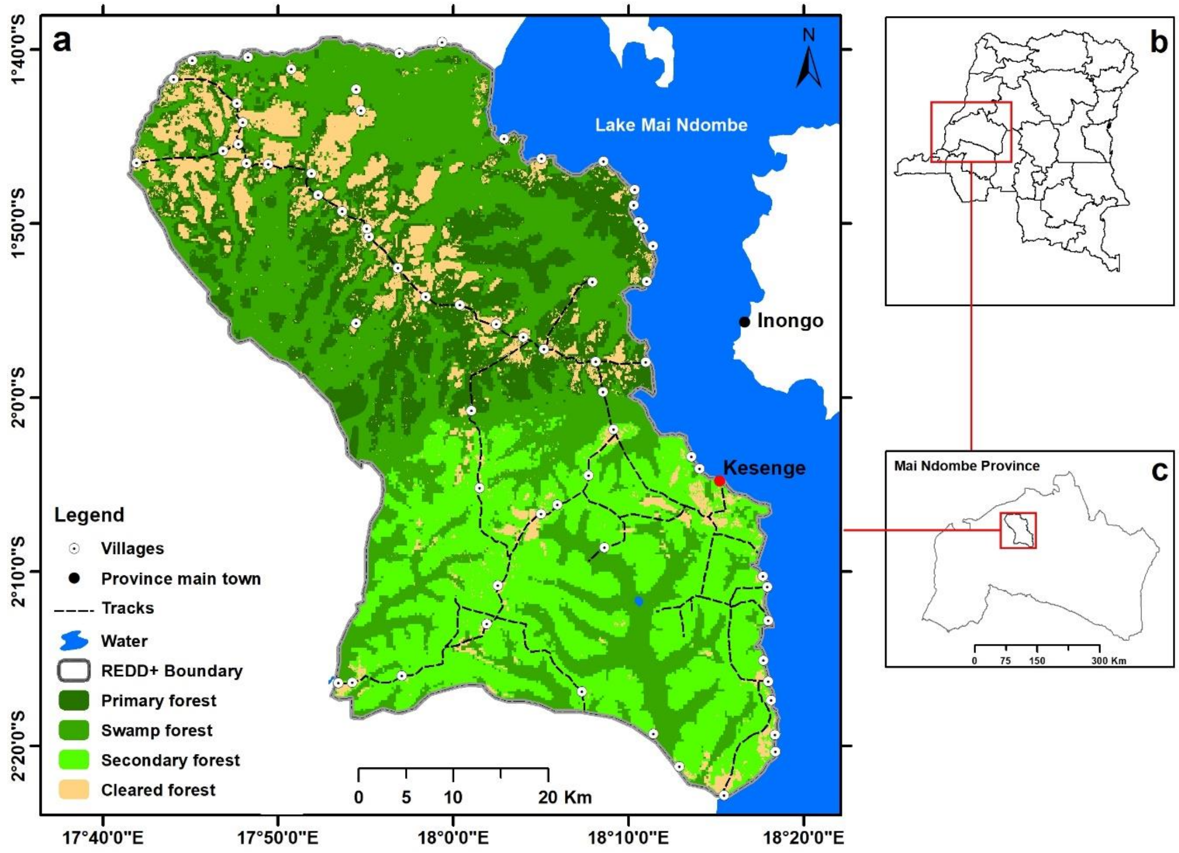

Figure 4.1 Map showing the location of the **(a)** Mai Ndombe REDD+ project study area and a highly generalized forest stratification derived by Landsat land cover classification (WWC-CCB, 2012) over the **(b)** Mai Ndombe Province of the **(c)** DRC.

Hereafter, for brevity and convenience, I refer to the Mai Ndombe REDD+ project, as the "forest reserve". It is located close to the equator in the tropical monsoon climate zone (Beck et al., 2018) with two wet seasons (March–May and September–December) and two dry seasons (June–August and January–February), a mean annual rainfall of 1,800 mm, and mean annual temperature of 76°F (i.e., 24°C) (Samba et al., 2007; Bwangoy et al., 2010). The regional soils are predominantly Ferralsols (Oxisols in the USDA soil taxonomy) associated with low fertility and low nutrient retaining capacity (Deckers, 1993). There may be peat deposits within the forest reserve (Crezee et al., 2022), but their extent has not yet been mapped.

The forest reserve is home to approximately 50,000 people (WWC-CCB, 2012) that live in villages (50-300 people, with a few larger villages >1,000 people) located along the shores of Lake Mai Ndombe and within the forest along tracks and unmetalled roads. The majority of people practice traditional shifting agriculture, hunting, and fishing, with few other alternative livelihood opportunities. Centuries of colonialism and exploitation have led to very poor transport infrastructure (external access to the forest reserve is only possible via the lake), and no economic support from the DRC government, resulting in the communities being highly marginalized and disenfranchised from their economic and political empowerment. The REDD+ project was established in 2011 with a 30-year contract with the DRC government, in partnership with the communities living in and around the reserve, to generate carbon offset revenue through the voluntary carbon market (WWC-PDD, 2012). The communities agreed, through the Free, Prior and Informed Consent (FPIC) process, to co-create strategies for improved food security, access to healthcare and education, while maintaining their tradition of living sustainably with the forest. Notably, under the agreement, community members agreed that no timber harvesting occurs within the forest reserve but that they can use the forest up to 1.2 km around their villages, typically for slash-and-burn agriculture - mainly cassava, also sometimes maize, sorghum, peanuts that is common in the western DRC (Molinario et al., 2015; Shapiro et al., 2016). To date, the Mai Ndombe REDD+ project has verified >31 million metric tons of CO_2e, i.e., avoided CO_2 equivalent emissions. The primary undisturbed forest occurs predominantly in the north (Figure 4.2 darker green tones) and is tropical evergreen terra *firma* rainforest with tracts of swamp forest that are seasonally or permanently inundated. The primary forest is characterized by high tree crown cover (typically 70% or greater), with 35-45 m mature tree heights, and heterogeneous shade tolerant tree species, including valuable hardwood species. Prior to establishment of the REDD+ project in 2011, the forest reserve has been logged three times: first in the 1920s by the Société Forestrière et Commercial du Congo (Forescom) that was established in 1912 in the colonial era, and then by commercial companies BIMPE AGRO in the 1980s and SOFORMA in the 2000s that were focused on logging hardwood species (WWC-CCB, 2012). In 2008, following a governmental revision of the DRC National Forest Code, most of the logging contracts in the province were suspended in an effort to address corruption in the forestry sector. The successional secondary forest (previously logged as described above) across the reserve is evident in Figure 4.2 (lighter green tones) as is the cleared forested land around villages (Figure 4.2 pink tones).

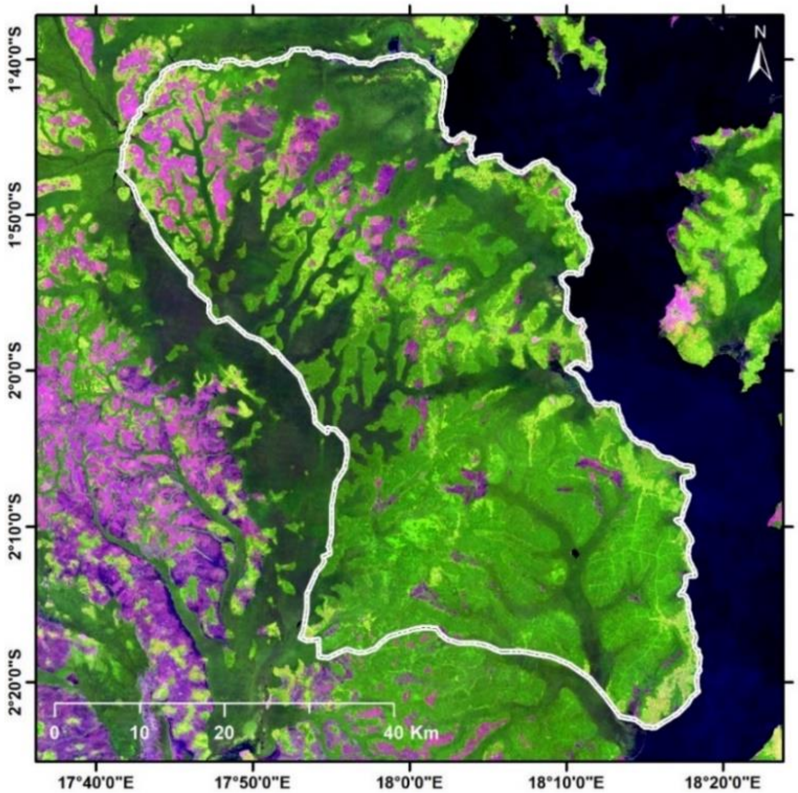

Figure 4.2 The Mai Ndombe REDD+ forest reserve study area (within the white boundary) superimposed on a 30 m top of atmosphere reflectance mosaic derived from year 2021 Landsat time series (Global Forest Change, 2021). Longer wavelength Landsat bands (1610, 865, 655 nm) are shown because the shorter wavelength bands are sensitive to atmospheric contamination.

Financial support for REDD+ projects requires measured, reported, and verified (MRV) estimates of forest above ground biomass to quantify avoided emissions achieved by conservation of forest biomass (Herold et al., 2019; Sandker et al., 2021). The REDD+ project has 463 15 m radius field plots across the reserve in the primary forest and secondary forest areas that have not been disturbed since 2008 (WWC-PDD, 2012). Annual inventories are undertaken by experienced DRC forest officers assisted by local technical support staff. About 20% of the 463 field plots are inventoried each year - measuring dbh for all trees with dbh > 10 cm and undertaking tree species inventory. A total of 313 species have been identified, the most common are *Millettia laurentii* (Figure 4.3a), *Entandrophragma meliaceae*, *Ongokeagore*, and *Tessmannia africana* (WWC-CCB, 2012). Tree species growing at locations that have been disturbed since the REDD+ project was initiated have *not* been surveyed.

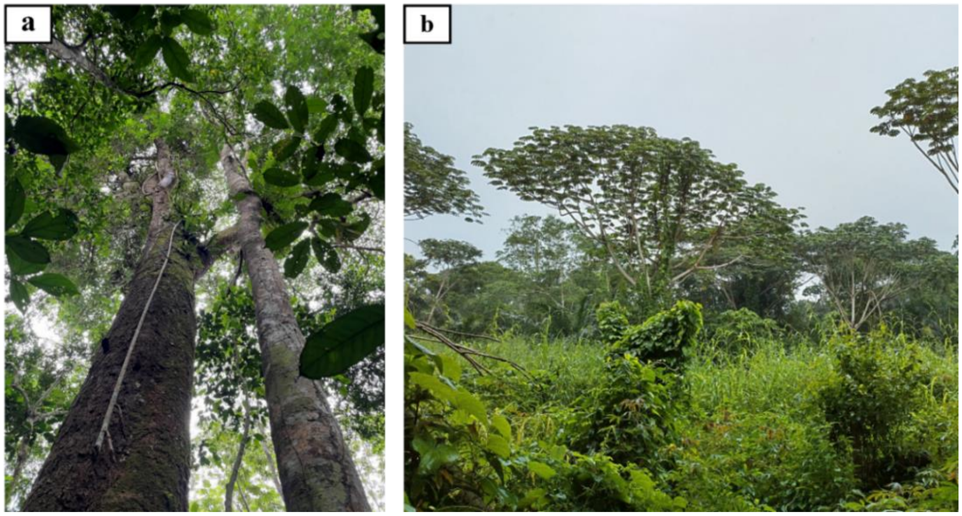

Figure 4.3 Photographs taken during our 2023 field work in successional secondary forest in the eastern part of the Mai Ndombe REDD+ projects showing **(a)** two ~30 m high *Millettia laurentii* trees (local name: Wenge, well-known for its valuable hardwood and designated "endangered" in the IUCN Red List), **(b)** ~20 m high *Musanga cecropioides* pioneer tree species (background) with cleared land covered by weeds, shrubs, and cassava (foreground) (photo credit: David Roy)

4.2.2 Data

4.2.2.1 Global forest cover loss product

The global forest cover loss (FCL) Landsat product (Hansen et a., 2013; Global Forest Change, 2021) was used to select locations where forest cover loss occurred from 2001 to 2018, i.e., up to the year before the first GEDI acquisitions. The FCL product is defined in geographic coordinates with 0.00025° resolution (about 30 m at the equator) hereafter referred to for brevity as being defined with respect to 30 m pixels. The product defines the location and the first year in the period 2001 to current when a Landsat 30 m pixel experienced significant canopy cover loss (Hansen et al., 2016). The FCL product was validated with reported 88% producer's and 87% user's accuracies (Hansen et al., 2013). The temporal reporting accuracy of the global forest change product has not been documented and is expected to be less precise where there are fewer annual surface observations. This is evident in Figure 4.4 that shows (a) Landsat imagery with intact forest on the east and west (dark green), degraded forest in the middle (light green), and cleared areas (orange), (b) the forest change product that classified most of the FCL in 2013 and incorrectly labeled loss in 2014 (yellow stripe) due to missing observations caused by the Landsat-7 Scan Line Corrector.

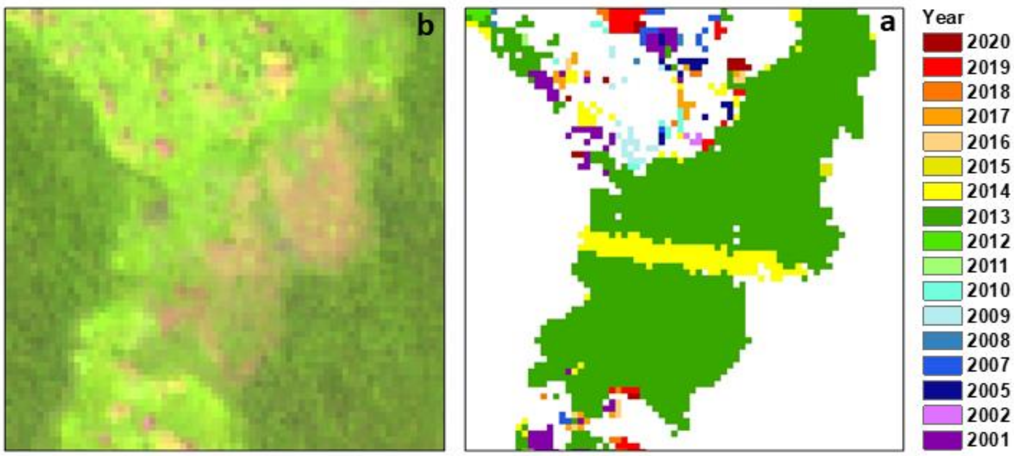

Figure 4.4 Study area 80 × 70 30m pixel subset: **(a)** Landsat-8 30m false color (1610, 865, 655 nm) surface reflectance, **(b)** global forest change product 30 m mapped year of FCL (17 different mapped years in the period 2001 to 2020).

4.2.2.2 GEDI relative height product

The most recently processed Version 2 GEDI Level 2A RH data product (Dubayah et al., 2021) acquired from April 18th 2019 to November 13th 2021 over the forest reserve was used. The GEDI RH product provides estimates of the canopy height relative to the ground (RH_p) in 25 m diameter footprints spaced 600 m and 60 m apart across track and along track (Dubayah et al., 2020). Different algorithm setting groups are used to control the GEDI waveform interpretation providing 7 RH_p values per GEDI footprint with p defined at 1% intervals from 0% (i.e., the surface) to 100% (assumed to be the top of the canopy). The optimal RH_p value is denoted in the product (Beck et al., 2021). In this study, unless otherwise stated, the optimal RH_p values were used. Only good quality GEDI footprints labelled as good waveform quality (quality_flag = 1) without potential for degraded geolocation under suboptimal operating conditions (degrade_flag = 0) were used. Across the forest reserve there were a total of 87,288 good quality GEDI RH footprints. Only GEDI RH footprint center coordinates falling within a 30 m study area 2001-2018 FCL pixel (Section 4.2.2.1) surrounded by at least 5 FCL pixels were retained resulting in 3,987 good quality GEDI RH footprints.

The RH_{95} GEDI heights were used as they typically represent forest canopy heights while avoiding noisy observations (Schneider et al. 2020; Potapov et al., 2021; Roy et al. 2021). The location of each footprint center is provided with quality and degraded geolocation flags (Beck et al., 2021). The recent Version 2.1 GEDI data product has a reported horizontal geolocation uncertainty of 10.3 m (1 σ) which has been shown to introduce canopy height biases over spatially heterogeneous canopies and along forest edges (Beck et al., 2021; Roy et al., 2021).

4.3 Methods

4.3.1 Tree Species Inventory

4.3.1.1 GEDI RH footprint selection for field work

All good quality GEDI RH footprints where field work could be undertaken to derive species specific tree height chronosequence were identified. Only footprints that met the following criteria were used: (i) fall within a study area 30 m 2001-2018 FCL pixel that is surrounded by at least 5 FCL pixels; (ii) be geographically close to each other and to a forest reserve village; and (iii) have different FCL years. Criterion (i) was to ensure that only footprints over post-forest cover loss areas were selected, (ii) was to reduce field work access issues and allow coverage of many footprints during the field work, and (iii) was to ensure to capture a range of regrowing tree heights. Maps of the selected GEDI RH footprints for field work that includes locations of proximate villages were prepared in the USA and used as a guide during the field work.

4.3.1.2 Field work to quality assess GEDI RH footprint and collect tree species inventory data

Fieldwork was undertaken in the Mai Ndombe REDD+ forest reserve study area from 25 February 2023 to 12 March 2023 to quality assess the selected GEDI RH footprints and undertake tree species inventory in them. The quality assessment of the selected GEDI RH footprints was carried out using a GPS to locate them and all footprints with unfeasible GEDI forest canopy heights due to GEDI footprint geolocation or FCL reporting errors were discarded. Tree species inventory was undertaken with experienced REDD+ forest officers for the remaining GEDI RH footprints that passed the quality check. The tree species inventory data were collected using 15 m radius plot sizes. The 15 m radius size forest inventory plot was chosen because it is the same plot size used in the MNRP REDD+ Monitoring Reporting Verification (MRV) system (WWC-PDD, 2012) that has been successfully certified three times by the Verified Carbon Standard (VCS, 2023) and because it is similar to the GEDI 25 m footprint diameter. The center location of each selected GEDI RH footprint was used as the center of the forest inventory plot and, and in each footprint every tree with dbh > 10 cm was identified and the number of trees with dbh > 10 cm per species were recorded in an identification sheet (see APPENDIX) to capture species dominance, diversity, and composition. Digital photographs were also taken from the center of each plot in the North, South, East, West directions and upwards toward the canopy tops.

4.3.2 Tree Species Composition Analysis

The tree species composition of the regrowing trees in each GEDI footprint was derived using the unique species count data (surveyed as described above). In addition, the dry wood density (*WD*) of every tree species inventoried was assigned using the Global ICRAF Wood

Density Database (ICRAF, 2023). The minimum (WD_{min}) and maximum (WD_{max}) of all the collected data was summarized. In addition, the study area species abundance weighted wood density ($WD_{sp.w}$) was derived as:

$$WD_{sp.w} = \sum_{i=1}^{n} p_i \times WD_i \tag{1}$$

where *n* is the total number of tree species, p_i is the proportion of a given species *i* in all the GEDI RH footprints inventoried, WD_i is the wood density of species *i*.

4.3.3 Forest Height Regrowth and AGB Accumulation Quantification

4.3.3.1 Forest stand age estimation

The forest stand age of all selected GEDI RH footprints was derived using the GEDI footprint acquisition year and the 30 m annual global FCL data. For each footprint, the forest stand age was calculated by subtracting the reported FCL year from the year of the GEDI footprint acquisition (2019, 2020, or 2021) and adding one.

4.3.3.2 Forest height regrowth quantification

The forest height regrowth rate was quantified in a manner similar to that developed by Milenkovic et al. (2022) for Amazon rainforest regrowth assessment. An exponential regression fit with the form $y(x) = a \times (1 - e^{-\frac{x}{b}})$ was used, where *x* is the forest stand age and *y* is the GEDI RH_{95}. Exponential regrowth models have been used to quantify secondary forest regrowth in the tropics (Poorter et al., 2021; Milenkovic et al., 2022). The *a* parameter corresponds approximately to the total height growth and 3*b* to when the model reaches 95% of its asymptotic value (Bivand et al., 2013) and so the average regrowth rate was estimated as *a*/3*b*. This was undertaken considering the raw (all selected GEDI footprint data) and also considering the data aggregated by forest stand age i.e., using the median GEDI RH_{95} values for each unique stand age. The model was fitted multiple times varying *3b* annually from 20 to 30 years and then the model that maximized the R^2 was selected and used in this study. The range of 20 to 30 years was chosen because it is the interval of time taken by trees in tropical Africa to reach full maturity (Swaine et al., 1983; White, 1983).

4.3.3.3 Forest regrowth AGB accumulation quantification

The fitted exponential regression line equation derived from the median GEDI RH_{95} grouped by forest stand age (section 4.3.3.2) was used to predict forest height growth values from 0 to 30 years. The AGB accumulation of each predicted forest height growth value was derived as:

$$AGB = 10.43\ (h \times WD)^{1.19} \tag{2}$$

where *AGB* is the forest stand age forest regrowth AGB accumulation amount in Mg ha^{-1}, *h* is the predicted forest height growth values derived using the fitted exponential regression equation (see section 4.3.4.1), and *WD* is the dry wood density. The allometric equation used was defined by Xu et al. (2017) by statistically fitting 92 pairs of dominant canopy heights with field AGB estimates in DRC. The tree forest regrowth AGB accumulation amount trajectory was generated to examine the sensitivity of forest regrowth AGB accumulation to variation in dry wood density (WD), by application of Equation (2) to the predicted forest height growth values and using separately the study area minimum (WD_{min}), and maximum (WD_{max}) dry wood density values, and the species abundance weighted wood density ($WD_{sp.w}$) derived as Eq.(1).

4.3.4 Species Diversity Analysis

The study area species diversity was estimated at each GEDI RH footprint where tree species inventory was undertaken using the Shannon diversity index (Shannon, 1948) derived as:

$$H = -\sum_{i=1}^{S} p_i \times ln\, p_i \tag{3}$$

where $p_i = n_i/N$, n_i is the number of individual trees of species i, N the total number of trees in the footprint, and S is the total number of species in the footprint. This analysis was undertaken to better understand the pattern of secondary forest succession through time over the study area. The Shannon diversity index has been widely used to measure species diversity of tropical forests (Ruiz et al., 2005; William et al., 2008; Féret and Asner, 2014; Ifo et al., 2016). The Shannon diversity index values derived for each GEDI footprint inventoried were then plotted against forest stand age and an exponential regression model with the form $H(x) = c + a \times (1 - e^{-\frac{x}{b}})$, where x is the forest stand age, and where a+c corresponds approximately to the total species diversity increase, and 3b to when the model reaches 95% of its asymptotic value (Bivand et al., 2013), to characterize the study area species diversity increase rate through time. The goodness of fit (R^2) of the exponential regression and the regression confidence (p value) statistics were also derived.

4.4 Results

4.4.1 Tree Species Inventory

4.4.1.1 GEDI RH footprint selection for field work

A total of 62 GEDI RH footprints that had met all the three selection criteria described in section 4.3.1.1 were selected and visited during the field work. All the selected GEDI footprints were in vicinity of the village Kesenge (red dot in Figure 4.1a), which is close to the city of Inongo (~ 40 min by boat).

4.4.1.2 GEDI RH footprints quality assessment and tree species inventory data collection

Of the 62 selected GEDI RH footprint locations scheduled to be visited during the field work only 51 were accessible and of these 49 were unambiguously in areas of regrowing forest, and 7 had obviously incorrect geolocations resulting in 42 GEDI RH footprint locations hereafter referred to for clarity as the secondary forest GEDI footprints. For each of the 42 secondary forest

GEDI footprints forest inventory data, including tree species and number of individual trees per species, were collected and in total 695 trees belonging to 66 tree species were recorded.

4.4.2 Tree Species Composition Analysis

The analysis of the tree species composition of the 42 secondary forest GEDI footprint inventoried showed that 18 species accounted for 86.4% of the total number of trees surveyed. Figure 4.5 shows the proportion in percentage of each of the 18 tree species. *Macaranga monarda* from the *Euphorbiaceae* family found in 18.1% was the most abundant dominant tree species followed by *Musanga cecropioides* from the *Cecropiaceae* family *and Macaranga spinosa* from the *Euphorbiaceae* family found in 15.5% and 11.3% of the selected sites, respectively. The dry wood density (*WD*) of the 66 tree species ranged from 0.220 g/cm^3 (WD_{min}; *Ricinodendron heudelotii*) to 0.842 g/cm^3 (WD_{max}; *Pentaclethra macrophylla*) (ICRAF, 2023) and the species abundance weighted wood density ($WD_{sp.w}$) derived as Eq. (1) was 0.412 g/cm^3.

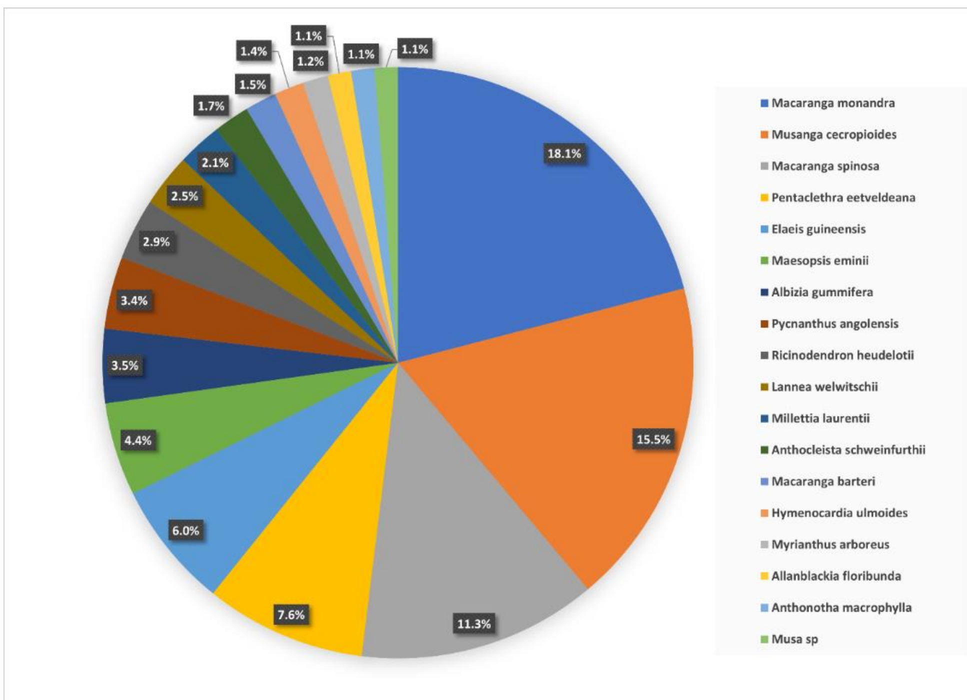

Figure 4.5 Tree species abundance of the 42 secondary forest GEDI footprint inventoried during field work in February 2023.

4.4.3 Forest Height Regrowth and AGB Accumulation Quantification

4.4.3.1 Forest Stand Age Estimation

The forest stand ages of the 42 secondary forest GEDI footprints were derived as described in section 4.3.3.1. There was a total of 14 forest stand age values that ranged from 3 to 21 years with a mean and median of 12 and 13 years, respectively. The forest stand age with the greatest number of sites was 5, 10, and 13 years with 7, 6, and 5 sites, respectively.

4.4.3.2 Forest height regrowth quantification

Figure 4.6 shows the GEDI tree height chronosequence derived for the 42 secondary forest GEDI footprints. The exponential regression fitted to the plotted GEDI tree height chronosequence data used to characterize the regrowth rate is shown in red. The regression was significant (p value < 0.001) with a R^2 value of 0.90. The total height growth (*a*) and the average height regrowth rate (*a*/3*b*) are 18.95 m 0.758 m/year, respectively. Regrowth rates are greater for younger trees.

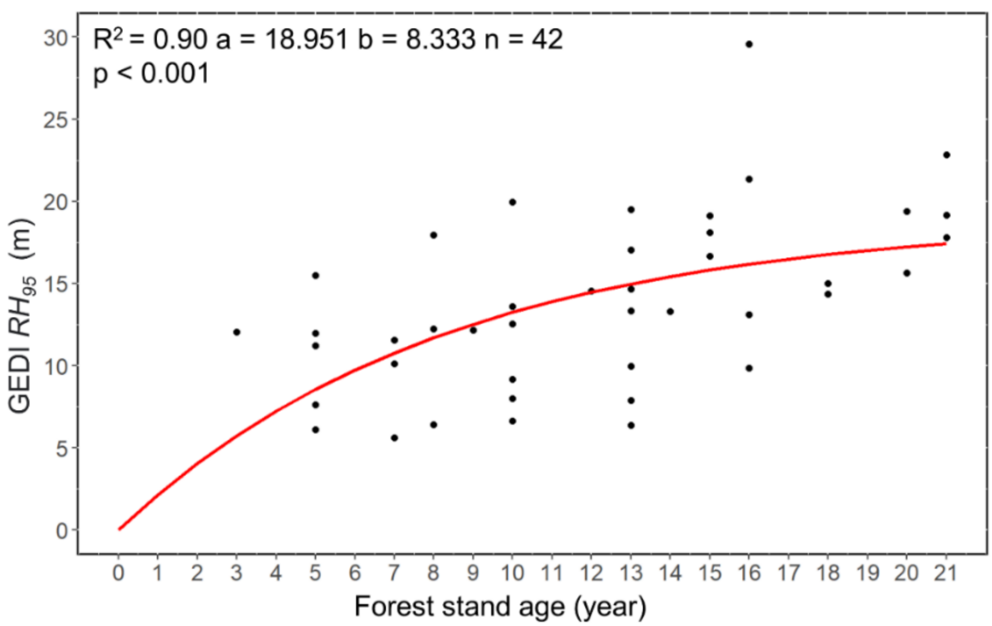

Figure 4.6 Tree height chronosequence showing *GEDI* RH_{95} values of the 42 secondary forest GEDI footprints acquired April 2019 to November 2021 where there was previous mapped 30 m forest cover loss.

Figure 4.7 shows the GEDI tree height chronosequence derived at 42 secondary forest GEDI footprints using the median values of GEDI RH_{95} data aggregated by forest stand. The exponential regression fitted to the plotted GEDI tree height chronosequence data used to characterize the regrowth rate is shown in red. The coefficient of determination (R^2) of regression is 0.97 which is higher than the raw data (Figure 4.6 as expected from statistical theory). The total height growth (*a*) and the average height regrowth rate (*a/3b*) are 19.08 m 0.763 m/year, respectively. Similarly, to Figure 4.6, as expected, height regrowth rates are greater for younger trees.

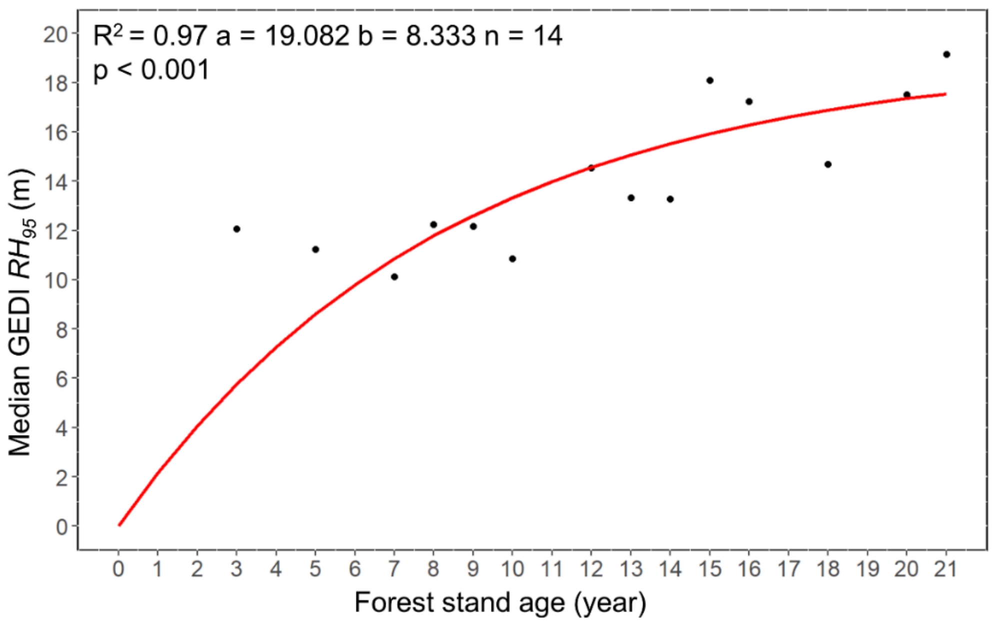

Figure 4.7 Tree height chronosequence showing the median GEDI RH_{95} derived from 42 secondary forest GEDI footprints acquired April 2019 to November 2021 across the forest reserve where there was previous mapped 30 m forest cover loss.

4.4.3.3 Forest AGB accumulation quantification

Figure 4.8 shows the three forest regrowth AGB accumulation trajectories derived by application of Equation (2) to the 30 year of predicted forest height growth values from the fitted exponential regression line equation (red line in Figure 4.7) plotted separately using the WD_{min}

(red line), $WD_{sp.w}$ (green line), and WD_{max} (blue line) to examine the sensitivity of forest regrowth AGB storage with respect to species dry wood density. The mature forest 30 year AGB was 55.76 Mg ha^{-1}, 117.52 Mg ha^{-1}, and 274.52 Mg ha^{-1} for the WD_{min}, $WD_{sp.w}$, and WD_{max} values, respectively.

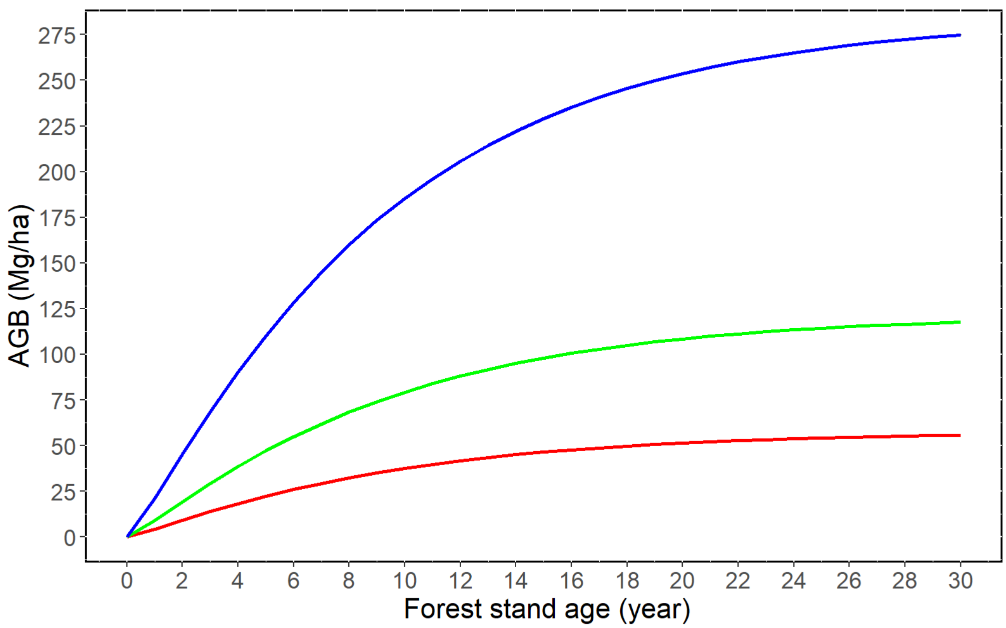

Figure 4.8 Forest regrowth AGB accumulation trajectory derived using 30 years of predicted height growth values generated using the fitted exponential regression model shown in Figure 4.7. Results are shown for AGB accumulation derived using the minimum (red line), maximum (blue line), and species abundance weighted average (green line) study area dry wood densities values.

4.4.4 Species Diversity Analysis

Figure 4.9 shows the Shannon diversity index (*H*) chronosequence derived at the 42 secondary forest GEDI footprints. The exponential regression model fitted to the plotted data are shown in red and was significant (p value < 0.001) with a 0.35 R^2 value. The *H* values ranged from 0 to 2.18 and the total *H* increase (a+c) was 2.07. This suggests that the tree species diversity increases with stand age.

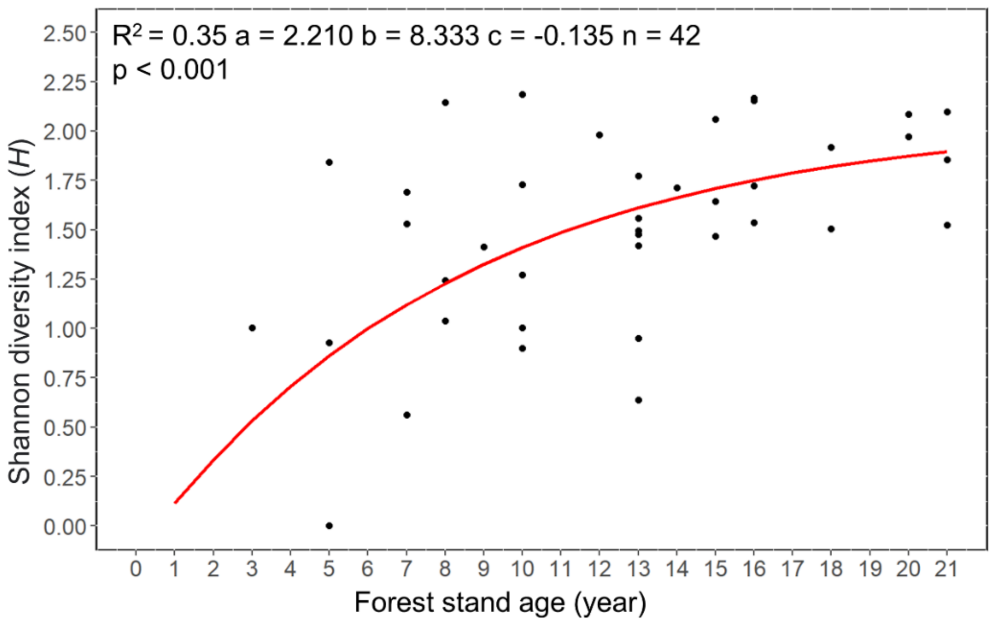

Figure 4.9 Shannon diversity index (H) of species diversity derived for each of the 42 secondary forest GEDI footprint inventoried plotted against forest stand age.

4.5 Discussion

Knowledge of regrowth rates and responses of tropical forests to past forms of disturbances may facilitate our understanding of the capacity of these ecosystems to respond to present and future events (Cole et al., 2014). Secondary forest regrowth rates in Central Africa are poorly documented mainly due to sparse forest inventory data (Baccini et al., 2017; Bauters et al., 2019). It is well established that LiDAR remote sensing can accurately measure forest canopy height and that multi-temporal airborne LiDAR measurements can provide forest regrowth estimation (Guerra-Hernandez et al., 2021; Riofrío et al., 2023). However, this approach to estimate forest growth has not yet been implemented in Central African tropical forests due to the lack of repeat airborne LiDAR measurements. GEDI is the first spaceborne LiDAR specifically designed to improve estimates of vegetation structure in the tropics including in Central Africa where those estimates are very limited. Such data underpin a new era of large-area approaches for measuring forest height in regrowing secondary forests of different ages (Milenkovic et al., 2022). Although there is high confidence that GEDI can retrieve measurements that allow estimation of forest canopy height in the tropics, its ability to quantify forest regrowth in Central Africa has not yet been tested.

In this study the ability of GEDI to quantify forest regrowth in a Central African tropical forest is assessed for the first time considering 42 regrowing forest locations observed by GEDI where previous satellite forest cover loss was mapped (i.e., secondary forest GEDI footprint locations). A GEDI tree height chronosequence was derived using Landsat based annual forest cover loss (FCL) data to locate and date forest disturbance from 2001 to 2018 (i.e., up to the year before the first GEDI acquisitions). The forest stand age was derived by subtracting the reported FCL year from the year of the GEDI footprint acquisition and adding one. GEDI RH_{95} values from the 42 secondary forest GEDI footprints were aggregated by forest stand age and the median values for each forest stand age derived. An exponential regression models was separately fitted to GEDI tree height chronosequence data generated using (i) the 42 GEDI RH_{95} values, and (ii) the median of the 42 GEDI RH_{95} values grouped by forest stand age. The exponential regression model derived using the median of GEDI RH_{95} values performed better (0.97 R^2) than that derived using the GEDI RH_{95} values (0.90 R^2). The average regrowth rate obtained from the median GEDI RH_{95} chronosequence data (0.763 m/year) was very similar to that derived using the GEDI RH_{95} chronosequence data (0.758 m/year) and is comparable to regrowth rates values found in other LiDAR-based studies in Amazon tropical forests. For example, Becknell et al. (2018) estimates forest regrowth using median airborne LiDAR heights and a Landsat-based disturbance history map in the Brazilian Atlantic Forest reporting a regrowth rate of 0.60 m/year on linearly modeled regrowth periods of 20 years. Similarly, Milenkovic et al. (2022) derived forest regrowth rate of Amazon rainforests and reported a forest height regrowth rate of 0.65 m/year on exponentially modeled regrowth derived using a saturation year ($3b$) of 30 years. The exponential regression derived with median GEDI RH_{95} values was used to predict forest height regrowth over 30 years (i.e., the typical time to maturity of Central Africa trees) and AGB accumulation amounts in forest regrowth were estimated using a DRC AGB allometric model parametrized with LiDAR canopy height and wood density. Tree species inventory data were collected over each of the 42 secondary forest GEDI footprints to help contextualize the results and understand the dynamics of AGB accumulation amounts in forest regrowth with respect to species wood densities variation. The tree species inventory results revealed that most of the sites were dominated by light demanding secondary forest pioneer tree species that are characterized by fast height growth rate mainly because of their light dependence. This is not surprising since the forest stand age of most of the secondary forest GEDI footprints considered in this study was lower than 20 years and therefore, they are expected to be dominated by early succession secondary forest tree species (Huntley, 2023). There were only 3 secondary forest GEDI footprints with forest stand age > 20 years. This was also expected as the light demanding secondary forest pioneer tree species commonly found in the study area (i.e., *Musanga cecropioides*, *Macaranga Monandra*, and *Macaranga spinosa)* usually start to die off after 20 years to giving way to shade tolerant and light demanding non-pioneer tree species. The dry wood density of the 66 tree species inventoried in the 42 secondary forest GEDI footprints ranged from 0.220 g/cm^3 (*Ricinodendron heudelotii*) to 0.842 g/cm^3 (*Pentaclethra macrophylla*) (ICRAF, 2023) with a species abundance weighted wood density ($WD_{sp.w}$) of 0.412 g/cm^3.

The modelled 30 year AGB varied significantly with the species dry wood density. Modelling the 30 year AGB using the WD_{min} and WD_{max} provided 30 year AGB estimates of 55.76 Mg ha^{-1} and 274.80 Mg ha^{-1}, respectively, that are considerably different to the 117.52 Mg ha^{-1} 30 year AGB estimates found using $WD_{sp.w}$. The 30 year AGB values (117.52 Mg ha^{-1}) derived using the $WD_{sp.w}$ that is representative of the 42 secondary forest GEDI footprint was further compared to the average AGB reported for primary and mature secondary forest in other Central African forest studies and is equivalent to about 49.39 % of the Xu et al. (2017) reported 237.93 Mg ha^{-1} mean AGB for Mai Ndombe province primary and mature secondary forests, 54% of the Baccini et al. (2008) reported 216 Mg ha^{-1} mean AGB for the evergreen rainforest of Central Africa, and 52.69% of the Silva et al. (2018) reported 223 Mg ha^{-1} mean AGB for Lope National Park in Central Gabon. These proportions are not negligible, and so demonstrate the importance of forest regrowth as an effective pathway to help mitigate climate change.

A species diversity analysis was also undertaken over the 42 secondary forest GEDI footprint using the Shannon diversity index (H) to understand the pattern of secondary succession over the study area and showed a gradual increase rate of species diversity rate with higher species diversity increase rate found for footprints with lower forest stand age (< 15) that start to stabilize at forest stand age > 15. Similar patterns have been observed by others in the tropics. For example, Ruiz et al. (2005) found in Colombian tropical moist forest a gradual increase in the Shannon diversity index from 2.31 for < 6 years old secondary forest to 3.35 for > 50 years old secondary forests. In another example, Lebrija-Trejos et al. (2008) also found a gradual increase in the Shannon diversity index from 1.36 to 3.34 during the 40 years of chronosequence considered. However, in contrast to these studies that have used more than 40 years of chronosequence data, this study only used 20 years of chronosequence with data available only for 14 forest stand age, which may not be sufficient to reliably capture the pattern of the study area secondary forest succession.

The 30 year AGB of forest regrowth reported in this study are not free of uncertainty. These AGB estimates reported in this study are dependent on the DRC tree height AGB allometric equation defined by Xu et al (2017) using undisturbed primary forest AGB field plots, and so may not be suitable for regrowing secondary forests. GEDI has geolocation errors of 10.3 m horizontally that introduces relative canopy height biases over spatially heterogeneous canopies and along forest edges (Roy et al., 2021). Validation of the GEDI relative height product accuracy is ongoing and requires contemporaneous ALS data, for example, Li et al. (2023) reported that for leaf-on southern Africa savannas, the GEDI RH_{98} product cannot reliably estimate canopy heights < 2.34 m (i.e., cannot estimate shrub heights) but can accurately estimate the canopy height of savanna trees. It should be noted that a geolocation error for GEDI tracks was neither estimated nor corrected in this study due to the lack of contemporaneous ALS data, yet it is potentially a significant source of uncertainty (Roy et al., 2021). The GEDI relative canopy height retrieval accuracy depends on several factors such as the GEDI beam sensitivity. GEDI observations with beam sensitivity > 0.9 (i.e., when >90% of the canopy cover is penetrated to the ground by the waveform energy) are more reliable than lower beam sensitivity (Duncanson et al., 2020). All 42 secondary forest GEDI footprints used in this study had beam sensitivity > 0.95. The beam sensitivity is dependent on the algorithm setting group used to control the GEDI waveform interpretation and although only optimal setting group data were used in this study this may not always be correctly defined. Calibration equations developed by OLS regressions between contemporaneous airborne laser scanner (ALS)-derived canopy height model data (y) and GEDI relative height data (x) and then applied to GEDI relative height data have been shown to improve GEDI relative height reporting accuracy and forest regrowth retrievals in the tropics (Milenkovic et al., 2022). However, because of the lack of contemporaneous ALS data over the study area, this process could not be undertaken in this study. Despite these issues, the 30 year AGB values derived using the study area lowest and highest wood density are 47.4% and 233.8% of 117.52 Mg ha^{-1}. This range of differences highlights the need to account for tree species type in carbon forest regrowth sequestration studies that has been ignored in carbon literature to date.

4.6 Conclusion

The ability of GEDI to quantify forest regrowth in Central African tropical forests was for the first time assessed considering 42 25 m GEDI footprints over Mai Ndombe secondary forest. A GEDI tree height chronosequence was derived using the Landsat based annual forest cover loss (FCL) data to locate and date forest disturbance from 2001 to 2018 (i.e., up to the year before the first GEDI acquisitions). Care was taken through field work data collection that only sites with evidence of forest regrowth were used. An exponential regression model quantified forest height regrowth. A DRC AGB allometric model was used to convert the modelled forest height growth into AGB. Tree species inventory was undertaken over all the 42 GEDI footprint locations and used to examine the impact of forest regrowth species differences on AGB accumulation amounts. The dry wood density of the 66 tree species inventoried in the 42 footprints ranged from 0.220 g/cm^3 to 0.842 g/cm^3 with a species abundance weighted wood density ($WD_{sp.w}$) of 0.412 g/cm^3. The 30 year AGB estimate derived using the least (55.76 Mg ha^{-1}) and most dense species (274.80 Mg ha^{-1}) are 47.4% and 233.8% of 117.52 Mg ha^{-1} 30 year AGB derived using $WD_{sp.w}$ (117.52 Mg ha^{-1}). The 117.52 Mg ha^{-1} 30 year AGB estimate is equivalent to about half of the average AGB reported for primary and mature secondary in other Central African forest studies. These results suggest that the AGB stored in forest regrowth is substantial, and that species information is vital for carbon forest regrowth sequestration studies and therefore needs to be taken into account.

REFERENCES

Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D., & Houghton, R. 2017. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science*. 358, 230–234.

Bauters, M., Vercleyen, O., Vanlauwe, B., Six, J., Bonyoma, B., Badjoko, H., & Boeckx, P. 2019. Long-term recovery of the functional community assembly and carbon pools in an African tropical forest succession. *Biotropica*, 51, 319–329.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., & Wood, E.F. 2018. Present and future Koppen-Geiger climate classification maps at 1-km resolution. *Nature Science Data*. 5, 180214.

Beck, J., Wirt, B., Luthcke, S., Hofton, M., & Armston, J. 2021. Global Ecosystem dynamics investigation (GEDI) level 02 user guide. Document version 2.0, April 2021. U.S. *Geological Survey, Earth Resources Observation and Science Center, Sioux Falls, South Dakota, USA.*

Becknell, J.M., Keller, M., Piotto, D., Longo, M., Nara dos-Santos, M., Scaranello, M.A., Bruno de Oliveira Cavalcante, R., Porder, S. 2018. Landscape-scale lidar analysis of aboveground biomass distribution in secondary Brazilian Atlantic Forest. *Biotropica* 50, 520–530.

Bennett, A. C., Dargie, G. C., Cuni-Sanchez, A., Tshibamba Mukendi, J., Hubau, W., Mukinzi, J. M., & Lewis, S. L. 2021. Resistance of African tropical forests to an extreme climate anomaly. *Proceedings of the National Academy of Sciences*. 118, e2003169118.

Bivand, R.S., Pebesma, E., & G´omez-Rubio, V. 2013. Interpolation and geostatistics. In: Bivand, R.S., Pebesma, E., G´omez-Rubio, V. (Eds.), Applied Spatial Data Analysis with R. *Springer New York, New York, NY*, pp. 213–261.

Boehm, H. D. V., Liesenberg, V., & Limin, S. H. 2013. Multi-temporal airborne LiDAR-survey and field measurements of tropical peat swamp forest to monitor changes. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 6, 1524–1530.

Brancalion, P. H., Niamir, A., Broadbent, E., Crouzeilles, R., Barros, F. S., Almeyda Zambrano, A. M., & Chazdon, R. L. 2019. Global restoration opportunities in tropical rainforest landscapes. *Science advances*. 5, eaav3223.

Brienen, R.J.W., & Zuidema, P.A. 2005. Relating tree growth to rainfall in Bolivian rain forests: a test for six species using tree ring analysis. *Oecologia*. 146, 1–12.

Bustamante, M. M., Roitman, I., Aide, T. M., Alencar, A., Anderson, L. O., Aragão, L., & Vieira, I. C. 2016. Toward an integrated monitoring framework to assess the effects of tropical forest degradation and recovery on carbon stocks and biodiversity. *Global Change Biology*. 22, 92–109.

Bwangoy, J.-R.B., Hansen, M.C., Roy, D.P., Grandi, G.D., & Justice, C.O. 2010. Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices. *Remote Sensing of Environment*. 114, 73–86.

Chazdon, R.L., Broadbent, E.N., Rozendaal, D.M., Bongers, F., Zambrano, A., Aide, T.M., Balvanera, P., Becknell, J.M., Boukili, V., Brancalion, P.H., & Craven, D. 2016. Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. *Science Advance*. 2, 1–10.

Cohen, W.B., Yang, Z., Healey, S.P., Kennedy, R.E., & Gorelick, N., 2018. A LandTrendr multispectral ensemble for forest disturbance detection. *Remote Sensing of Environment*. 205, 131–140.

Cole, L.E., Bhagwat, S.A., & Willis, K.J. 2014. Recovery and resilience of tropical forests after disturbance. *Nature Communication*. 5, 3906.

Crezee, B., Dargie, G. C., Ewango, C. E., Mitchard, E. T., Emba B, O., Kanyama T, J., & Lewis, S. L. 2022. Mapping peat thickness and carbon stocks of the central Congo Basin using field data. *Nature Geoscience*. 15, 639–644.

Deckers, J. 1993. Soil fertility and environmental problems in different ecological zones of the developing countries in Sub-Saharan Africa. *The role of plant nutrients and sustainable food production in sub-Saharan Africa*. 37–52.

Donovan, S.D., MacLean, D.A., Zhang, Y., Lavigne, M.B., & Kershaw, J.A., 2021. Evaluating annual spruce budworm defoliation using change detection of vegetation indices calculated from satellite hyperspectral imagery. *Remote Sensing of Environment*. 253, 112204.

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., & Silva, C. 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of remote sensing*. 1, 100002.

Dubayah, R., Hofton, M., Blair, J., Armston, J., Tang, H., & Luthcke, S. 2021. GEDI L2A Elevation and Height Metrics Data Global Footprint Level V002. *NASA EOSDIS Land Processes DAAC*. Accessed 2023-08-19.

Duncanson, L., Neuenschwander, A., Hancock, S., Thomas, N., Fatoyinbo, T., Simard, M., & Dubayah, R. 2020. Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. *Remote Sensing of Environment*. 242, 111779.

Egorov, A., Roy, D.P., & Boschetti, L. 2023. Generation and comprehensive validation of 30 m conterminous United States Landsat percent tree cover and forest cover loss annual products. *Science of Remote Sensing*. 7, 100084.

Fayolle, A., Engelbrecht, B., Freycon, V., Mortier, F., Swaine, M., Rejou-Mechain, M., Doucet, J.L., Fauvet, N., Cornu, G., & Gourlet-Fleury, S. 2012. Geological substrates shape tree species and trait distributions in African moist forests. *PLoS One*. 7, e42381.

Féret, J. B., & Asner, G. P. 2014. Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecological Applications*. 24, 1289–1296.

Ferraz, A., Saatchi, S., Mallet, C., & Meyer, V. 2016. Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment*. 183, 318–333.

Friedlingstein, P., O'sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., & Zheng, B. 2022. Global carbon budget 2022. *Earth System Science Data*. 2022, 1–159.

Gao, Bo-Cai.1996. NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Waterfrom Space. *Remote Sensing of Environment*. 58, 257–266.

Global Forest Change, 2021. Global Forest Change 2000–2021 Data Description Document. *University of Maryland, College Park, MD*. Retrieved from UMD GLAD web platform: https://storage.googleapis.com/earthenginepartners-hansen/GFC-2022-v1.9/download.html

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. 202, 18–27.

Gourlet-Fleury, S., Mortier, F., Fayolle, A., Baya, F., Ouedraogo, D., et.al. 2013. Tropical Forest recovery from logging: a 24 year silvicultural experiment from Central Africa. *Philosophical Transactions of the Royal Society Biological Science*. 368, 20120302.

Groenendijk, P., Sass-Klaassen, U., Bongers, F., & Zuidema, P.A. 2014. Potential of tree-ring analysis in a wet tropical forest: A case study on 22 commercial tree species in Central Africa. *Forest Ecology and Management*. 323, 65–78.

Guerra-Hernandez, J., & Pascual, A. 2021. Using GEDI lidar data and airborne laser scanning to assess height growth dynamics in fast-growing species: a showcase in Spain. *Forest Ecosystems*. 8, 1-17.

Hansen, M. C., Potapov, P., Moore, R., Hancher, M., Turubanova, S., Tyakavina, A., &Townsend, A. R. 2013. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*. 342. Data available on-line from: http://earthenginepartners.appspot.com/science-2013-global-forest

Hansen, M.C., Egorov, A.V., Potapov, P.V., Stehman, S.V., Tyukavina, A., Turubanova, S. A., Roy, D.P., Goetz, S.J., Loveland, T.R., Ju, J., Kommareddy, A., Kovalskyy, V., Forsyth, C., & Bents, T., 2014. Monitoring conterminous United States (CONUS) land cover change with web-enabled Landsat data (WELD). *Remote Sensing of Environment*, 140. 466–484.

Hansen, M.C., Krylov, A., Tyukavina, A., Potapov, P.V., Turubanova, S., Zutta, B., Ifo, S., Margono, B., Stolle, F., & Moore, R. 2016. Humid tropical forest disturbance alerts using Landsat data. *Environmental Research Letters*. 11, 034008.

Hardisky, M.A., Klemas, V., Smart, R., 1983. The influence of soil salinity, growth form, and leaf moisture on-the spectral radiance of Spartina alterniflora canopies. *Photogrammetric Engineering & Remote Sensing*. 49, 77–83.

Harris, N. L., Gibbs, D. A., Baccini, A., Birdsey, R. A., De Bruin, S., & Tyukavina, A. 2021. Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change*. 11, 234–240.

Heinrich, V. H., Vancutsem, C., Dalagnol, R., Rosan, T. M., Fawcett, D., Silva-Junior, C. H., & Aragão, L. E. 2023. The carbon sink of secondary and degraded humid tropical forests. *Nature*. 615, 436–442.

Herold, M., Carter, S., Avitabile, V., Espejo, A.B., Jonckheere, I., Lucas, R., McRoberts, R.E., Næsset, E., Nightingale, J., & De Sy, V. 2019. The Role and Need for Space-Based Forest Biomass-Related Measurements in Environmental Management and Policy. *Surveys in Geophysics*. 40, 757–778.

Hubau, W., Lewis, S. L., Phillips, O. L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., & Zemagho, L. 2020. Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature*. 579, 80–87.

Hughes, M.J., Kaylor, S.D., & Hayes, D.J. 2017. Patch-based Forest change detection from Landsat time series. *Forests*. 8, 166.

Huntley, B. J. 2023. The Guineo-Congolian Rain Forest Biome. In *Ecology of Angola: Terrestrial* Biomes and Ecoregions (pp279-304). Cham: *Springer International Publishing.*

ICRAF. 2023. ICRAF Wood Density Database. Available from ICRAF web platform at http://db.worldagroforestry.org//wd (Last accessed: 21 September 2023)

Ifo, S. A., Moutsambote, J. M., Koubouana, F., Yoka, J., Ndzai, S. F., Bouetou-Kadilamio, L. N. O.,& Joel, L. J. 2016. Tree species diversity, richness, and similarity in intact and degraded forest in the tropical rainforest of the Congo Basin: case of the forest of Likouala in the Republic of Congo. *International Journal of Forestry Research*. 2016.

Jin, S., & Sader, S.A. 2005. Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote Sensing of Environment*. 94, 364–372.

Lebrija-Trejos, E., Bongers, F., Pérez-García, E. A., & Meave, J. A. 2008. Successional change and resilience of a very dry tropical deciduous forest following shifting agriculture. *Biotropica*. 40, 422–431.

Lennox, G. D., Gardner, T. A., Thomson, J. R., Ferreira, J., Berenguer, E., Lees, A. C., & Barlow, J. 2018. Second rate or a second chance? Assessing biomass and biodiversity recovery in regenerating Amazonian forests. *Global Change Biology*. 24, 5680–5694.

Lewis, S. L., Wheeler, C. E., Mitchard, E. T., & Koch, A. 2019. Restoring natural forests is the best way to remove atmospheric carbon. *Nature*. 568, 25–28.

Li, X., Wessels, K., Armston, J., Hancock, S., Mathieu, R., Main, R., Naidoo, L., Erasmus, B., & Scholes, R. 2023. First validation of GEDI canopy heights in African savannas. *Remote Sensing of Environment*. 285, 113402.

Ligot, G., Fayolle, A., Gourlet-Fleury, S., Dainou, K., Gillet, J.-F., De Ridder, M., Drouet, T., Groenendijk, P., & Doucet, J.-L. 2019. Growth determinants of timber species Triplochiton scleroxylon and implications for forest management in central Africa. *Forest Ecology and Management*. 437, 211–221.

Madron, L.D. 2003. Accroissements Diamétriques du bété and iroko in Central African. Tropical rainforest. *CIRAD Bois et Forêts Des Tropiques.* 275, 83–87.

Madron, L.D., Nasi, R., & Detienne, P. 2000. Accroissements Diamétriques De quelque Essence en Forêt Dense Africaine. *CIRAD Bois et Forêts Des Tropiques.* 263, 63-74.

Milenković, M., Reiche, J., Armston, J., Neuenschwander, A., De Keersmaecker, W., Herold, M., & Verbesselt, J. 2022. Assessing Amazon rainforest regrowth with GEDI and ICESat-2 data. *Science of Remote Sensing*. 5, 100051.

Mills, M. B., Malhi, Y., Ewers, R. M., Kho, L. K., Teh, Y. A., Both, S., & Riutta, T. 2023. Tropical forests post-logging are a persistent net carbon source to the atmosphere. *Proceedings of the National Academy of Sciences*. 120, e2214462120.

Molinario, G., Hansen, M.C., & Potapov, P.V. 2015. Forest cover dynamics of shifting cultivation in the Democratic Republic of Congo: a remote sensing-based assessment for 2000–2010. *Environmental Research Letters.* 10, 094009.

Poorter, L., Craven, D., Jakovac, C.C., van der Sande, M.T., Amissah, L., Bongers, F., & Norden, N. 2021. Multidimensional tropical forest recovery. *Science*. 374, 1370-1376.

Potapov, P., Hansen, M.C., Kommareddy, I., Kommareddy, A., Turubanova, S., Pickens, A., Adusei, B., Tyukavina, A., & Ying, Q. 2020. Landsat Analysis Ready Data for Global Land Cover and Land Cover Change Mapping. *Remote Sensing*, 12, 426.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., & Hofton, M. 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*. 253, 112165.

Riofrío, J., White, J.C., Tompalski, P., Coops, N.C., & Wulder, M.A. 2023. Modelling height growth of temperate mixedwood forests using an age-independent approach and multi-temporal airborne laser scanning data. *Forest Ecology and Management*. 543, 121137

Ross, R. 1954. Ecological Studies on the Rain Forest of Southern Nigeria III Secondary succession in the Shashaa Forest reserve. *Journal of Ecology*. 42, 259–282.

Roy, D. P., Qin, Y., Kovalskyy, V., Vermote, E. F., Ju, J., Egorov, A., & Yan, L. 2014. Conterminous United States demonstration and characterization of MODIS-based Landsat ETM+ atmospheric correction. *Remote Sensing of Environment*. 140, 433–449.

Roy, D.P., Kashongwe, H.B., & Armston, J. 2021. The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. *Science of Remote Sensing*. 4,100024.

Ruiz, J., Fandiño, M.C., & Chazdon, R.L. 2005. Vegetation Structure, Composition, and Species Richness Across a 56-year Chronosequence of Dry Tropical Forest on Providencia Island, Colombia1. *Biotropica.* 37, 520–530

Shannon, C. E. 1948. A mathematical theory of communication. *The Bell system technical journal*. 27, 379–423.

Samba, G., Nganga, D., & Mpounza, M. 2007. Rainfall and temperature variations over Congo-Brazzaville between 1950 and 1998. *Theoretical and Applied Climatology*. 91, 85–97.

Sandker, M., Carrillo, O., Leng, C., Lee, D., d'Annunzio, R., & Fox, J. 2021. The Importance of High–Quality Data for REDD+ Monitoring and Reporting. *Forests.* 12,1–99.

Schneider, F. D., Ferraz, A., Hancock, S., Duncanson, L. I., Dubayah, R. O., Pavlick, R. P., & Schimel, D. S. 2020. Towards mapping the diversity of canopy structure from space with GEDI. *Environmental Research Letters*. 15, 115006.

Shapiro, A.C., Aguilar-Amuchastegui, N., Hostert, P., & Bastin, J.F. 2016. Using fragmentation to assess degradation of forest edges in Democratic Republic of Congo. *Carbon Balance Management*. 11, 11.

Silva, C.A., Saatchi, S., Garcia, M., Labriere, N., Klauberg, C., Ferraz, A., Meyer, V., Jeffery, K.J., Abernethy, K., White, L., Zhao, K., Lewis, S.L., & Hudak, A.T. 2018. Comparison of Small- and Large-Footprint Lidar Characterization of Tropical Forest Aboveground Structure and Biomass: A Case Study From Central Gabon. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 11, 3512–3526.

Swaine, M.D., & Hall, J.B. 1983. Early Succession on Cleared Forest Land in Ghana. *Journal of Ecology*. 71, 601–627.

Tyukavina, A., Hansen, M. C., Potapov, P., Parker, D., Okpa, C., Stehman, S. V., & Turubanova, S. 2018. Congo Basin forest loss dominated by increasing smallholder clearing. *Science Advances*. 4, eaat2993.

Vancutsem, C., Achard, F., Pekel, J. F., Vieilledent, G., Carboni, S., Simonetti, D., & Nasi, R. 2021. Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Science Advances*. 7, eabe1603.

VCS- Verified Carbon Standard. 2023. Verified REDD+ project database Available at: https://registry.verra.org/app/search/VCS/Registered. (Last accessed: 21 August 2023).

Verbeeck, H., Boeckx, P., & Steppe, K. 2011. Tropical forests: include Congo basin. *Nature*. 479, 179–179

Wang, Y., Ziv, G., Adami, M., Mitchard, E., Batterman, S. A., Buermann, W., & Galbraith, D. 2019. Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery. *Remote sensing of environment*. 221, 474–488.

White, F. 1983. The vegetation of Africa. *United Nations Educational, Scientific and Cultural Organization (UNESCO)*. 20, 1–354.

White, L. J., Masudi, E. B., Ndongo, J. D., Matondo, R., Soudan-Nonault, A., Ngomanda, A., & Lewis, S. L. 2021. Congo Basin rainforest—invest in science. *Nature*. 598, 411– 414.

Williams, M. R. C. M., Ryan, C. M., Rees, R. M., Sambane, E., Fernando, J., & Grace, J. 2008. Carbon sequestration and biodiversity of re-growing miombo woodlands in Mozambique. *Forest Ecology and management*. 254, 145–155.

Wimberly, M.C., Dwomoh, F.K., Numata, I., Mensah, F., Amoako, J., Nekorchuk, D.M., & McMahon, A. 2022. Historical trends of degradation, loss, and recovery in the tropical forest reserves of Ghana. *International Journal of Digital Earth*. 15, 30–51.

Worbes, M., Staschel, R., Roloff, A., & Junk, W.J. 2003. Tree ring analysis reveals age structure, dynamics and wood production of a natural forest stand in Cameroon. *Forest Ecology and Management*.173, 105–123.

Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., & Gobakken, T. 2012. Lidar sampling for large-area forest characterization: A review. *Remote sensing of environment*. 121, 196–209.

Wulder, M. A., Roy, D. P., Radeloff, V. C., Loveland, T. R., Anderson, M. C., Johnson, D. M., & Cook, B. D. 2022. Fifty years of Landsat science and impacts. *Remote Sensing of Environment*. 280, 113195.

WWC-CCB. 2012. The Mai Ndombe REDD+ Project Climate, Community, and Biodiversity Standards Report 2nd Edition (CCB_PROJ_DESC_934_31OCT2012). *Verified Carbon Standard web platform* available at https://registry.verra.org/app/projectDetail/VCS/934/ . (Last accessed: 21 August 2023).

WWC-PDD. 2012. The Mai Ndombe REDD+ Project Description and Design Report (PROJ_DESC_934_19NOV2012). *Verified Carbon Standard web platform* available at https://registry.verra.org/app/projectDetail/VCS/934/. (Last accessed: 21 August 2023).

Xu, L., Saatchi, S. S., Shapiro, A., Meyer, V., Ferraz, A., Yang, Y., & Ebuta, D. 2017. Spatial Distribution of Carbon Stored in Forests of the Democratic Republic of Congo. *Nature Scientific Report*, 7, 15030.

Zuideema, P.A, & Jakovac, C.C. 2023. Regrowing tropical forests are a sizeable carbon sink, *Nature*, 615, 398-399.

APPENDIX

LAND COVER STATUS IDENTIFICATION SHEET	
Site General Information	
Date	
Strat time / End time	/
Crew member name	
Village	
Site name	
GPS Latitude	
GPS Longitude	
Elevation	
Land cover status (circle one) If other describe the current land cover	(1) Forest
	(2) Regrowing forest
	(3) Other:
Digital Photo ID # **N:** **E:** **S:** **W:** **Canopy:**	
If Land cover status is (1) or (2) collect the site tree species composition over a 15 m radius plots around the site GPS Latitude and Longitude	
Tree Species Diversity and Composition	
Tree species name	**Number of trees**

CHAPTER 5. RESEARCH SUMMARY AND RECOMMENDATIONS

5.1 Summary of Research Hypotheses

The findings of the three research hypotheses are summarized below:

Hypothesis 1 (Chapter #2): Satellite derived reflectance has been used to predict tree height with models trained using LiDAR data. However, reflectance saturation occurs over dense vegetation conditions and factors, including the degree of canopy cover, the leaf area index, soil background reflectance, understory vegetation, and shadows, modify reflectance. Typically, mature tree stands in central African tropical forests have similar reflectance throughout the year whereas stands in secondary forests, that are characterized by trees with lower height, lower canopy cover and so more apparent understory vegetation, have more evident wet and dry season reflectance differences. **Therefore, using dry and wet season satellite reflectance imagery is hypothesized to significantly improve forest height prediction over Mai Ndombe province compared to using single season images.**

This hypothesis *was negated*. More accurate 30 m forest canopy height mapping results were obtained using both wet and dry season Landsat-8 OLI images, but the accuracy was only marginally improved over the mapping results obtained using single season images. Specifically, comparing predicted 30 m canopy heights with 2,639 independent 30 m heights provided Root Mean Square Error (RMSE) values of 3.84 m (both images), 4.17 m (dry season image), and 4.43 m (wet season image). This < 60 cm RMSE difference between image models is less than one percent of the 20.4 m (both images) 20.6 m (dry season image), and 20.8 m (wet season image) average mapped study area forest canopy height and so does not constitute a significant difference. Not surprisingly, when the canopy heights were converted to AGB there were also only small differences between the estimated mean study area AGB varying from 204 Mg ha^{-1} (both images), 206 Mg ha^{-1} (dry season image), 211 Mg ha^{-1} (wet season image).

These results reinforce the importance of GEDI for forest canopy height mapping, particularly as a recent GEDI relative height product validation study, undertaken in Southern Africa savannas, reported a 1.64 m RMSE (Li et al., 2023) that is less than the RMSE obtained using both Landsat-8 OLI images.

<u>Hypothesis 2 (Chapter #3):</u> The Global Ecosystem Dynamics Investigation (GEDI) LiDAR onboard the International Space Station (ISS) is a sampling instrument used to generate 25 m footprint relative height and aboveground biomass (AGB) products for nearly 4 years. Financial support for Reducing Emission from Deforestation and Degradation (REDD+) projects requires measured, reported, and verified estimates of AGB to quantify avoided emissions achieved by conservation of forest biomass in secondary forest dominated landscapes. The Intergovernmental Panel on Climate Change's (IPCC) good practice guidance recommends 10% uncertainty for REDD+ forest AGB reporting. Recent studies have used six months of GEDI data to produce wall-to-wall forest canopy height and AGB maps at national to near global scale. Although there is high confidence that GEDI can retrieve measurements that allow estimation of AGB at scale, less is known about how well its operational deployment performs for measurement of AGB to support REDD+ projects. **Using a six month time period of GEDI observations is hypothesized to be sufficient for forest AGB assessment with 10% uncertainty at REDD+ project scale in Mai Ndombe province.**

This hypothesis was *negated*. A total of 31 months of GEDI AGB good quality data selected across 15 REDD+ project scale (50 × 50 km) carefully selected forested sites in Mai Ndombe province found that >3% of the 30 m forest pixels across each site had GEDI good quality AGB footprint values. Modelling different combinations of GEDI orbits for the different sites indicated that 17, 41 and 57 GEDI orbits were needed to characterize the overall mean GEDI AGB (OMGA) to within ±5%, ±10%, and ±20%, respectively.

Considering 778 sites across the province, an average of 13.03 days was required to obtain a single GEDI orbit containing good quality forest AGB data. This value was used to convert the orbit count results to time units and indicated that 0.61, 1.46, and 2.03 years are needed to characterize the overall mean GEDI AGB (OMGA) to within ±5%, ±10%, and ±20% respectively. Thus, more than six months of GEDI observations are needed for forest AGB assessment with 10% uncertainty at REDD+ project scale in Mai Ndombe province. The analysis assumed that the GEDI AGB product was correct. Notably, considering all 31 months the mean site AGB varied from 89.7 Mg ha^{-1} to 197.1 Mg ha^{-1} among the 15 sites and these values are lower than many central African tropical forest AGB values estimated without using GEDI data (Baccini et al., 2008; Lewis et al., 2013). A new version of the GEDI AGB product is planned (Dr John Amstron – GEDI science team, Personal communication (email), 20 April 2023). This new product may capture higher and more AGB spatial variation and so a longer GEDI observation period may be needed.

As tropical forest AGB inventory data are expensive and hard to collect *in situ* (Chave et al., 2019; Köhl et al., 2020) these results are promising given the relatively short time periods found. In other central African tropical forest localities, the appropriate GEDI observation period may be different depending not just on the forest AGB and spatial variation but also on the cloud cover, ephemeral surface water presence, and GEDI AGB retrieval sensitivity to the forest conditions. Moreover, the appropriate GEDI observation period to characterize AGB when forest change (deforestation, degradation, afforestation, and post-disturbance regrowth) is ongoing is unresolved.

<u>Hypothesis *3 (Chapter #4):*</u> Forests recovering from deforestation and forest degradation provide a potential natural climate mitigation strategy that is unaccounted for in Reducing Emission from Deforestation and Degradation (REDD+) project reporting. The GEDI relative height product enables tree height chronosequences to be generated and has been used recently to estimate Amazonian secondary forest regrowth rates without considering tree species. This has not been undertaken in central Africa where secondary forests regrowth rates are particularly poorly documented. **The impact of forest regrowth species differences on AGB accumulation in mature trees (at 25 m GEDI footprint scale) is hypothesized to be less than the 10% REDD+ forest AGB reporting uncertainty in Mai Ndombe province.**

This hypothesis was *negated*. Analysis of a tree height chronosequence derived using GEDI relative heights at 42 25 m diameter footprints where tree species inventory was undertaken provided a 117.52 Mg ha^{-1} 30 year AGB estimate. This value was derived using the species abundance weighted wood density (0.412 g/cm^3) derived considering the 66 species inventoried in the 42 footprints. Modelling the 30 year AGB using the lowest species wood density (*Ricinodendron heudelotii*, 0.220 g/cm^3) and the highest species wood density (*Pentaclethra macrophylla,* 0.842 g/cm^3) provided 30 year AGB estimates of 55.76 Mg ha^{-1} and 274.80 Mg ha^{-1}, respectively, that are considerably more than 10% different to 117.52 Mg ha^{-1}.

These model estimates are dependent on the DRC tree height AGB allometric equation defined by Xu et al (2017), the accuracy of the GEDI relative height data, the temporal reporting accuracy of the global Forest Cover Loss product, the appropriateness of the exponential regrowth model fit to the median relative height footprint values, and the assumption that the trees reached maturity at 30 years. Despite these issues, the 30 year AGB values derived using the least and most dense species are 47.4% and 233.8% of 117.52 Mg ha^{-1}. This range highlights the need to account for tree species type in carbon forest regrowth sequestration studies that has been ignored in carbon literature to date.

5.2 Recommendations for Future Research

Recommendations for future work, building on this dissertation research, that could increase the level of scientific understanding of the carbon sequestration potential of Central African forest regrowth are discussed below.

- **Improved forest canopy height mapping using more and other medium resolution satellite data**

The Chapter 2 results showed that forest canopy height and so AGB estimates were improved when both wet and dry seasons Landsat-8 OLI images were used. The western part of Democratic Republic of the Congo (DRC) is one of the cloudiest tropical regions in the world (Dommo et al., 2018) and for example there were only two cloud-free (< 20% cloud cover) Landast-8 images sensed there in a 2-year period. Researchers classify multi-temporal metrics of Landsat single pixel time series to help reduce cloud issues (Hansen et al., 2016; Ergorov et al., 2023). For example, Xu et al. (2017) mapped tree height across the DRC using the medians (i.e., 50th percentiles) of the red, near infrared (NIR), and the two shortwave Landsat-8 Operational Land Imager (OLI) reflective wavelength bands acquired over three years, with airborne laser scanner (ALS) tree height training data. Using longer time periods of Landsat data may increase the possibility of obtaining cloud-free observations but unfortunately may also capture surface change. Therefore, further research using a richer time series of medium resolution satellite data such as Landsat-9 (Masek et al., 2020) and/or Sentinel-2A/B (Drusch et al., 2012) to provide more could-free observations is recommended.

- **Improved forest canopy height prediction model**

Only a limited amount of airborne laser scanning (ALS) data were available over the study area that were used to train the forest canopy height prediction model (Chapter 2). In general, classifiers perform better when trained with more data and using training data that are representative of the landscape (Wulder et al. 2018). A total of four 10 km × 2 km ALS transects, corresponding to 0.8% of Mai Ndombe, were used to derive tree height training data. Using more ALS data is recommended to better capture forest canopy height variation and so provide a more robust tree height prediction model.

- **Improved landscape scale forest cover loss information**

The forest stand age used in Chapter 4 to help quantify forest regrowth was derived using the reported year of forest cover loss (FCL) defined in the 30 m Global FCL product (Hansen et al., 2013). Frequent cloud cover means that the temporal reporting accuracy of the global FCL product, which is based on an examination of Landsat-7 and -8 time series, may not report the year of forest loss correctly, particularly for years before the 2013 Landsat-8 launch when the global FCL product was derived using only Landsat-7 due to missing observations caused by the Landsat-7 Scan Line Corrector (Markham et al., 2004). The FCL product has a globally reported 88% omission and 87% commission error (Hansen et al., 2013). Omission errors, i.e., locations where forest cover loss failed to be detected by the 30 m FCL product, may be significant in central Africa where deforestation is often due to small-scale agriculture and manual clearing of individual trees (Tyukavina et al., 2018) that may go undetected at 30 m resolution. However, omission errors are not an issue for forest regrowth studies as GEDI relative height values at missing FCL locations will not be included in the tree height chronosequence derived to quantify forest regrowth. Commission errors, i.e., locations where the 30 m FCL product incorrectly detected forest loss will introduce errors that will need to be accounted for in forest regrowth studies. Using high-resolution commercial satellites data such as PlanetScope (Roy et al., 2021a) and BlackSky (Vrabel et al., 2023) to identify commission errors in the 30 m FCL product is recommended.

- **Improved GEDI relative height product**

The GEDI relative height product was used in Chapter 4 to quantify forest regrowth in Mai Ndombe province. The product accuracy depends on several factors. GEDI coverage beam data, daytime GEDI data, and GEDI observations where < 90% of the canopy cover is penetrated to the ground by the waveform energy (i.e., beam sensitivity < 0.9), may be less reliable (Hofton et al., 2019; Beck et al., 2021). For example, GEDI beam sensitivity retrieval is dependent on the algorithm setting group used to control the GEDI waveform interpretation and although optimal setting group data were used in this dissertation this may not always be correctly defined. In addition, GEDI has geolocation errors (10.3 m) that introduce relative canopy height biases over spatially heterogeneous canopies and along forest edges (Roy et al., 2021b; Beck et al., 2021). Different strategies have been used to handle these issues. For example, Li et al., (2023) used the *collocateWaves* program described in Hancock et al. (2019) that maximizes the correlation between along-track segments of on-orbit and simulated GEDI waveforms to determine the horizontal and vertical offset between the GEDI and airborne laser scanning (ALS) datasets (Hancock et al., 2019) and correct GEDI geolocation accuracy across diverse savanna vegetation in southern Africa. In another example, Milenkovic et al. (2022) performed a linear regression analysis between GEDI canopy heights and their corresponding contemporaneous airborne LiDAR canopy heights to derive a calibration model that was used to correct GEDI canopy heights values in dense tropical forests. However, both methods require contemporaneous ALS data that are rare in central Africa. Collections of ALS data are increasing globally as the acquisition cost of ALS data decrease, so as airborne laser scanner data are becoming future research to test these two approaches in central African tropical forests, to see if the quality of GEDI data can be improved, is strongly recommended.

- **Improved GEDI footprint-level allometric AGB for central African tropical forests**

The GEDI footprint-level allometric AGB product is derived by applying an allometric model to relative canopy height estimates retrieved from each GEDI footprint waveform (Duncanson et al., 2022). The models were defined for each continent and up to four plant functional types by statistical comparison of AGB ground-based estimates with GEDI waveform relative heights simulated from ALS data (Duncanson et al., 2022). The GEDI AGB product available over central Africa was derived using an African Evergreen Broadleaf Tree (EBT) model with a low reported 66.9% RMSE fit uncertainty and no reported bias term. In addition, the GEDI AGB data may not be reliably defined for low AGB conditions, typically outside of the range of the ground-based forest AGB data used to derive the EBT model. Future research to improve the EBT GEDI allometric AGB model using more and representative training data (i.e., ground-based forest AGB data and contemporaneous ALS data) is recommended.

- **Use longer time period of GEDI observations to better capture regrowing forest locations.**

A total of 31 months (18 April 2019 to 13 November 2021) of GEDI foot-print level AGB data were considered in Chapter 3 to examine the number of GEDI orbits and amount of time required to characterize tropical forest AGB in Mai Ndombe province. The GEDI mission was planned to last for two years (Dubayah et al., 2020) but collected four years of data (18 April 2019 to 16 March 2023). The instrument was then stored in March 2023 on the International Space Station and is planned to restart collecting data in fall 2024. Considering four years of GEDI data and/or the 2024 onward data is recommended to better capture forest stand age variation and so provide more complete GEDI tree height chronosequence data.

- **Species specific ontogenetic model of forest regrowth**

GEDI tree height chronosequences data were used by Milenkovic et al. (2022) and in Chapter 4 to quantify forest regrowth in the Amazon and in Mai Ndombe province, respectively. In Chapter 4 an exponential regrowth model was fitted multiple times varying the model parameter that controlled the mature tree height age (when the exponential model reached 95% of its asymptotic value) from 20 to 30 years and then selecting the final model that maximized the regression coefficient of determination (R^2). Milenkovic et al. (2022) compared three types of models (spherical, exponential, and logistic) and found that all three models were very similar with R^2 differences < 0.05. Rather than using models based on the practical requirements of data fitting, it may be more appropriate to use models that have biological meaning. Therefore, further research to investigate different model forms guided by the ontogenetic tree growth literature (Zeide, 1993; Hérault et al., 2011; Scalon et al., 2022) is recommended. Given that species differences have an important impact on the AGB accumulation amounts in forest regrowth, species-specific forest regrowth studies are recommended.

- **Assessment of carbon losses from necromass and soil in central Africa regrowing forests**

Recent findings using forest AGB plots and an eddy covariance tower to quantify and partition net ecosystem CO_2 exchange in Borneo tropical forests revealed that regrowing forests from deforestation and forest degradation were not a carbon sink but a net carbon source instead due to significant carbon emissions from soil organic matter and deadwood in logged forests (Mills et al., 2023), so future research to assess this in central African tropical forests is recommended.

- **Global context**

Central African forests constitute the second largest continuous block of tropical forests in the world, and the majority are in the Democratic Republic of the Congo (Vancutsem et al. 2021). The contribution of central African forest regrowth in the global carbon cycle is not well known (Bauters et al., 2019) for several reason including difficulties to reliably date forest cover loss occurrence and obtaining sufficient tree height and/or stem diameter samples to quantify forest regrowth at landscape scale. The carbon sequestration potential of forest regrowth was quantified at 42 GEDI relative height footprints acquired over Mai Ndombe secondary forest (Chapter 4). A similar approach could be expanded to quantify the carbon sequestration potential of forest regrowth in all the DRC by deriving a GEDI tree height chronosequence for all the DRC and calculating the proportion of the total carbon stored in DRC forest regrowth relative to the global terrestrial carbon stock. This implies the need to expand this dissertation research at a national scale to quantify the carbon sequestration potential of all DRC forest regrowth.

REFERENCES

Baccini, A., Laporte, N., Goetz, S. J., Sun, M., & Dong, H. 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters*. *3*, 045011.

Bauters, M., Vercleyen, O., Vanlauwe, B., Six, J., Bonyoma, B., Badjoko, H., & Boeckx, P. 2019. Long-term recovery of the functional community assembly and carbon pools in an African tropical forest succession. *Biotropica*. 51, 319–329.

Beck, J., Wirt, B., Luthcke, S., Hofton, M., Armston, J. 2021. Global Ecosystem dynamics investigation (GEDI) level 02 user guide. Document version 2.0, April 2021. U.S. *Geological Survey, Earth Resources Observation and Science Center, Sioux Falls, South Dakota, USA*.

Chave, J., Davies, S. J., Phillips, O. L., Lewis, S. L., Sist, P., Schepaschenko, D., & Saatchi, S. 2019. Ground data are essential for biomass remote sensing missions. *Surveys in Geophysics*. 40, 863–880.

Dommo, A., Philippon, N., Vondou, D.A., Sèze, G., & Eastman, R. 2018. The June–September Low Cloud Cover in Western Central Africa: Mean Spatial Distribution and Diurnal Evolution and Associated Atmospheric Dynamics. *Journal of Climate*. 31, 9585–9603.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., & Bargellini, P. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*. *120*, 25–36.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., & Silva, C. 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*. 1, 100002.

Duncanson, L., Kellner, J. R., Armston, J., Dubayah, R., Minor, D. M., Hancock, S., & Zgraggen, C. 2022. Aboveground biomass density models for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar mission. *Remote Sensing of Environment*. 270, 112845.

Egorov, A., Roy, D.P., & Boschetti, L. 2023. Generation and comprehensive validation of 30 m conterminous United States Landsat percent tree cover and forest cover loss annual products. *Science of Remote Sensing*. 7, 100084.

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

Hansen, M. C., Krylov, A., Tyukavina, A., Potapov, P. V., Turubanova, S., Zutta, B., & Moore, R. 2016. Humid tropical forest disturbance alerts using Landsat data. *Environmental Research Letters*. 11, 034008.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., & Townshend, J.R.G. 2013. High-resolution global maps of 21st-century forest cover change. *Science*. 342, 850–853.

Hérault, B., Bachelot, B., Poorter, L., Rossi, V., Bongers, F., Chave, J., & Baraloto, C. 2011. Functional traits shape ontogenetic growth trajectories of rain forest tree species. *Journal of Ecology*. 99, 1431-1440.

Hofton, M., Blair, B., Story, S., & Yi, D. 2019. Algorithm Theoretical Basis Document (ATBD) for GEDI Transmit and Receive Waveform Processing for L1 and L2 Products. *Goddard Space Flight Center, Greenbelt, MD, USA.*

Köhl, M., Neupane, P.R., & Mundhenk, P. 2020. REDD+ measurement, reporting and verification – A cost trap? Implications for financing REDD+MRV costs by result-based payments. *Ecological Economics.* 168, 106513.

Lewis, S. L., Sonké, B., Sunderland, T., Begne, S. K., Lopez-Gonzalez, G., Van Der Heijden, G. M., & Zemagho, L. 2013. Above-ground biomass and structure of 260 African tropical forests. *Philosophical Transactions of the Royal Society B: Biological Sciences*. *368*, 20120295.

Li, X., Wessels, K., Armston, J., Hancock, S., Mathieu, R., Main, R., Naidoo, L., Erasmus, B., & Scholes, R. 2023. First validation of GEDI canopy heights in African savannas. *Remote Sensing of Environment*. 285, 113402.

Markham, B. L., Thome, K. J., Barsi, J. A., Kaita, E., Helder, D. L., Barker, J. L., & Scaramuzza, P. L. 2004. Landsat-7 ETM+ on-orbit reflective-band radiometric stability and absolute calibration. *IEEE Transactions on Geoscience and Remote Sensing*. 42, 2810-2820.

Masek, J. G., Wulder, M. A., Markham, B., McCorkel, J., Crawford, C. J., Storey, J., & Jenstrom, D. T. 2020. Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment*. 248, 111968.

Milenković, M., Reiche, J., Armston, J., Neuenschwander, A., De Keersmaecker, W., Herold, M., & Verbesselt, J. 2022. Assessing Amazon rainforest regrowth with GEDI and ICESat-2 data. *Science of Remote Sensing.* 5, 100051.

Roy, D. P., Huang, H., Houborg, R., & Martins, V. S. 2021a. A global analysis of the temporal availability of PlanetScope high spatial resolution multi-spectral imagery. *Remote Sensing of Environment*. 264, 112586.

Roy, D.P., Kashongwe, H.B., & Armston, J. 2021b. The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. *Science of Remote Sensing*. 4, 100024.

Scalon, M. C., Bohn, A., Coelho, G. C., Meister, L., Alves, R. D. F., Secco, R. T., & Marques, M. 2022. Relationship Between Growth Trajectories and Functional Traits for Woody Trees in a Secondary Tropical Forest. *Frontiers in Forests and Global Change*. 5, 754656.

Scalon, M. C., Bohn, A., Coelho, G. C., Meister, L., Alves, R. D. F., Secco, R. T., & Marques, M. 2022. Relationship Between Growth Trajectories and Functional Traits for Woody Trees in a Secondary Tropical Forest. *Frontiers in Forests and Global Change*. 5, 754656.

Tyukavina, A., Hansen, M. C., Potapov, P., Parker, D., Okpa, C., Stehman, S. V., & Turubanova, S. 2018. Congo Basin forest loss dominated by increasing smallholder clearing. *Science Advances*. 4, eaat2993.

Vancutsem, C., Achard, F., Pekel, J. F., Vieilledent, G., Carboni, S., Simonetti, D., & Nasi, R. 2021. Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Science Advances.* 7, eabe1603.

Vrabel, J. C., Anderson, C., Bresnahan, P. C., Christopherson, J. B., Clauson, J., Kim, M., & Sampath, A. 2023. System characterization report on the BlackSky Global multispectral sensor (No. 2021-1030-O). *U.S. Geological Survey, Earth Resources Observation and Science Center, Sioux Falls, South Dakota, USA.*

Wulder, M. A., Coops, N. C., Roy, D. P., White, J. C., & Hermosilla, T. 2018. Land cover 2.0. *International Journal of Remote Sensing*. *39*, 4254-4284.

Xu, L., Saatchi, S. S., Shapiro, A., Meyer, V., Ferraz, A., Yang, Y., & Ebuta, D. 2017. Spatial Distribution of Carbon Stored in Forests of the Democratic Republic of Congo. *Nature Scientific Report*. 7, 15030.

Zeide, B. 1993. Analysis of growth equations. *Forest science*. 39, 594-616.