Committee on Earth Observation Satellites

Working Group on Calibration and Validation

Land Product Validation Subgroup

Aboveground Woody Biomass Product Validation

Good Practices Protocol

Version 1.0 – 2021

Editors: Laura Duncanson, Mat Disney, John Armston, Jaime Nickeson, David Minor, Fernando Camacho

Citation: Duncanson, L., Armston, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., Carter, S., Chave, J., Herold, M., MacBean, N., McRoberts, R., Minor, D., Paul, K., Réjou-Méchain, M., Roxburgh, S., Williams, M., Albinet, C., Baker, T., Bartholomeus, H., Bastin, J.F., Coomes, D., Crowther, T., Davies, S., de Bruin, S., De Kauwe, M., Domke, G., Dubayah, R., Falkowski, M., Fatoyinbo, L., Goetz, S., Jantz, P., Jonckheere, I., Jucker, T., Kay, H., Kellner, J., Labriere, N., Lucas, R., Mitchard, E., Morsdorf, F., Naesset, E., Park, T., Phillips, O.L., Ploton, P., Puliti, S., Quegan, S., Saatchi, S., Schaaf, C., Schepaschenko, D., Scipal, K., Stovall, A., Thiel, C., Wulder, M.A., Camacho, F., Nickeson, J., Román, M., Margolis, H. (2021). Aboveground Woody Biomass Product Validation Good Practices Protocol. Version 1.0. In L. Duncanson, M. Disney, J. Armston, J. Nickeson, D. Minor, and F. Camacho (Eds.), Good Practices for Satellite Derived Land Product Validation, (p. 236): Land Product Validation Subgroup (WGCV/CEOS), doi:10.5067/doc/ceoswgcv/lpv/agb.001

Duncanson, L.[1], Armston, J.[1], Disney, M.[2,3], Avitabile, V.[4], Barbier, N.[5], Calders, K.[6], Carter, S.[7], Chave, J.[8], Herold, M.[7], MacBean, N.[9], McRoberts, R.[10], Minor, D.[1], Paul, K.[11], Réjou-Méchain, M.[5], Roxburgh, S.[11], Williams, M.[12,13], Albinet, C.[14], Baker, T.[15], Bartholomeus, H.[7], Bastin, J.F.[16], Coomes, D.[17], Crowther, T.[18], Davies, S.[19], de Bruin, S.[7], De Kauwe, M.[20], Domke, G.[21], Dubayah, R.[1], Falkowski, M.[22], Fatoyinbo, L.[23], Goetz, S.[24], Jantz, P.[24], Jonckheere, I.[25], Jucker, T.[26], Kay, H.[27], Kellner, J.[28], Labriere, N.[8], Lucas, R.[27], Mitchard, E.[12], Morsdorf, F.[29], Næsset, E.[30], Park, T.[31], Phillips, O.L.[14], Ploton, P.[5], Puliti, S.[32], Quegan, S.[33], Saatchi, S.[34], Schaaf, C.[35], Schepaschenko, D.[36], Scipal, K.[13], Stovall, A.[23], Thiel, C.[37], Wulder, M.A.[38], Camacho, F.[39], Nickeson, J.[23,40], Román, M.[41], Margolis, H[22].

1. Department of Geographical Sciences, University of Maryland, College Park, MD, USA
2. UCL Department of Geography, University College London, Gower Street, London, UK
3. NERC National Centre for Earth Observation (NCEO), UCL, Gower Street, London, UK
4. European Commission, Joint Research Centre, Ispra, Italy
5. AMAP, Univ Montpellier, IRD, CNRS, CIRAD, INRAE, Montpellier, France
6. CAVElab - Computational & Applied Vegetation Ecology, Dept. of Environment, Faculty of Bioscience Engineering, Ghent University, Belgium
7. Laboratory of Geo-Information Science and Remote Sensing, Wageningen University & Research, Wageningen, The Netherlands
8. Laboratoire Evolution et Diversité Biologique, UMR 5174, CNRS, IRD, Université Paul Sabatier, Toulouse, France
9. Department of Geography, Indiana University, IN, USA
10. University of Minnesota College of Food, Agricultural and Natural Resource Sciences, Department of Forest Resources, MN, USA
11. CSIRO Land & Water, Canberra 2601, Australia
12. School of GeoSciences, University of Edinburgh, Edinburgh, UK
13. NERC National Centre for Earth Observation (NCEO), University of Edinburgh, Edinburgh, UK
14. European Space Agency, Frascati, Italy
15. School of Geography, University of Leeds, UK
16. TERRA, Teaching and Research Centre, Gembloux Agro Bio-Tech, University of Liège, Liège, Belgium
17. Department of Plant Sciences, University of Cambridge Conservation Research Institute, David Attenborough Building, Cambridge, UK
18. Department of Environmental System Science, ETH Zurich, Switzerland
19. Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Washington DC, USA
20. Climate Change Research Center, University of New South Wales, Australia
21. Northern Research Station, USDA Forest Service, USA

22. NASA Headquarters, 300 E Street SW, Washington DC, USA
23. NASA Goddard Space Flight Center, Biospheric Sciences Laboratory, Greenbelt, MD, USA
24. School of Informatics, Computing, and Cyber Systems, Northern Arizona University, 1899 S San Francisco St, Flagstaff, AZ, USA
25. Food and Agriculture Organisation of the United Nations (FAO), viale delle Terme di Caracalla, 00153 Rome, Italy
26. School of Biological Sciences, University of Bristol, 24 Tyndall Avenue, Bristol, UK
27. Department of Geography and Earth Science, Aberystwyth University, UK
28. Institute at Brown for Environment and Society and Department of Ecology and Evolutionary Biology, Brown University, Providence, RI, USA
29. Department of Geography, University of Zurich (UZH), Switzerland
30. Norwegian University of Life Sciences (NMBU), Norway
31. NASA Ames Research Center, Mountain View, CA, USA
32. Norwegian Institute for Bioeconomy Research (NIBIO), Division of Forest and Forest Resources, National Forest Inventory, Høgskoleveien 8, 1433 Ås, Norway
33. School of Mathematics and Statistics, Hicks Building, University of Sheffield, Sheffield S3 7RH, UK
34. NASA JPL, 4800 Oak Grove Dr, Pasadena, CA, USA
35. School for the Environment, University of Massachusetts Boston, Boston, MA, USA
36. International Institute for Applied Systems Analysis, Laxenburg A-2361, Austria
37. German Aerospace Center (DLR), Institute for Data Science, Jena, Germany
38. Canadian Forest Service (Pacific Forestry Center), Natural Resources Canada, 506 West Burnside Road, Victoria, BC, Canada
39. Earth Observation Laboratory (EOLAB), Parc Cientific Universitat de Valencia, Valencia, Spain.
40. Science Systems and Applications, Inc., 10210 Greenbelt Rd, Lanham, MD, USA
41. Universities Space Research Association, 7178 Columbia Gateway Drive, Columbia, MD, USA

We gratefully acknowledge the following expert reviewers whose comments resulted in substantial improvements to the document:

Sebastien Bauwens, Masato Hayashi, Alex Held, Andrew Hudak, Nikolai Knapp, Stefan Maier, C. Patnaik, Paul Patterson, Ake Rosenqvist, Svetlana Saarela, Paul Montesano

Chapter leads: Laura Duncanson (Chapter 1), Keryn Paul, Jerome Chave, Kim Calders (Chapter 2), Maxime Réjou-Méchain, John Armston (Chapter 3), Ron McRoberts and Stephen Roxburgh (Chapter 4), Mat Williams, Martin Herold, Sarah Carter, Natasha MacBean (Chapter 5), Valerio Avitabile (Chapter 6), Mat Disney, Laura Duncanson, John Armston (Chapter 7)

List of Revisions

Version	Revision	Date	Author
0.1	Draft for community review	5th September, 2020	Duncanson, Armston, Disney et al.
1.0	CEOS WGCV LPV and community acceptance	5th March, 2021	Duncanson, Armston, Disney et al.

Editor's Note

This document reflects the view of the biomass focus area within the CEOS WGCV Land Product Validation sub-group. This focus area provides the community involved in the production and validation of satellite-based woody aboveground biomass products with a forum for documenting accepted good practices in an open and transparent manner, that is scientifically defensible. This 'title' document (V1.0) has undergone review by remote sensing experts from across the globe. This represents the current state of knowledge for satellite biomass remote sensing and includes a summary of current knowledge and data gaps toward operational validation of products at a global scale.

We note that currently (March, 2021), no globally representative, systematically collected reference system for biomass product validation is available. We make recommendations for new data collections specifically designed for this purpose, but acknowledge that these recommendations and the authorship for this document are biased toward tropical moist forests. While we attempt to make recommendations that are applicable to all ecosystems, we note that forest ecosystems are dynamic and structurally complex, and a single set of recommendations will not apply to all ecosystems. Additionally, existing reference datasets, such as National Forest Inventories (NFIs), provide critical data addressing many needs beyond the validation of biomass products. While our recommendations focus on ideal datasets for validation and inter-comparison at the plot and pixel scale, we stress that our recommendations may be impractical in some cases, and should not replace, but complement, existing datasets that are maintained by in-country organizations, such as NFIs.

This document is published in advance of forthcoming biomass products from a new generation of lidar and SAR sensors (e.g. NASA's GEDI, NASA-ISRO's NISAR, ESA's BIOMASS), and will undoubtedly be updated as we learn from the development and validation of these products. This protocol focuses on aboveground woody biomass stock, and thus future extensions are anticipated that may provide guidance on other biomass pools (e.g. non-woody), and on biomass change. It is therefore expected that this protocol document and recommendations will undergo subsequent regular iterations based on community feedback and scientific advancement.

Finally, we gratefully acknowledge several expert reviewers whose thoughtful comments substantially improved this document. We welcome experts to participate in the ongoing improvement of this document and invite the broader community to make use of it for their research and applications related to woody biomass products derived from satellite data. All contributors will be recognized as such in the document and on the CEOS WGCV LPV website.

Sincerely,

The Editors,

Laura Duncanson, University of Maryland, College Park
John Armston, University of Maryland, College Park
Mat Disney, University College London
Jaime Nickeson, SSAI, NASA Goddard Space Flight Center
David Minor, University of Maryland, College Park
Fernando Camacho, EOLAB, Chairperson of the CEOS WGCV Land Product Validation group

Table of Contents

List of Acronyms and Nomenclature

3D	Three-Dimensional
3DEP	Three-Dimensional Elevation Program
AGB	Aboveground Biomass
AGBD	Aboveground Biomass Density
ALOS	Advanced Land Observing Satellite (JAXA)
ALS	Airborne Laser Scanning
ARD	Analysis Ready Data
ATBD	Algorithm Theoretical Basis Document
BAM	Biomass Allometric Model
BEF	Biomass Expansion Factor
BIOMASS	BIOMASS satellite mission (ESA)
BMA	Bayesian Model Averaging
BRM	Biomass Reference Measurement
CBL	Compact Biomass Lidar
CBM-CFS3	Carbon Budget Model of the Canadian Forest Service
CEPF	Critical Ecosystem Partnership Fund
CEOS	Committee on Earth Observation Satellites
CHM	Canopy Height Model
CIFOR	Center for International Forestry Research
CONAE	Argentine Space Agency
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CV	Coefficient of Variation
DBH	Diameter at Breast Height
DEM	Digital Elevation Model
DGNSS	Differential GNSS
DLR	German Aerospace Center
DSM	Digital Surface Model
DTM	Digital Terrain Model

DVM	Dynamic Vegetation Model
ECV	Essential Climate Variable
EO	Earth Observing
EPA	Environmental Protection Agency (US)
ESA	European Space Agency
ESM	Earth System Models
EXP	Expansion Estimator
FAO	Food and Agriculture Organization (United Nations)
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FCPF	Forest Carbon Partnership Facility (World Bank)
FIA	Forest Inventory and Analysis (US Forest Service)
ForestGEO	Forest Global Earth Observatory
FRM	Fiducial Reference Measurements
GCOS	Global Climate Observing System
GEDI	Global Ecosystem Dynamics Investigation (NASA)
GFOI	Global Forest Observations Initiative
GHG	Greenhouse Gas or Greenhouse Gases
GIS	Geographic Information System
GLAS	Geoscience Laser Altimeter System (NASA)
GLONASS	Global navigation satellite systems
GNSS	Global navigation satellite systems
GOFC-GOLD	Global Observation of Forest Cover-Global Observation of Land Dynamics
GPG	Good Practice Guidance (IPCC 2003 Good Practice Guidance)
GPS	Global Positioning System
GREG	Generalized Regression
GSV	Growing Stock Volume
HPCM	High Priority Candidate Mission
IceSAT	Ice Cloud and land Elevation Satellite (NASA)
IceSAT-2	Ice Cloud and land Elevation Satellite-2 (NASA)
IIASA	International Institute for Applied Systems Analysis

IBM	Individual Based Model
ILAMB	International Land Model Benchmarking
INPE	Instituto Nacional de Pesquisas Espaciais (Brazil)
IPCC	Intergovernmental Panel on Climate Change
ISRO	Indian Space Research Organization
JAXA	Japan Aerospace Exploration Agency
LAI	Leaf Area Index
Landsat	Land Satellite (NASA)
LIDAR	Light Detection and Ranging
LPV	Land Product Validation (a CEOS WGCV subgroup)
LST	Land Surface Temperature
LVIS	Land, Vegetation, and Ice Sensor (NASA)
MAAP	Multi-Mission Algorithm and Analysis Platform (ESA/NASA)
MABEL	Multiple Altimeter Beam Experimental Lidar
MC	Moisture Content
MCH	Mean Canopy Height
MODIS	Moderate Resolution Imaging Spectroradiometer (NASA/NOAA)
MOLI	Multi-footprint Observation Lidar and Imager (JAXA)
MRV	Measuring, Reporting, and Verification
MSE	Mean Square Error
NASA	National Aeronautics and Space Administration (US)
NSF	National Science Foundation
NRCS	Natural Resources Conservation Service (USDA)
NEON	National Ecosystem Observation Network (US)
NetCDF	Network Common Data Form
NGGI	National Greenhouse Gas Inventory (IPCC)
NISAR	NASA-ISRO Synthetic Aperture Radar
NFI	National Forest Inventory
NULESU	National University of Life and Environmental Sciences of Ukraine
NGO	Non-governmental organization

NNRG	Northwest Natural Resource Group (US)
PALSAR	Phased Array type L-band Synthetic Aperture Radar (JAXA)
PFT	Plant Functional Type
POM	Point of Measurement
QSM	Quantitative Structural Model
QA/QC	quality assurance/quality control
REDD+	Reducing Emissions from Deforestation, plus
RMSE	Root Mean Square Error
RMSD	Root Mean Square Deviation
ROSE-L	Radar Observatory System for Europe in L-band
RS	Remote Sensing
SAOCOM	Argentine Microwaves Observation Satellite
SAR	Synthetic Aperture Radar
SD	Standard Deviation
SDG	Stepped Diameter Gauges
SDG	Sustainable Development Goals
SfM	Structure from Motion
TANDEM-L	TerraSAR-L add-on for Digital Elevation Measurement (DLR)
TERN	Terrestrial Ecosystem Research Network (Australia)
TLS	Terrestrial Laser Scanning
UAV	Unmanned Aerial Vehicle
UAV-LS	Unmanned Aerial Vehicle-Laser Scanner
UN	United Nations
UNFCCC	United Nations Framework Convention on Climate Change
UNSD	United Nations Statistical Division
UPM	Universidad Politécnica de Madrid
USFS	United States Forest Service
USGS	United States Geological Survey
US	United States
UTM	Universal Transverse Mercator

WD	Wood Density
WGCV	Working Group on Calibration and Validation (CEOS)

CEOS Biomass Protocol Executive Summary

- **Biomass maps are vital for understanding and mitigating climate change**
- **New biomass products are being produced from satellite missions**
- **Differences between products may lead to confusion and reduced confidence**
- **There is a need for clear guidance on generating and using these products**
- **The CEOS biomass protocol is aimed at both users and producers of biomass products**

Biomass maps are important for understanding and mitigating climate change

Forest biomass, defined as the dry-weight of the standing live or dead woody component of aboveground vegetation, has been recognized as a Global Climate Observing System (GCOS) Essential Climate Variable (ECV), a critical input to the United Nations' (UN) Reducing Emissions from Deforestation and Forest Degradation-plus (REDD+) program, as inputs reporting toward the UN's Sustainable Development Goals (SDGs), and an important input to Earth system models. Spatially continuous maps of forest biomass are therefore important inputs for decreasing the uncertainties in the global carbon cycle, underpinning forest management and climate mitigation strategies, and global carbon cycle science.

New satellite biomass products are being produced

To address the need to estimate biomass globally, several recent and upcoming Earth Observing (EO) missions will collect satellite data giving information on forest structure and aboveground biomass. We anticipate these datasets will drive development of many new global forest biomass products. Some of these products will be official mission products, but many will be generated independently by scientists, Non-Governmental Organizations (NGOs) and other interested parties, each potentially adopting a different combination of mission data, training data, and statistical algorithms according to needs and resources.

Multiple new products may lead to user confusion

New biomass products will not necessarily agree with each other, and will vary in quality, spatial resolution and date of prediction. Products that disagree with each other will likely cause confusion and without consistent validation users will not know which products to trust.

There is a need for clear guidance

This protocol will assist biomass map producers in good practices for estimating and reporting uncertainties in their products, and will inform users how to interpret products and conduct independent validation. The goal of this protocol is to facilitate consistent and transparent biomass product uncertainty estimation so that products can be used effectively for science, forest management and policy applications.

Anticipated protocol audience

This protocol is aimed at three primary audiences: 1) Biomass map product producers aiming to improve and standardize how uncertainties are reported for increased product utility, 2) Biomass map product users needing to understand how reported uncertainties are estimated, and 3) Those wishing to perform independent validation of biomass products, either with new or existing reference data.

Summary of Protocol Content and Recommendations

- **Common agreed-upon definitions**
- **The requirement for more high-quality reference data**
- **Uncertainty propagation from field measurements and biomass models to maps**
- **Independent validation: necessary but proceed with caution**
- **Recommendations for compilation of new reference datasets**

Definition of biomass

We define biomass as the dry mass of live or dead matter from tree or shrub (woody plant) life forms, typically expressed as a per area density (e.g. Mg of aboveground biomass per hectare). Thus, we do not include non-woody or belowground biomass. When discussing individual tree or plot total biomass (not density), the definition is Aboveground Biomass (AGB), whereas for plot or pixel level densities, as commonly estimated in mapped products, the definition is Aboveground Biomass Density, usually per hectare (AGB/ha).

High-quality reference data are required

High-quality biomass reference data are required both to produce accurate biomass maps, and to conduct product validation. This point may seem obvious, but the most direct high-quality biomass reference data, weighed tree biomass, are very difficult to acquire for anything other than small numbers of trees. Instead, we rely on easy-to-make tree size measurements (e.g. stem diameter, total tree height) and the use of statistical (allometric) models to translate these size measurements into estimates of biomass. Many errors and uncertainties are included in reference data, and many existing biomass measurements from field plots have such high uncertainties that they are inappropriate for generating or validating products.

This protocol makes specific recommendations for establishing reference measurements in forests, including the use of traditional measurements and both terrestrial and airborne lidar.

Airborne lidar biomass maps are key tools for scaling between field plots and satellite products

Airborne lidar is now operationally acquired throughout the world. These data can be used with local field plots to create reference maps at multiple spatial resolutions so that a single site can be used to validate the full suite of anticipated satellite biomass products, regardless of their spatial resolution. This protocol presents recommendations for the acquisition of airborne lidar (sensors, sampling designs, etc.) as well as guidance for linking field plots to airborne lidar data, fitting local biomass models, and propagating and reporting uncertainties in reference maps. There are also recommendations for how to subsequently link airborne lidar maps to satellite products.

Propagating uncertainties from field measurements and biomass models to maps

Existing biomass products often (but not always) estimate uncertainties from the statistical models used to link field biomass estimates to satellite remote sensing data. These uncertainties are typically underestimated when they do not include measurement, geolocation and allometric modeling uncertainty (i.e. calibration uncertainty). This protocol provides recommendations on how best to estimate and reduce errors, as well as how to propagate these to estimate uncertainties of biomass products at a range of scales.

Independent validation is necessary but should be conducted with caution

Using existing field plot biomass estimates to validate biomass products is important, but it can introduce so much uncertainty that the results may become meaningless for many applications. This is particularly true for small plots in heterogeneous forests, plots with high geolocation uncertainty, plots with a temporal lag between field and satellite data collection, or plots where individual tree information is not available. It is recommended to filter reference data to a minimum quality standard prior to conducting independent validation.

Recommendations for compiling new reference data

New reference data are required to update existing reference measurements (either to make them current, or improve them to meet the standards recommended in the protocol). There are many geographic domains with insufficient reference data, or where uncertainties in biomass

products cannot be accurately estimated due to a lack of data. It is recommended that new datasets are routinely collected, in collaboration with existing networks, that both update and gap-fill existing reference datasets. These recommendations are summarized in Table 0.1 (also in Tables 8.1- 8.3). Table 0.1 outlines ideal collection of new datasets, but recommendations differ by ecosystem and should be followed as far as practicable. Reference data that do not fulfill these criteria (e.g. small plots as routinely collected via National Forest Inventories (NFIs)) are still useful for wide area product validation (see Chapter 6), and the collection and sharing of these data are still highly encouraged and appreciated by the CEOS biomass community.

Table 0.1 Recommendations for collection of biomass reference measurements. Note that data collected following existing protocols are often useful despite not meeting all recommendations below. See Chapter 2 for details.

<table>
<tr><td colspan="2">Recommendations for all data collections
Data should be free and open access within at most 1 year after data collection
Data should be acquired in collaboration with long term field plot networks and local partners wherever possible</td></tr>
<tr><td>Field Plot Recommendations
Square plots
Easier to link to gridded products
Large plots (minimum 0.25 ha in tropics, ideally 1 ha plots with 0.25 ha or 0.04 ha subplots)
Minimizes edge effects and geolocation uncertainties
Smaller plots (<0.25 ha) are acceptable outside of tropics provided airborne lidar available
Stem-mapped where possible
Geolocated with high accuracy, and reported uncertainties
Trained botanist should be employed for species identification</td><td>Airborne Lidar Recommendations
Minimum ~4 pulses/m^2 with 4 returns /pulse, but minimum is ecosystem-dependent. Ideally ≥ 8 pulses/m^2
Preferably acquired same season as field plots
Acquired within 2 years, ideally 1, of field data acquisition
Repeated every ~5 years or when disturbance is detected
Wall-to-wall coverage of at least 10 km^2
Cover both the plots and local environmental and forest structure gradient
Smaller area of coverage acceptable if only UAV-LS lidar available</td></tr>
<tr><td>Spatial Distribution of Field Plots
Plots cover environmental gradients under airborne lidar collection that are locally or regionally correlated to biomass (e.g. topographic gradients)
Sufficient number of plots collected to train a lidar model (min approximately 30, depending on complexity of system)</td><td>Terrestrial Lidar Recommendations
Data collection in new or existing long-term plots
Data augments field measurements, does not replace them
1 ha plots preferable
Data acquired in a grid pattern
Spacing 10 m in dense forests, 20 m in open areas
Can be changed to ensure consistent sampling and minimize occlusion
Instrument must have ability to range tallest trees in 1 ha plots (150 m range)
Repeated ~ every 5 years or when disturbance is detected
Multiple scans need to be coregistered (either through use of targets or with sensor that has automatic coregistration)</td></tr>
</table>

Summary of Good Practice Recommendations, Current Data and Knowledge Gaps

Each chapter in the protocol is summarized in terms of good practice recommendations and community highlighted knowledge and data gaps. These gaps include both gaps in data acquisitions (as particularly highlighted in Chapter 8) but also areas where further research or tool development would help progress biomass validation activities. Table 0.2 provides a summary of these recommendations. Note that a more detailed summary of these recommendations is found in Chapter 7, and the reader should refer to the source chapter for complete details.

Table 0.2 Summary of protocol recommendations by chapter

Chapter	Good practice recommendations	Knowledge and/or Data Gaps
Collection of Reference Data (Chapter 2)	● Collect data following recommendations in Table 0.1 ● Report measurement and model uncertainty for plot-level AGBD estimation	● Standardization across field measurement protocols for AGBD estimation ● Tool development for automated TLS and UAV-LS processing ● More research on how to estimate wood density in the field ● Improvement of generalized allometric models using larger samples and TLS
Linking field, airborne and satellite data (Chapter 3)	● Collect large, well geolocated, preferably square plots (see Table 0.1) ● Develop local AGBD maps using high-quality wall-to-wall airborne lidar data and locally trained AGBD models ● Maps should be at the spatial resolution of plots or subplots and can be subsequently aggregated ● Estimate and report per-pixel uncertainties in lidar AGBD maps to aid validation	● A quality control framework for systematic airborne lidar AGBD maps following protocol recommendations ● Further research into spatial propagation of uncertainty for per pixel estimation ● Tools to facilitate lidar AGBD map production and comparison with satellite biomass maps
Error estimation and propagation (Chapter 4)	● Error reporting should comply with IPCC good practices guidelines ● Measurement and modeling errors should be estimated following appropriate inference methods and propagated to mapped products	● Online tools to facilitate error estimation and propagation for product producers and independent validation ● Extension of existing tools (e.g. R BIOMASS package) outside of tropics and from field plots to mapped products
Making products and validation useful for different communities (Chapter 5)	● Uncertainties should be clearly and consistently reported ● Adoption of consistent terminology and methods to error and uncertainty estimation by all map producers	● Biomass product harmonization or intercomparison activities would facilitate uptake ● Annually updated harmonized estimates would aid policy uptake ● Definitions of forest/non-forest by community require consensus for product intercomparison and change analysis
User-led Validation (Chapter 6)	● Only conduct independent validation with user provided data if data are appropriate for validation (contemporaneous, complete metadata, same resolution as satellite products) ● Conduct thorough data screening and filtering prior to conducting validation	● Simple online tools to facilitate data screening, filtering and error estimation
Protocol Implementation	● Work with existing plot networks to collect and	● New and updated BRMs data collection,

(Chapter 8)	curate reference data ● Follow measurement recommendations in Table 0.1 ● Independent validation should be transparent, consistent and repeatable ● Collection of new contemporaneous field and airborne lidar datasets over a set of globally representative Biomass Reference Measurement Sites (BRMs). ● Reference data should be made free and open	particularly in existing geographic gaps (continental asia, drylands, dense tropical forests) ● Independent validation should be conducted on an open access platform with publicly available user-friendly tools (e.g. on MAAP platform or similar)

Suggestions for Using Protocol by Different Groups

Of the three audiences anticipated for this protocol (biomass product generators, biomass product users, and those wishing to collect biomass reference data), the following flow charts may facilitate reading of the protocol. Naturally, many readers will occupy multiple categories, but the following flow charts were designed to guide specific readers to the chapters we deem most relevant for each of these three audiences.

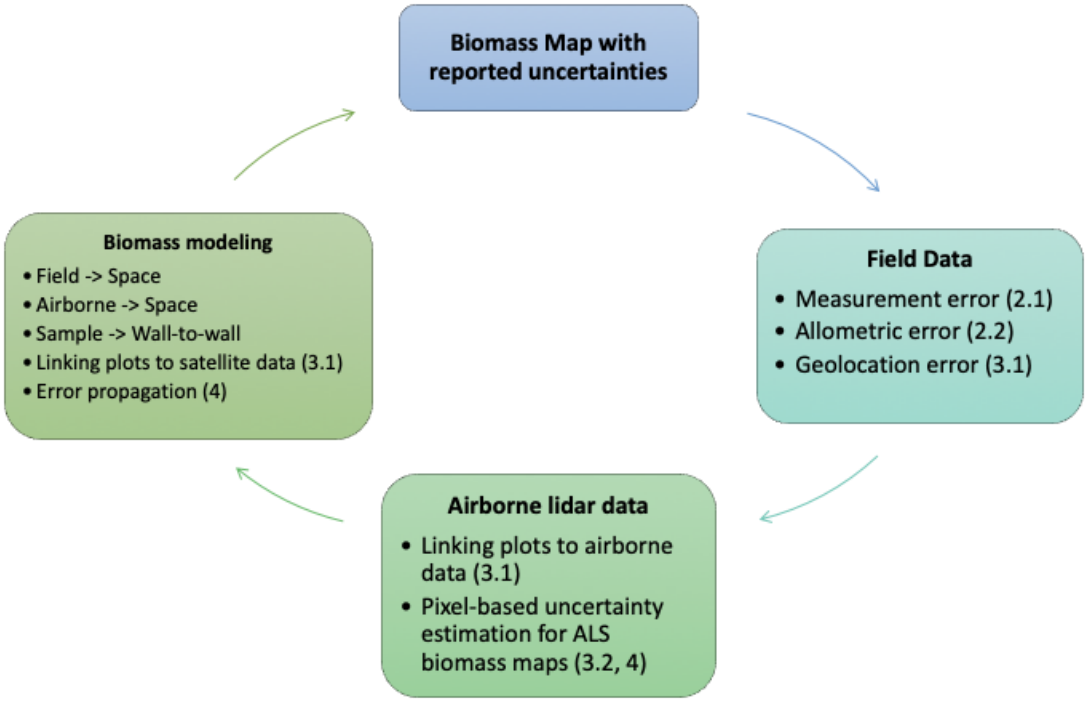

Fig 0.1 Guide for using the protocol to aid in generation of biomass products and uncertainty reporting.

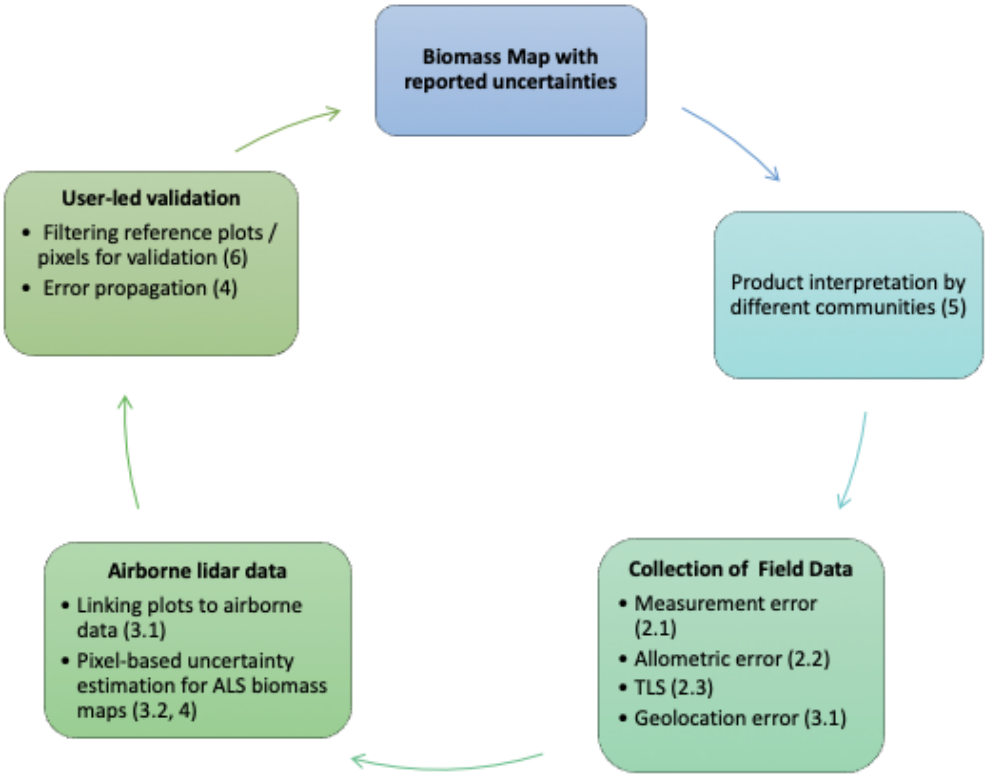

Fig 0.2 Guide for using the protocol to aid biomass product users in either interpretation or independent validation of biomass products.

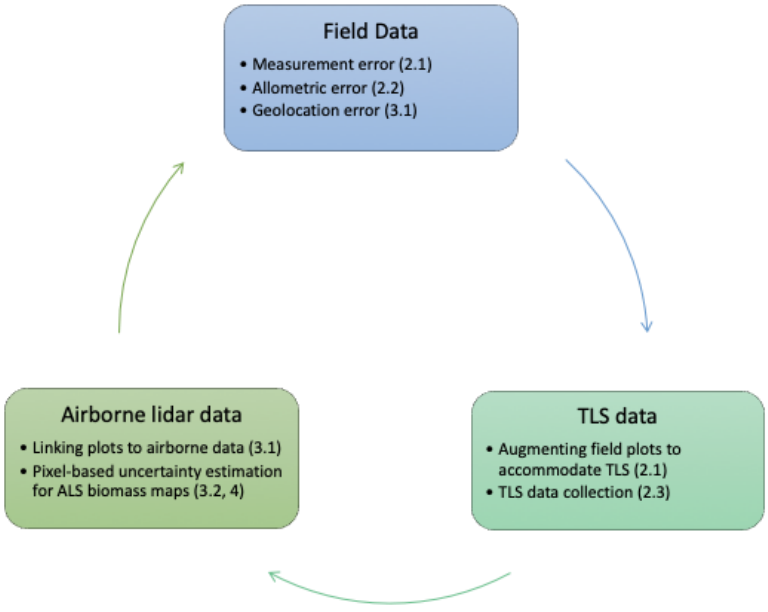

Fig 0.3 Guide for using the protocol to aid in the consistent collection of high-quality forest reference biomass measurements.

Chapter 1: Introduction

Forest biomass has been recognized as a Global Climate Observing System (GCOS) Essential Climate Variable (ECV), a critical input to the United Nations' (UN) Reducing Emissions from Deforestation and Forest Degradation-plus (REDD+) program, and an important input to Earth system models (Herold et al., 2019). Spatially continuous maps of forest biomass are therefore an important input for reducing uncertainties in global carbon stock and flux estimates from forests, particularly for areas where insufficient ground or airborne lidar data are available. Accurate biomass products are of great importance for forest management and climate mitigation. However, due to a previous dearth of satellite data specifically designed for producing accurate estimates of forest structure (Goetz et al., 2009), few global-scale forest biomass products are currently available, and the assessment of their accuracy is challenged by a lack of appropriate reference data. To overcome this critical carbon accounting gap, several upcoming Earth Observing (EO) missions will collect satellite data sensitive to forest structure and aboveground biomass, defined as the dry-weight of the live or dead woody component of aboveground vegetation. We anticipate a multitude of new global forest biomass products in the coming decade, but foresee challenges in their intercomparison and validation across biomass products. These challenges have already been highlighted by several studies comparing the few existing continental or global-scale biomass products (Avitabile et al., 2016; Avitabile & Camia, 2018; Baccini et al., 2012; Huang et al., 2015; Mitchard et al., 2013; Saatchi, Harris, et al., 2011; Santoro et al., 2015; Thurner et al., 2014), and may hinder the effective adoption of biomass products for various policy, management and science applications.

A specific example of the importance of independent biomass product validation comes from comparisons of two widely known pantropical biomass maps (Baccini et al., 2012; Saatchi, Harris, et al., 2011). By independent, we mean using reference data that were not included in the generation of products, and ideally conducted by a third party. Despite having been produced from the same core satellite datasets (the Geoscience Laser Altimeter System [GLAS] and the Moderate-resolution Imaging Spectroradiometer [MODIS]), these maps differ substantially in several tropical areas (Avitabile et al., 2016; Mitchard et al., 2013, 2014) potentially because they employed different empirical modeling approaches, calibration datasets, and extrapolation techniques. However, the exact causes of discrepancies between these products, or indeed a determination of the more accurate product for a given application, is not possible without common approaches to independent validation. Aboveground biomass product validation is challenging, primarily because of the paucity of high-quality, publicly available, and globally representative reference sites with well-characterized uncertainties, and challenges related to the fact that these reference data are not direct measurements but rather estimates based on tree-level allometric model predictions (D. B. Clark & Kellner, 2012). Indeed, in the pantropical case, the map producers themselves had limited available validation datasets, and Baccini et al. (2012) and Saatchi et al. (2011) performed cross-validation of their map products using a subset of GLAS data that were deliberately left out of their biomass model training, rather than validating with an independent dataset. While Saatchi et al. (2011) conducted an error propagation for the final estimated uncertainties associated with their pantropical product, and

Baccini et al. (2012) reported confidence intervals on their estimates per continent, the degree of accuracy of these products in geographic areas outside the calibration range, or at the various resolutions needed for policy implementation (Herold et al., 2019), was not possible. These products have been compared to the Intergovernmental Panel on Climate Change (IPCC) Tier 1 biomass estimates, following the 2006 IPCC Good Practice Guidance (GPG; IPCC, 2003, 2006). Although a composite of the two pantropical products was suitable to replace IPCC Tier 1 estimates when national inventories were not available, it was recommended that national estimates be favoured over these remote sensing-based estimates, given the large disparities between the products and national inventories (Langner et al., 2014).

This is not a criticism of past biomass products. Indeed, historic biomass products are a function of the data (ground through to satellite) available at that time. Global products are typically produced at a grid cell size relating to the spatial resolution of the global image data source utilized. Larger pixels can be expected to obscure local variability and result in lower variance outcomes. Comparison to higher resolution data, such as from samples of lidar data, illustrate these issues (e.g. Bolton et al., 2013). As the spatial resolution of global image data options have gotten finer, so too has the detail and quality of large area biomass products. The next generation of biomass products from new mission datasets are expected to improve upon these past examples, and the sheer number of missions will result in a larger number of products to compare and validate. Indeed, the issue of product validation will become even more pressing as the number of spaceborne datasets specifically designed to map ecosystem structure increases (e.g., the National Aeronautics and Space Agency's (NASA's) Global Ecosystem Dynamics Investigation (GEDI), the NASA/Indian Space Research Organization's (ISRO's) NASA-ISRO Synthetic Aperture Radar (NISAR), the European Space Agency's (ESA's) BIOMASS, and the Japan Aerospace Exploration Agency's (JAXA's) Advanced Land Observing Satellite (ALOS-4), and approaches to biomass estimation using these data diversify. Previous biomass products have varied in terms of the spatial and temporal resolution, modeling approach, geographic scope of calibration data, scaling, error propagation, and uncertainty reporting (Goetz et al., 2009; Huang et al., 2015; Mitchard et al., 2013). To effectively meet the goals of scientists and decision-makers, the global change community requires well-tested validation approaches that are transparent and flexible (with respect to geographic scope and spatial resolution).

The Committee on Earth Observation Satellites (CEOS) is an international body that works to coordinate Earth Observation programs and data collected by space agencies. For nearly two decades, the CEOS Working Group on Calibration and Validation (WGCV) has had a sub-group specifically focused on Land Product Validation (LPV). In close coordination with CEOS member agencies, the LPV subgroup has recognized the need for good practices and protocols to guide biomass product validation in advance of the expected suite of upcoming biomass products. This LPV subgroup launched the biomass focus area in 2017 to help gather community support in developing a validation protocol for the products that will be generated from the upcoming biomass-related missions.

1.1 Earth Observation Missions for Biomass Mapping

NASA's Ice Cloud and land Elevation Satellite (ICESat) was launched in 2003, and its Geoscience Laser Altimeter System (GLAS) instrument collected global waveform lidar measurements over vegetation that were used to estimate forest height and structure until the last ICESat laser failed in 2009 (Abshire et al., 2005). GLAS data were not designed to study forest structure, but these data have nevertheless become popular for biomass mapping. These data are relatively sparse in spatial sampling, and each lidar footprint illuminated a nominally 65 m diameter circle, which resulted in the mixing of reflected signal from ground and canopy surfaces, ultimately presenting challenges for estimating biomass in areas of high relief or structural complexity (Duncanson et al., 2010). Despite these challenges, many wide area biomass maps used GLAS data to map forest structure (e.g. Baccini et al., 2008; Margolis et al., 2015a; Saatchi, Harris, et al., 2011; Simard et al., 2011; Su et al., 2016).

Several current and upcoming missions (e.g. GEDI, BIOMASS, NISAR) should provide improved data for biomass mapping compared to those earlier sensors as they are designed with a primary science goal of mapping forest biomass. Official mission biomass products are expected from each of these missions. Still, because of the publicly available nature of these mission datasets we also expect a host of other new biomass products through data fusion and alternative algorithms, etc. We, therefore, anticipate the release of products with a range of spatial resolutions, geographic extents, and temporal domains. Table 1.1 shows the expected resolution of core biomass products from upcoming spaceborne missions themselves, but fusion products will likely present both coarser and higher resolution maps.

Many of these missions have specific biomass product accuracy requirements as part of the criteria by which mission success is determined (Table 1.1). Independent validation of these products at their nominal resolution would help demonstrate that requirements have been met. This is particularly useful if validation of biomass products from each mission or estimation approach can be conducted at the same set of sites, allowing direct comparisons between accuracies of each product in different forest types, environments, disturbance histories, etc. Comparisons between official mission products and other new biomass maps will also allow the community to appreciate the accuracy impacts of algorithmic improvements, data fusion approaches, etc., on product accuracy and ultimately reduce confusion and latency in the adoption of new biomass mapping approaches. This collaborative strategy will help to achieve higher CEOS validation stages which require spatiotemporally systematic reference datasets (Table 1.2)"

Table 1.1 Current and expected biomass-relevant missions. Note that only NISAR, GEDI and BIOMASS are approved missions with a formal requirement for biomass mapping accuracy. Missions with no official biomass product are marked Not Applicable (NA), and missions that have yet to publish requirements are to be determined (TBD).

Mission	Funding Agency	Launch Date (Expected)	Data Type	Measurement resolution	Biomass map Resolution	Geographic Domain	Accuracy Requirement
ICESat-2	NASA	09/2018	532 nm photon counting lidar	13m footprint aggregated to 100-m transect	NA	Global	NA
SAOCOM 1A	CONAE	10/2018	L band SAR	10-100m depending on mode	NA	Global	NA
GEDI	NASA	12/2018	1064 nm waveform lidar	25 m circular footprint	1 km^2	ISS (+/- ~51.6°)	<20% standard error for 80% of forested 1 km cells
SAOCOM 1B	CONAE	08/2020	L band SAR	10-100m depending on mode	NA	Global	NA
ALOS-4	JAXA	(2022)	L-band SAR	1-25 m depending on mode	NA	Global	NA
NISAR	NASA-ISRO	(2023)	L-band SAR	3 - 10 m (depends on mode)	1 ha	Global	<20% RMSE for <100 Mg/ha
BIOMASS	ESA	(2022)	P-band SAR	60 x 50 m with >6 looks	4 ha	Global except western Europe and North America	<20% RMSE for AGB >50 Mg/ha; 10 Mg/ha for AGB ≤50 Mg/ha
MOLI	JAXA	(2023)	1064 nm waveform lidar	25 m circular footprint	250 m	ISS (+/- ~51.6°)	20 Mg/ha for < 100 Mg/ha; 25% for > 100 Mg/ha
TanDEM-L	DLR	(2023)	L-band SAR	TBD	1 ha	Global	20% accuracy or 20 Mg/ha
Copernicus HPCM ROSE-L	ESA/EC	(2027)	L-band SAR	TBD	1 ha	Global	TBD

1.2 CEOS Validation Stages

Table 1.2 CEOS validation stage hierarchy. The eventual goal is to mature existing and forthcoming biomass products from stage 0 to 4.

Validation Stage - Definition and Current State		Variable
0	No validation. Product accuracy has not been assessed. Product considered beta.	
1	Product accuracy is assessed from a small (typically < 30) set of locations and time periods by comparison with in-situ or other suitable reference data.	**Snow** **Fire Radiative Power** **Biomass**
2	Product accuracy is estimated over a significant (typically > 30) set of locations and time periods by comparison with reference in situ or other suitable reference data. Spatial and temporal consistency of the product, and its consistency with similar products, has been evaluated over globally representative locations and time periods. Results are published in the peer-reviewed literature.	**fAPAR** **Phenology** **Burned Area** **Land Cover** **LAI**

3	Uncertainties in the product and its associated structure are well quantified over a significant (typically > 30) set of locations and time periods representing global conditions by comparison with reference in situ or other suitable reference data. Validation procedures follow community-agreed-upon good practices. Spatial and temporal consistency of the product, and its consistency with similar products, has been evaluated over globally representative locations and time periods. Results are published in the peer-reviewed literature.	**Vegetation Indices** **Albedo** **Soil Moisture** **LST & Emissivity** **Active Fire**
4	Validation results for stage 3 are systematically updated when new product versions are released or as the interannual time series expands. When appropriate for the product, uncertainties in the product are quantified using fiducial reference measurements over a global network of sites and time periods (if available).	

1.3 Definitions

This chapter provides definitions used in this protocol and relevant to global aboveground biomass map training and validation.

1.3.1 Definition of Aboveground Biomass

For the purposes of CEOS LPV the definition of aboveground biomass density is the above ground standing dry mass of live or dead matter from tree or shrub (woody plant) life forms, typically expressed as a mass at the individual tree level, or a mass per unit area (density) for mapped products. The most common unit of reporting is in Megagrams, or tonnes, per hectare, Mg/ha.

There is no single universally accepted definition of aboveground biomass across different user communities. Other definitions of aboveground biomass include: considering live wood only; including leaf mass; including woody and/or non-woody debris. The definition we use here is most relevant for earth observation: measurements respond to standing biomass only; optical observations, particularly lidar, are sensitive to living and dead standing biomass; radar observations are sensitive to mostly living woody biomass due to moisture content and the size of scattering objects relative to wavelength. We note also that allometric models that are used to translate EO measurements to aboveground biomass do not universally or systematically consider the fraction of live and dead wood in their calibration data.

1.3.2 Definition of Biomass Allometric Model

Aboveground biomass has been measured for several thousand trees globally, and related to tree dimensions (i.e. stem diameter, tree height and sometimes crown area) and wood density of those trees using regression techniques to produce so-called "allometric models". These models are used to estimate aboveground biomass from the dimensions and the species of individual trees recorded in forest inventory plots.

1.3.3 Definitions of Associated Physical Variables

Stem volume

Stem volume is the volume of the whole stem of an individual tree, measured in cubic meters (m^3).

Growing stock volume

Total volume of the stems of all living trees per unit area (m^3/ha).

Basal Area

Cross sectional area of all trees in a plot or stand measured at breast height, expressed as an area per area (m^2/ha).

Wood density

The density of a wood sample is its oven-dry mass divided by its fresh volume, given in units of g/cm^3. Wood density is the conversion factor used to translate between volume to biomass. Wood density is a key functional trait in plants, and a common explanatory variable for many allometric models. The wood densities of many tree species are held in a global database, allowing wood density to be estimated from taxonomic information collected during forest inventories.

Stem diameter

Stem diameter is a structural measure of the girth of an individual tree. Diameter is typically measured 1.3 m above the ground, known as diameter at breast height (DBH), or 50 cm above buttresses, stilt root, or deformities, if present. Stem diameter, along with tree height and wood density, can be used to predict aboveground biomass using allometric relationships. Note that while 1.3 m is the typical standard, this can vary by country, e.g. 1.2 m in Japan.

Tree height

Tree height is a structural measure of the vertical distance between the ground level and the tip of an individual tree. Tree height is often used in allometric models to estimate biomass, and is either measured in the field or estimated from trunk diameter through locally-calibrated or generic models rather than measured.

Canopy height

Canopy height is the vertical distance between the ground level and the highest point of tree tops, measured from any given location in a forest.

Canopy cover

Canopy cover is the proportion of ground area which overlaps with forest canopy in a vertical profile of a given local area.

Tree density
Tree density is the number of trees in a given area divided by that area, reported as stems/ha.

1.3.4 Definitions of Spatial and Geometrical Aspects

Coregistration
Coregistration is the recording of the spatial location of multiple samples to allow them to be paired. Coregistration is necessary for matching ground-based and remotely sensed measurements, and for pairing multiple adjacent or overlapping remote sensing measurements.

Incidence angle
The incidence angle is the angle between a ray approaching the surface and the direction perpendicular to the ground surface. The incidence angle can play a large role in spatial mismatch between remote sensing measurements and ground measurements.

Occlusion
Occlusion is the obstruction of distant objects by near objects from the perspective of a laser scanner.

Point density
Point density (m^{-2}) is the number of lidar pulses reflected and recorded per unit area by a laser scanner. The point density is determined by the frequency of the scanner, the velocity of the scanner motion, and the amount of overlap between coregistered samples.

Point cloud
A point cloud is a three-dimensional representation of forest structure via discrete measurements of surfaces. Point clouds can be generated through laser scanning or stereoscopic structure from motion images. Structural measurements such as stem diameter, tree height, and stem volume can be derived from point clouds.

Region of interest
An area (m^2) containing reference data.

Canopy Height Model (CHM)
Airborne lidar data are often used to produce high resolution canopy height maps. These maps can be used to correct plot geolocation error, check the accuracy of field measured heights, and as an input to local airborne lidar aboveground biomass density (AGBD) maps.

1.3.5 Definitions of Other Key Terms

Forest
The standard Food and Agriculture Organization (FAO) definition of forest is land with a tree canopy cover of more than 10 percent and area of more than 0.5 ha. The trees should be able to reach a minimum height of 5 m. Young stands that have not yet but are expected to reach a

crown density of 10 percent and tree height of 5 m are included under forest, as are temporarily unstocked areas (FAO, 2018). This definition results in large areas of dry and semi-arid tropics that are not tree-dominated being classed as forest, and current debate about the global extent of forests in dryland areas (as derived from high resolution optical imagery) hinges on details of definition (Bastin et al., 2017; Griffith et al., 2017). For the purposes of this overview we adopt as simple an operational definition of forests as possible, as 'land with largely continuous tree cover'.

Reference data
In this document, reference data are biomass estimates used during calibration and validation of remotely sensed biomass products, which are measured independently from the biomass products being assessed, and are not included in product training data. Reference data are ideally high-quality field measurements and/or airborne maps, following a standardised measurement protocol and traceable uncertainty estimates, at the same spatial resolution as the remotely sensed data source, and sampled across the full range of locations, site characteristics and the dynamic range that are present in the biomass product.

Calibration
Calibration in remote sensing is the process of assessing and correcting for systematic error in remotely sensed datasets, while for the purposes of this protocol it is used to mean fitting statistical models, and/or assessing and correcting for systematic error in biomass products through comparison with reference data.

Validation
Validation is the process of determining the quality of measurements and biomass model outputs through comparison with reference data.

1.3.6 Definitions of Validation Metrics

The validation process quantifies the following metrics in a statistically rigorous way over multiple locations and time periods representing global conditions. The definitions of these validation metrics have been sourced from the "Evaluation of measurement data: Guide to the expression of uncertainty in measurement" (GUM-2008) published by the Joint Committee for Guides in Metrology (JCGM, 2008).

Bias
Bias is the difference between this estimator's expected value and the true value of the parameter being estimated. Bias in an estimator or procedure may produce systematic error in the results (in this case the biomass map).

Measurand
The physical quantity of a measured property (i.e. the biomass of a tree).

Error

Error of measurement is the "result of a measurement minus a true value of the measurand" (JCGM, 2008, p. 36). The true value is not known, but typically "a number of results of measurements of a quantity is used to establish a conventional true value" (JCGM, 2008, p. 33). In the case of biomass measurements, which may depend on multiple "values of quantities other than the measurand, the errors of the measured values of these quantities contribute to the error of the result of the measurement" (JCGM, 2008, p. 37).

Error is composed of systematic error (bias) and random error. Systematic error is the "mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus a true value of the measurand" (JCGM, 2008, p. 37). Random error is the "result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions" (JCGM, 2008, p. 37).

Uncertainty

Uncertainty is the distribution of errors for an estimate. Components of uncertainty include systematic error (bias) and random error (precision). Uncertainty may arise from a range of measurement and model errors, and both systematic error and precision should be accounted for in uncertainty estimation and reporting.

Accuracy

Accuracy of measurement is the "closeness of the agreement between the result of a measurement and a true value of the measurand" (JCGM, 2008, p. 35). Measurements are more accurate when they have smaller error values.

Precision

Precision is the repeatability and reproducibility of the results of measurement. Repeatability is the "closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement" (JCGM, 2008, p. 35). Reproducibility is the "closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement"(JCGM, 2008, p. 35).

Chapter 2: Generation of Reference Datasets

This chapter provides an overview of errors and uncertainties in field reference datasets as well as recommendations for how to minimize these errors given the current state of knowledge. Section 2.1 focuses on forest measurements in the field (e.g. physical measurements of trees), Section 2.2 focuses on uncertainties in allometric models, which estimate field biomass from measured attributes. Section 2.3 focuses on good practices for using newer technologies (Terrestrial Laser Scanning (TLS) and to a lesser degree unmanned airborne vehicle (UAV) lidar) to augment more traditional measurements and reduce uncertainties. This chapter provides guidance on good practices for ground estimates of forest biomass, and can be used in combination with field measurement protocols from plot networks as a guide for those seeking to establish new plots for biomass reference, or for those seeking to interpret existing plot data.

2.1 Field Measurement Errors

Keryn Paul, Nicolas Barbier, Grant Domke, Laura Duncanson, Tommaso Jucker, James Kellner, Oliver Phillips, and Jerome Chave

2.1.1 Background

Ideally, remotely sensed data should be related to biomass measurements obtained by harvesting and weighing all woody vegetation within an area of interest to minimise measurement errors in ground-based biomass reference data (D. B. Clark & Kellner, 2012). However, such extensive harvesting is rarely possible nor desirable, due to insufficient resources and/or due to protection laws associated with the study site. Hence, operationally, we estimate rather than measure biomass of woody vegetation using inventory studies. Given results from inventory studies are only estimates of biomass, they encompass some degree of measurement error and modeling error (D. B. Clark & Kellner, 2012). This Chapter focuses on measurement errors in the generation of biomass reference data.

There has been much interest in identifying key sources of error in stand biomass estimates (e.g. Berger et al., 2014; Chave et al., 2004; Magalhães & Seifert, 2015b; Molto et al., 2013; Shettles et al., 2015), including for calibrations of emerging technologies in remote and terrestrial sensors (e.g. Kalliovirta et al., 2005). However, many error propagation analyses have either excluded measurement errors, or made broad assumptions regarding their magnitude (e.g. Gertner & Dzialowy, 1984; Magalhães & Seifert, 2015b).

This Chapter considers field measurement errors when applying either allometric models or terrestrial remote sensing (e.g. TLS or UAV) to provide plot- or transect-based estimates of biomass stored as woody vegetation. Plot selection or placement within a study site, and the size, shape, and geolocation are considered key sources of measurement error, however they are not considered here given this topic is comprehensively discussed in Section 3.1; '*Geolocation and spatial scale*'. The application of particular models or protocols for allometric- or TLS/UAV-based estimates of biomass also induce measurement errors, but these are considered in Sections 2.2

and 2.3; 'Allometric error' and 'TLS and UAVs'. Here we consider three other key areas of potential measurement error when applying either allometric models or TLS/UAV to provide plot- or transect-based estimates of biomass carbon of woody vegetation, and provide recommendations for minimizing uncertainties in plot estimates of AGBD. We touch briefly on a) plot geolocation and extent which is discussed in greater detail in Section 3.1, and provide a discussion of uncertainties associated with b) incomplete inventories, c) tree attribution (i.e. health, species), and d) destructive harvesting of trees for inclusion in allometric models.

Note that specific recommendations for measurements (e.g. tree height, diameter, species) are often ecosystem dependent, and thus there is no single recommended protocol for field measurements in the validation of global biomass products. However, we outline the sources of uncertainty in field biomass estimation in general, and make general recommendations for reducing these uncertainties. For specific measurement recommendations we refer practitioners to Table 2.1, and recommend that measurements are collected following the protocol of the plot network for the ecosystem in question.

Table 2.1 Plot networks to follow protocols for establishment of new sites

Ecosystem	Plot Network(s)	References
Tropical forests	ForestGEO, FAO, ForestPlots.net	https://forestgeo.si.edu/protocols http://www.fao.org/3/a-ap358e.pdf https://www.forestcarbonpartnership.org/sites/fcp/files/fcp-docs/2015/October/Forest%20Inventory%20%26%20Management_Manual.pdf https://www.forestplots.net/en/using-forestplots/in-the-field/
Temperate forests	NEON, ForestGEO, NNRG, IIASA, CEPF (Russia), NULESU (Ukraine), NRCS, TERN, USDA-FS	https://www.neonscience.org https://forestgeo.si.edu/protocols https://www.nnrg.org/wp-content/uploads/2015/02/NCF-Inventory-Monitoring-Guidelines.pdf https://directives.sc.egov.usda.gov/OpenNonWebContent.aspx?content=42554.wba https://www.tern.org.au/wp-content/uploads/TERN-Forests-Field-Manual_web.pdf https://www.fia.fs.fed.us/library/field-guides-methods-proc/index.php
Dryland forests	TERN	https://www.tern.org.au/wp-content/uploads/TERN-Rangelands-Survey-Protocols-Manual_web.pdf
Mangroves	Blue Carbon Initiative, CIFOR	https://www.thebluecarboninitiative.org/manual

		https://www.cifor.org/publications/pdf_files/WPapers/WP86CIFOR.pdf
Boreal forests	NEON, IIASA, CEPF (Russia)	https://doi.org/10.1038/s41597-019-0196-1 https://link.springer.com/article/10.1186/s40663-019-0165-3 Sukachev V.N., Zonn S.V., Metodicheskie ukazanija k izucheniju tipov lesa (Guidelines for the study of forest types), Moscow: Izd-vo AN SSSR, 1961, 144 p.

2.1.2 Incomplete Biomass Inventories

Plot-level biomass estimates often have errors associated with incomplete inventory of individual trees within a plot. These are both by design (e.g. not sampling trees smaller than a given diameter) or accidental (e.g. through uncertainty of plot boundaries leading to undersampling of trees at a plot's edge). These inadvertent and intentional incomplete inventories are discussed below.

2.1.2.1 Inadvertently incomplete inventories

Sources of error

Even if plot boundaries are accurate, inventories of individuals within the plot may be incomplete due to operators inadvertently excluding certain sources of biomass. For example, during the inventory assessments, new recruits such as re-sprouting trees or shrubs may be missed, or some larger trees may be inadvertently missed (e.g. O. L. Phillips et al., 2002). This will result in under-estimates of plot-level biomass. Similarly, if plot boundaries are inaccurate, 'edge effect' errors in AGBD estimates may be incurred through inconsistencies in the inclusion of individuals along the plot boundaries. Plot boundaries may be inadvertently over- or under-estimated such that individuals along the boundary may be incorrectly included or excluded from the sample which, in turn, results in over- or under-estimates of plot-level biomass.

Examples of quantification of measurement errors

To our knowledge, there are no previous studies that provide quantification of the error attributable to operators inadvertently either: (i) missing accounting for trees or new recruits such as re-sprouting trees or shrubs when undertaking plot inventories, or; (ii) incorrectly marking plot borders during inventories of stand biomass assessments, thereby assigning individuals measured within these plots to a plot area that is incorrect. In challenging terrains and/or dense vegetation that impede an operators 'line-of-sight', accurate placement of measurement tape or pegs indicating plot boundaries may be compromised. These conditions make it difficult to ensure each individual within the plot is accounted for during the inventory, and also make it difficult to ensure accurate plot layouts.

Recommendations for validation

In stands of relatively high stocking densities, **it is recommended** that operators minimise the risk of inadvertently missing individuals within the plot by visually marking (e.g., by tape or spray paint) each individual once it has been accounted for. This will make it easier for operators to check that *all* individuals within the plot were actually measured as expected. Under steep or complex topography, bearings and tape measurements will not coincide. If the reference area considered has to be horizontal (i.e., perpendicular to the local zenith), then the bearings have to be trusted when selecting which individuals to include in the inventory.

To minimise edge-induced measurement error, particularly in such challenging environments, **it is recommended** that small plots be avoided as these have relatively large perimeter-to-area ratios. Forest Observation System values of aboveground biomass (AGB) and canopy height estimates with their associated uncertainties are provided at a 0.25 ha scale from field measurements made in permanent research plots across the world's forests. For permanent sample plots with multiple re-measurements, edge-induced measurement error may also be minimised through clearly marking plot boundaries using permanent corner posts or metal stakes for plot and/or subplot boundaries. In the case where single measurement inventory plots are desired (which would not follow our primary recommendations but may be used to train local biomass maps or as part of a forest inventory program not primarily focused on biomass), edge-induced measurement error can be minimised through the use of circular plots or transects (i.e. long narrow plot, e.g. 100 m long by 5 m wide). In circular plots, once the central point is located, a tape measure of set length may be extended along the radius to a point along the perimeter to confirm whether or not an individual is included within the circular plot. Similarly, in transects, once the starting point has been located, a tape measure of set length (e.g. 100 m) may be placed on the ground along a set bearing. Then, the operator may confirm whether or not an individual is included in the transect by using a stick of set length (e.g. 2.5 m) to extend either left or right from the tape, thereby clearly indicating the boundary. Note that in areas with slopes, this measurement needs to be adjusted according to the slope.

2.1.2.2 Deliberate incomplete inventories

Sources of error

A second related source of error in stand-level biomass may be attributable to inconsistencies in the deliberate exclusion of certain sources of biomass. For example, it is relatively common in many biomass assessment studies that stems below a lower size class limit be deliberately excluded on the basis that they are laborious to measure and contribute relatively little to total stand-level biomass (e.g. D. B. Clark & Kellner, 2012). Similarly, relatively inaccessible species such as woody climbing plants (e.g. lianas or vines) are often deliberately excluded from inventories, thereby further contributing to under-estimates of biomass in some forest types (e.g. Chave et al., 2004; O. L. Phillips et al., 2002). Some forest inventory assessments of stand biomass also deliberately exclude understorey species (K. D. Johnson et al., 2017) and any individuals that are assessed as being predominantly dead (Chao et al., 2009).

Examples of quantification of measurement errors

Small individuals: Small trees (<10 cm diameter) tend to not be included in inventories for stand biomass assessments given the common assumption they contribute relatively little to total biomass. However, Schroeder et al. (1997) showed that their contribution depends on the successional stage. They found that in stands of <50 Mg/ha, biomass with trees of <10 cm DBH contributed ~75% of the biomass of trees >10 cm DBH, whereas the proportion dropped to 10% for stands with aboveground biomass >175 Mg/ha. Thus, for most temperate hardwood forests, ignoring the small diameter trees may significantly underestimate total carbon storage in live biomass.

Woody climbing species: Although lianas are commonly excluded from inventories, these have been shown to represent up to 5% of total biomass in some tropical forests (Hegarty & Caballé, 1991), and are expected to increase their proportional contribution under a changing climate and increased disturbance regimes (O. L. Phillips et al., 2002). Additionally, given their relatively higher leaf:mass ratio, liana signals will have a disproportionate impact on satellite remote sensing signals compared to trees.

Understorey species: Recent work by Johnson et al. (2017) indicated that in various forest types across the United States (US), the exclusion of understorey species would result in an under-estimate of total biomass, with understorey species comprising approximately 2% of the total aboveground live tree carbon pool.

Dead individuals: Clearly, as for small, climbing and understorey individuals, the errors in stand biomass estimates that result from exclusion of dead individuals in the inventory will depend on their relative contribution to the population of the total stand. In mature tropical forests necromass can form a large fraction of total aboveground biomass. For example, in Amazonia the mean ratio of necromass to aboveground biomass is 0.127, implying that mature Amazonian forests store a total necromass of 9.6±1.0 Pg C (Chao et al., 2009). But this varies substantially, with this study showing necromass is twice as great in forests with low stem mortality rates than in forests with high stem mortality rates (58.5±10.6 and 27.3±3.2 Mg ha^{-1}, respectively). The impact of excluding dead individuals is therefore substantial, and variable, and typically exceeds the impact of excluding understorey trees or lianas.

Recommendations for validation

It is recommended that inventories of AGB be complete, with all components of AGB being measured, even if resource limitations necessitate that components be measured with a precision that reflects their relative likely contribution to total stand biomass. For permanent biomass reference plots, in particular, measurements of small trees, standing dead trees, and commonly ignored species such (e.g. lianas, palms) should be measured. For example, for small trees (<10 cm diameter) or dead trees, **it is recommended** that they be included in the inventory, albeit with their stem diameters measured using highly efficient instruments such as the Stepped Diameter Gauge (Paul, Larmour, et al., 2017). **It is also recommended** that inventories include

woody climbing species using standardised protocols (e.g. O. L. Phillips et al., 2002). For non-permanent plots measured to train airborne datasets, where logistics for taking these additional measurements are challenging, estimates of the relative contribution of AGB from these other pools should be made (e.g. from smaller plots as in the US Forest Inventory and Analysis (FIA) program).

In summary, to minimize uncertainties regarding incomplete inventories, **we recommend**:

- Large plots (≥1 ha)
- Visually tagging trees that have been measured (e.g. with tape)
- Measurement of all trees that contribute to biomass (including small trees, lianas, palms, etc.), using subplots to estimate relative contributions where logistics prevent measurement of every individual.
- Measurement and recording the status of standing dead trees

2.1.3 Attribution of individuals within the plot

Sources of error

Individual trees are typically attributed with a botanical identification (species, family, plant functional type (PFT)), and as live or dead. All individuals within selected plots are generally assessed for these key attributes, as these are considered to have an important influence on their biomass. Different species or groups of species have differing structure and stem woody density, and hence, differing biomass. Another common attribute considered is a subjective assessment of the health of the individual (e.g. dead, or partially dead, etc.) given the extent of dead woody material or hollows that an individual contains will influence its average stem wood density. This is particularly important in stands with high rates of mortality. If the attribution of such characteristics is inconsistent between operators or studies, errors will result from incorrect assignment of allometric models or estimates of stem wood density.

In tropical forests, identifying trees is a challenge that integrated assessments of biomass have to address. This is because tree species composition regulates how much carbon forests store (Baker et al., 2004). The biomass contained in trees is governed by the volume and density of their wood, and while the former can be sensed remotely by satellite, the latter in general either cannot, or is estimated indirectly as a function of species / PFT (Ustin & Gamon, 2010). The density of tree wood is mostly determined by a tree's genetic identity, and so (except in rare occasions where wood density may be directly measured for every single tree) we need to identify the tree. This is ideally done to species, and certainly at least to genus with which most variation in wood density is associated (Baker et al., 2004). This challenge is greatest in tropical forests where there are often thousands of tree species and several hundred tree genera present locally. Even a single hectare of forest in Amazonia can have up to 300 tree species, some potentially unknown to science (Gentry, 1988).

Examples of quantification of measurement errors in tree attribution

Errors or inconsistencies in attribution between operators or studies may result in biased estimates of stand biomass. However, this is discussed only briefly here given that errors or inconsistencies in attribution of species within plots manifests through to possible biases via the application of inappropriate models using either allometrics or TLS/UAV based estimates of biomass, which is discussed in detail in Section 2.2 and 2.3, respectively.

As an example, in studies where allometric models based on PFT are applied (e.g. Paul et al., 2016), an incorrect attribution of a tree to a category of PFT will result in a 16-37% error in the individual's aboveground biomass, depending on the tree size. Incorrect attribution of a shrub or small multi-stemmed tree to another plant functional type will result in a 24-36% error in the individual's aboveground biomass, depending on its size (K. I. Paul, personal communication, 2019). Similarly, based on generic allometric models, incorrect attribution of a eucalypt tree or acacia shrub to a 'live' rather than 'dead' condition will result in errors of up to 36-39%, depending on the size of the individual (K. I. Paul, personal communication, 2019). Clearly, given the magnitude of these possible errors at the individual-level, errors at the stand-level will depend on: (i) how many individuals are inaccurately attributed, and; (ii) whether there is bias resulting from consistent incorrect attribution, e.g. of 'dead' individuals being attributed as 'live', etc.

With respect to tree species identification, identity can impact forest biomass via volumetric allometric differences and especially due to wood density. Species composition varies at all spatial scales, thus bias and uncertainty result if individual identity is ignored. A recent pan-tropical review found that compositional differences cause variation in forest biomass and carbon density of up to 20% locally, and that additional large-scale floristic variation leads to variation in tropical forest mean wood density of up to 30% (O. L. Phillips et al., 2019). Across the lowland tropics basal area-weighted wood density values range widely, from 0.467 to 0.728 g cm^{-3} at the regional scale (and is more variable still for individual forest plots). Consequently, tropical biomass assessment requires locally validated measurement of tree-by-tree botanical identity.

Recommendations for validation

To minimise errors resulting from inaccurate or inconsistent attribution of individuals within the inventory sample, **it is recommended** that; (i) standard protocols be made available to provide clear definitions of the different categories for guiding attribution to available allometric models or TLS/UAVs; (ii) these standard protocols include clear instructions on how to attribute individuals as either 'live' or 'dead', and; (iii) operators be well trained and experienced in identification of individual trees and shrubs.

Efforts to map and monitor tropical forest carbon using remote sensing techniques should be combined with tree-level measurement of species identity by botanists working in inventory

plots. This requires integrating botanical expertise and associated costs into the core processes of biomass assessment.

It is also recommended that the biomass community ensure that they work with trained botanists dedicated to making and curating herbaria collections (vouchers), and do not rely simply on the allocation of names in the field. This is because plant identification requires collection of material, making herbarium vouchers from these, identifying these in herbaria, and storing them permanently. Thus, with vouchers we have scientific reproducibility, and names become hypotheses which can be tested over time. Without vouchers, names are opinions, and their quality cannot be evaluated.

Further, **it is recommended** that biomass reference plots be permanent samples, not one-off forest inventories. This is because tropical trees are sterile most of the time and the lack of flowers or fruits degrades identification quality. Given the often short and unpredictable phenologies of many tropical trees, forest inventories miss the reproductive period of most species (Baker et al., 2017). Permanent plots that can be repeatedly collected provide greater long-term potential for reliable identifications, and so minimise uncertainty in wood density.

It is recommended to upload digital copies of collections to digital herbaria archives that are integrated with forest plot databases. This is because accurately assessing spatial variation in species (and hence wood density) and tracking long-term biodiversity change is challenging to do consistently for large numbers of plots. We also need to standardise the process of identification across time and space, requiring the ability to make side-by-side comparison of multiple voucher specimens digitally. High-resolution specimen images are increasingly available online: for example ForestPlots.net (Lopez-Gonzalez et al., 2011) is implementing workflows and software to allow on-screen comparison of multiple high-resolution plot voucher specimens and images of living plants.

We note that the benefits of working with established permanent plots goes beyond simply improving confidence in identifications. There are other benefits including being able to match individual trees from TLS surveys (Section 2.3) to tagged and censused individual trees (See Section 2.3), e.g. via post-hoc matching from the registered point cloud. Integration with repeatedly measured biomass reference sites where long-term botanical and ecological inventory is supported will provide the highest quality and greatest certainty to the biomass community (Chave et al., 2019).

In summary, to minimize errors in tree species identification and health attribution **it is recommended** that:

- Plots are collected following existing measurement protocols in an appropriate forest plot network for the ecosystem in question (e.g. ForestGeo, Forestplots, Afritron)
- New plots be integrated into an existing plot network
- Standard protocols are adopted for classification of trees as live vs. dead, again following existing measurement protocols in an appropriate forest plot network for the ecosystem in question (e.g. ForestGeo, ForestPlots.net, Afritron)

- Forest reference measurements are collected by individuals trained for attribution of trees as live/dead
- Well-trained botanists familiar with the regional flora are involved in plot collection
- Plots become permanent plots, integrated into an existing plot network and setup with the intention of long-term measurement
- Digital copies of plot data are uploaded to digital herbaria archives that are then integrated with established forest plot databases.

2.1.4 Measurement of stem diameter

Stem diameter (D) is one of the most easily measured variables commonly applied in allometric models predicting biomass or wood volume (e.g. Berger et al., 2014). There is variation between studies in the tools used for D measurement and the height of the stem at which these measurements are made. Tools used for D measurement include stem diameter tape, callipers, Biltmore stick, laser callipers, sector-fork, terrestrial laser scanner, laser-camera, laser-relascope, and stepped diameter gauges (SDG). Regardless of the tool used, for commercial tree species, D is traditionally measured at a height of 130 cm above the ground (diameter at breast height, DBH, or D130). Note that certain countries use different standards, e.g. DBH of 120 cm in Japan, which are also acceptable provided the height of DBH measurement is reported in metadata. This is ergonomic and avoids lower positions where the stem is often non-circular (Biging & Wensel, 1988; Gregoire et al., 1990). But for many shrubs and young or multi-stemmed trees, D measurement is impractical at this height, with measurements generally taken below the most common point at which the stem divides into multiple leaders (West, 2009), typically at 50, 30 or 10 cm above the ground; D50, D30 or D10. Another important protocol for D measurement entails the handling of multi-stemmed individuals, which is often ecosystem dependent. In many dry and temperate forest and woodland regions, all measurements of multi-stemmed individuals (D_i) are converted to a single value (equivalent stem diameter, $D_e=\sqrt{\Sigma D_i^2}$, cm), such that the total basal area (cm^2) for all stems is equal to the basal area of a tree with this equivalent single diameter. In other cases, each stem is considered as an individual, and ladders are used to measure D above the split. In other cases, again, the tree is treated as a single stem if the split occurs above 130 cm, or multiple stems if the split occurs below 130 cm. The selection of the method for determining whether a split tree is considered as one or two individual stems has large implications for biomass estimation. **It is recommended to follow** a consistent protocol with the appropriate field plot network for the ecosystem in question (see Table 2.1).

Sources of error

As for all measurement variables applied in allometric models, bias in D measurements can scale-up to bias in stand-scale estimates of biomass or volume (Methley, 2001). For example, a bias in D130 of 5% will cause the volume per hectare to be biased by 15% (Gertner, 1990). Increasing the sampling intensities will not decrease bias. The only means for reducing the bias is through the reduction of the bias in the actual D measurements.

Bias may result when operators are untrained (e.g. consistent incorrect placement of the tool above or below the specified height of D measurement), and when different tools are used to measure D. As discussed below, when selecting tools to measure D, trade-offs need to be considered between the time (and therefore resources) required to use them and the potential for some bias, particularly where this facilitates a greater accuracy of stand biomass estimates as more individuals can be measured (i.e. decreased sampling error) for the same level of resourcing (Paul, Larmour, et al., 2017).

Although random errors may cancel out at the stand scale, there will nonetheless always be some imprecision of D-derived stand biomass estimates due to random errors in these D measurements. Random errors in D measurement may result from a combination of recording errors (either in reading the tool or documenting results), divergences in tool placement with respect to the stem (e.g. around various deformities in the stem), placement of the tool above or below the specified height (e.g. often required in complex systems to avoid stem defects or branching), and if using the tape, differences in tape tension (e.g. often required for rough- or loose-barked trees) (e.g. Elzinga et al., 2005), or measuring in close proximity to irregularly shaped stems (e.g. Weaver et al., 2015). The extent of these errors is influenced by five main factors that influence the extent of bias and precision of D measurement: (i) operator, particularly with respect to their level of training; (ii) height of the stem at which D measurement are made; (iii) form of the individual, e.g. single- or multi-stemmed individuals; (iv) size of the individual, and; (v) tool used (Paul, Larmour, et al., 2017).

Examples of quantification of measurement errors

Clearly an untrained operator may propagate bias if they consistently fail to correctly apply protocols. One common example has been the tendency for some operators to measure trees at a standard height regardless of any stem deformity, thereby causing positive bias in estimated basal area and biomass (O. L. Phillips et al., 2002). If protocols for D measurement of buttress trees are not followed to measure D above buttress, then this will result in an overestimation of biomass via allometric model application (D. A. Clark, 2002; O. L. Phillips et al., 2002).

Among trained operators no bias has been found in D measurement errors (Kitahara et al., 2009, 2010; Paul, Larmour, et al., 2017). Here the main source of bias in D measurements relates to the tool used. Numerous studies have demonstrated that differences in D measurements between alternative tools (e.g. diameter tap, Biltmore stick, calliper and SDG) increases with tree size (Moran & Williams, 2002; Wilson et al., 2007), and when measuring D of individuals with single-stems when compared to multi-stemmed individuals (Paul, Larmour, et al., 2017). Nonetheless, differences in D measurement between tools were generally negligible; averaging <0.17 cm (Behre, 1926; Gregoire et al., 1990; Kalliovirta et al., 2005; Paul, Larmour, et al., 2017; Vastaranta et al., 2008; Weaver et al., 2015).

Although not well quantified, another source of bias may result from measurement of changes in D over time within permanent sample plots. Both positive and negative biases are possible, but effects are likely to be small on AGB. For example, small bias may result from rounding up

negative values in measured changes in stem diameters, and where nails were used to mark locations for D measurement, localized swelling of trees around nails may result in erroneous results (O. L. Phillips et al., 2002). Another issue with measuring changes in D is buttress creep, with bole irregularities moving up with time which requires movement of point of measurement (POM) higher up the stem (O. L. Phillips et al., 2002).

In contrast to bias, random errors in D measurement have been shown to be relatively large, with a mean standard deviation (SD) averaging <[1] cm, and a coefficient of variation (CV) of D measurements generally in the order of 2–8% (Auclair, 1986; Barker et al., 2002; Berger et al., 2014; Elzinga et al., 2005; Holdaway et al., 2014; McRoberts et al., 1994; Myers, 1961; Nester, 1981; Omule, 1980; Paul, Larmour, et al., 2017). The frequency and magnitude of these random errors are relatively consistent between the trained operators, with negligible differences in repeatability regardless of whether the repeat measures were made by different operators or multiple measurements from the same operator (Auclair, 1986; Paul, Larmour, et al., 2017). However, this may not be the case when untrained operators are employed. Larger SD (up to 2.6 cm) and CVs (up to 12.8%) have been observed when repeatability of D measurement was assessed amongst untrained operators (Kitahara et al., 2009, 2010).

When employing trained operators, the magnitude of random errors in D measurement will be predominantly influenced by: (i) height at which D was measured; (ii) size of the individual, (iii) form of the individual; and; (iv) tool used. These are discussed in turn below.

When estimating plot level biomass and associated uncertainties, statistical tools can be used to propagate measurement errors. For example, the BIOMASS package in R accounts for two types of measurement errors of D - small, frequent random errors and large, infrequent errors (e.g. due to misplacing the decimal place when transcribing values). Thus, **it is highly recommended** to report these types of errors in field surveys.

Height of the stem: Recent work (Paul, Larmour, et al., 2017) has shown that for large individuals (i.e. those having D > 10 cm), random errors are generally higher (often significantly) at D10, D30, or D50 when compared to D130. In contrast, for small individuals, random errors were generally the highest at D130 when compared to D10, D30, or D50. These findings may be attributed to large individuals having more random errors at lower heights due to an increase in the occurrence of irregular stem shapes at these relatively low stem heights (e.g. Weaver et al., 2015), while for small individuals, random errors at lower heights may be relatively small due to a smaller number of stems at these lower stem heights (see below).

Size of the individual: Although the standard deviation (SD) in D measurements increases with the size of the individual (Holdaway et al., 2014; McRoberts et al., 1994; Omule, 1980), when random errors are expressed in relative terms (CV), the influence of the size of the individual becomes less important. Indeed, Paul et al. (2017) found that large individuals (i.e. those having D > 10 cm) had slightly lower CV (1.9–4.9%) of D measurement when compared to smaller individuals (3.1–7.1%).

Form of the individual: Due to the possible accumulation of errors from multiple measurements, total random errors will be higher when measuring multi-stemmed individuals (and then calculating D_e) when compared to single-stemmed individuals. Paul et al. (2017) found that the mean CV of D measurement was 1.0 to 1.2% higher for multi-stemmed individuals when compared to single-stemmed individuals. It is likely that when compared to single-stemmed individuals, when measuring D of multi-stemmed individuals, operators have greater challenges accessing the stems to measure, resulting in a greater frequency of operators failing to judge the target height of D measurement, account for differences in tape tension, or place the instrument in its proper plane.

Tool used: Although previous comparison of tools has focused on bias rather than precision, the study of Paul et al. (2017) did compare the precision of D measurement between diameter tape and the SDG; a tool having, on the end of a measurement handle, a metal plate with stepped increments of 1 cm with gaps designed to slip over stems of varying D sizes from 2 to 16 cm. They found that regardless of the form of the individual, when measuring D of small individuals, random errors were generally larger when using the SDG compared to tape. By comparison, for large-multi-stemmed individuals, outlier measurements amongst repeat measurements (leading to particularly large SDs relative to the equivalent CV) were most common when measured with tape. It is possible that use of the SDG can avoid erroneous readings as it can be quickly and easily rotated around stems of irregular shape to ensure the operator is satisfied that the reading recorded is not affected by loose bark, stem defects, or other variations. Indeed, 75% of the observations exceeding the 95% confidence interval of SDs for large single-stemmed individuals (i.e. SD >1.57 cm) were estimated from measurements made with tape.

Recommendations for validation

To minimise bias, it is paramount that operators be well trained in D measurements. This will be particularly important for operators measuring stands with numerous buttressed trees (D. A. Clark, 2002; O. L. Phillips et al., 2002). In addition, it will also be important to ensure that the tools selected are appropriate for the level of accuracy required at the individual and stand scales. In permanent sample plots where detection of small changes in D is required, accuracy of D measurements should be maximised by avoiding tools which may introduce a small amount of bias, e.g. SDG (Paul, Larmour, et al., 2017).

To minimise random errors, D measurements of large individuals (D>10 cm) should be made at a height of 130 cm (D130) (or above, when the stem is buttressed), particularly when the species is single-stemmed. For such large, single-stemmed individuals, the tool used has little influence, although use of the SDG may minimise the frequency of reading errors. For smaller individuals that are either single or multi-stemmed, random errors may be minimised by using the tape at measurement heights <130 cm (e.g. D10, D30 or D50), but noting that additional care is required in irregular-shaped large single-stemmed individuals.

In summary, to minimize errors associated with stem diameter measurement **it is recommended**:

- Training practitioners to follow protocols in determining appropriate height to measure stems (e.g. above buttress, for multi-stems)
- For small or irregular stems, using stepped diameter gauge (SDG) instruments
- For large (>10 cm) and/or regular stems, diameter tapes, calipers and other diameter measurement tools are all appropriate.

2.1.5 Measurement of tree height

Sources of error

Tree height (H; the vertical distance between the ground level and the tip of the tree) is a primary variable used in the estimation of tree and stand volume, and hence biomass. It is therefore an important variable in many inventories. However, H measurements are usually one of the more time-intensive, and therefore expensive, components of forest inventory, especially in tropical forests where tall, closed canopies, and dense understory occur, limiting the sight of tree tops.

The most direct method for measuring H (up to 25 m) involves the use of height poles. But due to the practical difficulties in measuring H directly, indirect methods are often used, generally what is referred to as the "tangent method", which involves clinometer measurement of angles to the tree base and treetop, and the horizontal distance to the tree stem. Hand-held laser rangefinders (with electronic measurement of distances and angles) are increasingly being used to estimate H. However, rangefinders are difficult or impossible to implement in closed stands where the treetops are not easily visible.

The emergence of airborne and terrestrial lidar provides individual tree height measurements that are highly correlated with field-derived measurements (Andersen et al., 2006). These, and other alternatives such as application of stereo photogrammetry (St-Onge et al., 2004) are providing efficient alternatives to traditional field-based H measurement techniques.

Examples of quantification of measurement errors

Measuring H (up to 25 m) directly using height poles is reliable but susceptible to parallax error that can range as high as 10% (Schreuder et al., 1993). Larjavaara and Muller-Landau (2013) quantified systematic and random errors of the tangent method and laser rangefinder method as applied by five technicians to 74 trees between 5.7 and 39.2 m tall in a Neotropical moist forest in Panama. They found that the tangent method produced unbiased H estimates whereas the laser rangefinder resulted in systematic underestimates of 20% on average. However the reverse was true for random error. Random error was high using the tangent method, with overestimates of H >100% being found in about 2% of the H measurements made. The laser rangefinder method was faster to learn, displayed less variation in heights among technicians, and so had lower random error. Others have reported that laser rangefinders can yield H measurement errors of only 1%-2% (Wing et al., 2004). But regardless of whether the tangent method or laser rangefinder method is used, the precision of field H measurements will be impacted by including

the offset between measured distance and crown-top position, tree-top occlusion, ground slope, obstacles for distance measurements, and clinometer operator error (Hunter et al., 2013).

Underestimation of H has been found with airborne laser scanning (ALS), TLS and photogrammetry techniques, although in variable amounts according to local and instrument conditions (Laurin et al., 2019; St-Onge et al., 2004; Y. Wang et al., 2019). This is because compared to field measurements of H described above, these remotely-sensed measurements of H are less sensitive to stand complexity, crown classes, and species, with occlusion effects leading to omissions of trees in intermediate and suppressed crown classes (e.g. with ALS), and incomplete crowns of tall trees (e.g. with TLS) (Y. Wang et al., 2019). Hence, Wang et al. (2019) concluded that: (i) ALS-based H was most accurate for tall trees, and least accurate for trees in intermediate crown class, due to the difficulty of identifying treetops, and (ii) TLS-based H was reliable for trees lower than 15–20 m in height, depending on the complexity of forest stands. Laurin et al. (2019) found photogrammetry techniques resulted in lower estimates of H when compared to ALS and TLS, and that ALS and TLS gave similar estimates of H. Errors in the ALS-derived H measurement could be up to ±0.5 m (Leckie et al., 2003), or as low of 0.02-0.3 m (Andersen et al., 2006; Reutebuch et al., 2003), with high-density, narrow beam lidar being significantly more accurate (both in terms of bias and precision) than wide-beam lidar (Andersen et al., 2006). Simultaneous use of different methods may help minimise the uncertainty in H (Laurin et al., 2019). For example, using a combination of lidar and photogrammetry, St-Onge et al. (2004) found biases averaged 0.59 m, but with bias correction using a calibration subset, this was decreased to 0.02 m.

Recommendations for validation

The emergence of airborne lidar as a forest measurement tool has dramatically increased the efficiency of forest inventories. Although they may not always yield errors as low as field methods, lidar-derived H measurements reduce the cost and increase the efficiency of the inventory, thereby increasing overall accuracy at the stand-scale. **It is also recommended** that species-specific correction factors be applied to lidar-derived H measurements to maximise the accuracy of measurement (Andersen et al., 2006). The best method for producing unbiased measurements for H consists of first assessing the bias on a calibration subset of trees and then correcting all other measurements accordingly (St-Onge et al., 2004).

In summary, to reduce uncertainties associated with tree height measurements **it is recommended**:

- Collection of height data in the field using laser range finders, following plot network protocols
- Attribution of heights to tree locations from canopy height models derived from high resolution, high point density airborne or UAV-LS or
- Attribution of height to trees through the collection of TLS data

2.1.6 Importance of minimising measurement error: Allometrics vs. Inventory

The required precision of D or H of will depend on the type of study. When D or H measurements are used for model development (e.g. allometric or stem volume models, or training terrestrial or remote sensors), or for monitoring small changes in attributes over time (e.g. repeat D inventory measurements of relatively small permanent sample plots in National Forest Inventories), it is essential that errors are minimised (West, 2009). This is discussed in Section 2.2.

In contrast, for stand-scale inventory studies that estimate area-based AGB at the site-level, random errors in D or H will largely cancel out as the number of measured individuals increases (Auclair, 1986; Cunia, 1986; Gertner, 1990). Hence when resources are limited, the accuracy of inventories of AGB is maximised by measuring more individuals, rather than spending more time maximising the accuracy of D measurements of individuals (Paul, Larmour, et al., 2017). This is because the dominant source of random error in inventory-based estimates of stand AGB is the sampling error (i.e. error resulting from insufficient sampling of a population, (Dot, 2016)).

Inventory surveys for estimation of stand AGB often have limited resources. Therefore, maximising stand-scale accuracy by ensuring sufficient sampling intensity will require consideration for the efficiencies of measurement of D and H. For example, when measuring D, the SDG or the Biltmore stick are about twice as quick as the tape (Wilson et al., 2007). Efficiency gains for some other tools tested have been less promising. For example, Kallivirta et al. (2005) found that the relative time of laser-relascope to the callipers was 0.97.

2.1.7 Summary and current knowledge gaps

Field measurement errors need to be fully accounted for in the error propagation (Chapter 4). As outlined by Clark and Kellner (2012), a global database of biomass harvest plots will require a significant research commitment, but would improve our ability to quantify measurement errors resulting in errors in ground-based estimates of biomass, and how these may be minimised.

To minimize uncertainties in the collection of forest biomass plot measurements, the following are recommended:

For plot location:

- The use of differential Global Positioning System (GPS) as available, with <100 km distance from base stations
- The collection of >15 Global navigation satellite systems (GNSS) points per location at plot establishment, each acquired at least 30 minutes or 10 meters apart (separations in space or time)
- The collection of large (minimum 1 ha) plots
- Where possible, the correction of field plot locations through the use of UAV, TLS or airborne lidar height datasets

For tree species and status identification:

- Visually tagging trees that have been measured (e.g. with tape)
- Measurement of all trees that contribute to biomass (including small trees, lianas, palms, etc.), using subplots to estimate relative contributions where logistics prevent measurement of every individual.
- Measurement and recording the status of standing dead trees
- Plot data are collected following existing measurement protocols from an established forest plot network for the ecosystem in question (e.g. ForestGeo, ForestPlots.net, Afritron)
- New plots be integrated into an existing plot network
- Standard protocols are adopted for classification of trees as live vs. dead
- Forest reference measurements are collected by individuals trained for attribution of trees as live/dead
- Well-trained botanists are involved in plot collection
- Plots become permanent plots, integrated into an existing plot network and setup with the intention of long-term measurement
- Digital copies of plot data are uploaded to digital herbaria archives that are then integrated with forest plot databases.

For diameter measurement:

- Training practitioners to follow protocols in determining appropriate height to measure stems (e.g. above buttress, for multi-stems)
- For small or irregular stems using SDG instruments
- For large (>10 cm) and/or regular stems, diameter tapes, calipers and other diameter measurement tools are all appropriate

For tree height measurement:

- Collection of height data in the field using laser range finders, following plot network protocols
- Attribution of heights to tree locations from canopy height models derived from high resolution, high point density airborne or UAV lidar or
- Attribution of height to trees through the collection of TLS data

2.2 Allometric Errors

Jerome Chave, Keryn Paul, Stephen Roxburgh, Dmitry Schepaschenko

2.2.1 Background

Tree biomass is almost never measured directly, as this implies destructive harvesting of all trees within a stand. Rather, biomass inference is usually model based: a sample of trees is destructively harvested, and the data are used to relate tree aboveground biomass to predictive variables, such as trunk diameter or tree height, through a statistical model.

In woody biomes, biomass reference datasets are deduced from tree inventories using biomass allometric models (BAMs). In tree inventories, all stems above a conventional trunk diameter are mapped, tagged, identified, and measured within one or a collection of contiguous plots (see Section 2.1). Practical issues and proper quality assessment procedures for field measurements in tree inventories are covered in Section 2.1 and Chapter 3. This chapter is concerned with the estimation of live aboveground biomass from in situ reference datasets using BAMs (S. Brown et al., 1989).

Some guiding principles prevail during the construction of biomass allometric models. They are borrowed from the principles of allometric theory, briefly reviewed here. In all living organisms, larger linear organism size implies larger mass, and this relation is called an allometry (Calder, 1984; Huxley & Teissier, 1936; Niklas, 1994). Flowering plants vary in size over eight orders of magnitude, from minute organisms less than a gram in total weight to giant eucalypt trees that are reported to weigh in excess of 200 tons aboveground (Sillett et al., 2015).

Theory suggests that tree height is allometrically related to tree dimensions. The first theoretical consideration is the pipe model theory (Shinozaki et al., 1964): trees may be represented as an ensemble of vessels running from the leaf to the fine root, and operating independently. Trees may then be seen as a bundle of vessels constricted at the trunk and expanding towards the crown and the root system. Thus a tree could be theoretically folded into a cylindrical object, and its volume would then scale as the product of total tree height and trunk cross-sectional area. Allometry has also stimulated the development of the West-Brown-Enquist theory of power law scaling in organismic biology. Applied to plants, light-harvesting constraints imposed an isometric relationship between tree basal area (i.e. trunk cross-sectional area) and tree crown area (Enquist & Niklas, 2002). This theory also predicts that tree biomass should scale as the 8/3th power of trunk diameter.

Assuming a conical tree with basal diameter *D* and total height *H*, the aboveground tree volume is $f \times H \times \frac{\pi}{4} D^2$, where f is a form factor, equal to 1 if the cone is in fact a cylinder. The conversion factor from plant volume to oven-dry biomass is the wood density, ρ (in g/cm^3), defined as oven-dry mass divided by green volume. So, aboveground biomass (B) may be deduced from the following model $B = f \times \rho \times H \times D^2$. This is the simplest example of a BAM,

and it may be used to convert tree trunk diameter, height and wood density into an estimate of oven-dry biomass (in kg).

In general, however, the BAM used for biomass estimation are empirical functions, and are not deduced from theoretical considerations (S. Brown, 1997; D. B. Clark & Kellner, 2012). Errors associated with the design of BAM and their proper use are discussed. Biomass estimation from tree inventories with BAM produce an estimated value, which could be the mean estimate or the most likely estimate (depending on the inferential framework), together with a confidence associated with this estimate. Here we summarize the potential issues associated with this procedure.

Biomass includes both aboveground and belowground components. Aboveground biomass is much easier to estimate than belowground biomass from tree inventories, and therefore a large part of the carbon accounting literature focuses on aboveground biomass. Belowground biomass is often inferred from aboveground biomass using stand-scale relationships (Cairns et al., 1997; Mokany et al., 2006), or tree-level relationships (Ledo et al., 2018; Poorter et al., 2015). In many temperate and tropical forests, live belowground biomass accounts for ca. 20% of the total biomass, but can be as high as ca. 60% for some trees adapted for lower rainfall (Paul et al., 2019). Biomass is sometimes defined to include both living and dead biological matter, but here we refer to biomass as live biomass. Methods for estimation of dead biomass in forest stands differ from that used from live biomass, involving protocols of coarse woody debris sampling (Keller et al., 2004), and soil organic carbon sampling (Jobbágy & Jackson, 2000).

2.2.2 Reference tree biomass measurements

Sources of error

Development of BAM requires the measurement of biomass of selected individual trees. Given this generally requires field measurement of fresh weights of biomass, errors in dry weights may result from sampling errors when selecting representative individuals measured, accuracy of the load cells and balances used, and errors associated with sawdust and attached vegetation like epiphytes.

Fresh weight measurements may be converted to dry weight estimates using moisture content (MC) corrections. Because any errors in the estimation of the MC correction are translated proportionally to the biomass prediction of an individual tree or shrub, care is required to ensure MC estimates are unbiased and as precise as possible. Protocols currently applied to attain MC differ in bias and precision (Paul, Roxburgh, et al., 2017), but the main source of error is sampling errors when selecting sub-samples of components (e.g. branches, bark, leaves, etc.) to represent moisture contents of those components. Errors in MC determination based on precision of the instrument, or sampling error for the individuals selected to represent the plant functional type, were relatively minor when compared to sampling errors associated with sub-sampling components of the individual selected.

For large trees, biomass is deduced from an estimate of stem volume based on measurement of tree sections, rather than on fresh weights (Henry et al., 2010). Large trees are usually sawed into sections of 1-2 m intervals, and each component is measured, assuming that this component has a conical shape. Measurements are then used to calculate the respective branch and trunk volume. Dry mass of the plant components is then determined by multiplying volume and wood density measurements. There are measurement errors in the stem wood density, with these varying depending on the instrument used, height of measurement, and number of replicates sampled. As for errors in MC mentioned above, errors in wood density measurement are dominated by variation of wood density across sub-samples of components. Sampled wood disks are extracted from the components using a chainsaw, and are brought back to the laboratory for careful wood density determination (Fayolle et al., 2013). Stem wood density is not only a common explanatory variable for many allometric models, but is required to convert remote sensing (i.e. TLS or UAV)-based volume estimates into biomass estimates.

An additional source of measurement error is the measurement of the carbon concentration of biomass that is applied to convert biomass to carbon. This factor is commonly assumed to be a constant between 0.48 and 0.50, but recent work suggests this assumption leads to a positive bias in carbon stocks (Martin & Thomas, 2011). It has been recently shown than the C content in wood ([C]) is related to wood density (WD) through a negative linear relation: [C] = -3.5*WD+49.3, suggesting that [C] varies from 45.8% to 49.3% across woody plants (Martin et al., 2018). This is an important consideration when attempting to develop carbon allometric models. Since this protocol is concerned with biomass estimation, not carbon, this point is not further considered here.

Examples of quantification of measurement errors

Based on information on contributors to individual tree or shrub biomass collated by Paul et al. (2016), when measuring fresh weights in the field, the precision of load cells of varying capacity ranges from ±0.05 to ±1 kg, but most commonly on the order of ±0.1 kg. However, for scales used to measure sub-samples taken for moisture content determination, the precision is usually much greater; ±0.001 to ±0.1 kg.

Errors associated with MC correction have been quantified by exploring MC measurements of 1396 individuals (trees and shrubs) of various sizes (Paul, Roxburgh, et al., 2017). Using a Monte-Carlo analysis this study found MC estimates may be based on at least the bole and crown components, with bias resulting if MC is based on stem wood only, particularly in young (or small) individuals. Little gain in accuracy was attained with more intensive sub-sampling (e.g. into foliage, twig, branches, bark, and stem wood components). Variation in MC of aboveground biomass can be substantial (CV > 15%) when considered across individuals of various sizes (or age) from across differing sites (or climates). However, to minimise bias, Paul, Roxburgh et al. (2017) found it was important to undertake MC sampling at each study site, and to stratify sampling among-individuals by both appropriate taxonomic grouping (e.g. plant functional type) and age-class. For a given PTF-by-size (or age) stratum at a given site, a precision of about 4%

coefficient of variation of the average MC estimate can be achieved with intensive within- and among-individual sampling.

Existing reference datasets for BAM construction rely on legacy data, a compendium of tree harvest data collected over decades, from different operators, and with different field protocols. Thus, field measurement errors are not always well documented. For example, forestry projects sometimes report the merchantable biomass, rather than total biomass. Also, the tree stump may be ignored from the aboveground biomass total. As a result, error associated with legacy data is not always quantifiable. For recently developed reference datasets for BAM construction, error sources associated with field measurements are generally well documented. In particular, protocols for measuring wood density correctly have been published (Williamson & Wiemann, 2010) and studies have explored the minimal sampling size necessary to provide accurate estimates of wood density (da Páscoa et al., 2020; Picard et al., 2012).

Recommendations for validation

In the development of new BAM it is important that destructive harvest reference datasets be as accurate as possible. Adhering to a fixed set of measurement protocols ought to be a prerequisite for new BAMs. Destructive measurements of large trees in particular are at a premium to better-define these models. If trees, or tree components, are directly weighed, errors in fresh weight measurements of biomass should be minimized by ensuring that load cells used in the field have precision <±0.5 kg, and that operators are well trained in the use of these load cells. For the conversion of fresh weight into dry weight, **it is recommended** that MC determination be derived at each study site for each PFT-by-size stratum. **It is recommended** that six individuals be randomly sampled, and partitioned and sub-sampled into crown and bole components. For large trees, measurements are based on volume estimates and conversion into dry weight based on wood density measurements. Volume estimation should be conducted carefully on tree components. For conversion of wood volume into dry weight, **it is recommended** that the same subsampling be conducted as for MC determination.

Today, volume estimation based on TLS technology would be **a recommended alternative** to conical approximations (see Section 2.3). This could greatly facilitate the calibration of new allometric models, or even implement model-free methods of biomass estimation. However, TLS also suffers from practical issues in leaf-on conditions, with laser occlusion diminishing the quality of tree reconstructions. TLS cannot detect hollow stems, which are frequent in old-growth forests, and also TLS measures volume, not biomass, and wood density is a crucial parameter in biomass estimation. For these reasons, advances in the routine collection of TLS remains a crucial area of research, and TLS should be acquired in addition to, but not in replacement of, traditional mensuration and allometric models.

2.2.3 Tree biomass estimation from allometric models

Sources of error

Trees are structurally, widely variable, based on their growth form, life histories (e.g. due to diseases, storm damage, drought), and local growing conditions (e.g. proximity to neighbours). We do not expect that a model will precisely estimate the biomass of every single tree given this natural variability. However, if the model has random prediction error, then sample estimates of the plot biomass are likely to improve as the number of sampled trees increases. Thus, even if the estimation of individual trees may be imprecise, biomass estimates of stands may be precise.

BAMs are often reported in the form of regression models, and the fit of model parameters is obtained from a training dataset (reference tree biomass measurement, see Section 2.2.1). A classic formulation of an empirical biomass model is: $B = F(D, H, \rho)$, where F is a function of the three variables trunk diameter (D), total tree height (H) and wood density (ρ). Theory suggests that a realistic model is: $B = f \times (\rho \times D^2 \times H)$, where f is a parameter. Sources of error associated with this procedure depend on how comprehensive the reference tree biomass dataset is. More complex model formulations can always be designed, but are more parameter-rich and therefore less parsimonious.

Simpler models have also been proposed. Since total tree height is seldom measured, it has often been proposed instead to use a model of the form: $B = a \times D^b$, where tree B depends only on the trunk diameter, and where b is a parameter (power law exponent). Note that a mathematically equivalent formula is obtained by taking a logarithmic transform of both sides of this equation: $ln(B) = ln(a) + b \times ln(D)$. This algebraic log-transformation is the basis for the construction of many biomass models. Setting $Y = ln(B), X = ln(D)$, the above becomes a linear model:

$$Y = a' + b \times X \text{ , where } a' = ln(a)$$

This model has the desirable property that B goes to zero as D goes to zero (if b is a positive parameter). The AGB of a single tree must be associated with an error term:

$$Y = a' + b \times X + \varepsilon$$

The error term ε is often assumed to be a normal error with no bias, and is uniquely determined by its spread, called the standard deviation σ: a large standard deviation means that random error is large, and precision low. There is no *a priori* reason for assuming that different allometric models should be applied to trees of different size classes: piecewise models inevitably come with discontinuities at the domain boundaries. Examples of classic formulations for BAMs are reported in Table 2.2.

Table 2.2 Example formulations of biomass allometric models

Biomass allometric model	Reference
$\ln(B)=\ln(a_1)+b_1 \times \ln(D) + c_1 \times (\ln(D))^2 + d_1 \times \ln(\rho)$	Alvarez et al. (2012); Fayolle et al. (2013)
$\ln(B)=\ln(a_2)+b_2 \times \ln(D) + c_2 \times \ln(H) + d_2 \times \ln(\rho)$	Chave et al. (2005); Alvarez et al. (2012)
$\ln(B)=\ln(a_3)+b_3 \times \ln(\rho \times D^2 \times H)$	Chave et al. (2014)
$\ln(B)=\ln(a_4)+b_4 \times \ln(D) + d_4 \times \ln(\rho)$	Chave et al. (2005); Alvarez et al. (2012)
$\ln(B)=\ln(a_4)+b_4 \times \ln(D)$	Paul et al. (2016)

The linear model returns the log-transformed AGB of a tree. Back-transforming the model above for a given tree i yields the following result:

$$B_i = exp(Y_i) = exp(a' + b \times ln(X_i) + \varepsilon_i)$$

The complication is that, if ε_i has a zero mean (i.e., is unbiased), $exp(\varepsilon_i)$ is no longer unbiased and a correction factor is required. Clifford et al. (2013) explored the relative performance of nine different correction factors in terms of correcting for this bias, but a robust correction factor that is commonly applied is proposed; the Baskerville correction factor (Baskerville, 1972). This correction factor depends on the magnitude of random error, which is quantified by σ, and writing $a'' = a \times exp(\sigma^2/2)$, the biomass of tree i may be estimated using the following equation: $B_i = a'' \times (D_i)^b$.

For some applications, the biomass allometric model may be biased, that is, systematic under- or over-estimation (JCGM, 2008). This may be because the model is used outside of its normal range of calibration, or because the model has been calibrated using too few destructively sampled trees (Roxburgh et al., 2015). If an allometric model has been calibrated using a sample of small trees, using it outside its range of calibration, i.e. to estimate large trees, may lead to bias (e.g. Duncanson, Rourke, et al., 2015). If an allometric model has been calibrated using only a few destructively harvested trees, typically less than 20-50 trees, it is more likely to be biased than a model calibrated based on hundreds or thousands of destructively harvested trees (Roxburgh et al., 2015). Producing biomass estimates from a biased allometric model results in potentially large systematic errors in remote sensing products and carbon stock estimates.

Examples of quantification of measurement errors

The ordinary least squares regression permits the estimation of three model parameters: (a', b, σ). Any statistical software can accept an input dataset with measured values of *D* and *B*, take a logarithmic transform of these values, and construct a best-fit regression by a linear model. In the R statistical computing platform, the command for linear model construction is lm(); the output includes the model parameters (a', b) and the random error estimate σ (stored in the variable sigma). The standard deviation is the canonical metric to represent precision in a linear model.

Note that in larger trees, biomass estimation has a larger standard deviation. This is related to the multiplicative nature of allometry that is often used to represent the dependency between biomass and biometric variables. However, the relative error (standard deviation divided by the expected value, also called coefficient of variation) does not strongly depend on tree size. One analysis suggests that the relative error in biomass estimation of a single tree can be as large as 50% (Chave et al., 2014). However, detecting the magnitude of a size dependency in prediction of biomass is crucially dependent on the reference biomass data (Burt et al., 2020). However, because the errors are uncorrelated across trees, stand-level error declines rapidly as stand area increases (i.e., as the number of trees included in the sample increase), since biomass density is obtained by summing individual tree biomass over a given area; (Réjou-Méchain et al., 2014). Most carbon accounting programs do not demand that individual trees be estimated accurately, but that plot-level biomass density estimates be accurate.

To calibrate a BAM, an important practical question is, how many trees are needed to estimate the parameters to a good precision? A study of this question, based on empirical datasets (Roxburgh et al., 2015) suggests that the minimal number of trees is 25, but that good performance is achieved only if the destructive sample reaches or exceeds 100 trees. Under the assumptions of allometric model form, combining samples from studies that use different measurement methods into a single biomass allometric model can lead to large model uncertainty (Burt et al., 2020). Careful consideration should be paid to how many trees were destructively sampled to develop a model, how well the sampled trees used to develop the model match the diameter distribution in the population of trees to which it is applied, and how well the model can estimate the biomass of an independent set of sample trees. To avoid developing local allometric models, one can use already existing destructive biomass data. Much useful destructive biomass information is already available in the literature (Clough et al., 2016; Paul et al., 2016), (see also Chave et al., 2014; Radtke et al., 2015). The literature should be reviewed before new fieldwork is performed or further stratification is decided.

Recommendations for validation

Validation of BAMs requires the availability of separate biomass tree reference datasets to be used only at the validation stage. A common practice is to validate the model by partitioning biomass reference data into a training set and an independent validation set. **It is recommended** that the validation dataset has a similar tree-size structure as the calibration dataset, so users should ensure that validation includes both small and large trees. This is crucial because large trees hold a disproportionate share of biomass, and destructive harvest samples are often smallest for the largest tree size classes. Paul et al. (2018) showed that an independent sample size of N≤15 often (37–46% of the time) provides insufficient statistical power to avoid incorrectly accepting “validation” (type II errors). They recommended at least 50 trees be sampled for each species.

Validation of the model predictions should be based on clearly defined metrics. Paul et al. (2018) recommended that an equivalence test may be applied to determine if the minimum detectable negligible difference between the reference model and the new independent data is <25% (or

whichever threshold is deemed acceptable). If so, the new dataset can then be combined with existing data to refine a generalised model, that can then be applied with confidence. If not, then the resources expended need not be wasted as the sample size is sufficient to develop a new model suitable for application to the specific species sampled.

Therefore, it is essential to validate allometric models based on independent destructive harvest data. Unfortunately, many published biomass models do not follow any standard validation procedure and thus selecting the best performing allometric model for application in a given geographic domain is often left to speculation.

Information on how and why to select a particular BAM is seldom provided. Unless objective criteria are available, confirmation bias can result in cherry-picking of models. **It is therefore also recommended** that users report procedures and criteria used for selecting existing models, and model diagnosis and validation (Sileshi, 2014).

2.2.4 Selecting suitable biomass allometric models

Sources of error

Estimation of biomass from forest inventories is often based on locally derived BAM, because it is perceived that a local model is more accurate, reflecting the specifics of local forest type. Likewise, it is generally assumed that a species-specific BAM is a better model than a species-independent BAM. For instance, in the temperate zone, it has been customary to develop species-specific models, e.g. different models for oak, poplar, birch or beech. Generalized (regional to global) BAMs are often considered to be less accurate than local species-specific models. Is it preferable to use a few generic biomass estimation models, applicable to a wide range of forest types, or many biomass estimation models, each constructed to one plant functional type, one forest type or even a single species.

Examples of quantification of measurement errors

Tree biomass estimation based on trunk diameter, total tree height and wood density should in theory depend on a single estimated parameter, the form factor (see 2.2.1). The form factor is unlikely to vary greatly across forest types, because this factor is constrained mechanically, and it is related to the tapering of trunks, which has been classically measured in forestry (Radtke et al., 2015). This suggests that a single model could be used across tree species in the tropics. A test of this hypothesis has been conducted across tropical forests, and confirms the idea that a single biomass model can be used across the tropics (Chave et al., 2014). Recently, a test of this hypothesis has been conducted in four countries of tropical Africa (Fayolle et al., 2018). Results based on this large independent dataset (over 600 destructively harvested trees) confirm that the biomass model of Chave et al. (2014) shows a low bias. Although more comparable studies are desirable to assess model accuracy, designing destructive harvest experiments for validation of allometric models is difficult, time consuming and expensive, especially when large trees are

included in the sampling strategy (which should be the standard practice). Such experiments should be carefully designed and coordinated.

The web platform GlobAllomTree (http://www.globallometree.org/), launched by FAO in 2013, collates over 13,000 BAMs from all sorts of studies (local, regional, national), but often based on small destructive harvests (less than 25 trees, Henry et al., 2010), making it difficult to assess the accuracy of individual models. Thus, while the model root mean square error (RMSE) can be estimated from the modelled structure of the data (Henry et al., 2013), a validation of the prediction error is difficult for most of the reported models. If the set of species is limited, it is however possible to deduce generalized BAMs by combining many local and species-specific BAMs, as was done for European forest tree species (Forrester et al., 2017). However, how this approach is generalizable to species-rich tropical forests is currently unknown.

Recommendations for validation

It is recommended that allometric model quality be carefully assessed, and that models that do not specify the range of validity in the independent variables, size of calibration sample, geographic domain of training data, and the associated random error metric, should not be used. While one global allometric model may not be reasonable, given for example differences between tropical forests and temperate conifer forests, more research into the accuracy of generalized allometries is required. In particular, the extent to which broadleaf trees and conifers vary in their biomass allometry should be better studied. Once biomass models are carefully validated, their selection for specific sites should follow a rigorous procedure (Pérez-Cruzado et al., 2015).

For biomass estimation of tropical forest stands, **it is recommended** that the R package BIOMASS be used to predict AGB from inventory measurements (Réjou-Méchain et al., 2017). This package is called by several other packages dealing with different data formats. It possesses default BAM options for tropical forest AGB prediction that can be used to estimate biomass with traditionally measurable attributes (D, H, WD), and it also provides uncertainties based on knowledge of the main error components in AGB prediction for tree inventories. New BAMs could theoretically be provided by the user, such as, for instance, a BAM specific to Australian forests (Paul et al., 2016), for forests of North America (Chojnacky et al., 2014), or for forests of Russia (Schepaschenko et al., 2017). However, the BIOMASS package would need to be updated to include these. To implement the R package BIOMASS with other BAM's raw data used to produce the BAM should be provided as input.

2.2.5 Biomass expansion factors

Sources of error

Many methods in temperate forestry attempt to infer aboveground biomass from harvestable biomass or from harvestable volumes, the classic currency in forestry. The forestry sector has implemented allometric principles since the 18th century to establish stand volume tables, which

quantify merchantable volume of forests. Stand volume tables have been developed to a high level of detail, and published models depend on species, soil fertility, and climate (Jenkins et al., 2003; Schepaschenko et al., 2017; Zianis & Mencuccini, 2004), thus procedures for merchantable volume accounting are complex. Because much data is available on harvestable volumes from the forestry sector, valorising this information is desirable.

This volume tables method seeks to estimate the total AGB of a tree by measuring only its merchantable log volume, excluding branches and leaves (or needles). It is sensible to ask whether there should be a simple scaling relationship between branch biomass and stem biomass. From that knowledge one computes the ratio of AGB divided by stem biomass, also known as the Biomass Expansion Factor (BEF).

Examples of quantification of measurement errors

Measurements of trees across size classes suggest that the BEF is proportionally higher in saplings than in mature trees. This suggests that attempts to model the BEF should be size-dependent. This may not be a major issue in even-aged plantations, but could be problematic for mixed and/or mature forests.

Also, trees growing in open conditions (e.g. in open woodlands, or in agropastoral landscapes) tend to branch off more than trees in closed-canopy forests or in plantations. Trees that branch off earlier tend to have more of their biomass in branches proportional to the total biomass, hence a higher BEF.

For low-diversity forests, and where extensive prior knowledge is available on BEF, the BEF-based inference, which considers site index and tree density (e.g. Schepaschenko et al., 2018), has advantages compared to BAMs. For instance, the site index is related to wood density, which may vary across sites within the same tree species. Likewise, tree density relates to canopy features: sparse trees invest more biomass to branches, compared to dense forest trees.

Recommendations for validation

AGB appears generally to be more directly related to tree biometric variates and forest structure as captured by tree inventories than to merchantable log volume, implying that estimating tree biomass from a BAM and a tree inventory is less error-prone than the alternative based on biomass expansion factors. **It is therefore recommended** to predict AGB from forest inventories combined with BAMs, rather than based on BEF. However, in areas where extensive forestry expertise is available, and for even-aged plantations, and also if BEF are validated extensively across the range of forest types and conditions, the combination of merchantable log volume and BEF could be an alternative.

2.3 Terrestrial laser scanning & unmanned aerial vehicles

Kim Calders, Nicolas Barbier, Harm Bartholomeus, James Kellner, Felix Morsdorf, Crystal Schaaf, Atticus Stovall, Christian Thiel

2.3.1 Background

New developments in remote sensing, such as the use of Terrestrial Laser Scanning (TLS) techniques and the use of Unmanned Aerial Vehicle (UAV) platforms, can help us with the validation of AGB products at local to regional scales. The potential of TLS for forest monitoring was first demonstrated more than a decade ago, but has not yet reached its full potential, for the reasons outlined above. (Calders et al., 2020), Newnham et al. (2015) & Anderson et al. (2016) provide a full review of the development of TLS as a forest measurement tool. The utilization of UAVs for the acquisition of ultra-high resolution (<1 m) imagery and laser scanning data has heavily increased during the past five years.

TLS is a ground-based remote sensing system that can measure three-dimensional (3D) vegetation structure (i.e. the size and location of canopy elements) to centimetre or even millimetre accuracy and precision. The location of points in 3D space is generated by transmitting laser energy and analysing the reflected energy as a function of time. TLS measurements are essentially not limited in azimuth and zenith angles. This is an important advantage compared to most airborne and satellite-based instruments that generally have a restricted viewing angle range.

Point clouds from UAV platforms can be generated directly through laser scanning (LS) or indirectly through structure from motion (SfM) (Iglhaut et al., 2019) from image data (Figure 2.1). In the latter case the overlap between images enables stereoscopic image processing resulting in 3D point clouds representing the forest structure (Hernández-Clemente et al., 2014; Lisein et al., 2013; Puliti et al., 2015; Suomalainen et al., 2014; Zarco-Tejada et al., 2014). The motivation to use UAV-SfM 3D data is the rapid delineation of several forest parameters such as canopy height, canopy cover, tree location, number of trees, tree density, tree height, stem volume, and tree species for continuous forested areas of several hectares for a relatively low cost compared to the UAV lidar systems. However, the success of using UAV-SfM to monitor forest structure will depend on the openness of the forest canopy. UAV-SfM data are less precise over gaps and can only describe the top of the canopy, whereas UAV-LS has a greater canopy penetration depth and better precision (Roşca et al., 2018).

Traditionally, allometric models are used for estimating AGB at local levels (see Section 2.2.3). Typically, the volume or AGB of trees in the calibration data of allometric models is related to their DBH and height. The calibration data that underpin these allometric models is often skewed towards smaller trees, which can result in large uncertainties for bigger trees (Calders, Newnham, et al., 2015). Destructive harvesting is expensive and not always practical or desirable. In this chapter we describe how 3D data from TLS and UAV platforms can contribute to the generation of AGB reference datasets.

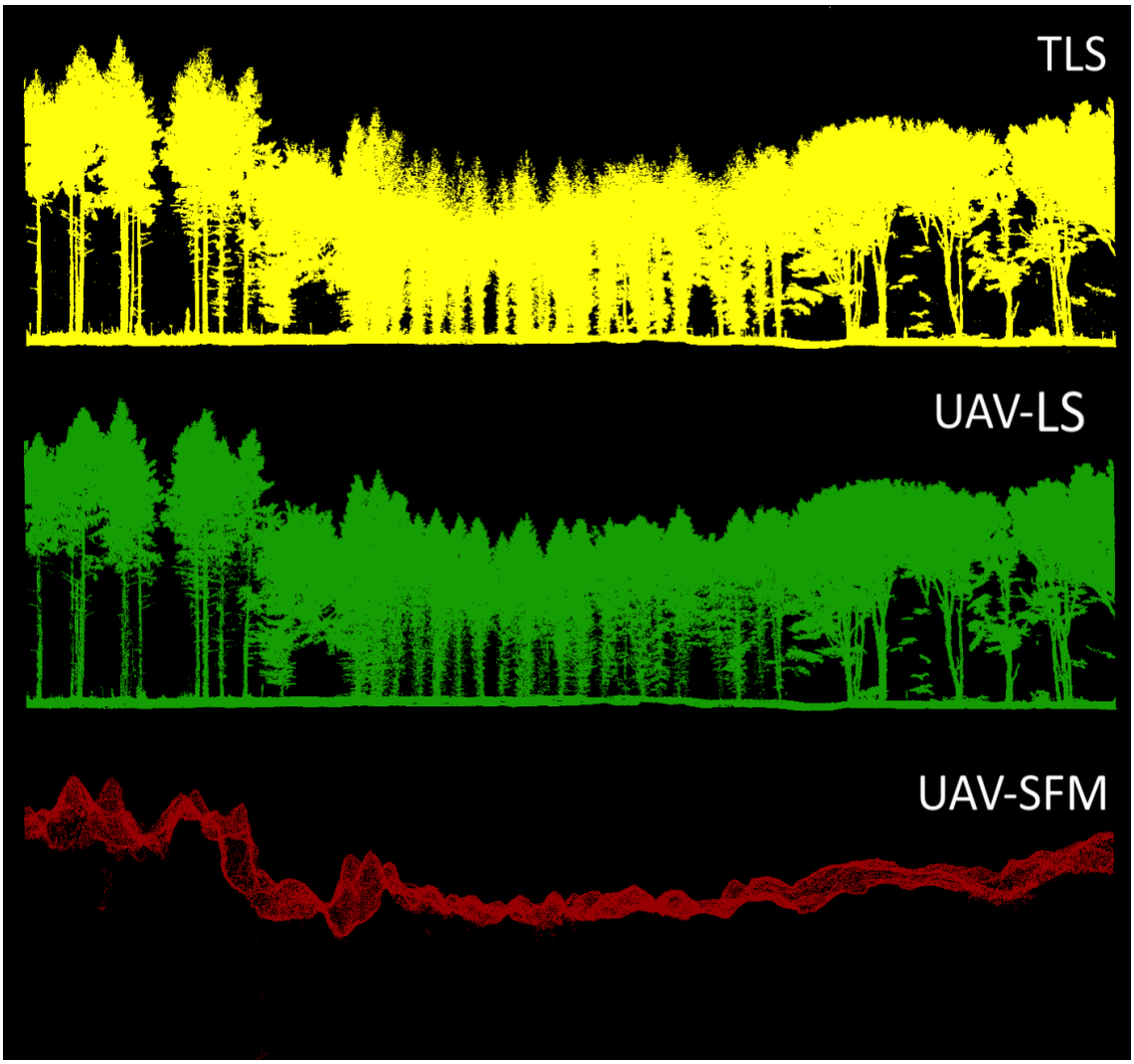

Figure 2.1 Slice of a forest with old Beech and Oak from the same perspective, acquired in the Netherlands with TLS (Riegl VZ-400)(top), UAV-LS (Riegl VUX-SYS) (middle) and UAV-SFM. The number of points in these slices are 69,547,639 (TLS); 5,396,682 (UAV-LS) and 96,229 (UAV-SFM)

2.3.2 Plot scanning protocols

2.3.2.1 Data collection

TLS sensors are usually tripod-mounted and record single scans from a fixed location. As such, scans are affected by occlusion, i.e., the near objects in the forest can obscure objects further from the scanner. The effects of occlusion can be significantly reduced by obtaining data from multiple scan locations. Multiple single scans made at different locations can be co-registered (to within mm accuracy depending on instrument and environment) using high reflectivity targets that act as tie-points between different scans (Wilkes et al., 2017). Recent work on reflector-less registration algorithms is promising (Kelbe et al., 2016), and a new range of commercial scanners (Leica BLK360 and RIEGL VZi-series) provide onboard registration without the need for targets. Currently, more testing (i.e., quantifying the effect of ecosystems, instrument characteristics and

sampling design on the performances of these algorithms) is required before it is recommended to use these in an operational context.

Since UAV-LS is a relatively new and rapidly developing technology, there is less expertise on the best acquisition settings. The point density depends on the frequency of the scanner in combination with the forward flying speed, number of flight lines and altitude, which can all be varied during flight planning. Therefore, acquired point densities vary from 50 pts/m^2 (Wallace et al., 2012), to 100-1500 pts/m^2 (Brede et al., 2017; Jaakkola et al., 2017), to >3000 pts/m^2 (Brede et al., 2017; Kellner et al., 2019; Morsdorf et al., 2017). Brede et al. (2017) showed that it is essential to include scan angles >30 degrees of nadir to get enough returns on the tree stem (see Figure 2.2). Further, using a system capturing multiple returns is essential to measure the below canopy structure and allow for individual tree modeling. This enhanced penetration depth due to multiple returns is a critical advantage of (some) UAV-LS systems over UAV-SfM in the context of generating AGB reference data from whole-tree volume reconstructions.

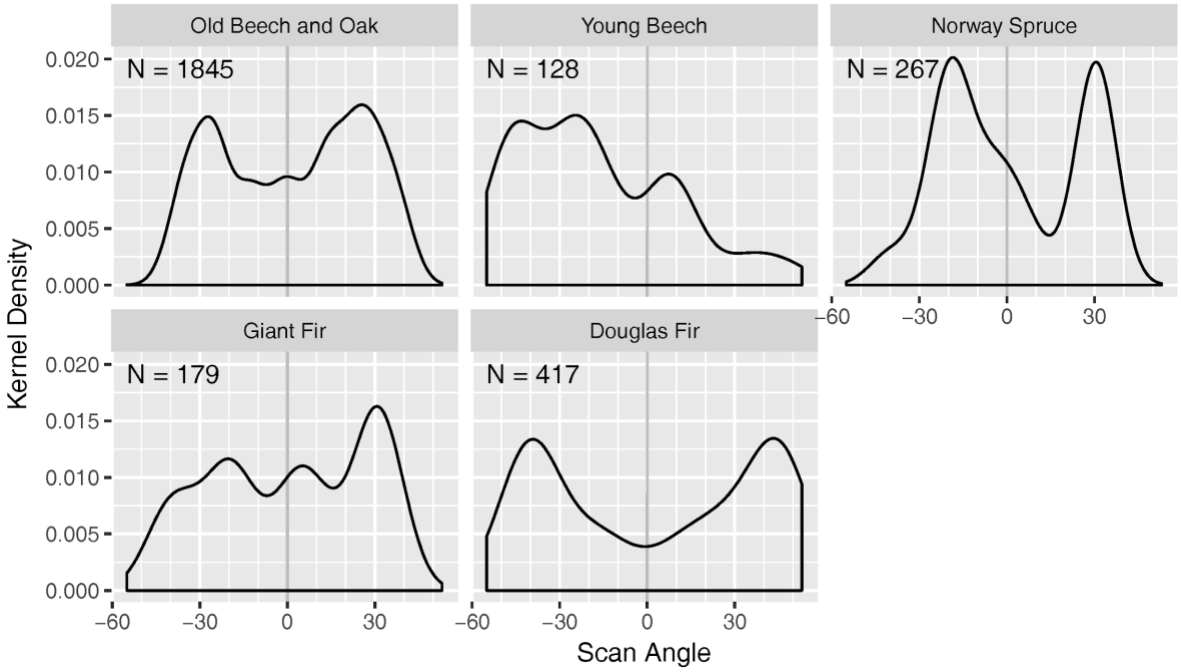

Figure 2.2 Scan angles distribution for points on the stem at DBH demonstrate that mainly scan angles of >30 degrees contribute to this part of the stem. (Brede et al., 2017)

2.3.2.2 Instrument requirements & limitations

A range of scientific and commercial scanners are currently available. Whereas airborne lidar systems have been used in forest measurements since the mid-eighties (R. Nelson et al., 1984), the first commercial terrestrial laser scanners came to the market in the late 90s with instruments such as the RIEGL LMS Z210 and CYRAX 2200. The first TLS instruments used a time-of-flight ranging principle, with phase-shift based ranging instruments following soon after. The commercial instruments were (and still are) generally developed for precision mapping and survey applications where structurally continuous surface targets dominate (e.g., urban areas and/or mineral and petrochemical exploration). This has implications for their use in forest

applications, where many laser hits are partial, and/or from softer targets (i.e. structurally fragmented or dispersed surfaces) with anisotropic reflecting surfaces such as leaves or needles and bark. Of the scientific (i.e. non-commercial) scanners, the Echidna Validation Instrument was one of the first laser scanners specifically designed to monitor vegetation (Strahler et al., 2008), more recently followed by the full waveform Dual Wavelength Echidna lidar (DWEL; Li et al., 2018) and the Salford Advanced Laser Canopy Analyzer (SALCA; Danson et al., 2014; Hancock et al., 2017). Table 2.3 provides a summary overview of commonly used TLS instruments in vegetation monitoring. Small robust instruments are increasingly being used in adverse environments, including the scientific Compact Biomass Lidar (CBL; Paynter et al., 2016) and the Leica BLK360 (Luck et al., 2020). Newnham et al. (2012) provide a detailed independent comparison between some commercial scanners and evaluate their performance for measuring vegetation structure.

Table 2.3 Examples of TLS instruments used to assess forest structure. Non-commercial instruments are shown in grey. Modified from (Calders et al., 2020).

Major Instrument Categories	**Short-range TOF + large beam divergence**	**Mid-range TOF + medium beam divergence**	**Long-range PS + small beam divergence**	**Long-range TOF + medium beam divergence + low noise**	**Mid-range Dual Wavelength + medium beam divergence**
Cost	$	$	$$	$$$	$$$
Ideal Forest conditions	+ Sparse/simple forests + Remote areas	+ Sparse/simple forests + Remote areas	+ Leaf-off or structurally simple forest stands + Remote areas	+ Best in tall/dense forests	+ Accessible, structurally simple forest stands
Optimal Forestry Applications	+ Rapid assessment + Robust + Cost-effective forest structural metrics	+ Rapid assessment + Cost-effective forest structural metrics	+ Finely resolving small branches	+ Finely resolving small branches + Potential for full waveform applications	+ Leaf-wood separation + Biochemical properties + Improved vertical foliage distribution + Potential for full waveform applications
Example Instrument	UMB CBL (SICK Lidar)	Leica BLK360	FARO Focus^{3D} X 330	RIEGL-VZ400i	SALCA
Ranging method	time-of-flight	time-of-flight	phase-shift	time-of-flight	time-of-flight
# returns	1st + 2nd	Single	Single	Multiple	Full waveform
Wavelength [nm]	905	830	1550	1550	1545.4 & 1063.4
Maximum Range [m]	40	0.6 - 60	0.6 - 330	1.5 – 250 (high speed) 0.5 – 800 (long range)	100 m
Samples/sec	11,000	360,000	122,000-976,000	42,000– 500,000	5,000
Beam Divergence [mrad]	15	0.4	0.19	0.35	0.56
Weight [kg]	3.9	1	5.2	9.7	17

Temperature range [deg C]	-30 to 50	5 to 40	5 to 40	0 to 40	5 to 30
References	Paynter et al. (2016; 2018)	Disney et al. (2019)	Liang et al. (2015); Pyörälä et al. (2018)	Bienert et al. (2018); Tian et al. (2019)	Danson et al. (2018); Schofield et al. (2016)

In recent years, the use of UAVs for the acquisition of (close range) remote sensing data has emerged (Kellner et al., 2019). Once the hardware is purchased, images can be recorded almost at any time and at low cost. The overlap between the images enables stereoscopic image processing (UAV-SfM), the delineation of high-density 3D point clouds and the generation of seamless image mosaics. Due to flight regulations and technical limitations at present, UAVs are used for the acquisition of local data only. Considering the maximum allowable distance of 500 m between pilot and UAV (legal requirement in many countries), an area of approximately 75 ha can be covered during one flight. A new development is the usage of UAVs as lidar platforms. One of the first off-the-shelf systems – the RiCOPTER – was released by RIEGL. As the drone can be operated at very low flight speed, great overlap between the tracks and variable flight altitude, the resulting sample point density is very high. Another interesting feature is the wide scanning angle of the small footprint lidar RIEGL VUX-1UAV sensor. Similar to the VUX-1UAV but more miniaturized survey-grade sensors are nowadays available such as the the RIEGL miniVUX series (miniVUX-1UAV, miniVUX-2UAV, miniVUX-3UAV, miniVUX-1DL). A number of alternative (lighter and cheaper) user-grade options are now available on the market, with increased operational versatility (e.g. Mdlidar1000 by Microdrones, DJI Matrice Series Livox with Mid-40 or Mid-100, Tron F9 VTOL fixed wing, Vapor 55 UAV helicopter, lidar hardware by YellowScan). It is important to mention that (currently) the decreased costs of the latter category of user-grade lidar sensors comes at the expense of poorer positioning, increased beam divergence (i.e. larger footprints), lower scanning frequency, and limited multi-return capabilities. All of these limitations can seriously hinder the use of the derived point cloud data for sampling the full 3D structure of trees. Vertical tree structures such as stems can (to some extent, depending on forest type) be sampled with a high density of pulses, which is of interest for forestry applications. Such systems allow users to obtain up-to-the-minute data which is of particular importance when solving specific local and regional issues in which a user defined spatial and temporal resolution needs to be met for the generation of reference datasets. Table 2.4 gives a comprehensive overview of typical UAV systems for monitoring forests.

Table 2.4 Selection of UAV Systems

Mission/Sensors	**DJI Phantom 4 Pro V2.0 Quadcopter (platform, gimbal, installed camera system)**	**Quantum systems Trinity F90+ VTOL fixed-wing (platform, RGB+ Micasense RegEdge)**	**DJI Matrice 600 (platform with integrated RIEGL miniVUX-1UAV lidar scanner)**	**RiCOPTER VUX-SYS (platform with integrated VUX-1UAV lidar scanner)**	**Quantum systems TRON VT LQuadrocopter fixed-wing hybrid (platform with integrated YellowScan "SURVEYOR" lidar scanner)**
Cost	$	$$	$$$	$$$$	$$$
Area coverage	++	+++	+	+	+++
Sensor type	Stereographic imaging system (discrete overlapping images)	Stereographic imaging system (discrete overlapping images)	Lidar (multiple return, echo intensity recording)	Lidar (multiple return, echo intensity recording)	Lidar (two returns)

Frequency	R, G, B (1'' CMOS)	R, G, B, Red edge, NIR (multiple sensors)	One laser (NIR), max. 100,000 shots/s	One laser (NIR), max. 500,000 shots/s	One laser (905 nm), max. 300,000 shots/s
References	(Puliti et al., 2019)	Quantum systems (2018)	Riegl (2018)	Riegl (2018) (Brede et al., 2019)	Quantum systems (2018)

2.3.2.3 Plot establishment

Sampling strategies for geometric modelling of AGB reference data should aim for minimal occlusion and uniform point density in order to provide consistent point cloud quality throughout the plot. This may require sampling over a larger area than the plot size, particularly for smaller plots (< 1 ha). When designing a sampling strategy for capturing data at plot level, it is important to consider that the resulting point cloud needs to:

1. Capture a large proportion of the target canopy to:
    a. Be spatially representative and account for occlusion; and
    b. Sample a wide range of view angles
2. Have a uniform point density across the scanning domain.
3. Be easily co-registered/aligned with the required degree of accuracy.

For terrestrial lidar data acquisition, **it is recommended** that the overall scan pattern forms a continuous "chain" where each scan location is linked to other (neighbouring) scan locations. A number of different configurations have been tested, dependent on stem and understory density. Different instruments and environments will require a custom-tailored plot setup. As a general rule, we advise to take into account the three recommendations from above to ensure good data quality. Wilkes et al. (2017) gives detailed recommendations for a high-end RIEGL VZ-instrument. For example, when understory vegetation is dense, a higher resolution sampling grid has been used, e.g. 10 m x 10 m, to ensure adequate sampling of the canopy through the understory, as well as occlusion of adjacent scan locations; whereas, if the understory is more open, a 20 m x 20 m sampling grid has been used. In general, the number of scan locations is an important consideration, particularly in tall or dense forest canopies. A sparse sample grid can save time and resources, particularly for large area field campaigns and monitoring programs. However, this may cause issues with co-registration and result in heterogeneity in point density across the plot. Furthermore, decreasing the sample density significantly decreases the fidelity of the branching structure towards the top of the canopy, potentially leading to more uncertainty in derived outputs sensitive to these canopy components (Figure 2.3). Recent efforts have been focused on analyzing data acquisition while in the field, to reduce the impact of occlusion by adapting the scanner placement dynamically (Paynter, Genest, Saenz, et al., 2018). In many National Forest Inventories, circular plots are a standard. In France, where TLS is increasingly integrated in the NFI, with ~2% of the plots scanned, about 10 scans are performed within the 15 m radius plots, following two concentric circles (at around 6 and 10 m from center), with the reference targets scattered within the inner circle (C. Vega, personal communication, n.d.).

To achieve good quality UAV-LS data, a wider range of view angles is more important than a higher point density (Brede et al., 2017). Building on the expertise with ALS and TLS systems, minimizing occlusion for UAV-LS is best done by using more flight strips (best in a crossed pattern)

to provide more observational angles into the canopy (Morsdorf et al., 2017). Furthermore, to reduce the occlusions in the lower parts of the canopy it is important to use sensors allowing for multiple returns and small footprints (Puliti et al., 2020). The latter is ensured by using sensors with narrow beam divergence (< 0.3 mrad), or by flying at very low altitudes with sensors with wider beam divergence (eg. Velodyne VLP-16) at the cost of a substantial loss in flight efficiency. Long-endurance platforms, such as the upcoming Aeroscout Scout B-330 (3 hrs flight time) allow for the sampling of 100 hectares. However, regulations are limited in many countries and often do not allow UAVs to fly too far from the launch spot. Another limitation is the need for a big enough gap in the canopy for take-off and landing, which is often not available in dense forests.

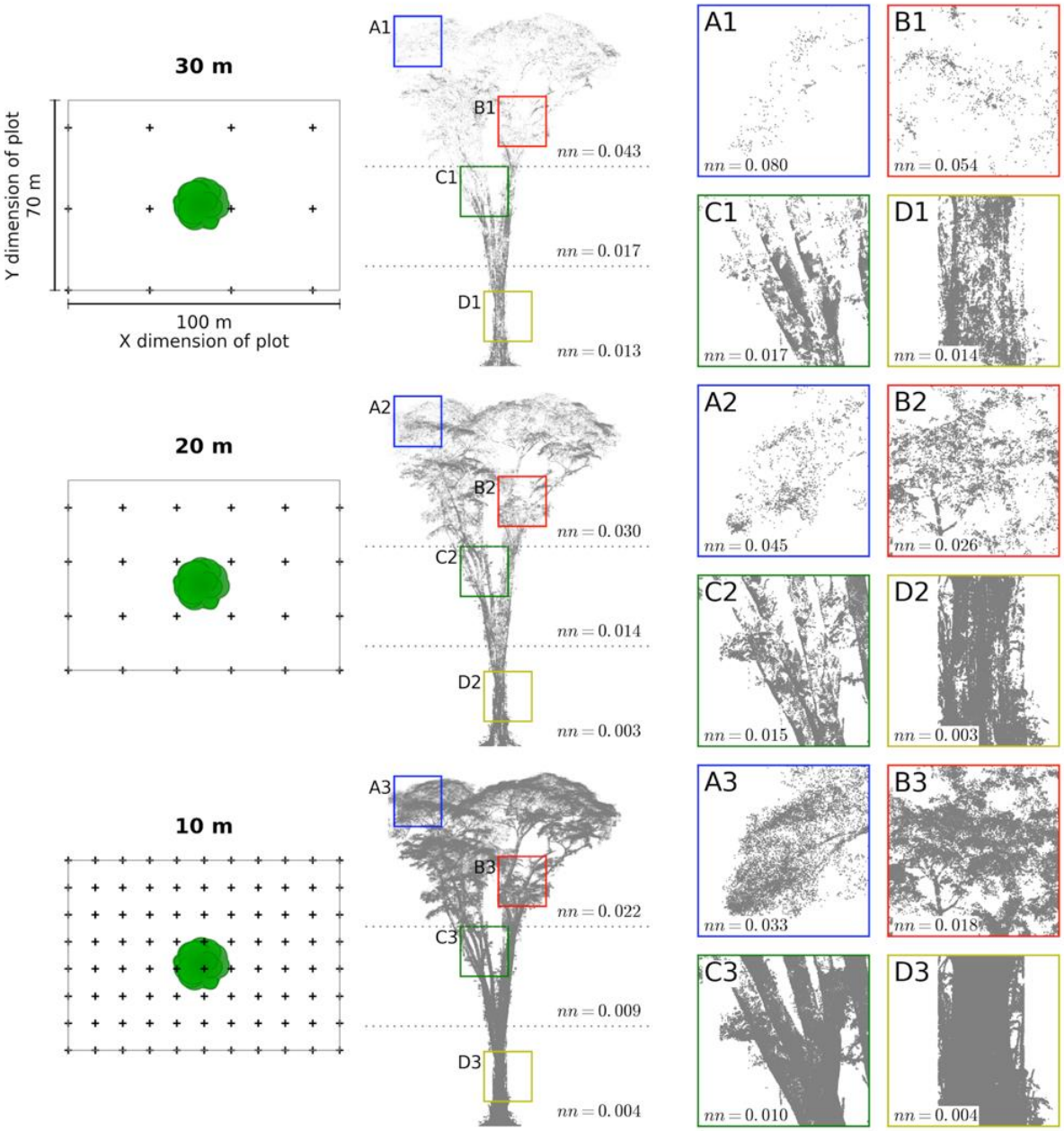

Figure 2.3 A comparison of mean nearest neighbour distance for points that comprise an individual tree derived using 10, 20 and 30 m grid sampling densities. (Left column) The location

of sample points and location and extent of the target tree. (Middle column) Point cloud representations of the sampled tree including mean nearest neighbour distance for different canopy heights. (Right column) Subset 3 m x 3 m x 3 m voxels for different areas of the tree, locations are identified in the middle column. The tree has been extracted from the Ankasa AfriSCAT plot in Ghana [Figure from Wilkes et al. (2017)].

2.3.3 Data processing protocols

2.3.3.1 Point cloud to volume estimates

The combination of multiple individual terrestrial lidar scans using registration targets significantly reduces occlusion. After the registration of individual scans to co-registered point clouds, estimating AGB requires:

(i) segmentation of the full point cloud into single trees (+ leaf vs. wood segmentation);
(ii) geometric modelling to estimate volume; and
(iii) conversion of volume to AGB using wood density.

Figure 2.4 Example of a co-registered and segmented TLS point cloud from Wytham Woods, UK. Single trees extracted for a 20 x 100 m transect. Data captured with a RIEGL VZ-400 instrument and tree extraction was done with the *treeseg* algorithm (Burt et al., 2019). (a) Top view; (b) Side view

Burt et al. (2019) describes an open-source software (*treeseg*) for the near-automatic extraction of tree-level point clouds from larger area point clouds (see Figure 2.4). They demonstrated an automated segmentation success rate of 70% in dense tropical rainforest and 96% in a more open forest. Other methods such as Raumonen et al. (2013), Trochta et al. (2017) or Zhong et al. (2017) offer alternatives. Generally, more manual intervention is required in complex ecosystems where crowns can interact with each other (Calders et al., 2020).

Several reconstruction algorithms have been developed to produce full 3D reconstructions of tree structure from single-tree point clouds. These so-called Quantitative Structural Models (QSMs, Figure 2.5) provide topologically-connected estimates of tree trunks and branches down to fine scale (cm), allowing for straightforward calculation of volume. Dassot et al. (2012) used simple geometric fitting to model the woody structure of individual trees, whereas Hosoi et al. (2013) used a voxel-based approach. Both these approaches require a substantial amount of manual input. Côté et al. (2009, 2011) developed an automated algorithm that models a free form circular cross-section woody model based on intensity filtering of the lidar returns. Fine branching and leaves are added to the woody structure based on the low intensity returns. All these methods suffer from the difficulty of assessing the accuracy of the resulting reconstructed QSMs. Disney et al. (2012) developed a 3D modelling approach to overcome this, by using 3D simulated TLS data from tree models whose structure and volume is known a priori. This approach was applied by Raumonen et al. (2013), who describe a reconstruction method based on local patch fitting to produce cylinder-based QSMs. Hackenberg et al. (2015) developed a slightly different approach to QSM reconstruction, but volume estimates agree well with Raumonen et al. (2013). Burt et al. (2021), Momo Takoudjou et al. (2018) and Gonzalez de Tanago et al. (2018) showed that the obtained QSMs could be used to obtain unbiased estimates of volume and biomass, even for very large tropical trees, allowing for the non-destructive calibration of allometric models. Care still needs to be taken to obtain unbiased estimates (Momo Takoudjou et al., 2018), as cumulating automated algorithms for tree segmentation, leaf v. wood segmentation and finally reconstruction can sometimes result in inconsistencies between reconstructed trees and the original point cloud. It is thus still recommended to maintain human supervision and quality control in the processing of dense forest stands.

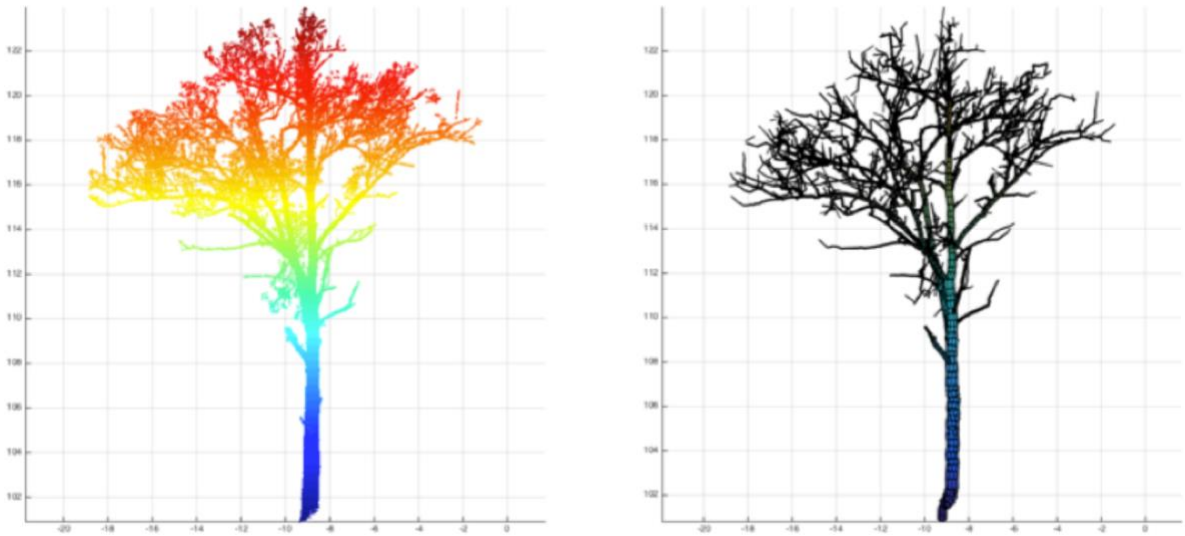

Figure 2.5 Point cloud (left) and corresponding QSM (right) using the method of Raumonen et al. (2013)**.**

For the operational delineation of UAV-derived point clouds several commercial and open-source software packages are available. This kind of software commonly comprises bundle adjustment

and SfM algorithms (Iglhaut et al., 2019), although, for commercial software, the code is not published. Based on the point clouds, Digital Surface Models (DSMs), and after vegetation filtering, Digital Terrain Models (DTMs) can be delineated by rasterizing the point clouds. Using UAV-based DSMs and DTMs, e.g. the canopy height can be precisely measured (Puliti et al., 2020; Thiel & Schmullius, 2017). For the delineation of elevation data products based on stereophotogrammetry and related approaches the same points or image objects must be identified and precisely detected in all overlapping images. Ground Control Points (GCPs) can be collected for ensuring a correct georeferencing of the resulting 3D products. In general, the number and positional accuracy of detectable points per unit area increases with increasing spatial resolution. A high point density permits small raster cells of the final elevation model. In areas with a very low point density, interpolation might be required as an intermediate step. Morsdorf et. al (2017) found that UAV-LS data has sufficient quality to model tree stems and large branches, but does not resolve fine branches. Using QSM algorithms to model individual trees from UAV-LS data and quantifying their uncertainty needs further testing, since these algorithms are developed for TLS data. While QSM approaches may be limited in UAV-LS data due to occlusions in the lower parts of the canopy, (Puliti et al., 2020) demonstrated the possibility to predict plot-, and stand-level timber volume using a sample of trees where DBH could be measured from the UAV-LS data.

2.3.3.2 Repeatability

One of the challenges of interoperability is the commercial confidentiality of instrument internal performance properties, making interoperability difficult. Calders et.al (2017) compared three TLS instruments from the same make and model (RIEGL-VZ400) and found that range accuracy between instruments is comparable, but that the radiometric calibration is instrument specific. Calders, Schenkels et al. (2015) reported a relative standard deviation in effective wood area index from repeated leaf-off scans of 0.72% for this TLS instrument (including removing and setting up the tripod and same instrument again over multiple days).

Both laser-based and stereogrammetry-based UAV approaches aim at a digital 3D representation of 3D real-world objects. Both approaches collect a finite number of discrete samples. In other words, the level of detail representing these real-world objects is limited. Moreover, repeated sampling will not result in the same samples, even for unchanged objects (e.g. vegetation has not moved etc.). Several factors are responsible for this shortcoming. The most obvious one is related to the varying acquisition geometry caused by varying acquisition positions of the UAV during the different campaigns. Therefore, the sample-based models of the real-world objects will have low reproducibility, deviating among measurements depending on the sample position, with the magnitude of this random error depending on the sampling rate.

However, depending on the research question or application, this drawback does not necessarily harm the usefulness or applicability of UAV-SfM or UAV-LS based point data. As found in previous studies (Brede et al., 2017; Morsdorf et al., 2017; Puliti et al., 2015; Thiel & Schmullius, 2017), a point density of several hundred to thousand points per m^2 can be achieved with UAV-SfM and UAV-LS techniques. In terms of the delineation of a detailed canopy height model for automatic tree detection the sampling density is more than sufficient, as all required elements (e.g. tree

tops, gaps between crowns) are captured with great detail. As in various (repeated) acquisitions the same level of detail is preserved, equal results can be expected.

Repeatability becomes a critical issue if the lidar point density is low. This can be true for airborne lidar systems which are flown with a typical nominal point density of below approximately 4 points per m² (Leitold et al., 2015). These systems may miss important details such as tree tops, or ground elevations under dense canopies, resulting in an underestimation of single tree or canopy height depending on sensor specification (Disney et al., 2010). Discrepancies in multi date airborne acquisitions can also lead to difficulty in interpreting change (e.g. Duncanson & Dubayah, 2018). Accordingly, the data acquired at two different UAV-LS campaigns might differ in a way that adds uncertainty to change estimation. This can include the detection of single (small) tree removal or the measurement of individual tree growth. For instance, Thiel et al. (2017) showed that a point density of 4 points per m² is not sufficient to automatically detect small trees, even in a simply structured forest. By using UAV-SfM data with a much higher point density the detection rate was obviously increased.

2.3.3.3 Stem map errors (linking to census measurements)

Linking existing census data to stemmaps derived from TLS or UAV data is important for determining the species of trees and using species-specific wood density for the conversion of volume to AGB. However, the accuracy of census coordinates is generally not as good as direct 3D measurements. Calders et al. (2018) manually compared census and TLS derived stem maps, achieving a link-success of 86% in a deciduous forest in the UK. Linking large trees is generally easier than smaller trees or multi-stem trees.

2.3.3.4 Characterization of errors and uncertainties

Methods have been developed to estimate more traditional forest monitoring properties such as DBH and tree height directly from the 3D point cloud. DBH and tree height are important structural measures (Section 2.1) that are historically used to predict AGB based on empirical allometric relationships (Chave et al., 2004, 2014). Numerous studies have estimated DBH using least square fitting from TLS point clouds. Tansey et al. (2009) reported a root mean squared error (RMSE) of 0.019 to 0.037 m (DBH range 16-40 cm) and analysis in Calders, Newnham et al. (2015) showed a RMSE of 0.0239 m (DBH range 11–62 cm). Stovall et al. (2017) estimated DBH using a convex hull fitting approach and reported a RMSE of 4.3% (DBH range 10.6-33.6 cm). Tree height estimates in the earlier days of terrestrial lidar reported large errors (RMSE 1.4 to 4.4 m) (van Leeuwen & Nieuwenhuis, 2010). Calders, Newnham et al. (2015) have shown that tree height estimates from TLS data captured with the RIEGL VZ-400 agreed closely (RMSE 0.55 m) with destructive harvest measurements, and showed closer agreement and less bias than traditional height measurements (RMSE 1.28 m).

The direct assessment of volumes through QSM reconstruction of tree point clouds from TLS allows for the calculation of AGB using wood density values. Hackenberg et al. (2015) developed a slightly different approach to QSM reconstruction compared to Raumonen et al. (2013), but found similar results when comparing TLS derived AGB estimates through cylinder fitting against destructively harvested measures, with prediction errors ranging from 2.75% to 7.30%. The AGB

of the single trees used in both studies was relatively small (Hackenberg et al. (2015) < 0.7 t, Calders, Newnham et al. (2015) < 3.4 t) compared to tropical trees, which can exceed an AGB of 75 t (Chave et al., 2014). Applying these methods in more challenging and complex tropical biomes with larger trees showed similar results (Gonzalez de Tanago et al., 2018; Momo Takoudjou et al., 2018). *TreeQSM* (Raumonen et al. (2013) computes the modelling error for QSMs by running multiple iterations of the same input QSM parameters. This generally results in small variations in the QSM due to different starting seeds for modelling. Momo Takoudjou et al. (2018) used *SimpleTree* (Hackenberg et al. (2015) and highlighted the importance of manual supervision and correction of the QSMs or the modelling input parameters. Smaller branches are generally not resolved in sufficient detail and will be modelled with higher uncertainty. This is especially true for smaller branches higher in the canopy due to the limitations of TLS instruments. Irregular shaped stems (e.g. due to the presence of buttresses) require mesh-fitting instead of cylinder fitting for these sections.

The conversion of tree point clouds to QSMs (see Section 2.3.3.1) requires leaf-off point clouds. Most QSM modelling approaches fit cylinders and presence of leaves would significantly overestimate volume and AGB (Calders, Newnham, et al., 2015). Typical approaches to leaf/wood separation include the inclusion of instrument-specific or intensity-driven methods, or using the geometric properties of the leaf-on point cloud (Boni Vicari et al., 2019; Krishna Moorthy et al., 2019; D. Wang et al., 2020), or a combination of these (Disney et al., 2018).

The use of wood density values that are representative of the whole tree is critical in biomass estimation from tree volume data. Global wood density databases often comprise wood density data sampled from the lower parts of trees (trunk or trunk base), however, wood density is known to vary vertically and radially (Chave et al., 2009). Using destructive data on 15 species in Cameroon, Sagang et al. (2018) showed that a correction was feasible through simple linear models, including a simple descriptor capturing the stem vs. crown relative proportions. A larger scale study on 52 species in Central Africa confirms this result (Momo et al., 2020)(Gonzalez de Tanago et al., 2018; Momo Takoudjou et al., 2018)(Momo et al., 2020). Another aspect lidar data cannot currently account for is the presence of hollow parts in the trees (Nogueira et al., 2006). In some locations and for some species, this phenomenon can lead to systematic error. Technical solutions, e.g. using the propagation or sonic or radar waves in the trunk are under study.

Trees moving due to wind have a negative effect on the data quality of all 3D sampling methods. In the most extreme case, the tree crowns are in different positions during different scans, e.g. for different lidar campaigns. Additionally, for UAV-LS campaigns commonly using slow vehicles and for TLS campaigns moving trees can cause a spread of points related to the differing tree crown/upper stem positions during the scan. Ultimately, wind might have a different impact for different campaigns. UAV-SfM based data is particularly impacted by the movement of objects during data acquisition, as this method relies on the recognition of the same objects in various images, assuming that these objects do not move. In general, the data quality required for good repeatability of all methods increases with decreasing wind speed. Ideally, measurement campaigns should be operated at calm conditions. As TLS and UAV-LS technologies evolve, scanning speed increases significantly (possibly to reach instantaneous scans with single photon

lidar technology), but the time taken to move the scanner around to multiply positions will remain commensurate with tree swaying, as is the case in SfM approaches.

2.3.4 Recommendations for validation

Calibration and validation of remote sensing data using 3D data from TLS and UAV platforms can be done through two different pathways: (1) a direct assessment of the volume of all the trees in a plot (and conversion to AGB); or (2) the development of local (nondestructive) allometric models based on a representative sample of trees.

The current state-of-the-art has demonstrated that TLS QSMs can be used as a replacement for destructive harvesting (Calders, Newnham, et al., 2015; Gonzalez de Tanago et al., 2018; Momo Takoudjou et al., 2018; Stovall & Shugart, 2018). TLS sampling methods can be leveraged for accurate AGB land product calibration and validation with full plot-level geometric reconstruction and targeted tree-level acquisitions for allometric model development. Plot-level reconstruction requires all trees in a plot to be modeled and is most feasible in less occluded forests, e.g. deciduous forests in leaf-off conditions. Tropical forests pose a major challenge due to high occlusion and large plot-size requirements, but TLS can clearly benefit the assessment and validation of satellite biomass products in high AGB density forests. Targeted tree-level geometric modeling can improve allometric models by sampling a full range of tree sizes and increasing the sample size for allometric model training. A greater immediate benefit from TLS is thus likely in the improvement of local and regional allometric models, as these can be built and amended from numerous tree- or plot-level TLS acquisitions and applied to existing forest plots (Lau et al., 2019).

Stovall & Shugart (2018) assessed both approaches for a temperate forest and outlined a framework for reducing uncertainty in biomass product calibration and validation. The key finding was clear TLS-driven improvement in AGB uncertainty by reducing allometric error and, compared with diameter-based allometric AGB equations, providing more direct estimates of standing AGB. Stovall and Shugart (2018) tested the TLS approaches against traditional allometric methods and showed a reduction in plot-level uncertainty from 81.1 to 62.4 Mg ha^{-1} (29.5% to 20.4%) RMSE for lidar-based mapping products. An adaptive approach, using both plot-level reconstructions and TLS allometry, **is recommended** to maximize the potential benefits from TLS.

Initial work on estimating volume from dense UAV-LS data is promising, and further development of UAV-LS methods can bridge the gap between airborne and terrestrial lidar (Puliti et al. 2020). Both pathways require high-quality data that represent the full structure of the tree. This might not be possible with UAV-SfM, but these data are still useful for upscaling at a relatively small cost.

2.3.5 Summary and current knowledge gaps

Studies estimating AGB from TLS are encouraging, as TLS does not only support relatively near-direct estimates of AGB (via volume, and wood density), but can also improve current allometric models. Destructively harvested reference measures of volume and AGB are expensive and time-

consuming and hence tend to be very limited both in terms of the number of trees and their size (very few large trees are sampled in this way). TLS data have the potential to provide similar volume information at a fraction of the cost, that are less biased in terms of tree size distributions. This is likely to reduce the uncertainty of the resulting reference AGB estimates compared with the use of existing generalized allometric models that underpin all current field-based and satellite-derived AGB estimates. Current geometric modeling methods provide clear, detailed and accurate characterizations of structure at an individual tree level, but more development is required to automate algorithms to provide efficient plot level AGB estimates. In dense tropical forests, individual tree and leaf vs. wood segmentation also remain challenging to automate.

One of the current limitations is that only the parts of the trees that are surveyed in the data will be modelled. Data gaps in the tree point clouds can be due to occlusion of other plant components, sensor characteristics, but also due to platform choice. Future research is needed to determine if UAV-derived AGB estimates are as accurate as TLS-based estimates. Of particular interest is the development of methods relying on sampling principles to infer population parameters using a sample of non-occluded trees (Puliti et al., 2020). Fusion of UAV and terrestrial lidar data can potentially be an important tool for the 3D mapping of forests at plot to landscape levels, which will be key for the reduction of large-scale measurement and scaling uncertainties in AGB and carbon estimates.

Chapter 3: Linking Reference Plots to Satellite Data

3.1 Spatio-Temporal Mismatches During Calibration/Validation Procedures

Maxime Réjou-Méchain, Michael A. Wulder, Valerio Avitabile, Laura Duncanson, Erik Næsset, Nicolas Barbier

3.1.1 Introduction

Remote sensing (RS) studies aiming at mapping forest biomass typically rely on ground-based estimates of biomass for calibration or validation. Whatever the biophysical variable of interest, the basic assumption is that the area measured on the ground at each sample point can be spatially and temporally matched with the same area as viewed via remote sensing. The magnitude of this approximation, and of the resulting spatial or temporal mismatch between ground and RS data, can be impacted by several factors, offering a source of uncertainty and/or systematic error in the final biomass map (Frazer et al., 2011; Gobakken & Næsset, 2008; Mascaro et al., 2011; Mitchard et al., 2014; Réjou-Méchain et al., 2014). A better understanding of the sources of ground and remotely sensed data mismatch provides an opportunity for improvement for RS-based biomass approaches.

Four main sources of discrepancy have been reported in the literature: i) mismatch between field plot and pixel sizes and shapes (e.g. Réjou-Méchain et al., 2014); (ii) spatial coregistration errors (e.g. Gobakken & Næsset, 2009); iii) lack of agreement between the forest components measured from the field (trunks) and from airborne/satellite signals (mostly crowns) (e.g. Mascaro et al., 2011) and iv) temporal difference between RS and field measurements (e.g. Avitabile & Camia, 2018). In this chapter, we review and describe these sources of mismatch and report new results from simulations to offer insight on how to understand or mitigate known issues. We finally discuss the implications of these errors for future satellite-based missions and provide practical recommendations for future work. For the purposes of forest biomass mapping in this chapter, we refer to coarse resolution as ≥500 m, medium resolution as 30-500 m, and high resolution as ≤30 m. It may be worth noting that the nominal spatial resolution of a sensor does not fully characterize image grain. For example, two sensors with 30 m spatial resolution may not convey the same level of information (e.g. one could be out of focus). The quality of an optical system (lens) should be more fully characterized via its optical transfer function.

3.1.2 Mismatch between field plot and pixel sizes

An obvious, though common, mismatch source is the use of field plots smaller than RS pixels. This represents a typical problem for large area biomass mapping using coarse resolution RS data

sources (pixel sizes ≥ 500 m) while field plots size often ranges from 0.04 to 1 ha, hence representing at best a 4% spatial sample of the larger pixels (e.g. Baccini et al., 2012; Saatchi, Harris, et al., 2011). Using large forest plots (8-50 ha in area), Réjou-Méchain et al. (2014) quantified the error due to such mismatches using plots of different sizes nested in a 4 ha RS footprint size (i.e., a 200 by 200 m pixel). They showed that this source of sampling mismatch can lead to errors of more than 30% on average, when using combinations of field plot and footprint sizes classically reported in the RS literature. Note that this error estimate was likely underestimated because the large plots on which the analyses were performed contained contiguous mature forest while large pixels often included a mosaic of forest and non-forest areas. Importantly, even at medium (≤ 50 m) resolution and within a forest context, this sampling error may still be large when smaller field plots are used due to a very high variability of forest biomass at a fine spatial grain (Réjou-Méchain et al., 2014).

Research using coarse to medium resolution RS products to extrapolate biomass measurements thus have to minimize this mismatch to reduce subsequent statistical errors that may propagate as a systematic error in the final map. **It is therefore recommended** to use a large field sampling area in general and, when possible, the calibration and validation of RS products use reference biomass values measured at the same spatial grain as that of the remotely sensed data source. When not possible, e.g. with coarse resolution products, several small plots may better capture the mean AGBD of a large pixel than a single large plot for a same total sampling area, even if this implies much more intensive field work, due to spatial correlation in AGBD distribution (Réjou-Méchain et al. 2014). More generally, two strategies may be adopted. First, a multi-step calibration approach can be implemented where intermediate high-resolution products, such as lidar, are calibrated from field plots and then used as reference to calibrate coarser-resolution remote sensing data (Asner et al., 2013; Baccini & Asner, 2013; Réjou-Méchain et al., 2019; Xu et al., 2017). However, much attention should be paid to appropriately propagate errors associated with every step up to the final product (Saarela et al., 2020). More details, and our recommendations for using airborne lidar biomass maps for satellite product validation, are presented in Section 3.2. The second strategy involves filtering plots that are not representative of larger pixels. This may be done through a screening procedure using higher resolution images and may be implemented to remove plots that are obviously not representative of the pixels. Such screening may tend to primarily filter low biomass plots located in open or fragmented forests, which are less likely to be representative of the larger, heterogeneous pixels. In such cases, to avoid creating a bias in the reference dataset by selecting only high-biomass plots, it is appropriate to further refine it to obtain a subset that also maintains the frequency distribution (i.e., histogram) of the original dataset in terms of number of plots per biomass class (Avitabile & Camia, 2018). Such an approach can be done visually using Google Earth imagery (Avitabile et al., 2016) or automatically using high resolution imagery (Mermoz et al., 2015) or, more likely, optical satellite-derived land cover. More details of this second approach are presented in Chapter 6.

3.1.3 Spatial co-registration errors

BOX 3.1
Satellite navigation basics
Global navigation satellite systems (GNSS) provide timing and positioning information based on the triangulation of signals sent by visible satellites of all currently available constellations (GPS, Galileo, BeiDou, GLONASS). The signals consist of radio waves at different frequencies in the L band (wavelength of 15-30 cm) and on the basis of the wave traveling time, receivers on the ground can compute a pseudo-range (comprising error due to clock uncertainty between the receiver and emitter). With signals from at least four non-aligned satellites, receivers can solve for XYZ position and clock uncertainty. Low grade receivers only use information contained in the code, while higher-end (survey grade) devices are also able to interpret phase shifts in the carrier waves to attain higher precision levels. Differential GNSS (DGNSS) uses a fixed ground station of known position (or a network thereof) to account for errors in pseudo-ranges obtained in a given geographic region at a given time, notably due to ionospheric and tropospheric perturbations of the radio-signal. These corrections can be applied to the pseudo range derived from the code, or from the phase. If the phase ambiguity can be solved for both the fixed and rover systems to within one wavelength, the solution is said to be fixed. Solutions solved +/- one wavelength are termed 'float'. These corrections can be used directly (real-time kinematic), or a posteriori (post-processed kinematic), to correct the positions acquired by a mobile DGNSS device (rover). Importantly, to achieve precise absolute positioning of the rover, the actual position of the fixed receiver (base) itself must be known accurately. Increasingly, for tropical regions where base station networks are sparse, alternatives exist, such as Trimble RTX subscriptions, or relatively cheap and easy-to-use devices (e.g. Reach RS2).

Spatial co-registration is required to assign a remotely sensed signal to a ground-based observation or prediction. Hence, Global Navigation Satellite System (GNSS, Box 3.1) instruments are generally used to geolocate ground observations (e.g., GPS, GLONASS, Galileo). GNSS accuracy is known to strongly vary depending primarily on the grade of the receiver, the sampling strategy, the number and positions of the satellites viewed, and on atmospheric, topographic and vegetation conditions (C. E. Johnson & Barton, 2004). For instance, Genova and Barton (2003) deployed two identical GNSS receivers side by side and registered the signal through time. They showed that some errors of up to 5 m were not synchronous, hence due only to the technical limitations of the instruments. GNSS technology was initially developed to be used in open areas, where acquisition conditions are much better than under forest canopies. However, the propagation of L-band waves is impacted by the water content in the troposphere and in plant leaves. Thus, several studies have shown that forest canopy strongly disturbs the GNSS signal (foliage can diminish the GNSS data logging efficiency by 47%; (Sigrist et al., 1999)), leading to an exponential increase of GNSS errors with canopy closure. Indeed, GNSS antennas are omnidirectional in order to capture satellite signals, thus they also receive "polluted" signals that have been reflected by e.g., the ground, the leaves or the trunks. This phenomenon, known as multipathing, is particularly important in dense multi-layered ecosystems such as forests, resulting in significant positioning errors. Using two precise GNSS receivers simultaneously located in an open-field and under a forest canopy in the same area, Johnson & Barton (2004) showed that while GNSS errors were less than ± 5 m for ca. 90% of the measurements in the open-field, the GNSS errors were larger than ± 5 m for ca. half of the measurements, and larger than ± 10 m for ca. 20% of the measurements under a forest canopy, with errors in some cases rising to more than 200 m under unfavourable satellite conditions. Topographical position is also expected to significantly influence GNSS accuracy with previous studies showing GNSS accuracy was higher in ridge locations than in valleys (e.g. Deckert & Bolstad, 1996). As a consequence,

previous work has indicated that accounting for GNSS-based plot location error could improve the root mean square error (RMSE) of AGBD in the order of 5-18% (Jung et al., 2013).

To determine whether technology has advanced since those studies, we repeated an experiment in a dense tropical forest in Gabon with a newer high-grade instrument, likely to have better performance than most devices currently used in forest inventories (see full methodology in **Appendix B**). Our results show an exponential increase in position accuracy under dense vegetation with the number of averaged GNSS points used (Figure 3.1). The accuracy associated with a single point is 5 m on average but can reach more than 70 m (confidence interval based on the 97.5th percentile), confirming that individual positioning measurements under close canopy cover have limited accuracy and may, in some cases, lead to extreme errors. Accuracy increases markedly when estimating the plot position using measurements from multiple locations, as the upper confidence interval goes under 5 m with 20 points and below 3 m for 50 GNSS points, confirming that georeferencing a ground forest plot requires dozens of measurements at different places or times (Sigrist et al., 1999). An important aspect to consider when sampling multiple points is to maximise the difference between the conditions of acquisition, i.e., accounting for the local forest structure and for the number of available satellites and their geometry. **It is therefore not recommended** to average positions over a few seconds or minutes in the same place when under dense canopies, but rather to average positions over several hours. Alternatively, several positions in a plot can be collected, spaced at least 10 m apart, if insufficient time at a single plot for several temporally spaced position measurements. Practical recommendations for georeferencing a plot are given in Box 3.2.

The state-of-the-art differential GNSS positioning in forests around the mid-1990s indicated an expected accuracy of around 3-4 m (e.g. Deckert & Bolstad, 1996) using C/A (course/acquisition) code observations. Since then, major improvements in code-based positioning have resulted in greater robustness due to the substantial increase in the number of satellites at any given place. With survey-grade receivers collecting carrier phase observations, there is evidence of greater opportunities for higher accuracy by combining code and phase measurements in real-time kinematic as well as static post processing modes. A challenge under forest canopies, when aiming to solve the phase ambiguity using the phase observations, is the frequent loss of signal from individual satellites due to branches and other obstacles. Solving the phase ambiguity requires continuous observation of the same satellites for multiple epochs. For instance, in our study case in Gabon, it was not possible to obtain fixed or float solutions over a few epochs for any of the 208 DGNSS points collected, so differential correction was only applied on the code and did not result in any improvement. This failure probably arose because phase shift information is very sensitive to multipathing effects and a fixed station can only account for atmospheric effects. By contrast, a series of experiments conducted in boreal forests around year 2000, which used survey-grade receivers under canopies of various densities, multiple available satellites (GPS and GLONASS), and multiple processing algorithms and software packages, concluded that it is possible to obtain an average accuracy of better than 0.5 m for plot positions under forest canopies within the time frame required to conduct biophysical observations (Næsset, 1999, 2001; Næsset et al., 2000). Based on these experiences, plot positioning with

survey-grade receivers and differential processing have been conducted successfully (as suggested by accuracy reported by processing software) in dry tropical forests as well as in dense rain forests (Ene et al., 2016; Hansen et al., 2015). Alternatively, a more traditional approach would be to use surveying instruments such as theodolites to traverse from open areas, where good DGNSS positions can be obtained. This, however, requires large openings in the plot vicinity, or access to the canopy.

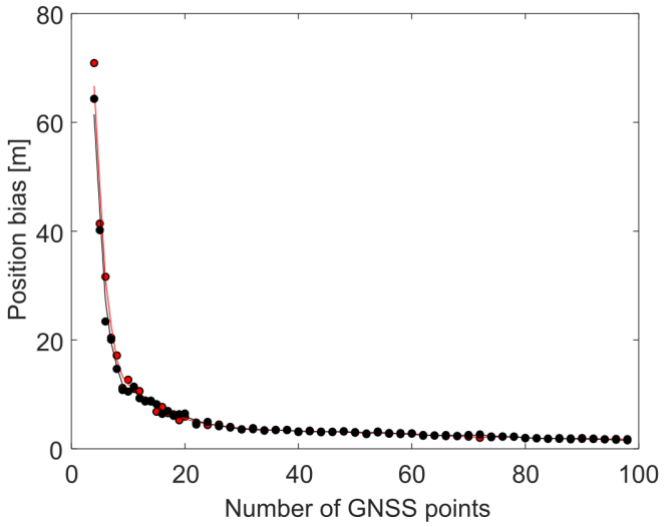

Figure 3.1 Geopositioning errors in a tropical dense forest from Gabon. Higher envelope of the confidence interval of plot positioning error (97.5th percentile of obtained position bias) in function of the number of GNSS points used, and the effect of differential correction. Black: no differential correction, Red: differential correction.

There are two major challenges associated with use of carrier phase observations under forest canopies in general and very dense tropical canopies more specifically. First, it is difficult to know how accurate a determined position is in a given case, and especially so if the estimated position is determined by a less robust technique. Most software used for differential processing estimates the precision of the coordinates. Næsset (2001) analyzed the relationship between precision of estimated coordinates reported by a particular software package and the true accuracy, and found a strong correlation. However, the estimated precision had to be multiplied by a factor of 2 to arrive at the true accuracy. Such relationships can be useful to assess if an estimated position is determined with satisfactory accuracy. Following such a strategy for accuracy assessment requires preparation in the form of establishing the empirical relationship between reported (estimated) precision and true (observed) accuracy. Second, solving the phase ambiguity (fixed solution) may give centimeter accuracy. However, under tree canopies, multipathing may result in false fixed solutions, with the consequence that positions are incorrectly determined with an offset (error) of a few meters and sometimes hundreds of meters. The processing software does not provide any "flag" for false fixed solutions. A more robust and safer strategy can therefore be to opt for a so-called float solution, which does not have the same potential for centimeter-level accuracy, but still can provide better accuracy than a pure code-based position, say, in the range of some decimeters and up to one or two meters. In post-

processing mode, the analyst determines the choice of solution (fixed versus float) during the processing, and need not decide on that during field work. It is also recommended to take measures to reduce the risk of detrimental effects such as multipathing and signal blocking by the trees and the vegetation. Elevating the antenna during observation from a commonly used height of around, say, 2 m, to 3-4 m is advisable, even if the duration of acquisition is likely more important to reduce multipathing effects in high canopy forests. Further, using an elevation mask for the satellites to be observed that is adapted to local conditions (higher mask for taller forests and denser canopies) may help reduce the aforementioned detrimental effects. Finally, in steep terrain where fewer satellites will be above the horizon, planning the field campaign at times when the satellite constellation is favorable can improve the chances of obtaining accurate coordinates.

BOX 3.2

Georeferencing a field plot in practice

GNSS devices are not precise enough to initially set up a plot in the field, hence, field plots should be ideally delineated with a theodolite or with a precise compass in the field for angles and with a laser or decameter (if the terrain is flat) for distances. Note that a magnetic declination exists in many areas (systematic difference between magnetic and geographic north) and should be accounted for during plot establishment and in the subsequent steps. When the plot is established, the georeferencing should start with a minimum of 15 GNSS points taken at different times (over several hours) in the plot center for circular plots or at the four corners for square plots. If not possible, GNSS points can be taken consecutively in different places where the relative coordinates of the plots are precisely known (which is challenging for circular plots). Mounting the GNSS field antenna on telescopic rods may improve acquisition conditions, and in some cases allows for fixed or float positions and for useful differential corrections. Finally, rigid transforms should be applied to convert the relative coordinates into absolute geographical coordinates. This step can be performed using the R statistical software with the correct CoordGPS function of the BIOMASS package (Réjou-Méchain et al., 2017). See also https://water.usgs.gov/osw/gps/.

Under dense canopies where collection of highly accurate geolocation information is particularly challenging, acquisition of UAV and/or airborne lidar can be alternatively used to correct for poor plot registration because their mounted GNSS sensors are not obstructed by the forest cover and thus provide good positional accuracy. These remote sensing data can be used in conjunction with stem mapped tree height measurements in the field to manually adjust plot corners (e.g. Labriere et al., 2018; Réjou-Méchain et al., 2015) or to automatically co-register forest inventory and lidar data (Dorigo et al., 2010). Matching between terrestrial and airborne lidar acquisitions, when both are available, is possibly the best-case scenario (see Chapter 2), as common canopy features are readily recognisable in both sources. This approach to geo-referencing ground plots by co-registering ALS and TLS data has been demonstrated to work well, at least in boreal forests, where an accuracy of 0.5-1 m was reported with ALS data with a point density <1 p/m^2 (Hauglin et al., 2014).

Additionally, satellite product geolocation or geometric accuracy is expected to vary by platform and instrument. However, different studies use different accuracy metrics making any comparison challenging. As a few examples, ALOS-2 Phased Array type L-band Synthetic Aperture Radar (PALSAR-2) data are expected to have a geolocation error of 6.2 m on average (Motohka et al., 2018), Landsat-8 data has an estimated geometric error of 12.6 m (90% circular error; Storey et al., 2014), the GEDI requirement is to have geolocation error of less than 10 m on average (Dubayah et al., 2020) and the NISAR calibration and validation plan (V0.9 JPL D-80829

5/14/2018) states geolocation accuracy of NISAR is expected to be better than 10 m. Thus, the spatial uncertainty associated with remote sensing data may easily be of the same order or higher than that of ground measurements, making any coregistration of small plots and individual pixels highly uncertain

3.1.4 Discrepancies between forest components estimated from field and remote sensing approaches

An important source of spatial mismatch is that forest biomass is "seen" differently by RS and field measurements. While field measurements are generally "trunk-based", considering a tree only if at least half the trunk base section is within the plot, RS sensors measure forests from an area- or volume-based perspective, considering only the plant material having a ground projection within the plot. Mascaro et al. (2011) showed that this representation mismatch leads to strong edge effects due to the bisection of tree crowns in lidar data, which then generates large errors in lidar-AGB models. Naturally, the impact of these edge effects on the global mismatch between field and RS measurements decreases with the spatial grain at which RS models are calibrated/validated (Mascaro et al., 2011). Given that this error is tightly linked to tree crown size, it also strongly depends on the forest type (e.g. needleleaf versus broadleaf forests) and is expected to increase with the forest successional status (Frazer et al., 2011).

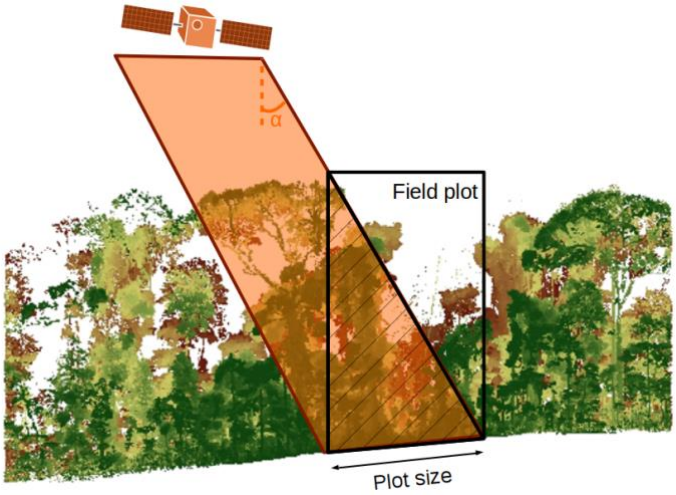

Figure 3.2 Illustration of the mismatch generated by an incidence angle (α) of ca. 30° and a field plot size of ca. 25 m. The hashed polygon illustrates the match existing between the forest volume measured by both the satellite sensor (in orange) and from the field (black). In this example, the large heterogeneity of the forest at the plot scale would result in significantly different biomass estimates from the satellite and ground measurement.

Another "representation" mismatch is that satellite instruments do not necessarily measure forests at nadir, i.e. from a purely vertical perspective. In particular, radar sensors are often used with an incidence angle typically > 30° in order to minimize the contribution of soil backscatter (Robinson et al., 2013). For instance, Saatchi et al. (2011) recommended incidence angles of 30°–40° for estimating AGBD in temperate forests using L-band radar. With such incidence angles, we may expect a relatively high mismatch between the volume of the trees measured from the ground and the volume intercepted by the satellite sensor, especially for small plots (Figure 3.2;

see also (Næsset et al., 2015)) or in topographically complex areas (Villard & Le Toan, 2015). Because this effect has been rarely quantified, we conducted simulations to quantify the effect of incidence angle on the spatial mismatch between ground and satellite-based measurements at different spatial scales, in addition to geolocation errors (see full methodology at **Appendix C**). The results show that an incidence angle of 30° leads to a large mismatch between canopy volumes measured from the ground and from this particular RS sensor for small plot sizes, with an error of more than 50% with 20 m plots (Figure 3.3). When geolocation errors are jointly simulated the error becomes larger, e.g. with an incidence angle of ca. 30° and a geolocation error of 15 m, the mean error was larger than 70%, 35% and 20% for plot size of 20, 50 and 100 m respectively. This indicates that whenever a signal is intercepted at a marked incidence angle, large field plots are needed for calibration and validation. However, if the forest structure is spatially homogeneous, the forests surrounding the plot are structurally similar to that inside the plot, errors associated with incidence angles are partly offset, with RMSE values peaking at only 30%. This also explains noisier error trends on the bottom row of Figure 3.3, due to a random contribution of trees located outside the plot.

There are two solutions to minimize uncertainties due to viewing geometry. First, the spatial resolution at which calibration/validation is performed should be large enough to maximize the overlap between RS and field observations. Second, terrestrial lidar point clouds may be used instead of traditional trunk-based measurements so that field data may match RS observations while explicitly accounting for RS acquisition parameters such as the incidence angle - if using synthetic aperture radar (SAR) data.

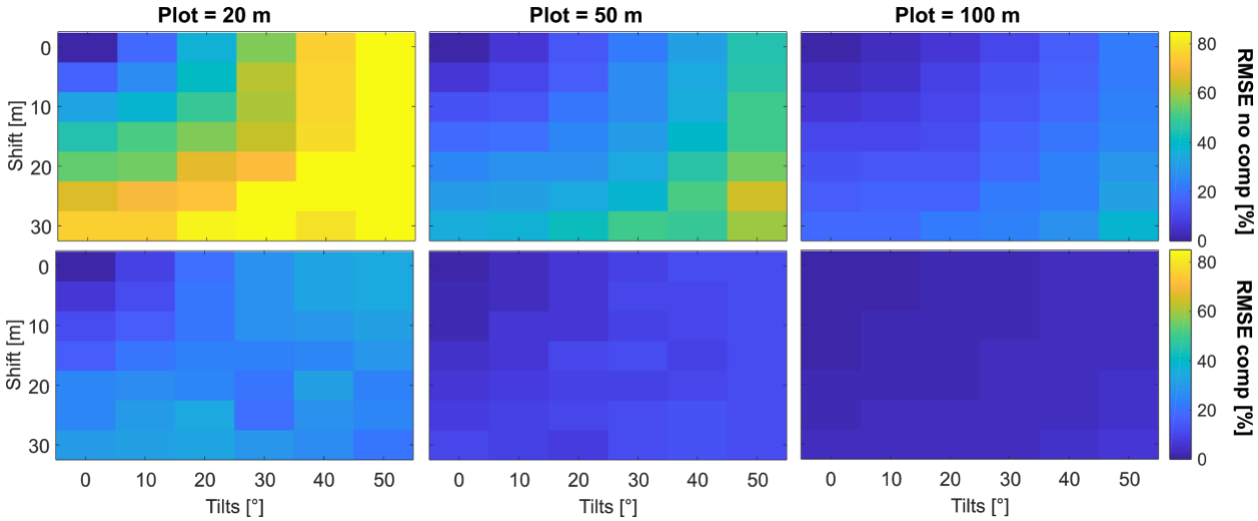

Figure 3.3 Joint effect of plot size (Plot), sensor angle incidence (Tilt) and geolocation error (Shift) on the difference in total crown volume estimated by an active remote sensor and the same volume as generally measured from the ground with trunk-based measurements. The sensor's footprint is identical to the plot size. The top row only considers the volumes of the trees measured within the plot and the bottom row also considers the volumes of the trees surrounding the plot, allowing for error compensation.

3.1.5 Temporal difference between remote sensing and field measurements

A common problem during RS calibration/validation is the temporal difference between RS and field measurements that are difficult to acquire simultaneously, particularly at large spatial scales. For biomass applications, increases in tree mass occur slowly over time, and thus ground measurements can often be used for calibration/validation of RS for a few years after acquisition, provided it is a relatively slow growing forest (i.e., not a recovering tropical forest), and there is no major biomass disturbance between the field and RS acquisitions. However, growth, turnover and disturbance rates vary through space and time, and cannot always be detected by optical sensors. The impact of this temporal difference is difficult to quantify due to the strong stochastic component in natural forest temporal dynamics. Chambers et al. (2013) showed, however, that in mature forests the error associated with temporal differences between measurements tends to decrease as plot size increases, as the AGB gains and losses at large scales average out. Thus, as the spatial resolution at which calibration/validation is performed gets coarser, the error due to temporal mismatches tends to be smaller. However, even large plots are subject to large scale ecological disturbances, such as hurricanes (Espírito-Santo et al., 2014), or to large scale mortality events during drought years (Feldpausch et al., 2016; O. L. Phillips et al., 2009). Thus, whenever a temporal difference exists between RS and field measurements, RS time series, such as Landsat (M. C. Hansen et al., 2013; Kennedy et al., 2010) or MODIS (Justice et al., 2002), may be used to discard any area that may have suffered from major disturbances, such as hurricanes or fires, during the intervening time. With the recent increase in the number of small satellite constellations designed to observe Earth (Boshuizen et al., 2014), high magnitude forest disturbances (e.g. wildfire, harvesting) may be monitored with a high temporal (within a month) and spatial (e.g. 5 m) resolution (Finer et al., 2018). Note also that whenever young, high productivity secondary forest areas are considered, growth is difficult to detect from time series data, such as from passive optical satellite data that are generally insensitive to slow increases in height beyond some saturation threshold. In this case, estimated growth rates may be used to correct and temporally align field and remote sensing measurements (Avitabile & Camia, 2018).

3.1.6 Discussion and recommendations

In this chapter, we have shown that several sources of spatial or temporal mismatch may occur during the calibration and validation steps, potentially leading to estimation errors. A common feature of these error sources is that they all decrease with increasing plot size. Indeed, the spatial overlap between field and RS observations increases with plot size in the presence of mismatch between field plot and pixel sizes (Réjou-Méchain et al., 2014), coregistration errors (Frazer et al., 2011), crown-bisection effects (Mascaro et al., 2011) or large incidence angles. Further, temporal differences are also expected to be better controlled with large field plots. Thus, the size of the plots and the spatial scale at which the RS product is calibrated or validated should be large enough (typically 1 ha in tropical forests) to minimize these errors.

As presented in this chapter, the minimum spatial scale that should be adopted during a calibration/validation approach is context-dependent, varying with sensor resolution, incidence angle and forest structure. For instance, the future BIOMASS mission, a P-band radar mission to be launched in 2022, will have a medium AGB product resolution of 200 x 200 m^2 with an

incidence angle of 23° to 32°, thus potentially generating large mismatch errors with small calibration/validation plots, especially in heterogeneous mature tropical forests. In such a case, the calibration/validation strategy should rely on either large plots, or use a two-step calibration strategy using intermediate high-resolution products, such as lidar (Réjou-Méchain et al., 2019). By contrast, the GEDI mission acquires data in a near-nadir mode and at a 25 m resolution, suggesting that smaller calibration plots may be used. One of the most important challenges for the GEDI mission is the co-registration accuracy and the edge effects that may, at 25 m scale, dominate all other errors.

In summary, **it is recommended** that calibration and validation are conducted at the resolution of the satellite product or sensor, and large well-georeferenced (see Box 3.2) field plots (1 ha where possible, and at least 0.25 ha under dense canopies) be used with subplot information, enabling multiple satellite products to be compared to a single plot. Smaller plots are suitable for areas with relatively sparse canopies and typically lower biomass densities (<100 Mg/ha), where geolocation accuracies are higher. Finally, where possible, **it is recommended** to acquire TLS data when comparing field plots to RS data acquired with high incidence angles.

3.2 Linking Field Plots to Spaceborne Data using Airborne lidar

John Armston, Laura Duncanson, Sassan Saatchi, James Kellner, Erik Næsset, Pierre Ploton, Ralph Dubayah

3.2.1 Background

Airborne lidar, or Airborne Laser Scanning (ALS), has been demonstrated as a technology to estimate forest canopy height, cover and vertical structure, and to produce high-resolution (sub-meter) map products over extensive areas (1000's of ha) (Dubayah et al., 2000; A. Fisher et al., 2020; Wulder et al., 2012). ALS estimates tree height using laser-based ranging from above and can resolve distance as accurately, if not better, than techniques used in the field (Asner et al., 2010). The precision of ALS systems and their ability to rapidly map a large area provides us with an alternative to direct estimates or measurements of vegetation structure and a pathway to scale ground-based forest inventory data to entire landscapes (> 1 km^2), which enables comparison to observations of vegetation structure from space (Zolkos et al., 2013). Additionally, ALS data have been used extensively to provide estimates of forest aboveground biomass density (AGBD) of all forest types and edaphic and climate conditions, covering the entire range and variability of live forest biomass across terrestrial ecosystems (Ene et al., 2013; Labriere et al., 2018; Wulder et al., 2012; Xu et al., 2017).

Along with field inventory data, airborne lidar is in many ways the gold standard reference dataset for a variety of applications including calibration and validation of EO data products and associated algorithms (Duncanson et al., 2019; Hancock et al., 2019; McRoberts, Næsset, Saatchi, et al., 2019; Xu et al., 2017). While for forest inventory, particularly in Europe, ALS acquisition and processing protocols are operational, important data acquisition, processing, AGBD modeling and mapping decisions can impact the accuracy of ALS AGBD reference maps (Tittmann et al., 2015). In this section, we summarize the general approaches to using ALS for EO biomass

map product validation, including decisions on data acquisition, processing, AGBD modeling, generation of AGBD maps, and comparison to EO map products.

3.2.1.1 Development of the multi-stage approach

We generally treat ALS data as more reliable for relating to biomass than spaceborne lidar data due to several sources of measurement and sampling uncertainty inherent in large footprint lidar observation from instruments such as ICESat/GLAS, ICESat-2 and GEDI (Hancock et al., 2019; Michael A. Lefsky et al., 2002; Popescu et al., 2011; Silva et al., 2018). A frequently used approach for wide area biomass mapping uses ALS data as an intermediary to link field plots and satellite data (e.g. R. Nelson et al., 2017). Generally, this involves clipping ALS data to match the spatial extent of field plots, developing empirical models to estimate field AGBD from ALS, application of that model to the full ALS dataset to map AGBD, and finally building an empirical relationship between ALS estimated AGBD and the EO measurements.

There are several advantages to this three-stage approach over linking field plots directly to EO data. First, it allows the use of field plots that are not of the same size and/or shape as EO measurements (e.g. for circular or near circular spaceborne lidar from GLAS or GEDI, or large pixels from SAR sensors or MODIS, as in Boudreau et al., 2008). Secondly, it allows training satellite models with a larger sample size from the ALS than is often available from the field data (e.g. Neigh et al., 2013). Similarly, reference maps of structure or biomass can be used for the validation of EO derived biomass maps by providing a large number of sample points across environmental gradients and the range of structure and biomass in the population (e.g. Duncanson et al., 2020). And finally, if the relationship between airborne estimates and satellite data are consistent through space and time, time-series satellite data may be used to map changes in AGBD (e.g. Margolis et al., 2015b).

It is recommended that these advantages of the three-stage approach are weighed against the impact of introducing a second model (field plot to ALS), which adds an additional source of uncertainty. It is important that the total uncertainty is smaller than that from the alternative approach – to go from directly field plots to EO data. For satellite lidar there is currently little choice as geolocation error and orbital patterns prevent direct linking of field plots and EO data (Patterson et al., 2019), however for satellite imaging EO data (optical, SAR) the choice may be less clear and both approaches should be considered (e.g. Saarela et al., 2016).

3.2.1.2 Reference sites and large area sampling

There are a wide range of ALS sampling approaches adopted in the literature, which reflects the diversity of applications, but also resources available to users. In general, three types of approaches to large area sampling exist: (i) local wall-to-wall ALS coverage over a reference site or set of reference sites (~100-10,000 ha) (Labriere et al., 2018); (ii) non-probability-based ALS transect sampling along environmental gradients or satellite ground tracks (10,000+ ha) (Dubayah et al., 2020); and (iii) probability-based ALS transect sampling for estimating biomass in an inferential framework over a large area (~100,000s ha), which is often applied in National Forest Inventories (e.g. Ene et al., 2016; Næsset et al., 2013; Xu et al., 2017). These three approaches to large area sampling each result in ALS data with different characteristics that

impact the calibration and validation of EO data. Local wall-to-wall ALS acquisitions are typically supported by field sampling to develop lidar-biomass model calibrations specific to local forest tree structure, species composition, and edaphic condition. This first approach is designed to minimize the model error and produce accurate biomass reference maps, however the lidar models may not be applicable outside the conditions it was trained on. The second approach has a larger spatial extent and may employ high altitude large footprint waveform ALS systems such as the NASA Land, Vegetation, and Ice Sensor (LVIS). However, lidar-biomass model calibrations may need to be developed for multiple prediction strata, which may not be cost-effective for individual flight lines, or a larger model error may need to be accepted to generalize to changes in vegetation composition and structure across the large transect. In the case of LVIS, such model development also needs to account for uncertainty associated with collocating large footprints with field plots acquired prior to the ALS acquisition. The third approach may also need to develop models for multiple prediction strata and is the most expensive, however if designed properly (see Chapter 4), can provide accurate local reference biomass, sampling across environmental gradients, and large area biomass estimates for comparison with regional to continental scale EO data products. A number of existing lidar campaigns have been conducted following each approach, with some examples provided in Table 3.1.

Table 3.1 Examples of ALS data collections acquired using local wall-to-wall, non-probability (NP) transect sampling, and probability (P) based transect sampling. This is by no means an exhaustive list as hundreds of campaigns have been conducted but highlights the diversity of ALS campaigns.

Collection	Sampling approaches	ALS instrument	Reference
TERN (Australia)	Local wall-to-wall	RIEGL LMS Q560	Cleverly et al. (2019)
AfriSAR (Gabon)	Local wall-to-wall Transects (NP)	NASA LVIS Facility	Fatoyinbo et al. (2021)
Arctic Boreal Vulnerability Experiment (ABoVE)	Transects (NP)	NASA LVIS Facility	Miller et al. (2019)
BIOMASS reference data (French Guiana and Gabon)	Local wall-to-wall	RIEGL LMS Q560	Labriere et al. (2018)
Trans-Canada	Transects (NP)	Optech ALTM	Wulder et al. (2012) Hopkinson et al. (2016)
La Selva, Costa Rica	Local wall-to-wall	NASA LVIS	Dubayah et al. (2010)
Tanzania Inventory	Transects (P)	Leica ALS70	Ene et al. (2016)
Norwegian Inventory	Transects (P)	Optech ALTM3100 NASA Portable Airborne Laser System (PALS)	Gobakken et al. (2012) Nelson et al. (2012)
Alaska Inventory	Transects (P)	NASA's GLiHT	Ene et al. (2018) Babcock et al. (2018)
NEON sites	Local wall-to-wall	Optech Gemini	Kampe et al. (2010)

There is great value in the use of large area ALS transects collected using non-probability or probability-based sampling designs (as outlined in Table 3.2) - these fill an important niche for national forest inventories and calibration and validation of equivalent scale estimates from EO maps. However, ALS transects acquired using non-probability sampling are often designed to capture contrasting conditions designed for hypothesis driven research or calibration and validation of EO instrument performance, rather than representative sampling of biomass and canopy structure. Different random and systematic ALS transect probability sampling designs have been discussed by Wulder et al. (2012) and VT0005 (Tittmann et al., 2015), however the aim of such acquisitions is typically to minimize uncertainty in the whole-region biomass estimate for a large area (e.g., a country) rather than the population unit of an EO map (i.e. a pixel). These exercises are expensive, and the optimal approach is often determined by regional constraints on the cost-efficiency of ground and airborne surveys and their importance in the inferential framework rather than the quality of high-resolution reference maps for pixel level validation of EO maps (E. Hansen et al., 2015).

Table 3.2 Advantages and challenges of local wall-to-wall, non-probability (NP) and probability (P) based ALS transect sampling.

Sampling approach	Advantages	Challenges
Local wall-to-wall	● Ensures complete detailed coverage over field plots ● Can be collected with UAV-LS ● Higher point density from overlapping swaths ● Can be used to build a global network coincident with existing plot monitoring networks ● Data processing and management packages can be consistently applied to produce local ALS AGBD maps trained with field data	● Requires consistent methodology to resolve differences between ALS instrument and survey specification ● Low intensity sampling across environmental gradients
Transects (NP)	● Large area and high intensity sampling across environmental gradients ● Cost-effective for high-altitude flights ● Flight-lines can be designed to sample different field plot collections across large areas ● Sampling design can be flexibly designed based on vegetation type, extent, environmental gradients, and target EO data acquisitions	● Potentially lower pulse density / ranging precision associated with higher altitude acquisitions ● Susceptible to changing weather conditions (potential loss of quality data over field plots) ● Accounting for disparate properties of different field data collections in model development ● Applicability of biomass models away from field plot collections often unknown
Transects (P)	● Unbiased estimators for whole-region biomass density and its uncertainty ● Can be used to assist and complement existing National Forest Inventory field data collections ● Established statistical inference methods for estimation of uncertainty	● High cost of developing repeatable, consistent and high-quality biomass density estimates across a range of ecosystems globally

As there is no ‘one size fits all’ transect sampling approach for validation of EO biomass products and to minimize pixel-level uncertainty in reference estimates of biomass, **it is recommended** to adopt local to regional wall-to-wall ALS coverage over permanent in situ reference plots. This will enable the generation of a global set of high-quality AGBD reference sites with consistent field and ALS acquisitions for validation of disparate EO biomass products, as well as leverage and expand existing monitoring networks that use this approach (e.g. TERN, NEON). It is important to

emphasize that local wall-to-wall lidar is not the only valid approach, or necessarily the optimal for regional validation strategies that need to make best use of all available reference data. ALS transect sampling is an increasingly important component of statistical inference frameworks used by NFI's, has in part driven the development of such statistical inference methods for estimating biomass density and its uncertainty (see Chapter 4), and they form an essential part of the overall calibration and validation strategy of EO missions such as GEDI (Dubayah et al., 2020).

3.2.2 Data acquisition protocols

3.2.2.1 Instrument specifications and limitations

There are three classes of lidar instruments commonly used in ALS systems:

1. **Discrete return lidar** - these systems use analogue detectors to record discrete, time-stamped trigger pulses from the received waveform in real time
2. **Waveform lidar** - these systems digitize the transmitted and received waveforms at a high frequency and records the energy returned over equal time intervals
3. **Photon counting lidar** - equipped with an array of receivers that are sensitive to individual photons, these systems use lower energy lasers and are capable of large area mapping from higher flying altitudes

The vast majority of research and almost all operational applications for estimation of biomass from ALS have used small footprint discrete return lidar systems operating in the near infrared wavelength (see Table 3.3). However, small footprint waveform lidar instruments are now commercially accessible and several other sensor types are available and are used for biomass product calibration and validation, including research instruments, such as NASA's LVIS, and photon counting instruments, such as the Multiple Altimeter Beam Experimental Lidar (MABEL), which was created for ICESat-2 testing and development.

For the purposes of AGBD product validation **it is not currently recommended** to use airborne photon counting instruments. Method development for photon counting lidar is advancing with the recent availability of ICESat-2 data (e.g. Narine et al., 2019), however airborne instruments are not as widely accessible as waveform or discrete return. There are two types of airborne photon-counting lidar systems available - Geiger-Mode Lidar and Single Photon Lidar (SPL). Recent research with SPL is showing promise for forest applications (e.g. Swatantran et al., 2016; J. C. White et al., 2021), particularly as photon counting lidar can efficiently map large areas (Stoker, Abdullah, et al., 2016). However further research is needed to use these for routine generation of high-quality reference forest structure and biomass maps across a wide range of ecosystems and observation conditions. Both full waveform and discrete return instruments are well established, and often used operationally, for large area mapping (Asner et al., 2010; Wulder et al., 2012). It is important to acknowledge UAV-LS, which has emerged as a new data stream that can collect high-quality, repeatable forest structure estimates that fill a niche between ALS and TLS for biomass calibration and validation activities (see section 2.3 for further details).

Discrete return lidar

- In the last three decades discrete return instruments have matured considerably, with reduced deadtime, increased number of pulses, and a proliferation of software to process point clouds. **It is recommended** to use instruments with ≥ 4 returns per pulse to ensure vertical structure is captured and avoid older instruments with only first and last returns (Gobakken & Næsset, 2008; Leitold et al., 2015).
- Has underpinned the development of operational forest inventory programs and algorithms for tree level attributes such as crown shape, height and biomass have almost been solely developed using these data (Holmgren & Persson, 2004; Vauhkonen et al., 2009).
- Rapid advances in discrete return lidar technology and the large number of different instruments available at any one time have resulted in disparate instrument characteristics. The impact of differences in analogue return detection methods, laser power, etc. on forest structure estimates are difficult to resolve (Disney et al., 2010; Næsset, 2009).

Waveform lidar

- Processing of waveforms to discrete returns is performed in post-processing. This enables additional points, and attributes of these points (e.g., apparent reflectance), to be derived, providing additional information on canopy properties. Benefits of small footprint waveform lidar for reducing uncertainty in biomass estimation is yet to be shown, however, national programs such as the Terrestrial Ecosystem Research Network (TERN) and the US National Ecosystem Observation Network (NEON) are making waveform datasets available for applications.
- Enables more direct retrieval of canopy cover and the vertical foliage profile that are less sensitive to survey configuration differences than discrete return lidar (Armston et al., 2013).
- Large footprint waveform lidar has a long history (> 20 years) of method development for canopy structure and biomass estimation, which has:
    - demonstrated vertical canopy structure can be rapidly and efficiently mapped over large areas when flown at higher altitude with larger footprint size (e.g. NASA's LVIS)
    - underpinned the development of the NASA GEDI mission (Dubayah et al., 2010, 2020).
    - exhibits greater uncertainty in height estimation on steep slopes because the ground and canopy signals blend, and with this, uncertainty increases with increasing footprint size (Hyde et al., 2005).

Table 3.3 Examples of ALS instruments used in the estimation of forest structure and AGBD. This list is non-exhaustive and new commercial instruments are often released by a number of manufacturers.

Instrument	Instrument type	Number of returns	Wavelength [nm]	Maximum Altitude [m AGL]	Maximum laser pulse repetition rate [kHz]	Beam Divergence [mrad]	Max FOV [deg]	References
NASA Land, Vegetation & Ice Sensor (LVIS) Facility	Large footprint waveform	N/A	1064	20000	4	0.75 to 3	12	Fatoyinbo et al. (2021) Blair et al. (1999)
Leica ALS50-II	Discrete return	4	1064	6000	150	0.22	75	Hudak et al. (2020)
Leica ALS70-HP	Discrete return / waveform optional	unlimited	1064	3500	500	0.22	75	Duncanson et al. (2020)
Optech Orion-M	Discrete return / waveform optional	4	1064	5000	500	0.25	50	Longo et al. (2016)
Optech Gemini	Discrete return / waveform optional	4	1064	4000	167	0.25 or 0.8	50	Goulden and Scholl (2019)
Optech ALTM3100	Discrete return	4	1064	3500	100	0.3 or 0.8	50	Gobakken et al. (2012)
RIEGL VQ-480i	Discrete return	6	1064	5000	550	0.3	60	Cook et al. (2013)
RIEGL LMS-Q560	Waveform	unlimited	1550	1200	200	0.5	60	Labrière et al. (2018)
RIEGL LMS-Q680i	Waveform	unlimited	1550	5000	400	0.5	60	Armston et al. (2013)
Leica SPL100	Photon counting	10	532	4500	60	0.08	60	White et al. (2021)

3.2.2.2 Survey requirements

ALS instruments vary in their laser wavelength, pulse repetition frequency (PRF), pulse length, pulse shape, signal triggering mechanism, detector sensitivity, beam divergence, vertical discrimination distance, and scan pattern (see Table 3.3). Surveys vary in their flying height above ground, flying speed, maximum off-nadir scan angle, scan rate and swath overlap. Changes in ALS discrete return instrument and survey characteristics can cause differences in canopy height and cover estimates that are independent of forest structure and difficult to quantify directly (Fisher et al., 2020). Individual instrument and survey properties are related to each other so it is difficult to isolate their impact, however ultimately their combined effect on individual laser pulse energy per unit area can lead to errors in canopy height and cover estimation (Armston et al., 2013; A. Fisher et al., 2020; Næsset, 2009).

Given the importance of accurate Digital Elevation Models (DEMs) for modeling canopy height, particularly in tropical forests, their accuracy requirements are directly relevant for ALS biomass map generation. For example, the US Geological Survey (USGS) Lidar Base Specification

(Heidemann, 2018) outlines quality levels that define minimum acceptable ALS acquisition parameters to ensure consistency between independent ALS acquisitions for the USGS 3D Elevation Program (3DEP) Program. Generation of accurate ALS biomass maps across multiple independent ALS acquisitions that follow such guidelines has been achieved (Hudak et al., 2020).

For forest applications a minimum density of 8 pulses/m^2 and non-vegetated vertical accuracy of ≤ 0.196 m (95% confidence interval) has been recommended for consistent estimates of forest canopy height and ground elevation (Heidemann, 2018). This is far greater than minimum density recommendations in the literature, which range from 0.25 pulse/m^2 in Norwegian coniferous forest (Gobakken & N Næsset, 2008), ≥ 1 pulse/m^2 in Tanzanian rainforest (Hansen et al., 2015), ≥ 2 points/m^2 in Japanese coniferous forest (Kodani & Awaya, 2013), and ≥ 4 points/m^2 in Brazilian Atlantic forests (Leitold et al., 2015). This latter finding is consistent with survey specifications used by the National Science Foundation (NSF) NEON ALS acquisitions, which also specify ≥ 4 points/m^2 (Goulden & Scholl, 2019).

For a consistent quality of estimated elevation, height and cover estimates across a wide range of forest environments, a minimum density of ≥ 4 pulses/m^2 **is recommended**. However, minimum pulse density recommendations are site and metric specific. Lower pulse densities ≥ 1 have been shown to be adequate for area-based biomass estimation in boreal environments and Hancock et al. (2019) found that ≥ 2 points/m^2 was sufficient for simulation of GEDI waveforms used to calibrate GEDI footprint biomass models, with some dependency on ALS instrument type. Pulse density may also be a far less important factor in minimizing biomass estimation uncertainty than the number and size of field plots (Gobakken & Næsset, 2008; Hansen et al., 2015), which may require greater investment to reduce overall uncertainty.

Within constraints of minimum pulse density and areal coverage requirements, **it is recommended** to use commercial survey and instrument configurations that maximize individual laser pulse measurement quality rather than even higher pulse density. Ideally, beam divergence, flight altitude, and pulse repetition rates are set to ensure footprint sizes < 30 cm and maintain sensitivity to canopy foliage. Also important is to ensure the maximum off-nadir scan angle is < 15° (Goodwin et al., 2007), and for local wall-to-wall mapping the swath overlap is 50% to ensure creation of coregistered flight lines, even sampling density, and minimisation of off-nadir scan angle. Existing protocols developed for TERN and NSF NEON outline these requirements in more detail, along with recommendations for quality assurance (QA) / quality control (QC) (Held et al., 2018; Krause & Goulden, 2015).

Wall-to-wall ALS reference sites can be selected in different ecoregions using a minimum area of 3 × 3 km (1000 ha) to capture landscape level variation in topography, forest structure, and composition within the same ecoregion. In addition to minimizing cost and logistics of airborne campaigns, a minimum area of 1000 ha will enable comparison with EO biomass map products from current and upcoming space missions with biomass requirements (NASA GEDI, NASA-ISRO NISAR and ESA's BIOMASS). TERN Australia (Cleverly et al., 2019) used 5 × 5 km and NSF NEON used 10 × 10 km (Kampe et al., 2010).

3.2.3 Data processing protocols

Data processing protocols are also variable, with providers using different detection thresholds, ground classification algorithms, noise filtering algorithms, and instrument calibration routines. The basic ALS processing workflow starts with basic QA/QC of lidar point clouds, point cloud classification, noise detection and removal. Again, existing QA/QC protocols developed for USGS 3DEP, TERN and NSF NEON outline these requirements in more detail. For the purposes of this protocol, we will not recreate one of the many excellent lidar processing protocols that are available for the community (Held et al., 2018; Krause & Goulden, 2015; Stoker, Brock, et al., 2016; Joanne C. White et al., 2013). We will, however, make general recommendations to minimize uncertainties in ALS AGBD maps associated with lidar processing.

Open-source, as well as commercial software, for the complete processing chain, from waveform processing (e.g. SPDLib; Bunting et al., 2013) to gridding of canopy metrics, including lidR (e.g. Roussel et al., 2018), PyLidar (Armston, Bunting, et al., 2020) and FUSION (McGaughey, 2020) are increasingly available. In the longer term, the development of open-source software that is suited to automated processing will ensure genuine transparency in the algorithms and data products generated from airborne lidar over large areas, and enable other researchers to replicate scientific results without incurring significant financial costs or encountering proprietary barriers.

3.2.3.1 Canopy height models

Ground classification is probably the most important consideration for AGBD reference map generation, since canopy height models are required to derive most ALS area and tree level metrics. However, details in processing algorithms (e.g. the selection of noise thresholds, map spatial resolution) can yield important differences in the final ALS metrics. Best practice is to interpolate ground elevation to the location of each point and then assign a height above ground, prior to gridding the data. These data can be clipped directly to field plot extents for ALS biomass model development (see Section 3.2.4.1), which minimizes uncertainty related to colocation of field plots and ALS metrics (see Section 3.1).

The minimum ALS product required for generation of canopy height models, which underpins the area-based approach, is the classified (unfiltered) ALS point cloud. The point cloud should contain all valid returns. Lidar points are typically classified as returns from the ground, canopy, or noise. At some sites, classification of buildings and other infrastructure is required. Most LAS (the industry standard binary format for storing airborne lidar data) processing software (e.g. LAStools) includes tools for classification, although the majority of commercially-acquired ALS data will already have a classification assigned. **It is recommended** to use existing protocols for QA/QC of point cloud classifications (Quadros & Keysers, 2018).

It is still recommended to check the classification visually through a point cloud visualization tool to ensure that noise above and below the canopy has been classified correctly, and that ground classification appears accurate, even under dense canopies. Occlusion mapping using voxel traversal of laser pulses can identify which parts of the canopy were acquired by different acquisitions (e.g. Kükenbrink et al., 2017). This basic QA/QC will help ensure errors in

classification do not propagate to errors in height and cover metrics, particularly in areas of high biomass or dense cover (e.g. > 400 Mg/ha or > 90% canopy cover).

Once a high-quality, classified point cloud is available, forest structure metrics (height and cover) are typically calculated by normalizing points to a height above the ground, binning the returns into a planimetric grid and then generating statistical summaries, or applying retrieval algorithms for biophysical parameters such as canopy cover. Decisions related to the selection of interpolation methods for DEM generation may impact the downstream canopy structure metrics and depend on canopy cover, slope, point density, and classification error. A guide for the generation of raster layers from point clouds is provided by NEON (Goulden & Scholl, 2019). Their algorithm theoretical basis document (ATBD) outlines the processing steps for the generation of NEON's analysis ready data (ARD) canopy height and DEM layers, among others.

3.2.3.2 Tree or area-based approach

In contrast to the relatively convergent ALS processing algorithms for areal-based or plot-level forest structure metric calculation, tree and crown extrapolation algorithms vary considerably in their methodology and accuracy. In the past three decades there has been an emergence of research focus on individual tree extraction from airborne lidar datasets (Ferraz, Saatchi, Mallet, & Meyer, 2016; Holmgren & Persson, 2004; Hyyppa & Inkinen, 1999; Koch et al., 2006; Popescu et al., 2003; S. D. Roberts et al., 2005). The ecological and management applications of individual tree crown information are obvious, and if individual crown information can be routinely and accurately extracted, individual AGB estimates could be made at a landscape scale. Additionally, individual tree based change detection and AGB dynamics could be monitored from multi-date lidar (e.g. Hyyppa et al., 2001; Kaartinen et al., 2012; K. Zhao et al., 2018).

Several papers have attempted to estimate individual tree based AGB using these approaches (Bortolot & Wynne, 2005; Dalponte et al., 2016; Duncanson, Dubayah, et al., 2015; Ferraz, Saatchi, Mallet, Jacquemoud, et al., 2016; Popescu, 2007; K. Zhao et al., 2009). This is attractive, as it is theoretically scale invariant (AGB is the sum of biomass of individual trees), and there are promising results showing crown width is an important allometric predictor of AGB (Jucker et al., 2017). However, there are a wide range of methods and associated uncertainties for the breadth of algorithms employed to extract individual tree crown information. Algorithm comparison studies (e.g. Aubry-Kientz et al., 2019; Kaartinen et al., 2012) demonstrate that the accuracy of approaches is both algorithm and ecosystem dependent, and challenges remain, particularly in multi-layer forests.

Specific comparisons of individual tree and areal based estimates have found that areal-based approaches have higher precision, lower bias and follow more mature established data processing methods (Coomes et al., 2017; Duncanson, Dubayah, et al., 2015; Maltamo et al., 2009; Yu et al., 2004). Tree-based approaches are presently most useful for open forests and conifers, and detection of understorey trees is not reliable (Duncanson, Dubayah, et al., 2015). **It is therefore recommended** that area-based approaches are used for the generation of ALS biomass reference maps, but as individual tree methods continue to mature this may be revisited in the future.

3.2.4 Biomass modeling and mapping

A huge number of studies have modeled AGBD from ALS structure metrics for purposes of local AGBD mapping, forestry and ecological studies (Zolkos et al., 2013). Models for converting lidar estimates of forest height or vertical structure into AGBD have been developed using large footprint waveform ALS (e.g. Drake et al., 2003; Dubayah et al., 2010; Fatoyinbo et al., 2021; Michael A. Lefsky et al., 1999; Means et al., 1999), discrete return airborne laser profiling (R. F. Nelson et al., 2007; e.g. Ross Nelson et al., 1988), and discrete return ALS (e.g. S. A. Hall et al., 2005; E. Hansen et al., 2015; Labriere et al., 2018; Lim et al., 2003; Næsset, 2003; R. Nelson et al., 2017; Omasa et al., 2003; Stephens et al., 2007; V. Thomas et al., 2006).

When developing a model for producing an ALS biomass map, one must make decisions about how to link the ALS and field plot data to train the model, the geographic specificity of the model (e.g. local to global), the spatial resolution of the model (and product), model specification (selection of model form, predictors, etc.), and how to characterize uncertainty. Here, we briefly summarize these considerations.

3.2.4.1 Linking airborne laser scanner and in situ reference data

For linking with ALS data and to minimize the impact of spatial and temporal mismatches (see Section 3.1) on ALS biomass modeling, **it is recommended** that field plot data be acquired using:

1. fixed area plots with a constant size and shape per ALS campaign (e.g. 0.25 ha squares) to avoid the assumption of scale invariance and for consistency between campaigns
2. a plot size large enough to minimize the impact of edge effects and collocation error (Gonçalves et al., 2017)

Depending on the forest types and size of trees, the plot size selected may vary. For boreal forests dominated by conifers, plots of greater than 0.1 ha may contain enough trees and have accurate ground estimates of biomass. The highest biomass model accuracies have been reported at the 1 ha scale with a model prediction error typically below 15%. Tropical sites with small plots (< 0.1 ha) are more challenging, with model prediction errors potentially above 40% (Zolkos et al., 2013). For tropical forests, **it is recommended** to collect plot sizes greater than or equal to 0.25 ha where ALS metrics typically represent forest structure at a scale larger than the crown of an individual large tree.

When interpreting these model prediction error statistics, it is also important to understand that there is often a trade-off between plot size and the number of plots (Gobakken & Næsset, 2008; Hansen et al., 2015). Model prediction errors (residual variance) are not necessarily a good measure of quality when the objective is to use field plot data to calibrate models for subsequent ALS based prediction and estimation over large areas (e.g. ≥ 1 km^2). **It is recommended** to ensure the range of sample values in the reference plot data and associated ALS predictors are representative, i.e. they are similar in the sample as in the population.

Fixed area square shaped plots **are recommended** in tropical forests since these are easier to establish and maintain (see Appendix A). Circular plots are often used in boreal and temperate ecosystems with lower tree density and AGBD, compared to tropical forests, and where they can be easier to establish and maintain due to their often smaller size (e.g. 0.1 ha; White et al., 2013) and have better line-of-sight to individual trees for stem mapping. Orientation of plots is less important, but may need to be specified if other imaging sensor data are being acquired in addition to ALS, for example where square plots aligned in the azimuth direction of SAR flight lines can ensure better collocation with individual pixels (see Section 3.1).

For linking with field plot data and to ensure spatial and temporal consistency in ALS biomass reference maps, **it is recommended** that ALS data be acquired:

1. in maximum leaf-on conditions, where deciduous species are present;
2. before establishing field plots where possible, since the ALS data can be used to select field plot locations for representative sampling and avoid redundant field effort;
3. ideally within one year of field data collection, to minimize the impact of change such as understorey recruitment events, tree fall, defoliation from insect damage; and
4. across a sufficiently large area (minimum 3 × 3 km) to capture landscape scale variation in canopy structure and AGBD, and enable development of more generalized models.

Temporal alignment between field and ALS acquisitions is important for spatial alignment of field and ALS data (Labriere et al., 2018; Chapter 3.1). Change between ALS and field acquisitions can have a large contribution to total model error with one study showing 7-17% error for 3-year misalignment in tropical secondary forest (Gonçalves et al., 2017). Different approaches are available to account for this error. One is to apply growth models to predict AGB stocks at the time of the ALS acquisition (e.g. Gobakken & Næsset, 2008). However, uncertainty in growth model results is not easily propagated through model-based inference frameworks used for characterizing uncertainty (Babcock et al., 2016). **It is recommended** that the time difference between the ALS and field plot acquisition is minimized, which will require budgeting for new ground surveys coincident with airborne campaigns.

3.2.4.2 Model specification

ALS biomass model specification requires the selection of model form and predictor variables, and assessment of model performance. While a wide range of approaches have been used for modeling the relationship between AGBD and ALS metrics, ALS biomass models can be generally categorized into two types of approaches:

1. Data-driven selection of ALS metrics using parametric or machine learning algorithms. Some recent examples include Duncanson et al. (2020), Esteban et al. (2019), and Chen et al. (2018).
2. Theory-driven algorithms that are constrained by allometric scaling laws using height, canopy cover and to a lesser extent crown diameter (e.g. Asner et al., 2012; Asner & Mascaro, 2014; Coomes et al., 2017).

Following the theory that underpins tree-based allometry (see Section 2.2.3), a number of studies have used a generalized non-linear model form that is a function of canopy top height (H), stand basal area (BA), and stand wood density (WD) (Asner & Mascaro, 2014):

$$AGBD = a \times H^b \times BA^c \times WD^d,$$

where *a*, *b*, *c* and *d* are parameters (power law exponents).

Stand basal area is also not directly observed from the vertical view of ALS instruments. Asner and Mascaro (2014) recommended using regional models that relate canopy height to stand basal area, however this relationship can decouple (Duncanson, Dubayah, et al., 2015), and other studies in both temperate and tropical forests have found canopy cover × height improves model performance (Coomes et al., 2017; Fatoyinbo et al., 2021; Jucker et al., 2018; Ni-Meister et al., 2010). Model performance is higher when using plot-based estimates of mean woody density (Fatoyinbo et al., 2021), however ALS metrics are poor predictors of mean wood density, therefore regional averages must be used (Coomes et al., 2017).

Data-driven approaches typically use multiple linear regression with a combination of ALS height and cover metrics as predictors (e.g. E. Hansen et al., 2015). Models are often linearized by using logarithm or square root transformations of AGBD (and the predictors) to improve the linear relationship between the response and the predictors, and to reduce heteroscedasticity in the models. Predicting AGBD on the original scale requires back-transforming the logarithmic or square root predictions. This transformation is nonlinear and will generate bias, therefore a number of correction factors have been developed to minimize this bias, as are often applied in the estimation of tree biomass (see Section 2.2.3). Such transformations of AGBD may be also avoided through the use of nonlinear models, **which are recommended** where practical. Statistical modeling procedures to select ALS metrics to use as predictors are well documented in the literature, but **the general recommendation is** to minimize the number of predictors in the models to avoid overfitting and multicollinearity. Model performance should be assessed using cross-validation, including cross-validation based on non-random (e.g. geographic) subsets of the data (e.g. Ploton, Mortier, Réjou-Méchain, et al., 2020).

One ALS metric that is commonly used is mean canopy height (MCH). MCH represents the average canopy height within an area and often shows strong correlation to basal area, and hence AGBD. It has been shown to be a robust metric for capturing biomass variation across the landscape (Asner et al., 2012; M. A. Lefsky, 2010; Meyer et al., 2013). Hansen et al., (2017) found MCH, when used as a single predictor of AGBD in a power law model, outperformed data-driven multiple linear regression in tropical biomes, but not in temperate and boreal biomes. Different ALS metrics will be more or less useful in different regions based on local environment (e.g. canopy closure, height distributions), and to some extent ALS instrument and survey properties, and thus we do not provide a specific recommendation for which canopy metrics to include.

The selection of the geographic domain of ALS biomass model calibration will depend in part on the geographic domain of ALS coverage. ALS studies have fit AGBD models to lidar height metrics at local, regional and continental scales. The concept of generalized models that can be applied over large areas is appealing, but results in the literature have not been consistent. For example the approach used by Asner et al., (2012), corrected for regional wood density, was not found to

be useful in temperate forests (Duncanson, Dubayah, et al., 2015). Regional differences in wood density and complex forest structure have indicated local training of models is necessary (Coomes et al., 2017), however Knapp et al., (2020) found only a limited increase in model performance when moving from generalized to site-specific models in the tropics.

While next generation satellite lidar may help inform the development of generalized models to relate ALS metrics to AGBD, local and regional models typically produce higher accuracies and are able to capture relationships between AGBD and ALS metrics, which are driven by local or regional life history, floristic makeup, climatic and edaphic conditions. While **it is recommended** to continue research into generalized models for application to both large area ALS and spaceborne lidar surveys, site-specific models remain more consistently reliable at present, and ALS instrument and survey properties can also potentially cause systematic measurement differences (see Section 3.2.2). Therefore, **it is currently recommended**, where possible, to use local to regional training of ALS biomass models for individual ALS campaigns to ensure they are locally unbiased.

3.2.4.3 Spatial resolution of airborne laser scanner maps

Another critical consideration for AGBD modeling and mapping is spatial resolution. Airborne discrete return lidar can be processed at any spatial resolution (i.e. from sub-meter to km depending on the acquisition specifications). However, there is a minimum resolution at which AGBD data are meaningful, as spatial resolution of field observations is constrained to the individual tree level. Therefore, the highest possible spatial resolution will be at the resolution of the largest tree crowns in a given area (which can span to approximately 0.25 ha in the tropics).

In dense tropical forests, where large plots (≥ 0.25 ha) are recommended, an ideal situation for biomass map calibration and validation is the development of multiple resolution AGBD models, such that a single site can be used to validate satellite products of differing resolutions (see Section 3.2.5). Modeling and mapping at high (< 0.25 ha) spatial resolution should be conducted with caution, largely because of the uncertainties outlined in Chapters 2 and 3. In general, models should be fit at the spatial resolution of field plots. Where only small field plots are available (e.g. 0.0625 ha plots) for ALS biomass model development, the models should only be applied to generate 0.0625 ha spatial resolution ALS biomass reference maps. If coarser resolution products are needed to match the spatial resolution of EO biomass maps for validation, native resolution AGBD estimates can be aggregated easily, however care must be taken in aggregating per-pixel estimates of uncertainty (see Section 3.2.4.4).

An example from Barro Colorado Island (BCI) in Panama is given in Fig. 3.4, where large (1 ha) field plots had subplot and tree location information recorded, which enabled models to be fit at 0.05, 0.25 and 1 ha spatial resolutions. This example particularly highlights the impact of crown size on the relationship between spatial resolution and model accuracy. We see increasing model prediction accuracy with decreasing spatial resolution from 0.05 to 1 ha, which is attributed to a combination of edge effects, geolocation uncertainty in the field data, and large crown sizes. Comparable relationships were observed more recently for both large footprint waveform data (LVIS; Armston, Tang, et al., 2020) and small footprint discrete return data (Labriere et al., 2018)

in the AfriSAR campaign, Gabon (Fatoyinbo et al., 2021). Similar trends but lower model prediction errors were observed for the BCI study site by Knapp et al. (2018), who used simulations to assess AGBD model prediction error at different spatial resolutions in the absence of geolocation error. This work highlighted the importance of accurately collocating ground plots with ALS observations for calibration and validation activities associated with the NASA GEDI, NISA-ISRO NISAR and ESA's BIOMASS missions.

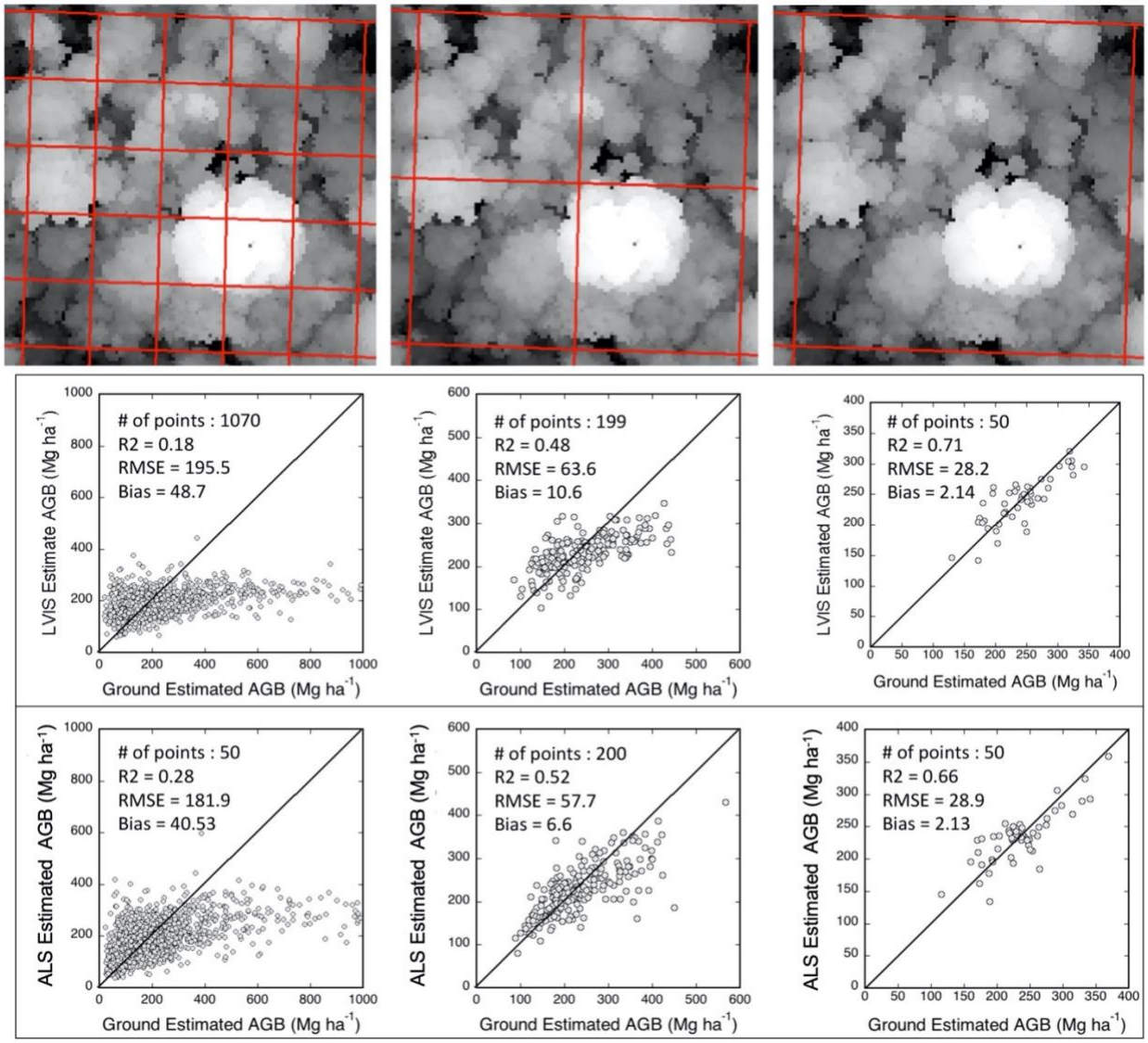

Figure 3.4 Example relationships between ground-estimated AGB from plots of different sizes (top) and ALS estimated AGB using LVIS (middle) and small footprint discrete return data (bottom) in tropical forests of Barro Colorado Island in Panama (adapted from Meyer et al., 2013)**.** These ALS models were non-linear and only used a single predictor (mean canopy height)**.** There is increasing model prediction error with plot sizes of (left to right) 0.05, 0.25 and 1 ha.

In general, model prediction accuracy will increase with decreasing spatial resolution. This is due to both decreases in uncertainty with respect to edge effects, and linking field plots to ALS data (see Sections 3.1 and 3.2.3.1 for details). However, these limitations are most prominent in dense

tropical forests where geolocation uncertainty and colocation errors tend to be higher and large tree crowns increase edge effects (see Sections 3.1 and 3.2.3.1 for details). In some temperate and most boreal forests, higher resolution ALS biomass modeling and mapping will yield highly accurate models given lower canopy cover and smaller tree sizes. **It is recommended** that ALS biomass models always be developed at the spatial resolution of field plots. For densely forested areas (particularly rainforest or areas with large crowns) where large plots (≥ 1 ha) are collected, these will ideally include sub-plot information (e.g. 0.0625 and 0.25 ha) so that AGBD models and maps can be produced at multiple resolutions.

3.2.4.4 Characterization of uncertainty

Propagation of uncertainty from field measurements through to satellite estimates is discussed at length in Chapter 4, but there are several considerations specific to the generation of ALS biomass maps that we discuss here to make practical recommendations on when uncertainty associated with an ALS biomass model can be assumed to be negligible and, when not, how to account for or mitigate its overall contribution.

Given the need to assess EO biomass products against reference estimates, rigorous approaches to characterization of uncertainty need to replace ad hoc approaches (Gregoire et al., 2016; R. Nelson et al., 2017). ALS biomass map uncertainty is typically assessed through comparison of model predictions with independent field estimates using techniques such as k-fold cross-validation using random or non-random (e.g. geographic transferability) subsets of the reference data (e.g. Dubayah et al., 2017; Duncanson et al., 2013; Ploton, Mortier, Réjou-Méchain, et al., 2020). Key recommendations for this type of assessment are outlined in Box 3.3. However, it is not possible to use these results to estimate the error of estimated AGB across an area, which can be as small as an individual EO biomass product pixel.

Some studies have undertaken more detailed error propagation in the estimation of AGBD from ALS, accounting for in situ measurement and allometric model errors (e.g. Asner et al., 2012; Jucker et al., 2018), however these still lack spatially explicit estimates of uncertainty. It is necessary to use a model-based inference framework to estimate uncertainty per pixel (Chen, Vaglio Laurin, et al., 2015), and the formal framework for this is outlined in Section 4.2.2. It has been demonstrated in several studies and data products in recent years (Chen et al., 2018; Esteban et al., 2019; Ståhl et al., 2016), but has not yet been adopted routinely by the remote sensing community due to a lack of software tools and/or statistical expertise.

The generation of reference ALS maps involves the use of multiple models. The first is the allometric model used to estimate plant biomass, which is described in detail in Section 2.2. The second is the model that relates the resulting estimates of plot AGBD to the ALS metrics as described in the preceding subsections of Section 3.2. Ignoring the uncertainty from one of these models can result in underestimation of the true uncertainty (Chen, Vaglio Laurin, et al., 2015; Saarela et al., 2016, 2020). The uncertainty of ALS AGBD estimates for individual locations on an ALS map (i.e., pixels) from the second model can be decomposed into three independent error terms in a model-based inference framework: (i) the error variance estimates related to

individual random errors of predictions; (ii) model parameter errors; and (iii) ALS measurement errors (Chen et al., 2016).

The mean AGBD when aggregating over small areas is straightforward and simply the mean of the ALS pixel estimates. Aggregating estimates of uncertainty over small areas to the spatial resolution of EO biomass products (0.0625 to 4 ha) is more complicated. This is because the spatial autocorrelation of model errors needs to be taken into account when estimating the term for individual random errors of predictions over such small areas (McRoberts et al., 2018; Saarela et al., 2020). Studies that account for spatial autocorrelation in the individual random errors of predictions are scarce - Chen et al. (2016) is an exception - and has been acknowledged as a challenging problem (Saarela et al., 2020). This error component can be ignored over larger areas (>= 1 km; Saarela et al., 2018), which is the current approach for coarse resolution remote sensing biomass products such as GEDI (Patterson et al., 2019). However, the spatial resolution of planned EO biomass products from the NASA-ISRO NISAR and ESA's BIOMASS missions are much finer (1 to 4 ha) and need to be considered.

Measurement error in ALS metrics is almost always assumed to be zero, although Asner et al. (2012) assumed 5% and Chen et al. (2016) assumed 10%, albeit using unrelated approaches to error propagation. Numerous studies have indicated that measurement error can be significant for ALS derived canopy height (Disney et al., 2010), canopy cover (Armston et al., 2013; A. Fisher et al., 2020) and other lidar perceived metrics (Næsset, 2009 and references therein), particularly from older discrete return lidar datasets with low pulse densities and variable footprint sizes (Roussel et al., 2017). Therefore, consistency in selection of sensor specifications and survey requirements **is recommended** as far as possible (see Section 3.2.2 for specific recommendations).

3.2.5 Comparison of EO products with airborne laser scanner reference maps

Assuming an ALS AGBD reference map has been produced following the recommendations above, including reporting of uncertainties (ideally per-pixel), it can be used for both training of EO AGBD models or validation of EO AGBD products. The same basic principles for training and validation using ALS AGBD maps apply:

1. ALS reference biomass maps need to be spatially linked to EO products.
2. Statistical modeling of the relationship between the ALS AGBD reference map and the EO product data will be conducted, through the development of predictive models (for training), evaluating the agreement between products, and for product validation.

It is critical to appreciate here that just as field plot estimates of AGBD do not represent true or direct estimates of AGBD (that would require plot level tree harvesting), ALS AGBD estimates will have even greater uncertainties since an additional modeling step is used in their generation. A key requirement here is that the ALS AGBD reference map is at least of greater quality than the EO map (GFOI, 2016, p. 125; Stehman, 2009). Thus, the uncertainty reporting recommendations for ALS AGBD reference products (described in 3.2.4.4) are particularly important.

3.2.5.1 EO map product requirements

ALS reference AGBD maps will typically need to be warped to the spatial resolution and map projection of the satellite product prior to validation, therefore **it is recommended** that the source ALS data are used to generate AGBD reference maps with the spatial resolution, map projection, and grid alignment that match each of the EO biomass map products. When assessing multiple satellite products, **it is recommended** to retain the EO biomass map products in their native map projection to avoid resampling, and therefore introducing another source of uncertainty, where possible.

Ideally, per-pixel uncertainties will be available for the EO map products, consistent with recommendations in Chapter 4, however these are rarely available for existing products and impractical to generate (McRoberts, Næsset, Liknes, et al., 2019). In any case, it is important to recognize that independent estimates of product uncertainty can be calculated using high-quality ALS AGBD estimates, that are still useful for identifying regions with large systematic deviations from truth, guide product selection when multiple EO data products are available for a region, and to prioritize regions for methodological improvement.

3.2.5.2 Recommendations for validation

Several studies have suggested or applied methods for validating EO AGBD products using ALS AGBD reference maps. However, at the time of this writing, there is no clear consensus or protocol to follow, and results from different studies are often difficult to compare, partially because different reference estimates of AGBD were used. As stated in Section 5.2, the IPCC (2003, 2006) guidelines recommend that uncertainty and stability of EO AGBD products are to be assessed and reported for relative and absolute systematic deviation and confidence interval or RMSE, overall and by biomass class/range estimated using reference data of better quality.

The typical metrics used for assessing the uncertainty and stability of EO AGBD products are systematic deviation (*SD*), relative systematic deviation (SD_R), *RMSE* and relative *RMSE*. Since the ALS biomass reference maps contain uncertainties, RMSE is referred to here as Root Mean Squared Deviation (RMSD). The relative metrics are expressed as a percentage and consider that uncertainty may be small relative to the AGBD estimated. One may calculate SD, SD_R, $RMSD$ and relative *RMSD* ($RMSD_R$) over a user defined spatial domain as:

$$SD = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)$$

$$SD_R = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{y}_i - y_i}{y_i}\right) \times 100$$

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}$$

$$RMSD_R = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\hat{y}_i - y_i}{y_i}\right)^2} \times 100$$

where $y_i$is the ALS reference estimate of AGBD and $\hat{y}_i$ is the EO map product estimate of AGBD. The calculation of these terms is simple and straightforward if you do not consider uncertainty in the ALS reference and EO biomass maps, however, **it is recommended** that care is taken to account for the effects of spatial autocorrelation (see Box 3.3 for specific recommendations). It's also important to note that under and over prediction at different biomass ranges may compensate each other in the calculation of SD and SD_R.

It is only possible to validate spaceborne estimates to the accuracy level of the ALS reference map (Duncanson et al., 2020; McRoberts, Næsset, Liknes, et al., 2019; McRoberts, Næsset, Saatchi, et al., 2019), however comparison of mean estimates are still useful to identify discrepancies in height and AGBD estimates that exist between different modeling approaches and underlying environmental conditions (Bolton et al., 2013; Duncanson et al., 2020; McRoberts, Næsset, Saatchi, et al., 2019; Joanne C. White et al., 2013). It is necessary that the ALS AGBD reference maps have absolute uncertainty that is small relative to the EO biomass products, therefore the ALS AGBD reference maps should be quality controlled and filtered such that only highly accurate products are used to estimate uncertainties.

There is no exact uncertainty requirement for ALS AGBD reference maps, and such a requirement would be undesirable as it will depend on spatial resolution and absolute value of AGBD. Presumably, all carefully constructed local ALS maps will be of higher quality than the EO products, but in particularly challenging areas (e.g. high biomass forests, areas with a dearth of training data, etc.), some local maps may have uncertainties approaching that of the EO product itself to be validated. In this case, the validation statistics would be meaningless. We suggest a comparison using SD, SD_R, $RMSD$ and $RMSD_R$ in EO biomass products over the domain of interest, as reported by the product authors, however it is necessary to account for the uncertainty of the ALS biomass reference maps in this process.

There are limited studies available that give insight into statistically rigorous methods to quantify the level of accuracy that can be determined by ALS biomass reference maps, with recent examples being McRoberts, Næsset, Liknes et al. (2019), McRoberts, Næsset, Saatchi et al. (2019) and Duncanson et al. (2020). McRoberts, Næsset, Saatchi et al. (2019) demonstrated an approach that took the form of a statistical hypothesis test that the mean AGBD estimates from the ALS reference and EO biomass map over a user defined spatial domain were not significantly different from zero. The approach used hybrid inference (see Chapter 4) to account for both sampling variability and non-negligible errors in the ALS biomass reference map, and made important recommendations on uncertainty reporting requirements for EO biomass maps.

In contrast to the approach of McRoberts, Næsset, Saatchi et al. (2019), Duncanson et al. (2020) compared confidence intervals between simulated EO AGBD estimates and reference ALS AGBD

estimates at the native resolution of each EO product (25 m - 100 m). In this study confidence intervals were available for both the ALS reference or EO biomass maps, which allowed large systematic deviations in the EO biomass map to be detected across environmental gradients, however only where the EO biomass estimates were outside the range of the ALS reference map confidence intervals. This highlighted the importance of using very high quality ALS AGBD estimates for validation of AGBD map products.

The generation of ALS AGBD reference maps and associated uncertainties at multiple spatial resolutions should account for differences in the spatial resolution of EO biomass products, such as those from GEDI, NISAR and BIOMASS, however their interpretation should be caveated by the spatial scale analyzed. It is well accepted that biomass model prediction errors reduce with increasing plot size or resolution (F. G. Hall et al., 2011; Labriere et al., 2018; Zolkos et al., 2013), largely because of the reduction in variance of reference AGBD estimates with increases in spatial scale (see Section 3.1). Calibration and validation results for EO biomass products at their native spatial resolution cannot be directly compared where different, therefore for EO biomass product intercomparison **it is recommended** to aggregate to the finest common spatial resolution possible.

Key to the application of ALS in validation of global EO products is consistency in the definitions, measurement methods, AGBD modeling techniques, and statistical inference frameworks used to characterize uncertainty. Given the diversity of ALS instruments and in situ protocols used in different biomes it is necessary to be adaptive, but care must be taken to ensure the integrity of core measurements (e.g. ALS canopy height and cover) are maintained over time as instrument specifications, survey configurations and data processing methods advance. Otherwise, validation efforts will continue to be one-offs, and ALS biomass models will not be spatially and temporally transferable or consistent.

We emphasize the need to minimize uncertainty in ALS reference maps as far as practical, but these uncertainty estimates need to be realistic. A tool for propagation of in situ and ALS measurement and model errors (see Chapters 2 and 4) to the resolution of EO AGBD products that is accessible and standardized is presently unavailable. Stratifying validation by land cover, geographic extent, environmental gradients (e.g. slope, canopy cover) and AGBD range/class will enable users to estimate the uncertainties of multiple available EO biomass products within their specific domain of interest. **It is therefore recommended** that a tool be developed with a user-friendly interface for comparison of local maps derived from ALS to global maps derived from EO at multiple spatial resolutions and by user defined strata.

Box 3.3

Assessment of model predictive performance

A central component of studies aiming to map any ecological variable, including AGB, is the validation of the mapping (or prediction) model. Indeed, relationships between wall-to-wall remote sensing data and AGB are complex (non-linear), weak (e.g. saturation) and often context-dependent (e.g. relationships climate-AGB). It is therefore crucial to provide a comprehensive, transparent and ideally reproducible assessment of model prediction error.

Model validation is commonly made using reference data that have been set aside at the model calibration stage to test model

predictive performance on "new" locations. While we recommend testing the model on "new" locations, to reflect the fact that most predicted pixels correspond to unsampled locations, care must be taken when defining the set of test data to ensure sufficient independence from model training data. Indeed, because of (i) the spatial autocorrelation in AGB and model predictors and (ii) the spatial proximity between reference data samples, a random split of the reference dataset into training and test sets may not result in two independent sets.

Failing to account for data spatial autocorrelation when validating the mapping model, and hence for potential dependencies between model training and test data sets, may result in a large overestimation of model predictive performance. This issue has been illustrated in Ploton, Mortier, Réjou-Méchain, et al. (2020), who used a machine-learning algorithm (Random Forest) to map forest AGB based on environmental and MODIS reflectance variables. When randomly splitting reference data into training and test sets and hence ignoring any potential spatial autocorrelation in the data, model validation statistics suggested that the model explained about half the variance in forest AGB. However, accounting for data spatial autocorrelation in model validation revealed that model predictive performance sharply decreased as the distance between training and test data increased (Fig. 3.5). For instance, model predictive power was down to c. 15 % when using a buffer of 50 km between validation data and the nearest training data point.

The effect of data spatial autocorrelation on model validation statistics is study-dependent, in that it will vary with the nature of the data (e.g. the ranges of spatial autocorrelation in AGB and auxiliary data, the number of auxiliary variables), the type of mapping model, the mapping spatial resolution, the spatial layout of reference AGB data, etc.

We therefore recommend that map producers:

- Assess and report on the range of spatial autocorrelation in forest AGB at their mapping resolution
- Perform a geographic validation of their mapping model, using e.g., a spatial k-fold cross-validation (Ploton, Mortier, Réjou-Méchain, et al., 2020) with a spatial cluster size substantially larger than the range of spatial autocorrelation in forest AGB (Ploton, Mortier, Réjou-Méchain, et al., 2020; D. R. Roberts et al., 2017)
- Provide validation statistics of a purely spatial model (e.g. a simple spatial kriging of reference data) as benchmark to help the interpretation of mapping model validation statistics and the assessment of their reliability
- Publish, together with the AGB map, the data used to train and validate the mapping model, to allow for independent examination of model predictive performance.
- If the mapping model is parametric, publish the model parameters and their variance-covariance matrix to allow for formal estimation and propagation of uncertainty (see Chapter 4).

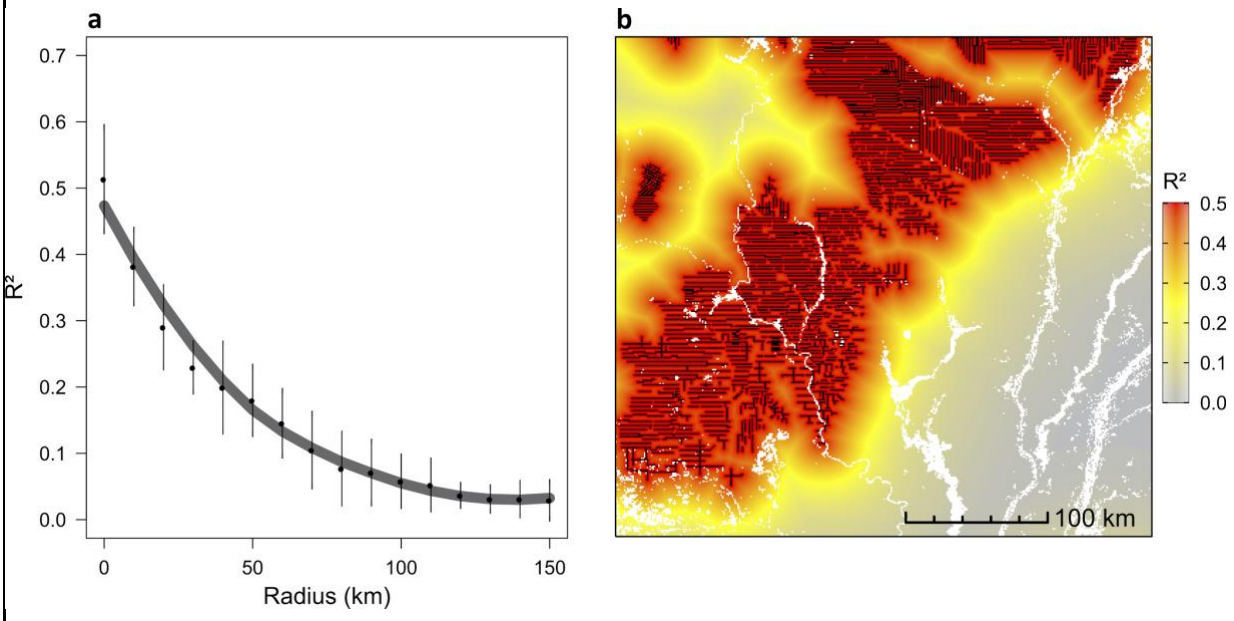

Fig 3.5 Model validation with buffered leave-one-out cross-validation (B-LOO CV). a. Change in the coefficient of determination (mean $R^2 \pm SD$) between predicted and observed pixel AGB as buffer radii for neighboring pixel exclusion increases in the B-LOO CV (see Ploton, Mortier, Réjou-Méchain, et al., 2020 for details). **b.** Projection of model prediction R^2 in the study area (with training pixels represented in black).

3.2.6 Summary and current knowledge gaps

Large and small footprint ALS have long been established as tools for near direct estimation of canopy height and cover, and subsequently through empirical modeling of AGBD. Both capture landscape scale variation in structure, providing a bridge between EO and local scale observations (UAV-LS, TLS, field). Large footprint lidar systems (e.g. NASA LVIS) enable rapid and large area wall-to-wall or transect acquisitions (e.g. entire countries in some cases) but are of limited availability, however small footprint ALS systems are now capable of this scale of acquisition as well.

A key issue for global validation of EO products are practical and logistical constraints on ALS data acquisition. Government restrictions on ALS data availability and flight authorization (e.g. some countries in continental Asia) can make access to existing data or acquisition of new data by international practitioners difficult. Depending on transit distance and areal coverage required, ALS data acquisitions can be expensive, with estimates quoted on the order of \$250 – \$600 km^2 (in US dollars) for small footprint ALS (Réjou-Méchain et al., 2019). However, it is the experience of the authors that costs are often a fraction of this, depending on in-country restrictions and instrument availability. Non-commercial instruments such as NASA facility instruments can acquire much larger areas than current commercial small footprint ALS systems for the same cost, however they are not as readily accessible and may incur large transit costs.

The following are recommendations when producing a reference ALS AGBD map for EO biomass map validation. These are in addition to recommendations on linking field plot data with ALS data that are aimed at minimizing the impact of geolocation error and plot size (Section 3.2).

1. Use recent, high-quality, well-vetted lidar instruments and survey configurations
    - Photon counting instruments are not yet mature enough for routine biomass estimation or tested across a wide range of ecosystems globally. They are not currently recommended, but show promise for cost-effective and large area acquisitions.
    - Small footprint waveform and discrete return instruments are both useful, provided sufficient point density (≥ 4 pulses m^{-2} in dense forest), small footprint size (< 30 cm), within a maximum off-nadir scan angle (< 15°), and preferability no deadtime between returns (≥ 4 returns per pulse)
    - Adhere to QA/QC protocols and data delivery specifications recommended by existing national programs - including TERN Australia, USGS 3DEP, and NSF NEON - for forest applications
2. Acquire wall-to-wall data that are spatially and temporally coincident with field plots
    - Transects still fill an important niche for large area sampling and validation of current spaceborne lidar observations
    - Wall-to-wall data are recommended coincident with field plot data over a large enough area (minimum 3 × 3 km, preferably larger) to capture landscape scale variation in canopy structure and AGBD and enable scaling from fine to coarse spatial resolution of EO biomass products

    - Acquisition should take place during the same season as field collection (preferably leaf-on in deciduous forests)
    - ALS acquisition should be as close in time as practicable to field acquisitions
        - In snow-free and maximum leaf-on conditions (where deciduous species are present)
        - Before plot acquisition where possible to aid in plot placement through stratified sampling by forest structure measurements
        - Ideally within one year of plot measurement to minimize change (although a longer time gap is acceptable in slow growing forests with no detected disturbance)
3. Develop local AGBD models using an area-based approach
    - Estimation of AGBD from ALS is indirect, therefore model calibration requires that *in situ* reference data are sampled across the full range of AGBD and the ALS metrics used as predictors in a single model
    - The use of a stratified sampling approach is recommended to achieve this sampling for a single model, preferably with strata that are defined using the ALS structure metrics themselves
    - Individual tree-based AGBD approaches currently introduce additional uncertainty and are less mature and generalizable than area-based approaches
    - Adopt simple parametric models that minimize the number of height, cover, and vertical structure metrics used, particularly when a limited number of field plots are available
    - Evaluate model performance using geographic cross-validation (e.g. Ploton, Mortier, Réjou-Méchain, et al., 2020)
4. Adopt a statistically rigorous framework for reference map generation
    - Ideally, use a formal statistical inference framework (see Chapter 4 for details) to propagate errors from allometric (see Chapter 2 for details) and ALS models
    - If possible, develop ALS biomass reference maps at the resolution of EO data products to simplify error propagation
    - Provide per pixel estimates of uncertainty in the form of mean squared errors to accompany the per-pixel mean estimates of AGBD
    - Clearly report the error sources that are included or omitted in any error propagation document that the process followed
    - Report the ground-to-ALS AGBD model used in full, including parameter estimates and their accompanying variance-covariance matrix

We have identified the following gaps in tools to enable systematic validation with a global set of linked field plots and ALS AGBD reference maps:

1. A quality control framework for AGBD map production that includes:
    - improved uncertainty estimates at the pixel level that account for multiple sources of error (see Chapter 4)
    - methods to aggregate AGBD uncertainties from ALS pixel level to the spatial resolution of EO map products

    - evaluation of ALS metric repeatability and biomass model performance when applied to ALS datasets from new instruments or acquisitions that are not coincident with field plots
2. A tool for propagation of *in situ* and ALS measurement and model errors (see Chapters 2 and 4) to the resolution of EO AGBD products that is accessible and standardized
3. A tool for automated comparison of local maps derived from ALS to global maps derived from EO at multiple spatial resolutions and by user defined strata

Chapter 4: Characterization and Propagation of Error

Stephen Roxburgh and Ronald E. McRoberts

Quantifying the uncertainty of estimates of forest aboveground (dry) biomass (AGB) is necessary for understanding and mitigating the impacts of climate change across a range of spatial domains, from individual forest plots or stands, through to regional, national or even global scales (GFOI, 2016). This chapter is concerned with the steps required to estimate the accuracy of map-based biomass estimates, inclusive of errors associated with reference data such as plot-level biomass (Chapter 2), and errors associated with the models that are used to combine that reference data with remote sensing products to construct maps of AGB (Chapter 3). Here we define reference data as the fundamental data from which statistical inference is made, typically plot-level AGB from forest inventories. Reference data are often assumed to be error free, but here we also consider the case of 'imperfect' reference data, that are themselves subject to uncertainty, which is always the case for forest AGB. This is an important distinction, as the nature of the reference data determines the statistical inference methods that can be used, with associated differences in the estimation of biomass and its uncertainty.

Error propagation is the term given to the identification and appropriate accounting for the effects of all relevant error sources contributing to an estimated quantity, with smaller levels of error associated with both greater accuracy (lack of systematic error) and greater precision (Figure 4.1). Estimating the bias of an estimator can be problematic in practice, as it requires knowledge of the 'true' underlying value (Figure 4.1), which in many cases is unknown. Sources of error that contribute to bias in AGB estimators that are amenable to estimation are instrument error, such as incorrectly calibrated weight scales that consistently under- or over report the mass of harvested tree components when developing allometric models (Section 4.1.1), and model predictions, with systematic error arising from incomplete or inappropriate model specification (Section 4.2.1).

The second component of uncertainty, precision, quantifies the spread around a given estimate, and is usually the primary focus of error propagation studies. It is useful to recognise that precision includes two broad components of uncertainty. The first is uncertainty arising from natural variability, such as genotypic or phenotypic variability in the size of individuals of the same age and species, or spatial variability in forest biomass due to e.g. edaphic factors or topography. The second source is often called random error, with the implication that it can potentially be reduced by steps such as improved measurement protocols, more efficient sampling strategies, or increased sample sizes. Uncertainty arising from variability and uncertainty arising from measurement errors can be difficult to disentangle, and thus they are often treated together. A wide range of analytical and Monte Carlo techniques have been applied to forest biomass error propagation (e.g. Breidenbach et al., 2014; I. F. Brown et al., 1995; Case & Hall, 2008; Chave et al., 2004; Chen, 2015; Chen et al., 2016; Chen, Laurin, et al., 2015; Cohen

et al., 2013; Gertner & Köhl, 1992; Holdaway et al., 2014; Keller et al., 2001; Lehtonen et al., 2007; Magalhães & Seifert, 2015a; Magnussen et al., 2014; Mavouroulou et al., 2014; McRoberts et al., 2015; McRoberts & Westfall, 2016; Melson et al., 2011; Molto et al., 2013; Ngomanda et al., 2014; D. L. Phillips et al., 2000; Picard et al., 2015, 2016; Réjou-Méchain et al., 2017; 2019; Yanai et al., 2010). With regard to international greenhouse gas accounting, the IPCC specifies two good practice guidelines for greenhouse gas (GHG) inventories: (i) "neither over- nor underestimates so far as can be judged" and (ii) "uncertainties are reduced as far as practicable" (GFOI, 2016, p. 15). A primary assumption underlying the second guideline is that uncertainty must be rigorously estimated before it can be reduced. A consequence when developing new or using existing biomass maps is that any estimates based on those maps are currently required to comply with these guidelines. In general terms, this means ensuring that estimates are either obtained or evaluated using unbiased estimators (satisfying criteria (i)), that all of the key error sources are recognised and accounted for, and that the methods applied are rigorous, appropriate, and quantifiable (satisfying criterion (ii)). In particular, the aim is to construct inferences in the form of confidence intervals for the estimates, for either whole-map biomass or for sub-map areas or specific locations. In a formal statistical sense, "inference" here means expressing a result in probabilistic terms, i.e., a 95% confidence interval in the form $\hat{\mu} \pm t_{0.95} \cdot SE(\hat{\mu})$, where $\hat{\mu}$ is the estimate of a parameter such as mean biomass, and $SE(\hat{\mu}) = \sqrt{\widehat{Var}(\hat{\mu})}$ if the estimator is unbiased (Section 4.2.1), or $SE(\hat{\mu}) = \sqrt{\widehat{MSE}(\hat{\mu})}$ if the estimator is biased (Section 4.2.2) (McRoberts, Næsset, Liknes, et al., 2019).

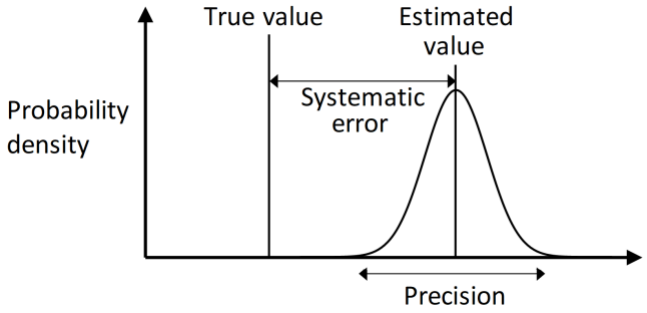

Figure 4.1 The accuracy of an estimate comprises two components. (i) Precision quantifies the variability around the estimated value. (ii) Systematic error (=bias) quantifies the degree to which the estimated value deviates from the true (but often unknown) population value.

Three broad forms of inference have been applied to estimate mean biomass ($\hat{\mu}$) and its uncertainty ($SE(\hat{\mu})$) (McRoberts, Næsset, Saatchi, et al., 2019). The first, design-based, requires a probability sample of reference data that is subject only to negligible uncertainty. The most straightforward application of design-based inference is where sample units are located independently and randomly in space, and where $\hat{\mu}$ and $\widehat{Var}(\hat{\mu})$ are calculated using the traditional simple expansion estimator (EXP) (also known as the simple random sampling (SRS) estimator). This, and other sampling designs such as systematic sampling and stratified sampling and their associated estimators, are discussed in Section 4.2.1. The main limitation when using the design-based expansion and stratified estimators for total-area biomass (or total-area

biomass density, i.e. biomass expressed on a per-unit area basis) is that the estimators do not produce spatially continuous maps of AGB, i.e. with sub-area resolution. To overcome this limitation, the model-assisted difference and regression estimators (Särndal et al. 1992; McRoberts, 2010; McRoberts et al., 2013) have been employed, whereby the probability sample data are combined via a model with auxiliary information such as remotely sensed data, from which AGB maps are constructed as by-products. Note that model-assisted estimators are still considered part of the design-based paradigm, because their validity rests on the underlying probability-based sample, rather than the validity of the model linking the probability samples with the auxiliary data. Model-assisted estimators are also discussed in Section 4.2.1.

The second form of inference that can be applied is model-based, which can use either probability or non-probability samples of the reference data, and when combined with spatial auxiliary data (such as obtained from remote sensing) can also produce maps of biomass. In contrast to design-based inference, the validity of the model-based estimator rests on the correctness of the model, rather than the sampling design (Section 4.2.2). Design-based and model-based estimators represent fundamentally different inferential paradigms, with different statistical procedures for estimating uncertainty. One similarity between them, however, is that both require that the reference sample data are characterized by biomass values that are error-free, or that are at least characterized by negligible uncertainty.

Almost always, however, plot-level reference data are subject to non-negligible uncertainty, for example when plot-level biomass is predicted using allometric models (Section 2.1). Although allometric models are often assumed to predict plot-level biomass with large accuracies, numerous studies have shown that the uncertainty of such biomass predictions may be non-negligible (Breidenbach et al., 2014; Chave et al., 2004; Réjou-Méchain et al., 2017; Ståhl et al., 2014). To account for such 'imperfect' reference data, hybrid inference (the third form of inference considered here) combines elements of design-based and model-based inference, and can be applied to AGB estimation and mapping to fully comply with the IPCC criteria. Hybrid inference is discussed in Section 4.2.4.

The chapter is structured into two sections. The first (Section 4.1) provides an overview and brief description of the sources of uncertainty that must be considered when estimating large area biomass parameters. The underlying assumption is that constructing a biomass map involves (1) the collection/assimilation of field data (typically plot-based inventory data) and other information for predicting stand-level biomass from allometric models - referred to as the reference data; and (2) the use of a model that combines auxiliary spatial data (typically derived from remote sensing) with the reference data to construct a biomass map. Uncertainties are present at both steps (1) and (2). Section 4.2 describes in greater detail the approaches that can be used to propagate those uncertainties throughout the estimation chain.

4.1 Sources of uncertainty

The effects of a wide range of uncertainties associated with the construction of biomass maps and the estimation of biomass parameters are manifest as either systematic error or reduced

precision (Tables 4.1a & 4.1b). Systematic errors can occur when making measurements, such as weighing whole trees or tree sections with instruments that have not been properly calibrated, or poor field techniques leading to consistent over- or under-estimation of quantities such as stem diameter during forest inventory. When detected, it is preferable to correct the affected data for systematic errors prior to further analysis, thus minimising the risks of measurement-based, systematic errors propagating through the estimation chain. Tables 4.1a & 4.1b therefore focus on sources of precision uncertainty, with the assumption that any systematic measurement error issues have been resolved.

This section describes the various sources of uncertainty that need to be considered, but does not provide specific details for the estimation and propagation of those uncertainties, which can be achieved by a range of methods such as analytical approximations, Monte Carlo methods, or a combination of the two. In Section 4.2 a generic Monte Carlo procedure is provided, in the context of hybrid inference, to illustrate how the various sources of uncertainty can be propagated through to the final map product and/or estimate. There are two main estimation steps (described in Sections 4.1.1 and 4.1.2), with uncertainties arising at several points.

4.1.1 Uncertainties in reference data

Reference data typically comprise values of biomass for local scales, such as forest inventory sample plots. When allometric models are used to predict the biomass for a sample plot there are multiple potential sources of uncertainty. The first are the errors associated with the development of the allometric model itself, which includes errors associated with measurements of the dependent variable (usually either individual tree dry mass, or individual tree volume; Table 4.1a, source A.1), measurements of the independent variables (e.g. tree DBH, height, wood density; Table 4.1a, source A.2), variances of the model parameter estimates (Table 4.1a, source A.3) and residual variance around model predictions. Of these, variances of allometric model parameter estimates have been most studied, with the key quantity for error propagation being the covariance matrix for the model parameter estimates (Cohen et al., 2013; McRoberts & Westfall, 2016). In general, increasing the sample size on which the allometric model is based is a sound strategy, as larger sample sizes increase the precision of the model parameter estimates (Chave et al., 2004; Roxburgh et al., 2015; van Breugel et al., 2011), noting that care must be taken to ensure the sampled trees used to construct the allometric model are representative of the broader population, and not, for example, all from the same stand or plot.

One overlooked and potentially important source of uncertainty during allometric model development involves errors that simultaneously affect multiple individuals, thus leading to non-negligible covariance, and increased uncertainty (Ståhl et al., 2014; Yanai et al., 2010). For example, when predicting dry biomass, moisture content correction factors to convert fresh mass to dry mass are usually based on oven-drying samples collected from a small subset of the population, but are applied to the population as a whole (Paul, Roxburgh, et al., 2017). Thus, any error in the moisture content correction factor will be simultaneously applied to multiple individuals. A Monte Carlo procedure for accounting for the errors associated with allometric model development is provided in McRoberts et al. (2016).

Table 4.1a Summary of sources of uncertainty when estimating plot-scale aboveground biomass through developing and applying allometric models to data collected from inventory sample plots.

Calculation step	Source of uncertainty	Brief description
A. Allometric model development	A.1. Dependent variables (volume, or dry mass; where dry mass = fresh mass x moisture content correction).	Allometric models predict either per-tree volume, or per-tree dry biomass. For either dependent variable there may be errors associated with the required measurements.
	A.2. Independent variables, e.g. stem diameter, height.	Similar to A1, there may be measurement errors associated with the independent variables that are used as biomass (or volume) predictors. Measurement errors may be either systematic errors that affect precision through such as random measurement errors, or natural variability.
	A.3. Allometric model parameter uncertainty.	Uncertainty arising from model parameter estimation as part of the model fitting procedure.
B. Allometric model use	B.1. Independent variables, e.g. stem diameter, height.	As per A.2., there may be errors in the measurement of the model independent variables. Forest inventory typically involves the measurement of 100s-1000s of individuals. Although large sample sizes can reduce random sampling error, often there will be natural variability that cannot be reduced by increased sampling effort.
	B.2. Allometric model prediction (residual) error.	Even if term A.3. can be minimised through e.g. increased sample sizes on which the allometric model is based, there is often large residual variation around the estimated relationship, simply because of natural variability (e.g. trees with the same stem diameter can have widely differing biomass).
	B.3. Conversion of volume to biomass (and biomass to carbon).	Volume estimates need to be converted into dry mass, typically by multiplying by wood basic density, which is variable within-individuals, within-species, and among species. Similarly, if biomass needs to be expressed in units of carbon, then this requires multiplying by tissue carbon concentration, which is usually treated as a constant (typically in the range 0.45 – 0.55) that may also be uncertain.
	B.4. Alternative allometric model choice error.	In some circumstances, especially when using multi-species (or generalised) models, there may be a choice of more than one model for a given situation, with often differing predictions between models.
C. Sample plot area	C.1. Plot area	For each sample plot the summation of individual tree biomass predictions and the propagation of the uncertainties A.1. – B.4. yields plot-level estimates of total biomass and uncertainty. To convert this to per-unit-area requires division by the plot area, which itself may be uncertain through e.g. errors in locating plot boundaries.

Table 4.1b. Summary of sources of uncertainty when estimating large-area or mapped biomass

Calculation step	Source of uncertainty	Description
D. Sampling variability	D.1. Spatial variability in biomass.	With design-based inference, sample plots can be spatially distributed using any of multiple probabilistic sampling designs including e.g. simple random, stratified with simple random sampling within strata, or using schemes that seek to distribute plots more uniformly (e.g. systematic sampling, restricted random sampling). Because forests are spatially variable (at all scales) and the sampling intensity is typically low relative to the total areal extent of the population, sampling variability is often an important source of uncertainty.
E. 'Predictive biomass model', to predict biomass from auxiliary data.	E.1. Independent variable errors	Statistical models can be used to predict measured biomass (the reference data) from auxiliary data, which typically comprise spatial layers such as raster data, with pixel-to-pixel variability. E.1 captures errors in the independent variables (i.e. the auxiliary data) of the predictive biomass model (e.g. remote-sensing-derived metrics; spatially interpolated climate data). There may also be location/georeferencing errors in the auxiliary data.
	E.2. Predictive biomass model parameter uncertainty	In the same way that there is uncertainty in estimated parameters of the allometric model (A.3), there will also be parameter uncertainty in the predictive biomass model.
	E.3. Predictive biomass model prediction (residual) error.	As with the allometric model prediction error (B.2), the biomass model makes predictions only imperfectly, which is quantified by the model residual error in combination with the uncertainty of the model parameter estimates.
F. Total (regional) biomass	F.1. Uncertainty in mapped total spatial extent	To convert estimates of large-area carbon density to total carbon stock requires multiplying by the regional area (e.g. the area within which the reference data plots were located). If there is uncertainty in the total area (e.g. through mapping errors of boundaries) this should also be included in the conversion. When areas are calculated from the summation over pixels/map units from an existing mapping product, then this error is likely negligible.

The next group of uncertainty sources are those associated with application of the allometric models for biomass prediction. The most familiar and most dominant of these are parameter estimate covariances and residual variance (Table 4.1a, source B.1). These sources of uncertainty characterize all studies of the propagation of errors using allometric models (Réjou-Méchain et al., 2017). Increased allometric model sample sizes can help reduce the covariances of model parameter estimates, but there will always be natural variability in the mass of individual trees of the same diameter that manifests as residual variability around model predictions. The effects of this latter source cannot be reduced by increased sampling effort and should therefore be included in any analysis.

Measurement errors are associated with the independent variables that are used to predict individual biomass, such as DBH and tree height (Section 2.1, Table 4.1a, source B.2). Field experiments have suggested minimal systematic error in diameter measurements, but with average precision (expressed as coefficient of variation) in the range of 3-7% (Holdaway et al., 2014; Paul, Larmour, et al., 2017). Errors in the values of independent variables of a simple linear

regression can cause 'dilution bias' (Réjou-Méchain et al., 2014), which reduces the estimate of the regression slope; for nonlinear or more complex models, such as random forest, the extent of this bias is difficult to predict and may require Monte Carlo or other randomisation techniques to assess.

If the allometric model predicts total tree volume, then conversion to biomass requires an estimate of the species wood basic density, which is usually imperfectly known given that wood density varies within individuals, within species, and among species (Table 4.1a, source B.3). Although the focus is the estimation of biomass, greenhouse gas accounting usually requires conversion of AGB to carbon (or carbon-dioxide equivalents). The conversion from dry mass to carbon is usually accomplished by multiplying dry mass by approximately 0.5, under the assumption the mass fraction of carbon in dry wood is constant. However, Martin et al. (2018) in a global review have suggested the carbon fraction is more variable than often assumed, and they provide data that could be used to assign uncertainty values to this estimate (Chave et al., 2009 also reviewed this).

The final source of error associated with the use of allometric models to predict biomass is due the mathematical form of the selected allometric model, when multiple competing allometric model forms are available for a given region (Table 4.1a, source B.4). Studies that have investigated the uncertainty that arises due to allometric model selection have shown this to be a potentially significant source of error (e.g. Chave et al., 2004; Mavouroulou et al., 2014; Melson et al., 2011; Picard et al., 2015, 2016). Uncertainty due to allometric model selection can be estimated by repeating the biomass estimation using each alternative model in turn, or if multiple model choices are available across a number of species, by selecting models at random within a Monte Carlo analysis. A more sophisticated approach weights the contribution of each model relative to the likelihood that it best represents the underlying data, using the method of Bayesian Model Averaging (BMA) (Picard et al., 2015). A comprehensive description of BMA in the context of biomass prediction is given by Mavouroulou et al. (2014), including R code listings for conducting the analyses.

The sources of error considered thus far impact the biomass prediction of an individual shrub or tree, and its uncertainty. Aggregation of the individuals within the sample plot simply requires summing each individual's predicted biomass, and for the error propagation, summing each individual's error variance, to yield the total biomass variance for the plot. During this process, the effects of correlation among observations and predictions for trees of the same plot are typically ignored. For most mapping purposes the biomass density of the plot is required (e.g. t AGB ha^{-1}), rather than the total biomass (e.g. t AGB). This requires dividing the total biomass estimate by the plot area, which itself may be an uncertain quantity, as marking out plots can lead to errors associated with plot boundaries being obscured by vegetation or topography, or GPS location errors (Table 4.1a, source C.1). Plot area errors can also arise due to topography, when plot boundaries are projected horizontally (Gertner & Köhl, 1992).

4.1.2 Uncertainties in constructing maps of biomass and estimating large area biomass parameters

A statistical estimator must be applied to the reference data (from Section 4.1.1) to estimate biomass and its uncertainty for the region from which the sample plots were located to construct maps of biomass. As noted in the introduction, three broad inference methods can be applied: design-based, model-based, and hybrid. Design-based inference includes a wide range of estimators, such as the EXP estimator for a simple random sample (Section 4.2.1), however only the design-based model-assisted estimators offer the possibility of constructing biomass maps as by-products of the estimation procedure with sub-area resolution. Similarly, model-based and hybrid inference also produce biomass maps, with hybrid inference combining elements from both the design-based and model-based methods. The appropriate inference method to use for a given situation depends upon the nature of the reference data and the details of how they were obtained, specifically, whether the sample was probability-based (e.g. as part of a designed sampling program) or non-probability-based (such as opportunistic or non-randomly located sample plots); whether the reference data can be considered to have non-negligible uncertainty or not; and whether suitable auxiliary data are available to be combined with the reference data to facilitate construction of a biomass map via a spatial predictive biomass model. Design-based, model-based and hybrid inference methods are described in greater detail in the next section, together with an explanation of the different assumptions underlying estimation of both biomass and its uncertainty. A decision-tree to guide users on the most appropriate inference method to use based on available information is provided in Figure 4.2.

The primary source of uncertainty associated with a design-based inference is sampling variability (Table 4.1b, source D.1). Sampling variability arises because, for most populations of interest, only a relatively small sample of individuals can be measured: thus, a hypothetical experiment with repeated sampling of the same population under the same randomisation scheme, but varying only the plot locations, will yield variable results, simply because each sample experiment will, by chance, tend to include a different set of individuals. The key contributors to sampling variability are: the spatial extent of the target population; the spatial variability in biomass over the population; the plot configuration (the plot components, size and shape); and the sampling design which includes the total number of plots and the method used to distribute plots in space (e.g. simple random, or other sampling methods designed specifically to provide a more representative coverage). Because spatial variability is ubiquitous and the sampling intensity is typically small relative to the total areal extent of the population, sampling variability is almost always a non-negligible source of uncertainty, and is often the dominant source.

Whilst biomass maps with sufficient resolution cannot be constructed from the reference data alone, the reference data can be combined with auxiliary data to construct maps of biomass and biomass uncertainty; this forms the basis for the design-based model-assisted regression, model-based and hybrid estimators (Sections 4.2.1 - 4.2.3).

Combining spatial auxiliary data, such as those derived from remote sensing sources, with the reference data within a predictive model provides the mechanism for constructing maps of

biomass and their uncertainty. Many types of predictive techniques have been used, such as linear and nonlinear regression models (e.g. McRoberts et al., 2016), or non-parametric machine learning algorithms such as random forests (e.g. Roxburgh et al., 2019). Similarly, many different types of spatial auxiliary data can be used as predictor variables, including mapped climate data, topographic data, reflectance data from a range of passive aerial and satellite sensors, and data from active sensors such as lidar and radar. Whilst there are many choices for the form of the model and for the choice of predictor variables, the basic goal is the same: to use the model and the auxiliary data to predict biomass for map units by training (= calibrating) the model using the available reference data. Regardless of the complexity of the modelling approach adopted or the range and type of auxiliary data used, three sources of uncertainty must be considered during the development of the predictive biomass model (Table 4.1b, sources E.1 – E.3), in addition to any uncertainty in the reference data used as the model dependent variable (Section 4.1.1).

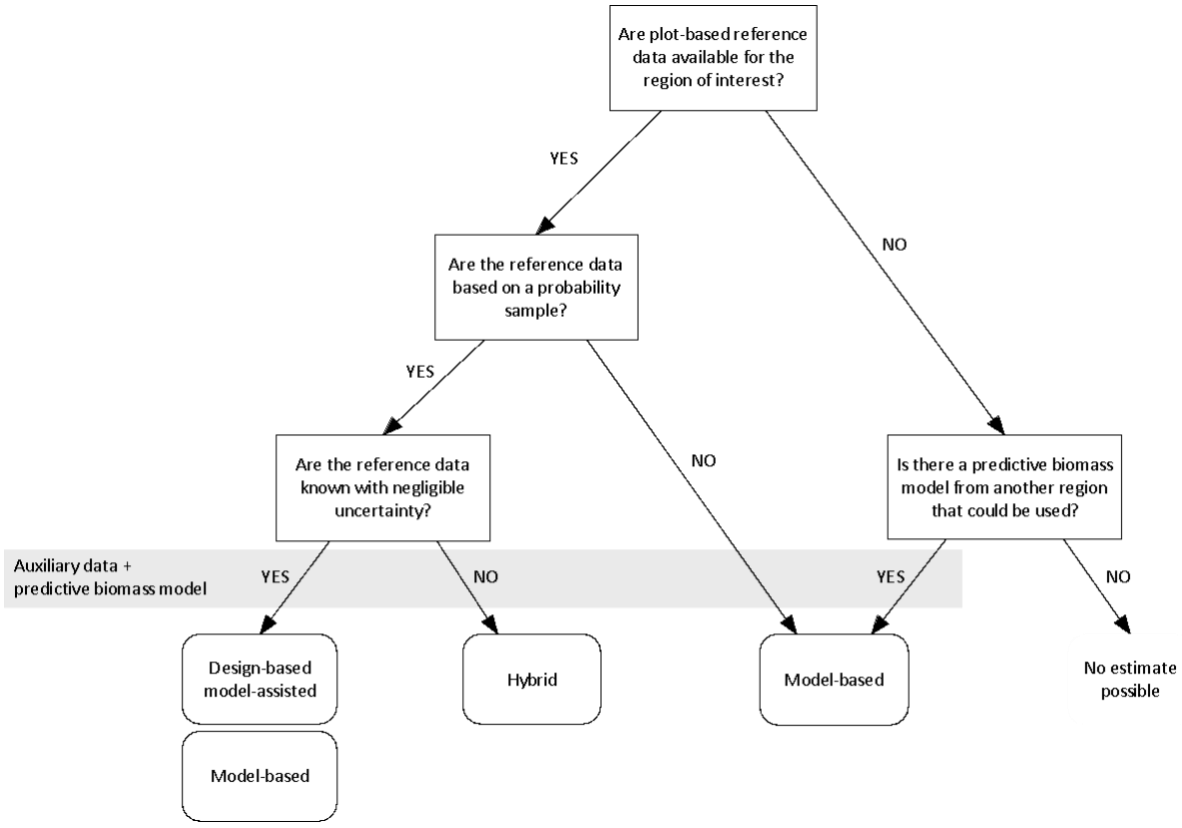

Figure 4.2 Decision tree to determine the most appropriate inference method for estimating biomass and its uncertainty. The grey bar denotes the point at which auxiliary data (e.g. remotely sensed data) are combined with the reference data to create a predictive biomass model that can then be used to predict sub-regional (e.g. pixel-level) biomass. Note that model-based methods can utilise either probability or non-probability sampled reference data. The inference methods (Model-based, Hybrid, Design-based model-assisted and Model-based are described in Section 4.2.

The first source of uncertainty is the auxiliary data themselves, such as the EO-derived metrics or other spatial information (Table 4.1b, source E.1). Whilst a wide range of remote sensing-based products can be related to biomass, such as spectral signals from optical sensors, or tree heights

predicted from lidar or SAR, these data may themselves be uncertain. Potential sources of error include those of registration and alignment (Saarela et al. 2016), uncertainties in the data products themselves, for example, in the prediction of tree height from lidar returns due to spatial variability in the returned points (in three dimensions), and errors associated with determining the ground level (GFOI, 2016).

The second and third sources of uncertainty arise because the relationships between the reference data and the auxiliary data are imperfect, and thus in the context of parametric regression, contribute to uncertainty in the model parameter estimates (Table 4.1b, source E.2) and model prediction errors (Table 4.1b, source E.3). These are analogous to sources A.3 and A.4 in Table 4.1a that were associated with the development of the allometric model. When using non-parametric modelling, the effects errors from these sources may need to be estimated via randomisation or Monte Carlo techniques.

The final estimation step involves estimating total biomass (or biomass density) of the region as a whole. If the reference data are in units of biomass density (i.e. t AGB ha^{-1}), then the absolute total biomass for the region requires multiplying the regional-scale mean biomass density by the regional area (ha). In cases where auxiliary data are derived from other mapping products, the total areal extent may be known with greater accuracy, such as summing over pixels in a raster image, it is then possible this uncertainty source could be assumed to be negligible. However, if the total area estimation was based on ground surveys, or manually translated from e.g. aerial photography or printed maps, then this error source may be non-negligible, and should be included in the estimation procedure. This can be addressed in an analogous way to the handling of individual plot area errors described in relation to error source C.1 in Table 4.1a.

4.1.3 Assessing the importance of error sources

Given the wide range of sources potentially contributing to total uncertainty, it is natural to ask which sources have negligible effects and could be excluded from analysis and which sources have non-negligible effects and should be included. Answering this question is problematic as it depends upon the information that is available (e.g. whether there are sufficient statistics describing the auxiliary data to allow model-based uncertainty estimates to be made, Section 4.2.2), or whether uncertainty information obtained from other studies could be used (such as applying field-based measurement errors obtained in one forest type or region to the error analysis of another; Figure 4.2). It also depends on whether the negligible/non-negligible effects of sources of uncertainty are being assessed in either absolute or relative terms. In particular, the effects of allometric model prediction uncertainty (source B.2) can be considered non-negligible in absolute terms whenever they are non-zero. However, even if they are non-negligible in absolute terms, they may still be negligible relative to other sources in application, when assessed against some standard. For example, the effects of allometric model prediction uncertainty have on occasion been found negligible relative to the effects of sampling variability (Chave et al., 2004; McRoberts et al., 2015); uncertainty sources may also be negligible relative to the biomass quantity being estimated, which can be quantified as the coefficient of variation (i.e. the square root of the error variance divided by the mean biomass).

The IPCC (2003, 2006) provides some guidance here, making it clear that setting hard requirements for accuracy (such as requiring estimates to be within ±10% of the mean 95% of the time) is neither practical nor desirable, and that the purpose of undertaking uncertainty analyses is not to dispute the validity of the estimates, but to help prioritise efforts to improve the accuracy and precision of future inventories, and to guide decisions on methodological choices (IPCC, 2000, p. 6.5). Within this context, and to be consistent with the IPCC guidance of providing neither over- nor underestimates, as can be judged with uncertainties that are reduced as far as practicable, the practical approach is to first identify the appropriate inference framework (Figure 4.2), and then to identify all error sources associated with the available data and models (Section 4.1), and to be as comprehensive as possible in including them within the error propagation analysis (Section 4.2).

4.2 Propagating errors from measurements to maps to estimates

A summary of the inference methods that can be used to combine reference data and spatial auxiliary data to produce biomass maps and estimates and their associated uncertainties is provided in Table 4.2. The first two methods (design-based, model-based) assume reference data (e.g. plot-level biomass) have negligible uncertainty, which may or may not be the case when biomass is predicted by allometric models (McRoberts & Westfall, 2016; Réjou-Méchain et al., 2017 and Section 4.1). The third (hybrid inference) combines elements from both design-based and model-based inference and can accommodate 'imperfect' reference data with non-negligible uncertainty.

Table 4.2. Summary of three inference methods for biomass estimation.

Inference method	Reference data	Estimator	Notes
	E.g. plot-scale biomass, with plots distributed within a finite region.	A procedure that uses reference data to infer the biomass of the full region.	
1. Design-based inference	A probability sample, with values observed with negligible uncertainty.	Statistical inference using e.g. simple expansion (EXP) estimators for mean biomass and its variance (reflecting sampling variability; Source D.1, Table 4.1b).	Can provide a biomass estimate of the region total and its uncertainty but cannot construct a map without auxiliary data.
		Design-based model-assisted estimators. These estimators use auxiliary data, typically in the form of EO-derived spatial layers, to refine the mean biomass and variance estimates from 1.	This method can construct a biomass map with sub-area resolution, by using the auxiliary data to predict biomass for all map units.
2. Model-based estimator	Either a probability or non-probability sample, with values observed with negligible uncertainty.	Estimation using a model that combines the reference data with auxiliary information.	Can generate a biomass map. In common with 1 the reference data is used to calibrate a model for predicting biomass, but with the mean biomass and mean squared error (MSE)[1] based on attributes of the model. For this estimator, the validity of the estimates are based on the validity of the model, rather than the validity of the probability-based reference data sample.
3. Hybrid (of design-based & model-based estimators).	A probability sample, with values observed with non-negligible uncertainty. Non-negligible uncertainty typically arises from the use of allometric models to predict plot biomass.	Estimation using a model that combines the reference data with auxiliary information, in addition to the sampling error derived from the probability sample. Overall uncertainty is a combination of model prediction uncertainty, reference data uncertainty, and sampling variability.	Construction of a biomass map, with comprehensive accounting of contributing sources of uncertainty

4.2.1 Design-based estimators

Design-based estimation and inference are based on probability samples of reference data that are collected within the region of interest. A probability sample requires that some form of randomisation was involved in the selection of the sample units, and that the 'inclusion probabilities', i.e. the probability of the population units being included in the sample, are positive and known. An additional requirement is that each population unit is assumed to have only a single value, apart from negligible measurement or other errors.

Multiple sampling designs are possible, the most familiar of which is simple random sampling, for in which population units are selected independently and completely at random. This is often considered the baseline methodology because it is easy to apply and intuitive, and the simple

[1] A mean-squared error (MSE) is calculated here rather than variance, as there is potential for the model to be biased. If the model is unbiased then the MSE equals the variance (McRoberts, Næsset, Saatchi, et al., 2019).

expansion estimator (EXP), described below, and sometimes called the simple random sampling estimator, is well-known to every student of statistics. However, in the context of locating sampling units in space, simple random sampling can be inefficient, in the sense that other methods for spatially locating sampling units can yield the same or greater precision with smaller sample sizes. These include stratified random sampling (where spatial strata are identified before the sample units are located), and systematic sampling (where sample units are located on a regular lattice, but with the lattice randomly positioned in space). Compared to simple random sampling, these methods seek to locate samples more representatively across space and are more efficient when there is spatial structure in the underlying response variable being measured. Statistical estimators for these alternative designs are readily available in standard statistical texts (e.g. Cochran, 1977), noting that the EXP estimator is usually applied to systematic designs, even though it is known that the variance estimator may be conservatively biased, i.e. slightly over-estimated.

The EXP estimator is described below. Importantly, the EXP estimator (and those for the other design-based sampling designs) cannot generate biomass maps from the reference data alone (Figure 4.3A), although when reference data are combined with auxiliary data as part of a design-based model-assisted regression estimator, biomass maps can be constructed (Figure 4.3B).

A simple illustration of the EXP estimator is where plot-level biomass (t ha^{-1}) is measured at n locations that are located completely at random within a region. In this case the mean biomass density estimate for the region as a whole is given by

$$\hat{\mu}^{EXP} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad (1)$$

and the variance of that estimate as

$$\widehat{Var}\left(\hat{\mu}^{EXP}\right) = \frac{1}{n \cdot (n-1)}\sum_{i=1}^{n}\left(y_i - \hat{\mu}^{EXP}\right)^2. \qquad (2)^2$$

Here y_i denotes the value for the i^{th} sample plot.

[2] Equation 2 excludes the 'finite population correction factor' $(1 - f)$, where f is the fraction of the total population that is included in the sample (Cochran 1977), as in most biomass mapping analyses f is small and typically close to zero.

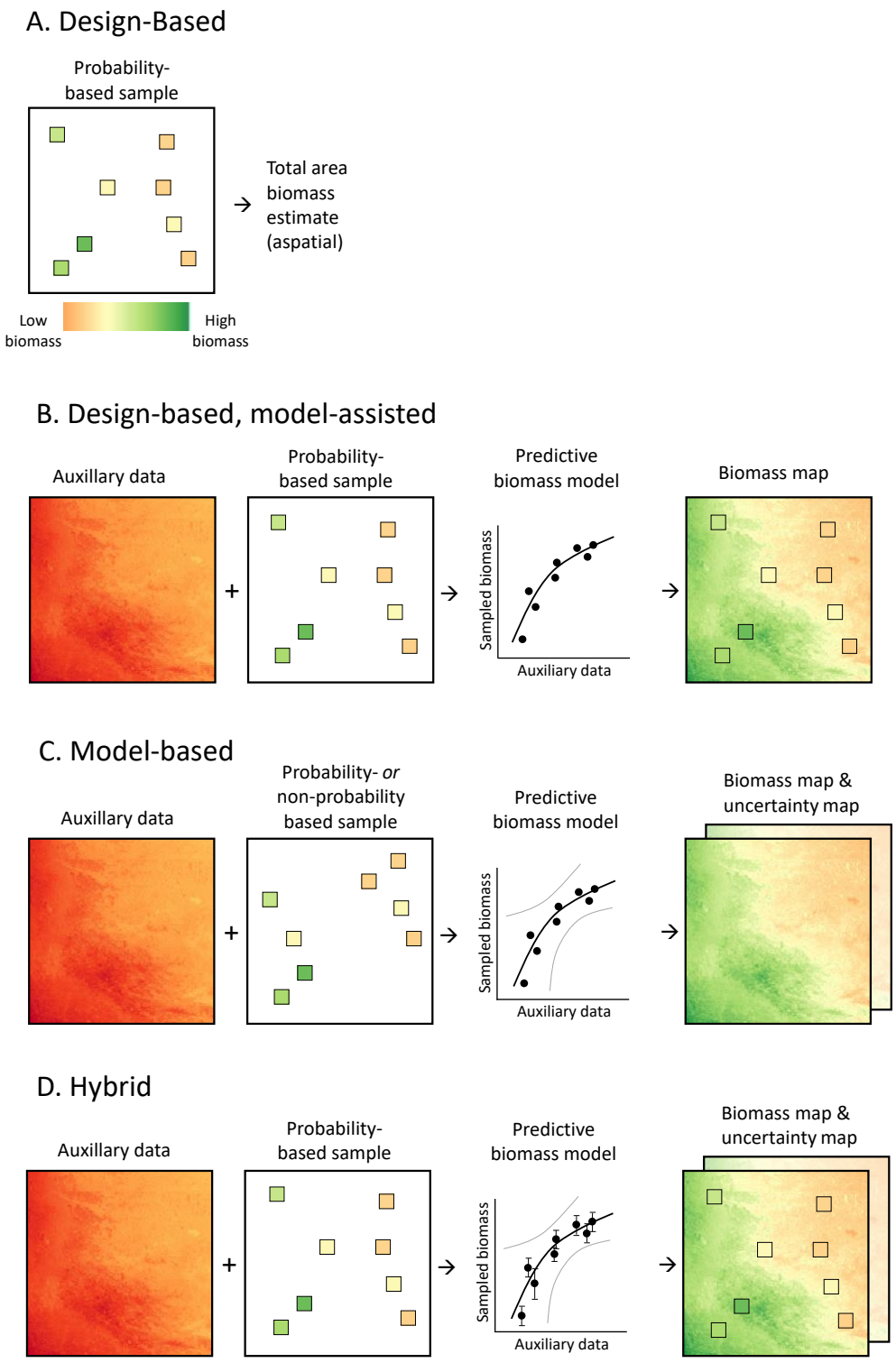

Figure 4.3 Illustration of the inference methods discussed in the text for estimating biomass and its uncertainty. (A) Design-based. (B) Design-based model-assisted. (C) Model-based and (D) Hybrid of design-based and model-based. The data points in the predictive biomass model represent the biomass estimate from the reference data (y-axis) and the auxiliary data value for the same location (x-axis).

To construct a biomass map using a probability sample of reference data requires spatially continuous auxiliary data, usually of the form of remote sensing-based or other mapped products. The design-based model-assisted estimators provide a formal statistical framework for using this additional spatial information to provide an alternative estimate of whole-region biomass density and its uncertainty (Figure 4.3B) (e.g. Särndal et al. 1992; McRoberts et al 2010, Næsset et al 2011; Saarela et al. 2015). With the model-assisted difference estimator, an existing map can be used, whereas with the model-assisted regression estimator map is constructed by using the probability sample to calibrate a 'spatial predictive biomass model' that predicts biomass from the auxiliary data. For simplicity the depiction in Figure 4.3B shows only one source of auxiliary data, but often multiple sources of spatial data are used, and the forms of the predictive model include simple linear regression, multiple linear regression, nonlinear regression, or non-parametric techniques such as random forests (Breiman 2001).

With the design-based model-assisted regression estimators, estimates of total region biomass and its uncertainty can be estimated, and a map can be constructed as a by-product of the analysis. Note that the construction of the biomass map using this method does not necessarily generate a corresponding map of biomass uncertainty, as the primary use of the model is to provide a single prediction of biomass for all map units, but with no associated estimate of uncertainty.

For large, equal probability samples, both the model-assisted difference and regression estimators of mean biomass are given by,

$$\hat{\mu}^{\mathrm{MA}} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}} \hat{\mathrm{y}}_{\mathrm{i}} - \bar{\varepsilon}, \qquad (3)$$

where MA denotes the model-assisted estimator. As per Eq. (3), this estimator has two terms: the first term is the average model-predicted biomass over all N map units with $\hat{\mathrm{y}}_{\mathrm{i}}$ as the prediction for the i^{th} map unit, and the second term ($\bar{\varepsilon}$) corrects for possible systematic prediction error in the model predictions and is the average difference between the map prediction of biomass and observed biomass for the sample units:

$$\bar{\varepsilon} = \frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}} \varepsilon_i, \qquad (4)$$

where $\varepsilon_{\mathrm{i}} = \mathrm{y}_{\mathrm{i}} - \hat{\mathrm{y}}_{\mathrm{i}}$. For large, equal probability samples, the variance of $\hat{\mu}^{\mathrm{MA}}$ is given by

$$\widehat{\mathrm{Var}}(\hat{\mu}^{\mathrm{MA}}) = \frac{1}{\mathrm{n}\cdot(\mathrm{n}-1)}\sum_{\mathrm{i}=1}^{\mathrm{n}} (\varepsilon_{\mathrm{i}} - \bar{\varepsilon})^2. \qquad (5)$$

where n is the size of the reference sample. The advantages of the design-based estimators are that (i) they are at least asymptotically unbiased, which arises from the probabilistic nature of the sample; (ii) they are relatively straightforward to apply; and (iii) for the model-assisted regression estimators, when auxiliary data are available maps of biomass can be produced. The disadvantages are that (i) implementing large-scale design-based monitoring programs with

sufficient numbers of reference sample locations can be costly to set up and maintain; (ii) fulfilling the requirement that reference data biomass values are obtained with negligible error can be prohibitively expensive and/or impractical. Achieving this either requires destructive harvesting and weighing at a large number of locations, or the development of allometric models of sufficient accuracy and precision that their prediction uncertainty is negligible. Often these conditions cannot easily be met; and (iii) whilst biomass maps result from the process, there are no corresponding maps of biomass uncertainty that would otherwise allow sub-regional (e.g. pixel-level) estimates of both biomass and uncertainty to be reported.

4.2.2 Model-based estimators

Design-based inference methods require a probability sample of reference data on which to base the inference. This contrasts with model-based estimators, where non-probabilistic reference data, e.g. collected on an ad-hoc basis, can be validly used. In this case, the validity of the inference rests on the correct specification of the model, although model-based inference can also be based on data collected from a probability sample (Figure 4.2), which is advantageous as it likely increases the model application, and decreases the likelihood of systematic prediction error. In common with design-based inference, the reference data should have negligible error.

A simple example is where plot-level biomass (t ha^{-1}) is measured at multiple locations, but where those plots do not constitute a probability sample, e.g. they were subjectively located such as within easy reach of roads, etc. If additional predictor variables at those locations are available (e.g. elevation, soil type, canopy structure indices from lidar, etc.) then a predictive biomass model can be fit to the data, as described for the design-based model-assisted regression estimator. If the predictor variables exist as spatial information, such as geographic information system (GIS) raster or polygon layers, then once calibrated with the reference data the model can be applied spatially to construct a map of biomass (Figure 4.3C).

Although the design-based model-assisted regression estimator and the model-based estimator may use the same predictive biomass model, there are important differences in the way the model is used, in the assumptions underlying the model fitting, and most importantly in the estimation of biomass and its uncertainty.

With model-based inference, a population or sample unit observation is expressed as

$$y_i = \mu_i + \varepsilon_i \quad (6)$$

where i indexes the sampling units, μ_i is the mean of the distribution of possible observations for the i^{th} unit, and $\varepsilon_i$is the deviation between the observation for the i^{th} population unit and its mean. Note this is fundamentally different from the design-based paradigm for which the i^{th} unit has one and only one value.

A regression model or a non-parametric technique such as random forests can be used to predict the population unit mean as

$$\hat{\mu}_i = f(\mathbf{x}_i; \hat{\boldsymbol{\beta}}) \quad (7)$$

where *i* indexes the population units, $\mathbf{x}_i$ is a vector of predictor variable observations, $\widehat{\boldsymbol{\beta}}$ is a vector of parameters to be estimated or selected, and $f(\mathbf{x}_i; \widehat{\boldsymbol{\beta}})$ denotes the relationship between the predictor variables and the parameters. The model-based (MB) estimator of the population mean is then simply the mean of model predictions over all N map units

$$\hat{\mu}^{MB} = \frac{1}{N}\sum_{i=1}^{N} \hat{\mu}_i^{MB} \tag{8}$$

where $\hat{\mu}_i^{MB}$ is the model prediction at location i. An additional key difference between design-based and model-based inference is that, unlike design-based estimators, the MB estimator of the mean cannot be guaranteed to be unbiased. This feature of model-based estimation necessarily entails additional validation checks to ensure that the model is fit for purpose. This can be achieved by comparison of the model predictions with independent data not used in the original model fitting (e.g. Roxburgh et al., 2019; see Box 3.3 in Chapter 3.2), or through checks to ensure the model shows no systematic lack of fit to the sample data, and that the range of values in the auxiliary data is similar in the sample as in the population, thereby supporting the contention that the sample data are representative of the population of interest. As indicated in Figure 4.2, it is also possible to apply an existing model-based estimator to a novel area that lacks reference data. In this case, extra care is required in the validation step to ensure that the novel area is similar to the area for which the model was developed (e.g. in close proximity, or with similar species composition and forest structure, etc.).

The model-based estimator of the uncertainty is expressed as a mean squared error ($\widehat{MSE}$) rather than a variance due to the fact that the estimator may be biased. The uncertainty can be approximated analytically using a relatively complex term that involves components of variability that capture sampling error, residual error, and spatial covariance.

The term that captures the effects of sampling variability on model parameter estimates and model predictions (pre) is denoted

$$\widehat{\mathrm{Var}}_{pre}^{MB} = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N} \widehat{\mathrm{Cov}}\left(\hat{\mu}_i^{MB}, \hat{\mu}_j^{MB}\right). \tag{9}$$

Recalling that $\hat{\mu}_i^{MB}$ represents the estimated mean of a distribution of possible of values for map location i, $\widehat{\mathrm{Var}}_{pre}^{MB}$ therefore quantifies the mean degree of relatedness between predictions for all possible pairs of map locations (*i*,*j*), and takes account of the fact that the model predictions at each point are based on the same sample data, hence the model predictions are correlated.

The residual (res) error term is given by

$$\widehat{\mathrm{Var}}_{res}^{MB} = \frac{1}{N}\sum_{i-1}^{N} \hat{\sigma}_i^2 \tag{10}$$

where $\hat{\sigma}_i^2 = \widehat{\mathrm{Var}}(\hat{\varepsilon}_i)$, the model residual variance for location i.

The third and final term quantifies covariances resulting from spatial correlation among the residuals (spa) and is given by

$$\widehat{\mathrm{Var}}_{\mathrm{spa}}^{\mathrm{MB}} = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N}\widehat{\mathrm{Cov}}\left(\hat{\varepsilon}_i, \hat{\varepsilon}_j\right). \quad (11)$$

where $\widehat{\mathrm{Cov}}\left(\hat{\varepsilon}_i, \hat{\varepsilon}_j\right)$ is the estimate of the spatial covariance among the residuals.

The total model-based error term (tot) includes the previously defined terms as

$$\widehat{\mathrm{MSE}}_{\mathrm{tot}}^{\mathrm{MB}} = \widehat{\mathrm{Var}}_{\mathrm{pre}}^{\mathrm{MB}} + \widehat{\mathrm{Var}}_{\mathrm{res}}^{\mathrm{MB}} + \widehat{\mathrm{Var}}_{\mathrm{spa}}^{\mathrm{MB}} \quad (12)$$

Equations 9 – 12 provide the generic recipe for calculating model-based uncertainty. Recent studies indicate that $\widehat{\mathrm{Var}}_{\mathrm{pre}}^{\mathrm{MB}}$ often dominates the other terms. The details for calculating $\widehat{\mathrm{Var}}_{\mathrm{pre}}^{\mathrm{MB}}$ term will vary by the type of model fitted. For regression models $\widehat{\mathrm{Cov}}\left(\hat{\mu}_i^{\mathrm{MB}}, \hat{\mu}_j^{\mathrm{MB}}\right)$ can be estimated analytically from the regression fit statistics.

The advantages of model-based inference are (i) the method can use data not collected as part of a probability-based sample design, and hence can be used in areas where data is sparse and/or where data collection is difficult or impossible; (ii) with careful validation a model developed in one region may be successfully applied in a region that does not have sufficient reference data from which to build a novel model; and (iii) the method can produce both maps of biomass, as well as maps of biomass uncertainty, at the resolution of the auxiliary data.

The disadvantages of model-based inference are (i) as with design-based inference, the reference data must have negligible error; (ii) there is no guarantee that the estimator of the mean will be unbiased, requiring additional validation steps to check for possible bias; and (iii) the estimation of the uncertainty term can become computationally unwieldy, and not all of the required information may be available from the auxiliary data from which to calculate all of the $\widehat{MSE}$ components.

4.2.3 Hybrid estimators

Both design- and model-based inference require that the reference data have zero or negligible uncertainty; however, the reference data on which biomass maps are often constructed are model predictions based on field-based measurements and almost always have non-negligible uncertainty arising from a range of measurement and model errors and natural variability (Chapter 2, Table 4.1a).

One solution for dealing with imperfect reference data is a hybrid inference method that combines elements of both design-based and model-based inference (Corona et al., 2014; Fattorini, 2012; McRoberts et al., 2016; Ståhl et al., 2016) (Figure 4.3D). The main elements of hybrid inference are that (i) the reference data have uncertainty that can be quantified, e.g. through propagating the errors in Table 4.1a, to give plot-level estimates of mean biomass, and its variance; (ii) the design-based component captures the sampling variability using the mean

plot biomass estimates within a design-based estimator, such as EXP; (iii) the model-based component captures the plot-level uncertainty by embedding it within the fitting of the predictive biomass model; and (iv) the model-based uncertainty is based only on the locations at which the reference data were collected, not all of the N map locations. The hybrid estimator of mean biomass is given by the appropriate design-based estimator using the plot biomass predictions as if they were error-free. For simple random sampling and the EXP estimator for the design-based component, the hybrid (HYB) estimator of the mean is

$$\hat{\mu}^{\mathrm{HYB}} = \frac{1}{n}\sum_{i=1}^{n}\hat{y}_i \qquad (13)$$

where n is the reference sample size, and $\hat{y}_i$ is the predicted plot biomass.

The estimator for the uncertainty additively combines the design-based and model-based components

$$\widehat{\mathrm{MSE}}\left(\hat{\mu}^{\mathrm{HYB}}\right) = \widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{HYB}}\right) + \widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{tot}}\left(\hat{\mu}^{\mathrm{HYB}}\right) \qquad (14)$$

Under simple random sampling, the design-based uncertainty component is

$$\widehat{\mathrm{MSE}}^{\mathrm{DB}}\left(\hat{\mu}^{\mathrm{HYB}}\right) = \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\left(\hat{y}_i - \hat{\mu}^{\mathrm{HYB}}\right)^2 \qquad (15)$$

and the model-based term, $\widehat{\mathrm{MSE}}^{\mathrm{MB}}_{\mathrm{tot}}(\hat{\mu}^{\mathrm{HYB}})$, can be estimated using the methods in the previous section (Equation 12), noting that the summations in Equations (13) – (15) for hybrid inference are over the number of observed sample points (n), and not the total number of map units (N).

4.2.4 Error propagation by Monte Carlo simulation

The effects of some of the error sources involved in the propagation of errors for estimating biomass uncertainty can be calculated analytically (See Box 4.1), but for clarity and simplicity, a generic Monte Carlo procedure to propagate uncertainty from all sources is described below. Monte Carlo methods for propagating uncertainty are also necessary for many of the error terms in Tables 4.1a and 4.1b, particularly measurement errors. Monte Carlo propagation of uncertainty is also the recommended method by the IPCC (IPCC, 2000, 2006).

The Monte Carlo example is based on the following assumptions:

- For any observations that are known to be measured with systematic error, and that error has been estimated, the required correction for estimated error has been applied to the measurement value prior to error propagation analysis. As an example, for a measurement X that is known to have an error of e, the error correction (X') is calculated as X' = X – e.
- The reference data are calculated from forest inventory sample plot data and are based on a probability sample – here assumed to be a simple random sample.

- Allometric models are used to convert the field measurement data to biomass, yielding plot-level biomass predictions. These reference data are assumed to be 'imperfect', in the sense that plot biomass predictions have non-negligible uncertainty arising from the sources described in Table 4.1a.
- Auxiliary data, in the form of remotely sensed information, e.g. a raster, are available upon which a predictive biomass model can be constructed, with the model calibrated to the reference data. Uncertainty of model predictions at a given point is a combination of the model fit errors (Sources E1 and E2, Table 4.1b), model prediction errors (Source E3, Table 4.1b), and additionally embeds the uncertainty in the reference data. The procedure described therefore represents a form of hybrid inference.
- The predictive biomass model is applied to all map units (e.g. pixels), to yield a map of modelled biomass.

Box 4.1 Error sources involved in the propagation of errors when estimating biomass uncertainty. Step 1 - allometric model development Two alternative options are provided, depending on whether information on measurement errors is available		Error Source from Table 4.1a & 4.1b
Either Option 1 - Including measurement errors		
Step 1a.	1a.1. *For models predicting dry mass, and where moisture content correction uncertainties are known* For each required allometric model, and for each individual tree (or shrub) on which the model is being constructed, select from the required probability distribution a value for the moisture content correction factor (MC) to be applied to that individual. Note: the same moisture content factor is typically applied to multiple individuals, and this needs to be recognised in the Monte Carlo procedure to preserve this source of covariance. If there is uncertainty in the fresh mass weights (e.g. through instrument errors), then select a fresh mass for this individual from the required probability distribution. Calculate the individual dry mass as fresh mass x MC. 1a.2. *For models predicting volume, or for models predicting dry mass and where moisture content correction uncertainties are unknown* Select the dependent variable directly from the required probability distribution.	A.1.
Step 1b.	For each individual, and for each independent variable (e.g. stem diameter, height, etc.), take the observed value and add to it random error drawn from an appropriate probability distribution.	A.2.
Step 1c.	Fit the allometric model(s) to the selected dependent and independent variables from Steps 1a and 1b. For each required allometric model, draw model parameter values at random from the covariance matrix for the model parameter estimates, typically assuming a multi-normal distribution, thus ensuring correlations among parameter values are respected. Alternatively, if the covariance matrix is not accurately estimated, such as may be the case for some nonlinear models, a bootstrap approach can be used to generate the distribution of model parameter estimates. GOTO Step 2.	A.3.
Or Option 2 – Excluding measurement errors		
Step 1a'.	For each required allometric model, draw model parameter values at random from the covariance matrix for the model parameter estimates, typically assuming a multi-normal distribution, thus ensuring correlations among parameter values are respected. Alternatively, if the covariance matrix is not accurately estimated, such as may be the case for some nonlinear models, a bootstrap approach can be used to generate the distribution of model parameter estimates. GOTO Step 2.	A.3.

Step 2. Reference data (= plot-level biomass)		Error Source from Tables 4.1a & 4.1b
Step 2a.	For each individual, and for each independent variable (e.g. stem diameter, height, etc.), take the observed value and add to it random error drawn from an appropriate probability distribution.	B.1.
Step 2b.	For each individual within each sample plot, and at the current values for the independent variables (from Step 2a), select a value for allometric model predicted biomass for each individual from a normal distribution centered on the model prediction value, and with a standard deviation based on the model's prediction (residual) error, taking note to apply appropriate bias correction if moving from a transformed scale to a natural scale. Under an assumption of homoscedasticity, the prediction error is given by, $$s_\varepsilon[1 + V_i^T(Z^TZ)^{-1}V_i]^{0.5}$$ where Z is an n x m matrix containing the first derivatives of the model with respect to the parameters for each of the n individuals for with the allometric model was fitted, and $\mathbf{V}_i$ is an m + 1 vector of the independent variables for the new individual i (Chatterjee & Hadi 2012), and where T is the transpose of the matrix. For linear models, $Z_i = X_i$, the matrix containing the predictor variables.	B.2.
Step 2c.	If the allometric model-predicted quantity is tree volume (requiring multiplication of AGB by wood density), or if predicted dry mass is to be reported in units of carbon (requiring multiplication of AGB by tissue carbon fraction), then select values for wood density (and/or carbon fraction) from the appropriate probability distribution. Note: a single randomly selected value for either variable should be applied to all individuals that are being estimated by the same allometric model, to retain the covariance.	B.3.
Step 3. Plot-level biomass and sampling variability		Error Source from Tables 4.1a & 4.1b
Step 3a.	Sum the biomass for all individuals within all n plots, represented symbolically as $[\mathrm{B}_1^{\mathrm{ref}}, \mathrm{B}_2^{\mathrm{ref}}, \mathrm{B}_3^{\mathrm{ref}}, \cdots \mathrm{B}_\mathrm{n}^{\mathrm{ref}}]$, i.e. a list of predicted biomass for each sample plot, where 'ref' denotes they are reference data predictions, where e.g. $\mathrm{B}_1^{\mathrm{ref}}$ is the predicted biomass (e.g. t Dry matter) for reference plot 1. If expressing plot-level data as a biomass density (e.g. t AGB ha^{-1}) then for each survey plot, draw from the appropriate probability distribution a value for the plot area, reflecting uncertainty in plot area due to e.g. errors in locating plot boundaries, etc. Divide the total plot biomass $\left([\mathrm{B}_1^{\mathrm{ref}}, \mathrm{B}_2^{\mathrm{ref}}, \mathrm{B}_3^{\mathrm{ref}}, \cdots \mathrm{B}_\mathrm{n}^{\mathrm{ref}}]\right)$ by the plot areas to yield plot-level biomass on a per-area basis $[\mathrm{B}_1^{\mathrm{ref}'}, \mathrm{B}_2^{\mathrm{ref}'}, \mathrm{B}_3^{\mathrm{ref}'}, \cdots \mathrm{B}_\mathrm{n}^{\mathrm{ref}'}]$.	C.1.
Step 3b.	To simulate the effects of sampling variability, randomly select, with replacement, n plots from $[\mathrm{B}_1^{\mathrm{ref}'}, \mathrm{B}_2^{\mathrm{ref}'}, \mathrm{B}_3^{\mathrm{ref}'}, \cdots \mathrm{B}_\mathrm{n}^{\mathrm{ref}'}]$, and then save this re-sampled list of plots for later analysis. Note for any given iteration some plots will likely occur in the list more than once, and some plots will not occur at all, although an even representation of plots is expected over a large number of iterations.	D.1.

	Repeat Steps 1a – 3b nRep replicate times, to give nRep sets of biomass predictions for each plot. The reference data biomass estimates for the n plots, for the k = 1..nRep replicates, can therefore be given by $[B_{1,k}^{ref'}, B_{2,k}^{ref'}, B_{3,k}^{ref'}, \cdots B_{n,k}^{ref'}]$. **Therefore** $B_{1,k}^{ref'}$ **is the predicted biomass density (e.g. t Dry matter ha^{-1}) for the kth Monte Carlo replicate for reference plot 1.**	
Step 4. Predictive biomass model - Development		Error Source from Tables 4.1a & 4.1b
Step 4a.	If there is error in the auxiliary data on which the predictive biomass model is to be based, then generate nRep random realisations of the data, through adding to the value of the auxiliary data at each location a random error drawn from an appropriate probability distribution. Typically, the auxiliary data will comprise spatial GIS layers, or other spatial information, hence care may be required to ensure the spatial autocorrelation in randomly generated layers is conserved.	E.1.
Step 4b.	For each of the k = 1..nRep replicates fit a predictive biomass model to the data, using the auxiliary data from 4a as the independent variables(s), and $[B_{1.k}^{ref'}, B_{2,k}^{ref'}, B_{3,k}^{ref'}, \cdots B_{n,k}^{ref'}]$ as the dependent variable.	-
Step 5. Predictive biomass model - Use		Error Source from Tables 4.1a & 4.1b
Step 5a.	For each of the nRep predictive biomass models, select model parameter values at random from the predictive biomass model's parameter variance-covariance matrix, as per Step 1.b' of the allometric model analysis. If the model form is other than regression, then drawing parameter values may require other methods, such as bootstrapped re-sampling or other non-parametric techniques.	E.2.
Step 5b.	Following a similar approach to Step 2b for allometric model prediction error, for each of the nRep predictive biomass models, and for all i = 1..N map units (including the n plot locations), select a value of predicted biomass from a normal distribution centred on the model prediction, using the auxiliary data from Step 4a as the independent variables, and with a standard deviation based on the model's prediction (residual) error, to give $B_{i,k}^{pre'}$ for plot i in replicate k. For each of the k = 1..nRep replicates calculate the mean biomass per unit area using the design-based model-assisted difference estimator (Equation (3)): $\hat{\mu}_k^{MA} = \frac{1}{N}\sum_{i=1}^{N} B_{i,k}^{pre'} - \bar{\varepsilon}$ where $B_{i,k}^{pre'}$ is the predicted biomass for the ith map unit for the kth replicate, and $\bar{\varepsilon} = \frac{1}{n}\sum_{j=1}^{n}\left(B_{j,k}^{pre'} - B_{j,k}^{ref'}\right)$. with $B_{j,k}^{ref'}$ and $B_{j,k}^{pre'}$ being the plot biomass for the kth replicate from Step 3a, and the model prediction of that biomass, respectively. The variance of the mean is calculated using Equation (5) $\widehat{Var}(\hat{\mu}_k^{MA}) = \frac{1}{n\cdot(n-1)}\sum_{i=1}^{n}\left(B_{i,k}^{pre'} - B_{i,k}^{ref'}\right)^2$ Save $\hat{\mu}_k^{MA}$ and $\widehat{Var}(\hat{\mu}_k^{MA})$ for later analysis.	E.3.

	For each of the i = 1...N map units (e.g. polygons, pixels) calculate the bias-adjusted predicted biomass for that location, $B_{i,k}^{pre'} - \bar{\varepsilon}$, and save the predicted biomass map for later analysis.	

Following Rubin (1987) and McRoberts et al. (2016), the population (i.e. whole map) mean biomass density can be calculated as:

$$\hat{\mu} = \frac{1}{\mathrm{nRep}} \sum_{k=1}^{\mathrm{nRep}} \hat{\mu}_k^{MA}, \tag{16}$$

and the variance of that estimate as

$$\widehat{\mathrm{var}}(\hat{\mu}) = \left(1 + \frac{1}{\mathrm{nRep}}\right) \cdot W_1 + W_2 \tag{17}$$

where W_1 is the between-replicate variance in predicted mean biomass, capturing the net effects of the various uncertainty sources that were included at each step in the calculation, and is given by

$$W_1 = \frac{1}{\mathrm{nRep} - 1} \sum_{k=1}^{\mathrm{nRep}} \left(\hat{\mu}_k^{MA} - \hat{\mu}\right)^2 \tag{18}$$

and where W_2 is the mean within-replicate variance, given by

$$W_2 = \frac{1}{\mathrm{nRep}} \sum_{k=1}^{\mathrm{nRep}} \widehat{\mathrm{Var}}\left(\hat{\mu}_k^{MA}\right) \tag{19}$$

If the total absolute biomass of the region ($\widehat{B}$) is required, it can simply be estimated as the estimate of mean biomass density multiplied by the total area, A.

$$\widehat{B} = \hat{\mu} \cdot A \tag{20}$$

with the variance scaled accordingly

$$\widehat{\mathrm{Var}}\left(\widehat{B}\right) = \widehat{\mathrm{var}}(\hat{\mu}) \cdot A^2. \tag{21}$$

If the total area A is known to be uncertain, with an error variance of $\widehat{Var}(A)$, then under the (reasonable) assumption that uncertainty in is *A* independent of the error sources contributing to mean biomass density, $\widehat{Var}\left(\hat{B}\right)$ can alternatively be calculated via Goodman (1960). This corresponds to error source F.1. in Table 4.1b.

$$\widehat{\mathrm{Var}}\left(\widehat{B}\right) = \widehat{\mathrm{var}}(\hat{\mu}) \cdot A^2 + \widehat{\mathrm{Var}}(A) \cdot \hat{\mu}^2 + \widehat{\mathrm{var}}(\hat{\mu}) \cdot \widehat{\mathrm{var}}(A) \tag{22}$$

Finally, the biomass map for the region can be created by calculating the mean model-predicted biomass across the k replicates of the analysis. For each map unit z (e.g. pixels) the biomass is:

$$\hat{\mu}_i = \frac{1}{nRep} \sum_{k=1}^{nRep} B_{i,k}^{pre'} \quad (23)$$

with an estimate of the variability of any given map unit z given by

$$\widehat{Var}(\hat{\mu}_i) = \frac{1}{nRep \cdot (nRep - 1)} \sum_{k=1}^{nRep} \left(\hat{\mu}_i - B_{i,k}^{pre'}\right)^2 . \quad (24)$$

4.3 Summary

In this chapter three broad inference methods for estimating biomass and its uncertainty from field-based data and for constructing maps of biomass from auxiliary spatial data were described. With the *design-based* inference EXP estimators (Figure 4.3a) only total regional biomass can be estimated, with no information available on the patterns of sub-regional (e.g. pixel-level) variability; thus, on its own, the design-based EXP estimators cannot be used for constructing maps of biomass. However, the *design-based, model-assisted* estimators (Figure 4.3b), combine spatially continuous, georeferenced auxiliary data with the sample (reference) biomass data in a predictive model to allow a biomass map to be constructed. In both design-based variants, the validity of the estimate rests upon the probability sample. *Model-based* inference (Figure 4.3c) is superficially similar to *design-based, model-assisted* inference, in that a predictive biomass model is constructed from a relationship between sample biomass and auxiliary data and used to construct a biomass map, but there are fundamental differences. Under model-based inference the sample need not be probability-based, and the validity of the estimate rests solely upon the validity of the model, not the validity of the sampling design. In model-based inference uncertainty is estimated quite differently from designed-based inference and is based on the characteristics of the model, with an underlying assumption that the population values are inherently variable (as opposed to design-based methods, where it is assumed that population units can only have a single value).

Both design-based and model-based inference require the reference data to be known with negligible uncertainty, which is almost never the case when plot-level biomass data are predicted using allometric models whose predictions are themselves subject to uncertainty. In this situation, *hybrid* inference can be applied when there is non-negligible uncertainty in the reference data, and that data are collected as part of a probability sample (Figure 4.3d). In this case, design-based inference (which captures the effects of sampling variability) is combined with model-based inference (which captures the uncertainty in the sample unit predictions). In general, the auxiliary data on which predictive biomass models are based can be multi-dimensional, and the predictive model can include traditional linear and non-linear regression, or alternative methods such as non-parametric or machine learning techniques. A decision tree is provided (Figure 4.2) that allows users to select the correct inference method that should be used, based on basic knowledge of the user's available data, in addition to the fundamental estimators for estimating biomass and its uncertainty under each of the three inference methods.

A total of 13 possible sources of error were identified, spanning the breadth of estimation steps from initial construction of allometric models through to the construction of the map and estimation of the biomass parameter and its uncertainty. A generalised Monte Carlo procedure is described to illustrate how these error sources can be combined or 'propagated' through the stages of biomass map construction, including the total map biomass estimate and its uncertainty, i.e. summed over all map units. Because the procedure incorporates both design-based and model-based elements, and there is an assumption that the reference data are subject to uncertainty, this example falls within the class of hybrid inference. Whilst the procedure is comprehensive, in that it accommodates all 13 of the identified error sources, there is scope, and indeed perhaps a need, to tailor the algorithm for each particular circumstance. In particular, information on some of the uncertainty sources may not be commonly available (such as those associated with the development of allometric models, where often only summary statistics such as parameter estimates, the regression residual error and sometimes the parameter covariance matrix might be the only information available). In such situations there is little choice but to assume those error sources are negligible, and exclude them from analysis. To recognise this, the described Monte Carlo procedure provides two options, one with a workflow that includes measurement errors during allometric model development, and one that includes only allometric model parameter uncertainty.

Whilst the classification of inference methods into design-based, model-based and hybrid is convenient and fits most situations, these categories are fuzzy at the edges. For example, when conducting analyses by Monte Carlo simulation, there is scope to adjust the procedure to suit the desired need. A simple example is a design-based survey with reference data that have a known and quantified uncertainty. In a Monte Carlo simulation, it is possible to create replicate reference datasets incorporating that uncertainty (such as done at Step 3a above), and to then calculate appropriate uncertainty metrics based on design-based estimators. However, such an analysis is neither pure design-based, as the reference data are imperfect, nor is it model-assisted or model-based, as no auxiliary data were used and no predictive biomass model was developed. This flexibility in the way Monte Carlo simulations can be formulated is also evident in the procedure described above, where the analysis is hybrid in nature, but the model-assisted estimator is used rather than the relatively more complex model-based estimator.

With the increase in availability of remote-sensing information at a range of spatial scales, from plot-scale terrestrial laser scans through to space-borne optical, lidar and radar products, there is increasing interest in developing methods to combine these multiple sources of information for the purpose of biomass estimation, and the creation of biomass maps (Saarela et al. 2016; 2018a). A key challenge for the future is therefore the further development of statistical methods and associated computational tools for integrating this information with available ground data, and for creating maps of biomass and related metrics, and their uncertainty. Multiple computational tools already exist for biomass estimation and the propagation of uncertainty, such as the 'HMR' R package of Saarela et al. (2018b) for undertaking generalised hierarchical model-based estimation with nested data; the 'BIOMASS' R package of Rejou-Mechain et al. (2017) for estimating above-ground biomass and its uncertainty in tropical forests; and the Bayesian Model Averaging (BMA) R code listings of Mavouroulou et al. (2014) for combining

alternative allometric models. Each of these tools provides targeted solutions for a subset of the steps described in Tables 4.1a and 4.1b, but there does not yet exist an integrated set of algorithms for biomass estimation and uncertainty propagation that spans the full range of calculations, from initial field-biomass estimation through to final map creation. The development of such a comprehensive tool-set would streamline the process of biomass map creation, and would provide a set of standardised methods and algorithms to facilitate the ongoing improvement of biomass maps and related products.

4.4 Recommendations

It is recommended, where possible and practicable, that both measurement and allometric model errors be propagated to estimate field plot AGB with associated uncertainties both in the development of remote sensing biomass products, and in their independent validation. Further, while we acknowledge that some flexibility in the selection of inference methodology is necessary, considering the breadth of forest inventory sampling approaches that we anticipate will be used for product development (and even validation), **it is generally recommended** to use hybrid inference for biomass products. This is because it does not assume negligible errors in reference datasets, which is almost never true for biomass. We also endorse the use of Monte Carlo approaches for error propagation, as they have the flexibility to account for differences in field and remote sensing datasets, and different flavours of empirical modeling.

Chapter 5: Considerations for the Utility of Protocol

There are a wide range of potential users of biomass products that have different needs in terms of spatial and temporal resolution, uncertainty requirements, and validation plans. This chapter summarizes the primary anticipated user types, written by researchers from three communities, in an attempt to summarize the community-specific considerations of these groups. Section 5.1 focuses on the modeling community, 5.2 focuses on policy applications, and 5.3 focuses on scientists interested in non-forest biomass. This Chapter focuses on recommendations for producers of biomass products rather than product users, and hopes to inform the next generation of remote sensing products as to the needs of the wide and disparate user communities.

5.1 Utility of protocol by modeling communities

Mathew Williams, Natasha MacBean, Martin De Kauwe

5.1.1 Model types

A variety of process models could benefit from interacting with and extracting value from biomass data products. These models range in complexity from intermediate complexity ecosystem C cycle (C-pool) models; to more complex forest cohort/age-class models; to highly resolved forest gap/individual based models (IBMs). All these model types typically simulate half-hourly to daily dynamics, driven by meteorology and other external factors such as fire disturbance or management interventions. Key differences among models are:

(i) C-pool models resolve the biogeochemistry of photosynthesis, C allocation to live C pools and their C turnover, representing biomass as the bulk mass of all live C pools. These models operate at a defined pixel scale, which may vary from 1 ha to 1° depending on the application, and are typically parameterised for plant functional types (D. B. Clark et al., 2011).

(ii) Cohort models track biogeochemical fluxes as above, but include more detail on the distribution of biomass among trees of different ages or sizes. Thus, these models can resolve differences between plantations and naturally regenerated forests and can predict forest height. These models have similar flexibility in spatial scale to C-pool models, but to-date are typically run at site, or sub-continental scales (Moorcroft et al., 2001).

(iii) Gap models, or IBMs, model populations of multiple individual stems within the forest patch scale (<1 ha). The model simulates individual stems of various species competing for resources (predominantly light), resolving the shifting steady state of natural forests at the scale of canopy dominant stems, with stochastic mortality. By summing individual stem dynamics, biomass at the patch scale, and the patch stem height, are determined (Ryan & Williams, 2011).

These models are generally used for different purposes. C-pool models are the focus for carbon cycle analyses. They are typically used within Earth System Models (ESMs) to represent terrestrial C cycling and water/energy flow. Model parameters are generally assigned from fixed 'plant functional types'; competition can be determined by comparing simulated growth between different PFTs. Cohort models can simulate C dynamics with more complexity around processes of growth and mortality and include the potential for simulating successional processes. Some ESMs are now beginning to include cohort modelling, but none are operational as part of the latest Coupled Model Intercomparison Project (Jones et al., 2016). Gap-models are high resolution simulators to explore ecological processes and interactions at fine scale (<100 m). They are not currently suitable for regional or global analyses due to the computational costs.

The allocation of carbon to different plant tissues, and the timescale for turnover of these tissues, remain key uncertainties in vegetation modelling and projections of the terrestrial carbon sink (Friend et al., 2014). The processes of respiration, allocation and turnover are afforded simple, largely empirical representations based on limited field studies, contrasting with the complexity in which a key process, such as photosynthesis, is represented. Most models represent net primary production as the outcome of the difference between photosynthesis and autotrophic respiration. However, recent experimental and observational studies highlight a lack of correlation between photosynthesis and growth, which points to a more complex control on plant production, including C storage pools within plants (Richardson et al., 2013; Rowland et al., 2014). The processes controlling autotrophic respiration are also poorly understood, from plant to ecosystem scales, compared to photosynthesis (R. Q. Thomas et al., 2019). Accurate and repeated estimates of biomass are likely to refine knowledge of the processes of growth and mortality of plant biomass and their variation across biomes, particularly when combined with other information, such as rates of photosynthesis.

Forest biomass is a dynamic outcome of biological processes of growth and mortality, and exogenous disturbance effects, both natural and anthropogenic. There are enormous challenges to modelling mortality and disturbance, as these processes are stochastic, and hard to observe (McDowell et al., 2013). For instance, we have little information on intrinsic rates of tree mortality across the globe (M. O. Johnson et al., 2016). We lack local information on fire impacts on forest biomass. Forest degradation (partial removal of biomass) is poorly quantified globally. Biomass time series can also provide key information on disturbance effects on forests. At landscape scales these data can resolve the effects of fire, extreme weather or pests and diseases, particularly when linked to other data such as burned area, or spectral monitoring of canopy changes linked to diseases. At fine enough scales (~10 m), biomass time series may even provide insight into natural mortality processes, and the death of large stems.

5.1.2 Integrating biomass data into models

To be useful, models need to represent real systems reliably. Reliability is generally ensured through the processes of calibration and independent validation against data with robust error estimates. Some portion of data is used to calibrate model parameters and set initial conditions. Remaining data are used to test model outputs, thereby evaluating model bias and random error. Biomass data are only recently available, and largely in temporally-ill-defined maps. Errors in biomass products have not always been clearly defined. Thus, biomass data have not typically been used in the calibration and validation process, which has instead focused on using flux data time series. Biomass data provide a significant opportunity to constrain the internal dynamics of C/biomass models, which have been shown to be key drivers of uncertainty among models used in forecasts of terrestrial C cycling. Biomass data can also constrain external forcing on models, particularly related to disturbance processes such as degradation.

Clear metadata are required for data products. Hence, **it is recommended** to distribute observations in a NetCDF format. NetCDF files are self-describing and allow significant metadata to be embedded alongside observations. This format is widely used by global modelling groups and would therefore facilitate the widest possible use. In the medium term it is further **recommended** to undertake an effort to combine biomass observations from different satellite sensors within a remote sensing data assimilation system to produce total aboveground biomass, and its components. These different sensors could resolve different parts of the vegetation (e.g., canopy, branches, trunks) linked to different (e.g., radar) wavelengths. The benefit of this integration would be that a) any assumptions made for retrieval algorithms for individual datasets could be made consistent with each other (e.g., underlying canopy structure, parameter values) and b) that more information (from more observations) could be incorporated in the biomass retrieval. The outcome of this combined retrieval should also have a more robust error estimate.

The simplest model-data interaction is biomass validation. Forest models produce time series of biomass or C stock maps. Information on biomass stocks and biomass change can be used for validation of all model types, to test steady-state predictions, and representation of dynamical processes. These data can also be used for setting or constraining initial conditions. To facilitate the validation process, **it is recommended** that biomass products are integrated within community land model benchmarking schemes such as the International Land Model Benchmarking (ILAMB) project (Collier et al., 2018). ILAMB is a software platform used for model evaluation by the land modelling community, and therefore provides a means for the biomass community to interact directly with modellers.

To strengthen the validation process, **it is recommended** that uncertainty metrics are linked to all biomass and height data products. Metadata should explain the nature of the uncertainty estimate and its derivation (e.g., following recommendations in Chapter 4). Clarity is required on the quantification of random error at pixel scale, spatial correlation of random error (i.e. sensitivity of random error to aggregation), and measurement bias at pixel-to-regional scales. Further, the error on biomass and height change estimates (i.e. pixel differences between

successive biomass/height maps) must be clearly specified and included at the time of product release. Information on local variability, i.e. biomass uniformity sub-pixel, is also valuable for the modelling process. While this protocol focuses on recommendations for biomass stock product validation, uncertainties on change information are equally important.

Data on height, and height change can be used to validate cohorts and IBMs. Cohort and gap models are increasingly being implemented in ESMs to improve the mechanistic realism of their dynamic vegetation model (DVM) component. DVMs are important as they are used for predicting long time scale changes in species distributions as a result of climate and land cover change. To maximise value for DVM validation **it is recommended that** clarity on the relationship between height and biomass estimates is provided (i.e., whether these measurements are correlated or independent). This information will help to interpret the model-data conformity or mis-match. Information on vertical variability of biomass from tomography (ESA Biomass) could verify height/biomass outputs of cohort models and IBMs. Full waveform lidar estimates of vertical canopy structure – as are obtained by the GEDI mission – will also be useful in this regard. Additionally, canopy structure estimates will be beneficial in parameterizing the multi-canopy layer energy schemes that are currently being implemented in some ESMs.

The geographic coverage and detail required by modellers is dependent on model domain and resolution. Products >=0.5° resolution serve for Earth System Model validation. Products at resolution of ~1 ha can serve cohort and gap models for decision support, interaction with local stakeholders, and understanding of fundamental ecological processes. Annual products allow for the resolution of landscape dynamics, including growth, disturbance and management. Sub-annual products allow further resolution of seasonal process variation, including growth and phenology. **It is recommended** that annual 0.5° products are provided as a standard for evaluation against global models. The latest CMIP6 simulations are targeting 1° and this resolution fits ILAMB benchmarking. **It is recommended** that the highest resolution products possible are provided with the fewest gaps and over the largest domain possible, to allow more detailed model and process evaluation for regional studies. The mapped domain, if limited spatially, could be decided in consultation with *in situ* biomass researchers, to target study regions linked to product calibration and validation.

5.1.3 Model data fusion

Biomass data can also play an important role in model calibration. The calibration process uses biomass maps, and other data such as satellite LAI, to constrain and calibrate model parameters and minimize model-data mismatch (Exbrayat et al., 2019). Calibration activities require georeferenced data sets with clear error characterization, as outlined above, and so are consistent with our previous recommendations. The production of biomass maps globally has already provided a basis to constrain estimates of wood residence time (Bloom et al., 2016; Forkel et al., 2019). Repeated biomass data evaluated at site scale has been shown to provide constraints on estimates of woody increment, allocation to wood, and on autotrophic respiration (Smallman et al., 2017). As repeated biomass data become available at high resolution across

broader domains, there is a clear opportunity to constrain C cycle processes in ecosystem and global C cycle models.

5.2 Policy-relevance of protocol

Martin Herold and Sarah Carter

Biomass information, including uncertainty estimates, are useful for a number of purposes from national to international reporting related to forest change, and to monitoring of climate variables and sustainability goals. This information can also be used to support local actors in implementing improved forest and land use management. They are central to the achievement of international goals and national commitments related to forest conservation and management, climate change, and sustainable development. Most prominently, the Paris Agreement on Climate Change requires that progress on the implementation of the United Nations Framework Convention on Climate Change (UNFCCC) National Determined Contributions are transparently reported, where improvement on the quality of national greenhouse gas (GHG) inventories is key. Biomass data are also relevant for Goal 15 of the UN Sustainable Development Goals (SDGs) that aims by 2020 to promote the implementation of sustainable management of all types of forests, halt deforestation, restore degraded forests and substantially increase afforestation and reforestation globally.

This diversity of users brings different requirements in terms of the data needs (including the scope and scale of biomass products), and in terms of considerations on uncertainty (see Table 5.1):

Table 5.1 Summary of key policy initiatives, user groups and requirements for biomass mapping (Herold et al., 2019)

Policy initiative	Key user groups	Biomass data needs	Considerations on uncertainty
IPCC/GCOS	Global climate data users and global assessments (i.e. IPCC WG 1 and WG3)	- Aboveground biomass as one of the Essential Climate Variables (ECV) - Optimal: regular global forest biomass monitoring at 50-100 m resolution at annual intervals - Focus on long-term consistency as input to global and regional climate and vegetation models	- Uncertainty and stability of ECV products are to be assessed and reported for relative and absolute systematic deviation and confidence interval or RMSE, overall and by biomass class/range estimated using reference data of better quality - Coordinated international activities for uncertainty characterization and reference data collection are required

UNFCCC national reporting, transparency framework and global stocktake	National GHG inventory experts and UNFCCC roster of experts for technical assessments and review	- Estimation of carbon emissions and removals from forest changes; also including non-forest areas with significant woody biomass (i.e. agriculture/ grasslands) - Reporting done by countries and all countries need to report regularly (at least a summary of the GHG inventory bi-annually) - Various options of integrating biomass maps in national GHG inventories - Greatest need in tropical countries where availability of plot data and sustainability of NFIs remains an issue - Regular global stocktake by UNFCCC may require data for reconciling country reporting to provide global estimates of progress	- IPCC GPG: estimates should be unbiased (on the national level) and uncertainties reduced as much as practicable - IPCC Tiers: national data needed for Tier 2 - Uncertainty estimates required on national estimate - Continuous improvement process as part of national GHG inventories - Need for consistent (global) data sources (including uncertainty characterization) for technical assessment and global stocktake
REDD+ activities	National REDD+ agencies, forest monitoring experts, REDD+ donors/investors and independent assessment bodies	- Estimation of emission factors and forest carbon stock changes for establishing forest reference level and REDD+ results to access result-based finance - Assessment of potential forest mitigation activities - Requirements can vary according to the financial initiative (i.e. Green Climate Fund, Forest Carbon Partnership Faciliy, bilateral organisations, voluntary markets) - Independent data sources that verification teams can use for comparison	- Uncertainty requirements vary by performance scheme - Consistency, accuracy and transparency needs to be sufficient to be convincing for technical assessments and verification of the results of REDD+ activities - Stepwise improvements possible and encouraged - Discounts in payments due to high uncertainty or conservative adjustments possible
Sustainable development goals	National statistical office responsible for SDG indicator reporting, UNSD	- Biomass changes as one sub-indicator of SDG 15.3.1 but also important input for others that are required for all countries - Detailed requirements are still developing but reporting is to be done by countries on annual basis 2015-2030	- Not yet known but reporting of SDG indicator should allow for the assessment of trends 2015-2030
(Sustainable) Forest management	National forest institutions, forest industry, certification schemes	- Management plans - Forest laws and strategies - Fire prevention plans - Bioenergy plans,	- Emphasis on stand-level/spatially explicit information provision - Quantification and understanding of uncertainties should be incorporated in management decisions
Stakeholder engagement and transparency	Societal actors and broader set of forest-related stakeholders, including local communities	- Free and open biomass data accessible to non-technical stakeholders - Biomass sometimes used for independent assessments, as a simple proxy for forest health, intactness and habitat quality, forest resources assessments, etc.	- Transparent information and uncertainty characterization required by all stakeholder groups - Uncertainty requirements vary by stakeholder group

A summary of the key initiatives, user groups and requirements for biomass mapping in Table 5.1 emphasizes that biomass data derived from space-based data are useful for a wide range of policy-relevant applications. Uncertainty information is required by all stakeholders interested in

using biomass maps (Romijn et al., 2018). This protocol document is described as the first agreed "good practices" document to provide a proper independent uncertainty assessment to be used for biomass mapping products generated using space data. **It is recommended** that independent uncertainty assessment, as described in this protocol, should be treated as a benchmark that essentially caters to two communities: (1) the data producers that require independent validation to understand the quality of their products and can help to improve their efforts over time, and (2) the user of the product that needs to understand whether to use a certain biomass product for a specific purpose and why; including consideration of how uncertainties affect the uptake of the product for a certain application. In that context, it is important to consider that different applications have specific requirements on what they expect from biomass and uncertainty characterization (see Table 5.1) in terms of:

- Geographic scope: ranging from global scale (i.e. for use in climate and vegetation models) to national level estimates (most relevant for country GHG inventories)
- Spatial resolution: from spatially-explicit/stand-level information of less than one hectare to multiple kilometres for global vegetation models
- Which final biomass metrics are needed by the users: can be the total or average biomass per country or forest type or a local (i.e. a forest management unit) /pixel-specific value
- The uncertainty estimates required: which can include relative and absolute systematic deviation (or bias) and confidence interval (or RMSE). They can be provided for the overall estimate or by biomass class/range
- Timing: whether a biomass estimate is needed for one time (i.e. one year) or for multiple years. Some users require data that are most recent (near-real time information) or available consistently for longer time series
- Thematic content: users often ask for one-time maps, but applications increasingly require biomass change information
- Definition: there is variability in needs in terms of which vegetation compartments are considered (i.e. minimum tree size) and whether or not dead biomass is included
- Compliance with IPCC good practice guidelines: (1) "neither over- nor underestimates so far as can be judged," and (2) "uncertainties are reduced as far as is practicable" (IPCC, 2006).

All these choices influence the way the biomass map product should be generated and how the uncertainty of the products is to be assessed for the purpose of the different user/policy requirements (see Section 5.2). This means that one biomass product and uncertainty characterization approach will not work for all users. **It is recommended** that any effort of assessing biomass products should be co-developed by both producers and users. Often user organizations have their own forest and biomass reference data that they would like to integrate with biomass map data. Such efforts are thus, not only aiming at the technical credibility of a

biomass map product, but also toward creating ownership and saliency that are often most important for ensuring user uptake. For example, an internationally-produced biomass map can only become part of the IPCC Tier 2 or Tier 3 national emission estimation if they are integrated with yield data acquired in the country. It is also often the user alone that can fully judge what implications the estimated biomass and uncertainties have on their application. This information can be fed back to the producer on where to improve. In summary, biomass map production for a specific application is an interactive process of co-creation with both producers and users, and uncertainty assessments are most useful if performed in partnership.

5.3 Utility of Protocol by non-forest research communities

Natasha MacBean, Oliver Philips, Michael Falkowski, Patrick Jantz, Scott Goetz, Tom Crowther

5.3.1 Context

The boundaries – conceptual, spatial, temporal – between 'forest' and 'non-forest' are complex and often contested by science and society. These boundaries matter. For example, the standard FAO definition of forest, having 10% tree cover or more (FAO, 2018), results in large areas of dry and semi-arid tropics that are not tree-dominated being classed as forest. Current debate about the global extent of forests in drylands, derived from high-resolution optical imagery, hinges on details of definition (Bastin et al., 2017; Griffith et al., 2017). For the purposes of this overview, we adopt as simple and operational definition of forests as possible, as 'land with largely continuous tree cover'. Most trees that grow *outside* of forests thus experience different growing conditions from most growing within them.

Given this, the user base for this biomass protocol is potentially very wide-ranging, and at least three distinct potential communities of "non-forest" user communities can be identified. These include, (1) users whose main focus is the state and trajectory of non-forest biomes themselves, as well as (2) forest-oriented users in academic and policy-maker communities who need to understand and quantify biomass dynamics and stocks as a result of land-use change processes that can take land out of and back into, the forest class. A third group of users (3), are focused on tracking carbon stores and fluxes below-ground; while this group is relatively smaller, it represents a chronically under-researched topic, and one that, for obvious reasons, is particularly challenging for space-based Earth observation to impact.

5.3.2 Significance

Each of these research communities study systems and processes of major global significance. Thus:

(1) *While non-forests contribute less than 20% of global biomass, they represent half of terrestrial productivity and cover ca. 70% of the Earth's land surface* (Pan et al., 2013). The largest terrestrial tropical biome, the savannas, are traditionally thought of as grass-dominated, yet the 'savanna' concept has proven even harder than that of 'forests' to operationalise consistently across continents (cf. Torello-Raventos et al., 2013). Much vegetation considered by most ecologists as

savanna is structurally dominated by trees, with, on average, at least 50% tree cover coexisting with C4 and C3 tropical grasses (Lloyd et al., 2008).

(2) *Approximately 200,000 km^2 of forest is converted each year to non-forest states as a direct result of human activities* (M. C. Hansen et al., 2013). Additionally, elsewhere the planet is gaining tree cover and biomass as a result of successional processes, active planting, and indirect anthropogenic drivers including longer growing seasons (e.g., at the boreal / tundra margin), changing fire regimes, and increasing carbon dioxide (e.g., woody encroachment at the tropical forest / savanna interface) (Mitchard & Flintrop, 2013; Schimel et al., 2015). However the boundaries between "forest" and "non-forest" are defined, it is important to be able to track biomass dynamics across these boundaries in as consistent and unbiased a way as possible.

(3) *Carbon stocks below-ground are easily overlooked, extremely poorly constrained, and yet can exceed aboveground values - and of course are not confined simply to biomass*. For example, there are huge but uncertain stores of organic carbon in Congo peatlands (6-46 Pg C, Dargie et al., 2017) and Amazon tree roots (13-26 Pg C, Malhi et al., 2006), and across all savannas most ecosystem carbon is below the surface, in soil and roots (e.g. Scharlemann et al., 2014). Global soil carbon has been estimated in excess of 2,000 Pg in the top 1 m of soil, much of that outside of forests (Batjes, 1996). While below-ground stores and processes are largely invisible from space, there is a strong need to link below-ground stock and flux measurements with the aboveground focus of most Earth observation and forest inventories.

5.3.3 Woodlands, savannas, and grasslands

If we adopt an operational definition of forests as land with continuous tree cover, then it follows from ecological first principles that non-forest trees will have different allometric relationships than forest trees, for at least three reasons. First, plants growing in non-forests experience less light competition once they exceed the ground layer. The reduced above-ground competition drives different responses to the trade-off between investing in stem height versus lateral growth, resulting in altered ratios of crown:stem biomass investment (cf. e.g. Bonser & Aarssen, 1994 for such plasticity in response to light competition). Likewise, in relatively short, disturbance-affected forests in parts of western Amazonia, crown:stem ratios for large trees are much greater from those in taller rainforests, with consequences for biomass (R. C. Goodman et al., 2014). Second, plants growing in non-forests will be limited more by deficiencies in the availability or seasonality of soil water or nutrients. This implies that different trade-offs exist between investing in growth above-ground and below-ground in forests and open systems, with the consequence of altered (increased) root:shoot ratios. Third, disturbance regimes that dominate open systems more than forests, notably fire and grazing, also heavily affect the allometry of trees in open ecosystems, helping drive above-ground ontogenetic changes and plastic responses to episodic disturbance (e.g. Archibald & Bond, 2003; Moncrieff et al., 2011), as well as enhanced allocation below-ground. Overall, field estimates of root:shoot ratios of tree-dominated neotropical vegetation summarized by Hoffman, et al. (2003) range from 1.0-2.9 in savanna, to 0.42–0.84 for dry forests, and as low as 0.1-0.17 in rainforests. Similar large differences apply globally, although savannas remain chronically data-poor in terms of detailed

biomass assessments (e.g. Mokany et al., 2006). In short, because forests and non-forests are so different ecologically, allocation of biomass above- and below-ground and to different above-ground compartments, also differs greatly. **It is recommended** that protocols devised to infer tree biomass in forests are evaluated for application to their non-forest equivalents, since it is expected that they will need substantial modification.

Very-high-resolution optical observations potentially permit large-scale, highly resolved mapping of large numbers of individual trees. In open systems, tree-centric biomass estimation is particularly promising as the scarcity of overlapping crowns reduces confusion, permitting accurate measurement of crown projections (cf. Bastin et al., 2017). By combining field data with crown metrics derived optically (or crown and height metrics from lidar point clouds (whether UAV, airborne, or spaceborne)), it is, in principle, possible to estimate the diameter and aboveground biomass of remotely sensed trees - if the species of each tree is also known (as identity is the main driver of wood density, which strongly determines AGB). Nevertheless, allometric prediction of biomass from crown - or crown and height - dimensions developed for forest trees (Jucker et al., 2017), entails systematic error if applied to estimate tree biomass outside forests, as allometric relationships differ. Improved estimation of non-forest tree aboveground biomass therefore requires developing bespoke allometric relationships for non-forest trees. **It is recommended** that for aboveground biomass estimates, these bespoke allometric relationships for non-forest trees be generated via a combination of TLS assessment and destructive harvesting. Both methods are needed because TLS can estimate wood volume but not wood density. Wood density, which also controls biomass, varies among species, within species, and within trees.

Regarding uncertainty, woodland biomass products would provide area and biomass estimates that would then be used in conjunction with models and other data to estimate total C stock changes in biomass, dead organic matter (woodland floor, coarse woody debris, etc.), and soil C. Uncertainty is generally propagated through inventory calculations, and optimally would be provided as variances or covariances in total biomass C for the spatiotemporal product (covariances in time and space).

5.3.4 Management in woodland ecosystems

In the US, most land management agencies with tenure over woodland systems (e.g., Bureau of Land Management, Natural Resources Conservation Service, USFS) are not concerned with changes in biomass or carbon. They are, however, interested in tracking changes in the spatial extent and canopy cover in these systems to support grazing and wildlife-related management goals. Given the strong linkage between tree cover and extent to biomass in woodland systems, such data products derived from remote sensing could be used to support these goals. The EPA's annual National Greenhouse Gas Inventory (NGGI) effort would be a primary end user of biomass or carbon data products in woodland systems in the US (S. Ogle, personal communication, n.d.). Biomass and carbon data products in woodlands could be used to address C stock changes due to woody encroachment, retraction of woodlands, agroforestry and other tree plantings outside of forest lands. Analyses from the national level (NGGI), to local municipal planning, would

certainly benefit from better data on biomass C in woodlands. In the US state of Arizona, the Altar Valley Conservation Alliance is using satellite data to understand the effects of prescribed fire on forage availability – biomass estimates would be a useful addition to this study.

In the US example, the geographic domains and spatial scale required by the NGGI or rangeland managers would depend on the analyses, but typical national scale analyses for the NGGIs evaluate changes from the 1980s through the most recent years (S. Ogle, personal communication, n.d.). An annual time scale is generally the focus of these types of assessments because smaller scale changes such as seasonal or monthly, are not used as the basis for carbon credits. Moreover, longer term changes can be derived from the annual data.

5.3.5 Soil/belowground biomass

The soil represents the largest terrestrial repository of carbon in the biosphere, storing ~3 times as much carbon as terrestrial vegetation. As the primary source of most macronutrients that are necessary for plant growth, the soil is critical for regulating aboveground productivity. But the soil also directly contains up to 50% of the plant biomass in most terrestrial ecosystems in the form of roots. In addition, soils also support the largest living and dead components of the organic biomass pools on Earth, as millions of species are responsible for decomposing and processing the dead biomass (primarily plant biomass) entering from the aboveground system. As such, soils form a highly dynamic component of the biosphere and a central component of Earth system dynamics.

The biochemical characteristics of the soil determine various ecosystem services including soil fertility, nutrient cycling and water filtration, while the biophysical characteristics determine soil quality, stability and structure. High-resolution maps of these soil characteristics are therefore critical for guiding targeted land management decisions that are aimed at improving food production, carbon sequestration or city/town planning capabilities. A detailed understanding of the spatial patterns in soil carbon storage is also necessary for scientific efforts to understand the global carbon budget and its influence on the future climate.

One key feature of the soil system is the challenges inherent in mapping soil organic features across space: the soil is opaque. The fact that ecosystem biomass and carbon is not directly visible from above is clearly a significant challenge for the Earth observation age. In particular, most biomass and carbon in non-forest ecosystems is belowground. The root:shoot ratios in these ecosystems are often high, but vary greatly and remain poorly constrained.

To address the challenge of mapping soil organic features across space, geologists and soil scientists have traditionally made use of physical soil sampling and ground penetrating radar approaches to detect a range of soil characteristics including soil depth (Wollschläger et al., 2010), carbon storage (Hruska et al., 1999) or roots (Comas et al., 2017). However, these approaches are time consuming, expensive and spatially restrictive, limiting their applicability across broad spatial scales. As a result, most high-resolution, broad-scale soil mapping efforts

have been based on inferences that can be drawn from spectral information and vegetation mapping approaches (J. B. Fisher et al., 2016; Hengl et al., 2014, 2017). Moreover, the soil classes which largely control variation in below-ground carbon (at least soil carbon), are variable at multiple scales and in many tropical regions require higher observation density than currently available for reliable interpolation (e.g. Moulatlet et al., 2017).

The approaches for inferring soil information from satellite data can vary, depending on the characteristics in focus. Some soil information can be inferred directly from observations of the aboveground system. For example, given the well-characterized obligate associations between specific tree taxa and their symbiotic mycorrhizal fungi or nitrogen fixing bacteria, it is possible to estimate information about the soil microbiome based on tree dominance from spectral approximations of plant taxonomic identities (J. B. Fisher et al., 2016). Specifically, nearly all tree species form symbiotic relationships with one of two types of mycorrhizal fungi – arbuscular mycorrhizal) and ectomycorrhizal fungi – so the detection and mapping of mycorrhizal status over large areas can be possible using high-resolution spectral estimates of plant groups (J. B. Fisher et al., 2016). As arbuscular mycorrhizal- and ectomycorrhizal-dominated forests often have distinct nutrient economies, this information can provide valuable insights into fundamental soil processes such as nutrient cycling and fertility (Averill et al., 2014). In addition, root:shoot ratios are well characterized for many tree species. As such, belowground plant biomass can be approximated in many forest communities using information about aboveground biomass in combination with tree taxonomic information (Cao et al., 2014).

Although some soil information can be approximated directly from aboveground vegetation information, the vast majority of belowground information can only be inferred indirectly. Most spatially-explicit information about soil characteristics are generated using globally fitted models using spectral information as covariate data, e.g. SoilGrids (Hengl et al., 2014, 2017), the Harmonized World Soil Database (HWSD; Fischer et al., 2008), or the FAO Global Soil Organic Carbon map (GSOCmap; FAO and ITPS, 2018). “Raw” soil data are also available from the World Soil Information Service (http://www.isric.org/explore/wosis/accessing-wosis-derived-datasets). Of all these products, SoilGrids is the most recent and highest resolution product.

Given that plants are primarily responsible for the organic matter inputs to soil, vegetation characteristics can explain a considerable proportion of the soil carbon and nutrient accumulation across landscapes. In addition, plants interact with geological processes and climate conditions to determine soil physicochemical properties. Using this approach, the latest SoilGrids database draws heavily on MODIS 250 m data to generate global maps of most soil physical and chemical characteristics (including organic carbon, pH, texture, etc.) across multiple depths (up to 2 m deep). This information can provide a valuable tool for linking belowground and aboveground components of the ecosystem, and for predicting the ecological context that might be necessary for shaping plant communities across the globe.

Just as with above-ground carbon, below-ground pools are sensitive to climate-change, to land-use change, and to change in ecosystem productivity. In the twenty-first century the need for large-scale, integrated, multi-technique, long-term, whole ecosystem assessments is stronger

than ever. A model for this already exists. Between 1970 and 1985 the Brazilian government set about the daunting task of mapping its remote Amazon and other territories from the point of view of geology, geomorphology, soil, vegetation, and agricultural capacity. This project - Projeto Radar da Amazônia, known since 1975 as Projeto RADAMBRASIL – combined cutting-edge remote-sensing technology of its time (airborne radar) with an extraordinary ground-based effort which included thousands of soil samples and inventory plots across an area of several million km^2. This led to what are still by far the world's best resource maps for a large tropical territory. Half a century on, the challenge now to the global community has remarkable parallels to that faced by Brazil in 1970 – how can we combine a growing array of new and old technologies in a concerted way to see as precisely as possible, and as deeply as possible?

It is recommended that satellite-derived aboveground biomass estimates are checked for consistency with available data on soil properties and belowground carbon storage, including those products listed above such as SoilGrids (Hengl et al., 2014, 2017).

5.3.6 Biodiversity

Biodiversity researchers frequently use biomass estimates for improving ecological understanding of how biomass relates to biodiversity and for assessing the potential of ecosystem service co-benefits, where investment in biomass protection or restoration is expected to provide incidental protections or gains for biodiversity (Brancalion et al., 2017; Van de Perre et al., 2018; Venter et al., 2009).

The ecological question of how biodiversity influences biomass focuses on carbon and nutrient cycling by animals and how biomass production, more so than total aboveground biomass stocks, is enhanced and appropriated by herbivores (Duffy et al., 2017; J. Liang et al., 2016). Effectively estimating biomass production from space requires precisely geolocated, repeat measurements of vegetation structure. Such measurements are currently not well quantified using existing spaceborne lidar but are likely to be improved with the next generation of instruments, particularly the GEDI mission. The question of how standing carbon stocks relate to biodiversity is more tractable from a remote sensing perspective and, with the negotiation of REDD+, has yielded a large volume of studies (reviewed by Goetz et al., 2015; Mulatu et al., 2017). Much of the focus of this literature has been on tropical forest ecosystems because of their contribution to global biodiversity, large aboveground carbon stocks, and high rates of forest conversion to agriculture (Gaston, 2000; Watson et al., 2016).

There is still debate about how biodiversity is related to biomass. We know that old growth forests contain unique species assemblages (Barlow et al., 2007; Watson et al., 2018) and that as regrowing forests accumulate biomass, their species assemblages become more like those found in old growth (Barlow et al., 2007; Letcher & Chazdon, 2009). Because of this, biomass estimates have frequently been used as indicators of forest condition and biodiversity (Bustamante et al., 2016 Figure 1). However, this relationship is likely to result mostly as a consequence of biomass and tree species richness covarying as a result of succession advancing, more than any causal relationship between the two. When forests are studied in detail, it is clear that the biodiversity-

AGB association only applies at some scales (Di Marco et al., 2018) or is restricted to a specific range of biomass values (Ferreira et al., 2018).

In tropical forests, it is notable that (1) very high biomass forests can have remarkably low diversity (e.g. monodominant systems in central Africa and southern Brazilian Amazonia, (Marimon et al. 2020)) and (2) the tropical continent with the most species-rich forests (South America) has the lowest per area biomass forests of all (ForestPlots.net, 2020). It is clear too that across all old-growth tropical forests there is no strong association between biomass and tree species richness (Sullivan et al., 2017).

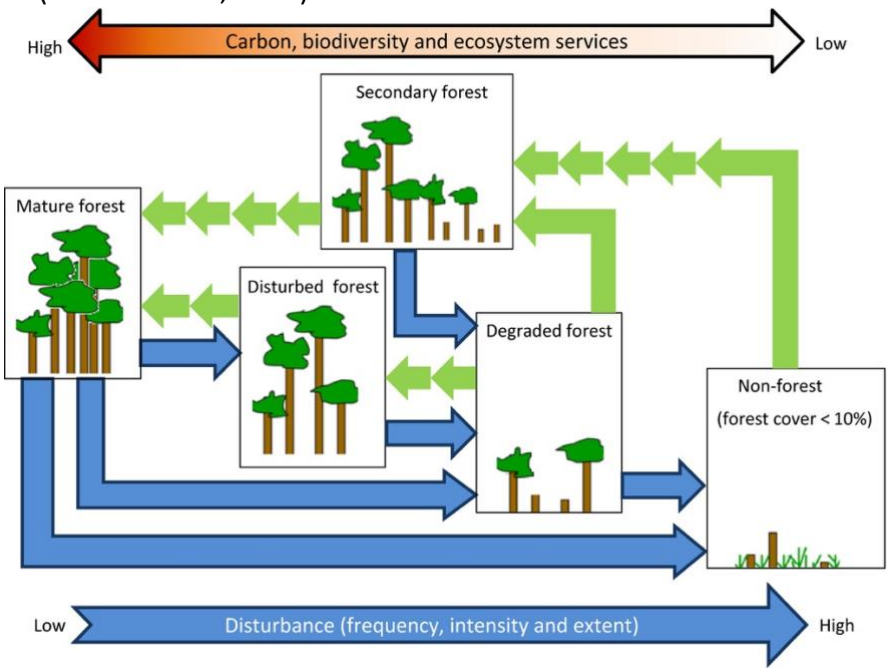

Figure 5.1. Conceptual model of the relationship between biomass and biodiversity (from Bustamante et al., 2016)**.** Note that the model does not apply to variation within old-growth forests, where the association between biomass and biodiversity breaks down.

Lasky et al. (2014) reported early successional forest plots average biomass of 76.4 mg ha^{-1}, mid-successional 116.5 mg ha^{-1}, and old-growth 198.3 mg ha^{-1}. Similar classes have been used to assess the regrowth potential of secondary forests in the Latin American tropics (Chazdon et al., 2016) and the potential for biodiversity-biomass co-benefits in Colombia (Gilroy et al., 2014). This implies that a biomass precision on the order of 10's of megagrams per hectare could be useful to distinguish age-classes relevant to biodiversity studies and conservation priorities. GEDI is expected to achieve accuracies of 20 Mg/ha (Goetz & Dubayah, 2011), suggesting a level of precision that will be useful for informing biodiversity-carbon studies and related decision-making needs. It is important to note again that the AGB-species richness relationship in the tropics breaks down in old-growth forests. Hence, AGB mapping can be a useful predictor of secondary tropical forest biodiversity, but AGB *per se* tells us little about biodiversity in old-growth systems.

It is recommended that satellite-derived aboveground biomass estimates be used for first-order prediction of tropical tree biodiversity, as long as old-growth forests can be first masked out using a reliable independent data source, such as data from INPE's PRODES project for Brazilian Amazon (http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes).

The spatial scale at which biomass information is needed for biodiversity assessments varies by application. Biodiversity sampling is time consuming and expensive (Bustamante et al., 2016) and scale sensitivity has been observed in the relationship between biodiversity and biomass (Chisholm et al., 2013). In many cases, establishing many small plots (< 0.1 hectare), may be the most efficient choice for multi-taxa biodiversity surveys, especially in heterogeneous landscapes (Gilroy et al., 2014). For example, once installed, camera trap grids are considerably more cost effective than repeat surveys. However, even one of the most extensive camera trapping efforts, the Tropical Ecology Assessment and Monitoring network, only deployed traps at a density of 1 trap/2 km^2. Such constraints could lead to scale mismatches between biodiversity observations and remotely sensed biomass estimates. Even so, with appropriate sampling design, the area represented by biodiversity plots can be large, on the order of several hectares (van der Sande et al., 2017), making comparisons with biomass estimates feasible. For larger payment for ecosystem services projects, such as those envisioned as part of a successful REDD+ strategy, the expected spatial scale of biomass estimates from spaceborne instruments may be commensurate with project needs. In a comprehensive assessment of restoration projects in the Atlantic Forest of Brazil, the project size averaged 10 hectares (Brancalion et al., 2018). Finer scale biomass estimates would allow for scaling effects to be explicitly considered, and may be desirable. It is important to recognise that human-impacted tropical forests are often extremely heterogeneous in space because human impacts (deforestation, fire, logging, subsequent use) are also very heterogeneous. This spatial variation adds challenges to the task of matching single and repeat satellite-observations to fine-grained on-the-ground processes.

Temporal scales needed by the biodiversity community also vary by application. Forest clearing results in immediate biomass decreases. In places where land use change occurs rapidly, annual observations would be most effective in documenting such changes. Biomass increases as forests regrow are much more gradual. Estimates of carbon accumulation of ~4 Mg of carbon per hectare per year in the western Andes (Gilroy et al., 2014) suggest that at expected accuracies of GEDI data products (Dubayah et al., 2020), a minimum of 5-10 years of regrowth would be needed to confidently detect change in biomass. Tropics-wide, the average annual rate of AGB increase in tropical forests varies widely, from 3.4 (Asia) to 7.6 (Africa) Mg/ha/yr in younger secondary forests, from 2.3 (North and South America) to 3.5 (Africa) Mg/ha/yr in older secondary forests, and from 0.7 (Asia) to 1.3 (Africa) Mg/ha/yr in old-growth forests (Requena Suarez et al., 2019). Clearly, AGB recovery rates decline with forest age. Performance periods for payment for ecosystem services projects may be anywhere from a few years to 30 years, thus a temporal resolution of 5-10 years may be acceptable and achievable in some forests.

It is recommended that satellite-derived aboveground biomass estimates report the uncertainty in (1) absolute values of AGB and associated metrics, such as canopy height, that are likely to change with forest age, (2) change over time in these metrics, and (3) the associated footprint

area and its precise location. Users will need to know the precise date when each metric was acquired, and for estimating biodiversity recovery rates, are likely to benefit most from sensors that measure the same location repeatedly over many years.

In summary, several existing biodiversity-biomass applications, especially those operating at scales > 1 ha, may be enhanced by next generation mission observations. Applications that require higher temporal or spatial resolutions than those provided by GEDI will likely need to rely on fusion approaches with planned spaceborne radar systems and/or existing high-resolution optical systems.

Chapter 6: User-led Validation with Pre-existing Reference Data

Valerio Avitabile, Jean-François Bastin, Sytze de Bruin, Laura Duncanson, Martin Herold, Inge Jonckheere

6.1 Introduction

This CEOS protocol has focused on a discussion of ideal good practices for spaceborne aboveground biomass product validation. However, oftentimes there are insufficient funds/resources to collect new data following the recommendations in Chapter 2. Existing and available data may provide reference information for biomass predictions that are not necessarily derived for the purpose of calibrating and validating biomass maps, and in this sense are different from the recommendations provided in Chapters 2 and 3. These data can be useful as reference data if their quality and characteristics allow for comparison with biomass map predictions (see e.g. Ploton, Mortier, Barbier, et al., 2020 for an example of valorisation of commercial forest inventories). These data can provide not only a quality assessment for the map producers but are often essential to build trust and to support the uptake of maps for certain end users and (local) applications.

Users of biomass maps include national inventory experts, global climate modelers, local climate change mitigation project implementers, sustainable supply chain managers, or environmental watchdog organizations. Each user category has specific needs when it comes to biomass estimation and accuracy requirements with respect to spatial resolution, geographic extent, timing, thematic content and definitions, as well as the type and standards of uncertainty reporting (see chapter 6; Herold et al., 2019). However, such requirements often cannot be assessed without considering additional existing reference data. In fact, it is often the comparison of the biomass map predictions with the user-owned reference data that builds confidence in the biomass map and determines whether and how satellite-based biomass data can be integrated into specific applications. These types of comparisons are often performed and this Chapter discusses the related methodological challenges and proposes good practices for doing so.

Existing reference data refer to quality forest data collected from various sources, produced for different purposes and using various methods. The main sources of forest biomass data include field plots, local maps and regional statistics produced by (regular) National Forest Inventories (NFIs), research forest plot networks and operational monitoring stations established for forestry, ecology or environmental purposes, as well as remote measurements of forest canopy and structure from satellite, airborne and/or UAV sensors. These data sources offer different types of validation opportunities – either through a comparison of regional averages from biomass maps with averages from statistics or gleaned from plot samples, or at the pixel-level, when high-quality ground plots and/or airborne lidar maps are available.

Existing datasets can be used as reference for the validation of biomass maps, assuming they are properly screened, processed and mutually harmonized with the biomass maps. Indeed, validation can be performed at pixel-level when the plot data satisfy the quality requirements indicated below (see 6.2.1), otherwise the validation should be performed using aggregated data at a regional scale (see 6.2.2.d).

It is recommended that the compilation and use of existing reference data is performed with caution and that careful attention is paid to their characteristics and specific uncertainties. The reliability of the validation results is affected by the uncertainty of the reference data themselves and the uncertainty due to the harmonization with the biomass map. Since the existing reference data have been produced for different purposes, they need to be carefully screened and harmonized to take into account their reliability and representativeness of the study area.

The reference data are estimates of biomass, and therefore are affected by errors, but they should be of higher quality than the map being validated, as recommended by the IPCC Guidelines (Herold et al., 2019; IPCC, 2019). In particular, imperfect conceptualization, measurement techniques and models of reference data are especially important, as they may cause systematic errors that are difficult to identify and quantify. Similarly, the lack of representativeness of data, in terms of spatial and thematic coverage (e.g., some regions or biomass classes may be missing or under-sampled), may lead to bias in the validation results.

The remainder of this chapter provides recommendations on screening, harmonizing and validation procedures for three types of reference data: field plots (Section 6.2), statistics (Sections 6.3) and local maps (Section 6.4). The following recommendations are mainly aimed at checking the quality of the reference data and harmonizing them with the map of interest to perform a proper comparison that minimizes differences due to variable definitions, spatial scale, temporal domain, etc. The Chapter does not include approaches to assess uncertainties of the reference data. All reference data contain uncertainties, and there is an increasing amount of literature and approaches available to quantify and propagate those. If practicable, an uncertainty assessment of the reference data should be conducted and included in the assessment of the quality and uncertainties in the biomass map and the related estimations (see Chapter 2 and Chapter 4). A case study showing the importance of reference data harmonization prior to product validation is presented in Box 6.1.

6.2 Recommendations when using field plot data

Forest plots provide estimates of forest biomass density at the sample location and represent the most common source of reference data for calibrating and validating biomass maps. Biomass is estimated at the tree level using tree measurements and allometric models, and the sum of tree biomass divided by the plot area provides the biomass density of the plot. Errors in the plot estimates vary depending on the field protocol and the appropriateness of the wood density and allometric models (see Chapter 2). If error sources are not independent, the plot estimates may present systematic errors that propagate to the validation of the biomass map. Further errors may originate when the field plots are compared directly to the biomass map due to various mismatches (e.g., spatial, temporal, definition) (see Chapter 3). Hence, it is essential to carefully screen and harmonize the field plot data with the map to obtain reliable validation results. An overview of the screening and harmonization process for field plots, in the form of a decision tree, is provided in Figure 6.1.

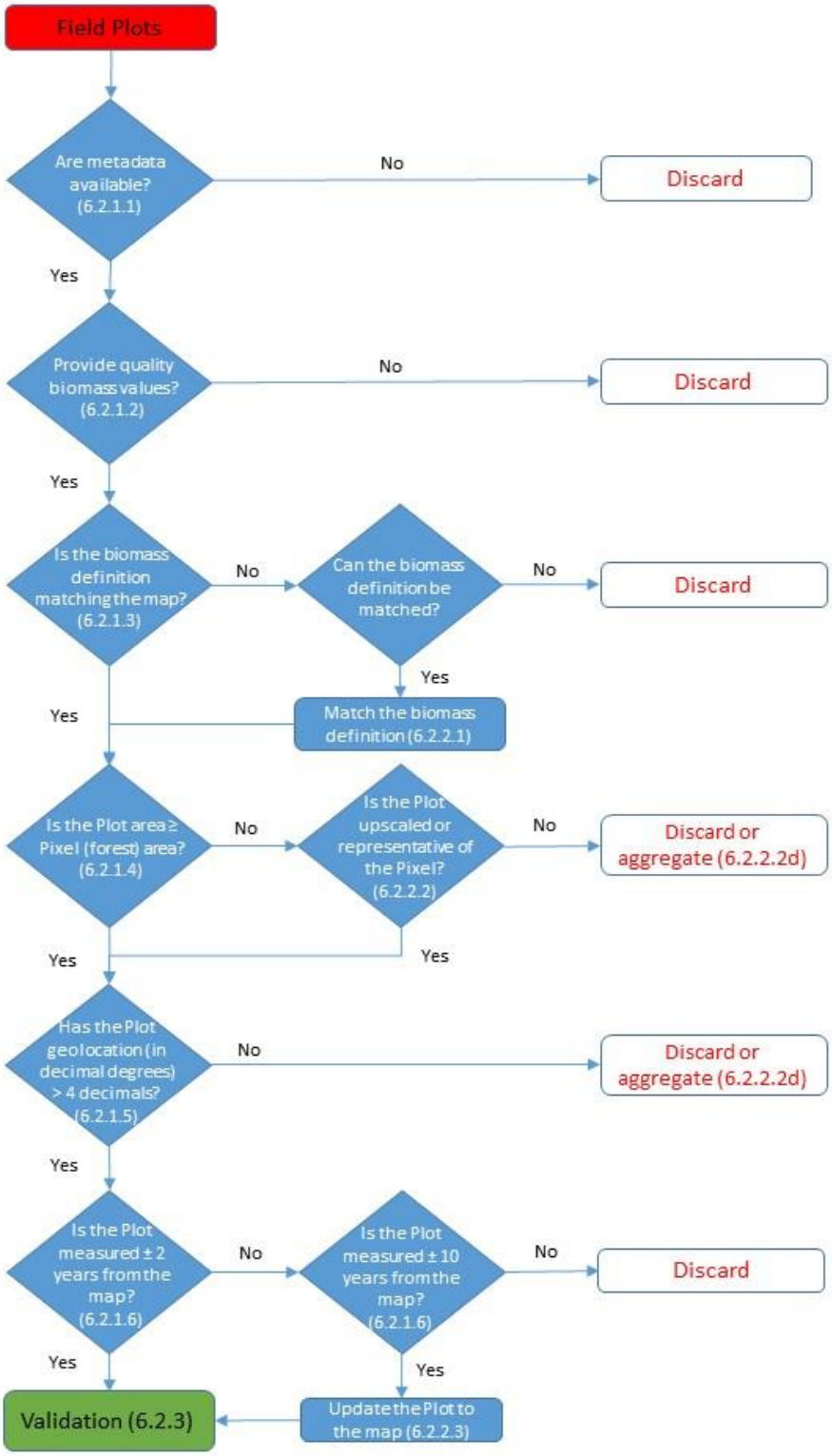

Figure 6.1 Decision tree for using field plots in the validation of a biomass map. The numbers refer to the sections of this chapter.

6.2.1 Screening of the reference plots

The recommended minimum criteria for using reference plot data for validation purposes are:

1. Metadata are available, including, at minimum, the definition of biomass (which parts of the tree are included), the method used to estimate biomass, the reference year(s) and the plot size. Optimally, the metadata also provides the tree parameters measured, the forest definition, the allometric model for biomass estimation, the plot shape, and the uncertainty of the biomass estimates are provided or can be estimated from the metadata.

2. The plot provides quality biomass estimates, meaning that the allometric model is considered appropriate for the forest type to which it is applied, and uses sufficient parameters (DBH and wood density and/or height).

3. The biomass definition matches, or it can be adjusted to, the definition employed in the biomass map being validated (see 6.2.2.1).

4. Plot area is larger, or can be upscaled, or is representative of the forest area within the pixel. The area of multiple plots located within the same map pixel can be summed. When the plot size is smaller than the pixel size (or, with pixels not fully forested, smaller than the forest area within the pixel), the plots need to be either upscaled using airborne lidar or other high-resolution data, or screened to assess if they are representative of the pixel (See 6.2.2.2). Even though NFIs usually employ design-based sampling designs and measure numerous but small plots (< 0.1 ha), larger plots (> 0.25 ha) are preferred for map validation because small plots present more plot-to-plot variation in biomass density, and errors associated with edge effects can be considerable, as the inclusion or exclusion of a single large tree can strongly influence the plot estimate. This is especially relevant in dense tropical forests where most of the biomass is found in a few of the largest trees (Bastin et al., 2015; Slik et al., 2013), which can easily be over or under-represented by small plots, causing large sampling errors (Réjou-Méchain et al., 2014). Moreover, geolocation errors cause larger discrepancies between reference plots and biomass maps in areas with high spatial heterogeneity in biomass (see Section 3.1).

5. Coordinates were acquired with GPS/GNSS devices with approximation at meter or sub-meter level (i.e., 4-6 decimals for coordinates in decimal degree). Plot geolocation uncertainty should be acceptable in relation to the size of the plot and of the map pixels (i.e. higher accuracy is needed with small plots and small pixels to reduce co-registration errors). Ideally, information on the number of GPS/GNSS measurements and how they were taken in the field (e.g. at the plot center or corners) should be available (see Section 3.1).

6. The year of the plot data collection is within ± 2 years of the biomass map reference year, or up to ± 10 years if the plot biomass is updated to the map using growth and mortality

rates (see 6.2.2.3), and change datasets derived from multi-temporal high resolution satellite observations indicate that no relevant forest change process occurred in the meantime. Forest change (e.g., deforestation, degradation or afforestation) can be identified through visual analysis of recent very high-resolution images (e.g., Google Earth) and/or from forest change maps and disturbance datasets (e.g. Hansen et al., 2013). The plots not representative of the state of the land at the time of acquisition of the satellite data used to produce the map should be discarded.

6.2.2 Harmonization of reference plots with maps

The following aspects of existing reference field plots (that pass the screening phase) need to be harmonized with the biomass map being validated:

1. Biomass definition

Field plots collected from various sources may provide estimates of different biomass compartments, and need to be harmonized to the same biomass definition of the biomass map using conversion or expansion factors (see Neumann et al., 2016; Vidal et al., 2016). For example, plot data may exclude the stump, the small branches and the foliage, or the trees below a certain DBH but may include palms or dead trees, while the map of interest may refer to the total aboveground biomass of all living trees. This also applies to reference data of Growing Stock Volume (GSV), which can be converted to biomass if appropriate conversion and expansion factors are available. However, the conversion of the GSV at stand level is less accurate than using appropriate allometric models at the tree level, and therefore **it is recommended** to perform the conversion at the tree level where possible.

2. Spatial overlap

Field plots and map pixels may present substantial spatial mismatches due to co-registration errors, the effect of incident angle, different plot sizes and shapes, and alignment between field plots and map pixels (see Section 3.1). When the plots are larger than the map pixel size, the map can be aggregated and aligned to match (as closely as possible) the plot area coverage. When the plots are smaller than the pixels, the use of field plots for the validation of the biomass map needs to be carefully evaluated. If available, airborne lidar or similar high-resolution data can be used as an intermediate layer to upscale the plot estimate to the pixel area using a multi-step approach (see Section 3.2). Otherwise, in order to perform a proper comparison of plots and pixels, the representativeness of the plots should be carefully evaluated (see below).

In any case, if plots and maps are provided in different geographic reference systems, first they need to be converted to the same reference system before comparison. Field plots can usually be treated as vector data and re-projected without deformation, and therefore it is preferable to use this option whenever possible. If the map needs to be projected (e.g., projection to an equal area reference system necessary for quantifying the biomass stocks), it is important to apply the most accurate approach (e.g., oversampling the map to a finer spatial resolution) because re-projecting a raster dataset always introduces some approximation in the resampling process.

Such approximations become more relevant with large map pixels (> 100 m) and when there is a large difference with the area of the field plots.

Then, the plot representativeness of the biomass density of the corresponding pixels should be assessed considering the following elements:

a) Alignment. When the plot is overlapping two or more pixels, the biomass density of the overlapping pixels should be computed as an area-weighted average, where each pixel contributes according to the fraction of the plot area located in the pixel.

b) Forest cover. If the pixel is not fully forested and the biomass outside forest is negligible, the biomass density of the forest plots should be multiplied with the fraction of forest cover (%) within the pixel to obtain the reference biomass density of the forest area within the pixel, which can be compared with the respective value provided by the biomass map. In case the biomass outside forest is not negligible, it should be estimated using a local or default biomass density value for the non-forest land cover classes and combined with the biomass density within forest using an area-weighted approach to estimate the reference value for the pixel(s). This step requires a forest map with higher spatial resolution than that of the biomass map. This correction is necessary only if the map provides the biomass density of the land area (i.e., the biomass density over the entire pixel) while it is not necessary when the map already applied a forest mask and provides the biomass density only of the forest areas within the pixel (i.e., non-forest areas are set to “no data”).

c) Representativeness. The plots should be screened to assess if they are representative of the forested areas located within the pixels. In fact, the biomass variability within a forest can be very large, and the small plots may not be representative of the mean biomass density within the forested area of the pixel. Representativeness may be assessed using a higher resolution map (Avitabile & Camia, 2018), radar images (Mermoz et al., 2015) or visual analysis of optical images (Avitabile et al., 2016) able to describe the biomass variability within the larger map pixels. The higher resolution product can be a map or satellite image representing or sensitive to a forest structural parameter related to biomass such as tree height or tree cover. The plots shall be discarded when they are located within pixels with substantial variability (e.g., large standard deviation) of the proxy variable (e.g., tree cover or height), and when the higher resolution map or images show that the plot is located in an area of the pixel that is substantially different from the mean conditions of the pixel. If multiple plots are located within the same pixel, the plots can be averaged. If these plots have different size, their biomass density should be combined with an area-weighted average, while a simple average can be used if the plots have the same size. If a plot overlaps multiple pixels, the plot representativeness should be assessed for all pixels where a substantial part of the plot is located (e.g., the pixels with > 20% of the plot area). Note that in a high biomass, closed canopy forest, multispectral data may appear relatively homogeneous despite considerable variability in biomass on the ground, and validation becomes particularly challenging. In such cases

with map biomass density higher than ~130 Mg/ha (P. Zhao et al., 2016) **it is recommended** to use plots that follow the recommendations from Chapter 2 and 3, if possible.

d) Spatial aggregation. In case the plots are not representative of the biomass density at pixel level or the representativeness cannot be properly evaluated due to the lack of higher resolution data, the plots should be either discarded, or aggregated and compared with the respective pixels at a coarser spatial resolution (e.g., 0.1° grid cells). The aggregation consists on computing the mean biomass density of all the plots located within the coarser grid cell and compare it with the mean biomass density of the corresponding pixels (i.e., not all the pixels within the 0.1° grid cell but only those where the plots are located). If the uncertainty of the plot biomass estimates is available, it can be used as weight to average the plots within the grid cell. In general, spatial aggregation allows reducing the uncertainty due to the spatial mismatch by comparing local average values (and thus, assessing systematic differences in mapped predictions of biomass at a coarser resolution), but does not produce an assessment of native pixel-level uncertainties.

3. Temporal domain

Field plots collected from different sources may refer to different periods and may differ by several years with the reference period of the map. Therefore, the plot data need to be adjusted to the year of the map considering the natural growth and mortality of the trees, given that plots affected by substantial forest disturbance were already identified and removed in the screening phase (see 6.2.1.6). Natural tree growth and mortality is relevant when a considerable period (> 2 years) occurred between the acquisition of the ground observations and the satellite data used to produce the map. However, the modification of the biomass values provided by the reference plots is likely to introduce substantial uncertainty and should be avoided, discarding the reference data from the validation exercise, unless it can be performed with local and/or reliable information. This temporal difference can be quantified and corrected using net growth rates (including mortality), or modelling approaches such as forest growth models. In absence of local/national and species-specific growth rates, the IPCC Tier 1 aboveground net biomass growth rates in natural forests (IPCC, 2019) may be used, if considered appropriate to the local forest conditions.

6.2.3 Validation approaches with reference plots

1. Inference methods

The choice of the validation approach should consider the statistical design of the reference data. Reference plots acquired by a NFI for biomass estimation are commonly based on a probability sampling design, and can be used for the validation of a biomass map for the same forest area represented by the NFI, applying conventional design-based statistical inference methods if the plots have negligible uncertainty, or hybrid inference methods if the plots have non-negligible uncertainty (see Chapter 4). However, the validation of biomass maps for areas not covered by (or not matching the area of) a NFI requires the collation of different plot datasets, that were likely acquired using a variety of plot sizes and sampling designs. Such complex amalgamated

samples should be treated as opportunistic (or ad-hoc) samples that require model-based inference frameworks. The validation approaches also need to take into account that NFIs are concerned only with the forest area of the given country or region while a biomass map may also include areas beyond forests and borders.

2. Validation metrics

The validation metrics used to assess the map accuracy can vary based on the user requirements (e.g. Chapter 5). In general, the comparison of the map and reference data should provide the uncertainty related to systematic deviation (or bias) and precision for the complete dataset (i.e., using all validation data) and by biomass class. Ideally, the validation metrics should provide a full error distribution as a function of reference biomass. Common metrics, such as bias and Root Mean Square Error (RMSE), assume unbiased reference data, and since this is difficult to prove, it may be more appropriate to refer to systematic deviation instead of bias, and Root Mean Square Deviation (RMSD) instead of RMSE. Systematic deviation refers to the systematic difference between the map predictions and the plot reference values, and equals bias when the plot values (which themselves are estimates) are unbiased. Systematic deviation is often the most significant error, and since it varies across the biomass ranges, it should be reported by biomass class. The RMSE/RMSD is a very commonly used metric, but when reported alone is not a strong indicator of uncertainty as it mixes systematic deviation and precision.

3. Propagation of uncertainties

Reference plot data are themselves not error-free and therefore comparisons between maps and plot data should be accompanied by an uncertainty analysis that takes into account the various sources of error (see recommendations from Chapter 4). The first step in such analysis is definition of the error model. The various error sources can be considered with an additive model expressing the unknown residual between a map prediction and plot value at location *x* as the sum of five components: the map biomass error at location *x*, the plot measurement error (which includes a plot size effect, see Réjou-Méchain et al., 2017), the positional error component, the within-pixel sampling error component (when the plot size is smaller than the map pixel) and an error introduced at a data harmonization step. Map pixels partly covered with forest undergo a harmonization procedure to scale to forest areas. Note that these five components are random variables whose values are unknown but can be described by probability distributions (Heuvelink & Snepvangers, 2005). Assuming the error terms in the error model are mutually uncorrelated, the total error variance equals the sum of the variances of the individual error terms.

6.3 Recommendations when using regional statistics

Reference data for the validation of biomass maps can be obtained from NFIs or forest management plans in the form of statistics (e.g., the mean biomass density of an area). It is important to consider that the statistics provided by the inventories are derived from field plots using an estimator appropriate to the probability sampling design of the plots, while the values from the map are obtained by averaging (in the case of the mean) all pixels within the region of interest. As in the case of field plots (Section 6.2), reference statistics also need to be screened to select reliable and accurate data and to be harmonized with the map to validate. Moreover,

when combining different reference statistics, such as NFIs from different countries or NFIs and forest management plans, the statistics need to be harmonized among each other if they present differences regarding the forest definition, the biomass definition and the temporal domain. An overview of the screening and harmonization process of reference statistics is provided in Figure 6.2 in the form of a decision tree.

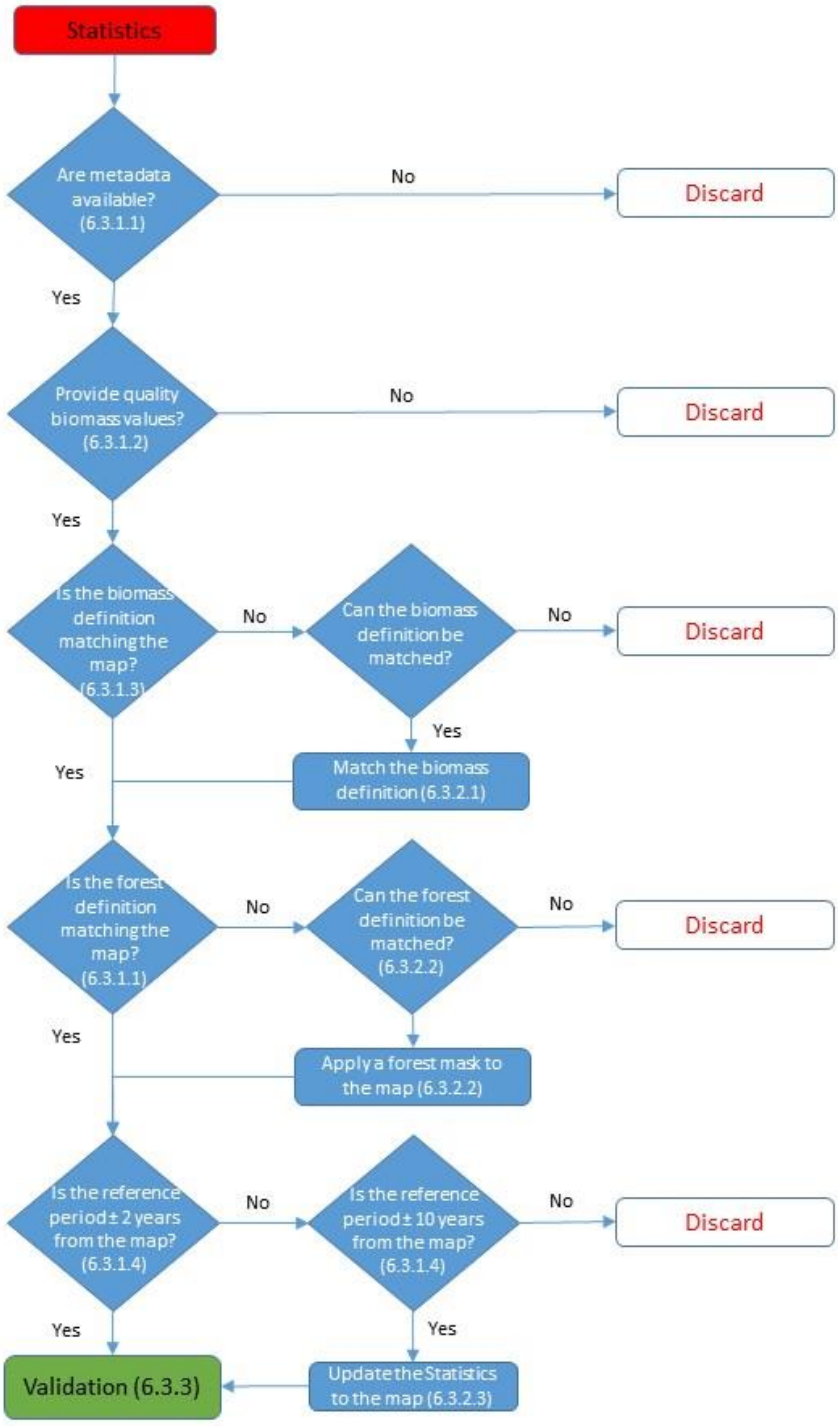

Figure 6.2 Decision tree for using regional statistics in the validation of a biomass map. The numbers refer to the sections of this chapter.

6.3.1 Screening of the reference statistics

The minimum criteria for using reference statistics for validation purposes mostly correspond to those indicated for reference plots:

1. Metadata are available as described in Section 6.2.1.1, and include also the number of plots used to develop the statistics.
2. The statistics provide quality biomass estimates, as described in Section 6.2.1.2.
3. Biomass definition matches, or it can be adjusted to, the definition employed in the map being validated (see 6.3.2.1).
4. The reference year of the statistics is within ± 2 years from the map reference year, or up to ± 10 years if the biomass statistics are updated to the map for natural growth, mortality and forest change, as described in Sections 6.2.1.6 and 6.3.2.3.

6.3.2 Harmonization of reference statistics with maps

Existing reference statistics that pass the screening phase may need to be harmonized with the map for the following aspects:

1. Biomass definition

Reference statistics, especially when coming from different sources, may refer to different biomass compartments, and need to be harmonized to the same biomass definition used in the biomass map as described for reference plots (see 6.2.2.1).

2. Spatial domain

Reference statistics typically refer to only forested areas, and thus a forest mask needs to be applied to the biomass map to compare the biomass stock and density of the map over the same (or comparable) forest area indicated by the statistics. The forest mask shall employ the same forest definition used in the reference statistics, and present a comparable forest area in the spatial units of the statistics (usually, sub-national administrative units).

3. Temporal domain

Reference statistics, as described for reference plots (see 6.2.2.3) may need to be adjusted to the year of the map considering the natural growth and mortality of trees, using net growth rates (including mortality), or modelling approaches such as forest growth models or carbon budget models. In the absence of local/national and species-specific growth rates, the IPCC Tier 1 aboveground net biomass growth rates in natural forests (IPCC, 2019) may be used, if considered appropriate to the local forest conditions. If forest change processes (i.e., deforestation, degradation or afforestation) are identified from existing forestry statistics, forest change maps and datasets, or through visual analysis of very high-resolution imagery (e.g., Google Earth), the reference statistics should be discarded, unless they can be corrected with local and/or reliable information. This correction requires the knowledge of the type of change that occurred (complete versus partial removal, or increment rate for new forests), the forest type where it occurred and the amount of biomass affected by the change. If the forest loss occurred after the reference year of the statistics, but before the reference year of the map, the statistics of the

corresponding forest type may be adjusted using forest growth models or simply in proportion to the forest area change.

6.3.3 Validation approaches with reference statistics

The recommendations related to the inference methods and validation metrics provided for field plots are also applicable to the reference statistics (see 6.2.3). In addition, the biomass map needs to be aggregated for each spatial unit of the NFI or management plan, such as a sub-national administrative unit or a forest stand, and compared with the respective reference value. However, these spatial units may have very different area coverage, which should be reflected in the validation approach by weighting the reference value according to the forest area of the spatial unit.

For example, the RMSE/RMSD of the biomass map can be computed as a weighted mean of the differences between the statistics and the map for each spatial unit, where the weights correspond to the forest area percentage of each spatial unit to total forest area. In this way, the reference statistics representative of larger areas will be given a corresponding weight in the computation of the map accuracy.

6.4 Recommendations when using high-quality local maps

Currently, there are large areas of the globe where forest plots and statistics are either missing or scarce due to the inaccessibility of the forest or lack of forest management and inventory systems (e.g., in some areas of the tropics). In other cases, reference data may exist but they are outdated (e.g., in boreal forests of Eurasia), or not accessible to researchers for various reasons. In such cases, a possible source of reference data may be found in local biomass maps that are considered of greater quality than the map being validated. Such datasets consist of biomass maps that usually have national or subnational coverage, are produced using local field plots and/or airborne lidar data, and can provide reference biomass values at a spatial scale intermediate between field plots and regional statistics.

These local maps can be used to obtain reference data using a sampling approach (e.g., systematic, random, stratified, etc.) to extract a certain number of reference map pixels (see 6.4.3). However, if the reference map has a substantial uncertainty at the pixel-level, but is assumed to provide estimates without systematic deviation at regional scale, it can be used to compute reference statistics (e.g., mean biomass density) over appropriate areas, such as administrative units stratified by ecoregions or forest types, and use these reference values for the map validation as indicated in Section 6.3.

An overview of the screening and harmonization process of reference maps is provided in Figure 6.3 in the form of a decision tree.

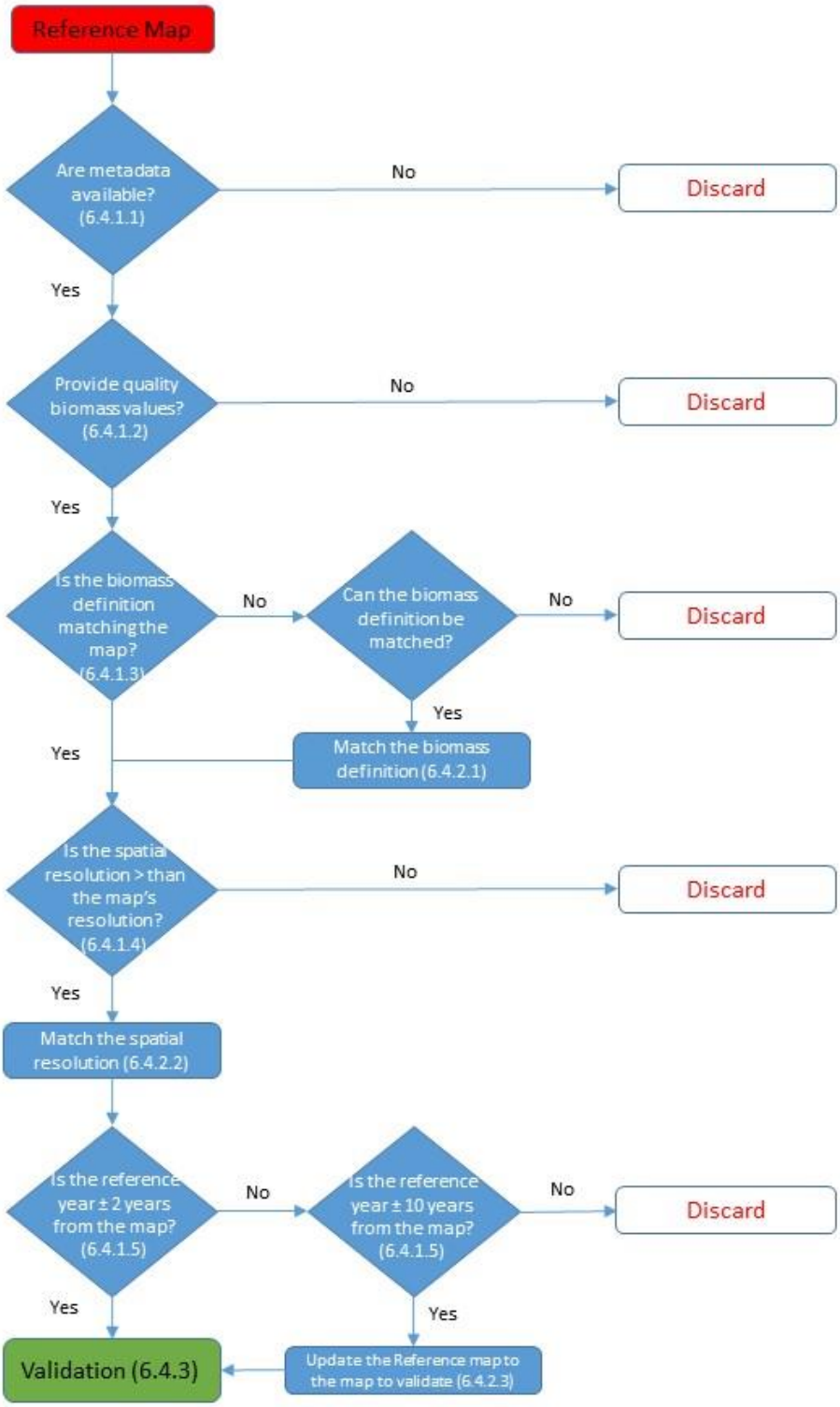

Figure 6.3 Decision tree for using local high-resolution biomass maps in the validation of a large-scale biomass map. The numbers refer to the sections of this chapter.

6.4.1 Screening of the reference maps

Local, high-quality biomass reference maps should present the following characteristics:

1. Metadata are available, including at minimum the definition of biomass employed, and the reference year(s). Optimally, the metadata also provide the forest definition, the allometric model for biomass estimation, the number of plots, the plot size, the remote sensing data (and the mapping approach used to develop the map), the uncertainty of the biomass estimates at pixel level and the accuracy at map level.

2. The reference maps provide quality biomass estimates, meaning that the map is calibrated with a sufficient number (e.g., > 30) of quality ground plots, possibly combined with airborne lidar data and/or high-resolution satellite data, and the input data and the mapping approach are appropriate to the forest type(s) to which they are applied (e.g., the maximum biomass in the area is below the saturation level of the remote sensing data).

3. The biomass definition of the reference map matches, or it can be adjusted to, the biomass definition employed in the biomass map being validated (see 6.4.2.1).

4. The spatial resolution of the reference map is higher than (or equal to) that of the map being validated.

5. The reference period of the reference maps is within ± 2 years of the map being validated, or up to ± 10 years if the reference map is updated to the biomass map being validated using growth and mortality rates (see 6.4.2.3).

6.4.2 Harmonization of the reference maps

Existing reference maps that pass the screening phase need to be harmonized with the biomass map being validated for the following aspects:

1. Biomass definition

The reference maps may refer to different biomass compartments, and will need to be harmonized to the same biomass definition as that used in the biomass map being validated using conversion or expansion factors (see Neumann et al., 2016; Vidal et al., 2016). For example, the map being validated may refer to the total aboveground biomass of live trees while the reference maps may exclude the stump, small branches and foliage, and the trees below a certain DBH.

2. Spatial overlap

The reference maps are expected to have a higher spatial resolution than the biomass map being validated, and will need to be aggregated and aligned to match the map being validated (as close as possible). If the maps have different reference systems, one will need to be reprojected before the aggregation. Since reprojecting a raster dataset always introduces some approximations in

the resampling process, and such approximations become more relevant with large map pixels, it is advisable to reproject the map with the higher spatial resolution.

3. Temporal domain

The reference maps may differ by several years from the reference period of the map being validated and will need to be adjusted to the year of the biomass map, considering the natural growth and mortality of trees, and the occurrence of forest change processes. However, such modification of the reference maps is likely to introduce substantial uncertainty and should be avoided, discarding the reference data from the validation exercise, unless it can be performed with local and reliable information and up to a temporal difference between the maps of ± 10 years. In such cases, it should follow the same approach indicated for the reference statistics (see 6.3.2.3). If the temporal difference between the maps is within ± 2 years, the temporal adjustment may be avoided, unless the forests are dominated by fast-growing species.

6.4.3 Validation approaches with the reference maps

1. Selection of the reference pixels

Only the cells with largest confidence (i.e., smallest uncertainty) should be selected from the reference maps and used as reference data. If the reference maps are based on empirical models, the map cells with greatest confidence are assumed to be those corresponding to the training data (field plots and/or lidar data). Uncertainty maps can also be used to identify the map cells with the smallest uncertainty. For maps based only on radar or optical data, whose signals saturate above a certain biomass density value, only pixels below such a threshold should be considered.

2. Amount of reference data for validation

If multiple reference datasets are available (i.e., including both field plots and high-resolution maps), in order to compile a reference database that is representative of the area of interest and well-balanced among the various reference datasets, the amount of reference data extracted from the reference biomass maps should be proportional to their area and not greater than the number of samples provided by the field datasets representing a similar area (e.g. Avitabile et al., 2016).

3. Statistical inference for validation

When using a biomass map as a source of reference data, it is important to first assess the accuracy of this map. McRoberts, Næsset, Liknes et al. (2019) provide guidance on using a local biomass map as source of reference data to assess the accuracy of a large-scale map. They considered that, even if the local map is of greater quality, it provides reference data that are likely to have non-negligible errors, and could be affected by systematic errors. Since the design-based estimators assume that reference data have, at most, negligible uncertainty (Snedecor & Cochran, 1967), they recommend the use of hybrid inferential techniques for assessing the uncertainty of the reference data extracted from a map. The hybrid inference consists of methods that combine model-based inference to estimate uncertainty due to the errors in the reference data, and design-based inference to estimate uncertainty due to sampling from the

finer resolution map (Ståhl et al., 2016). McRoberts, Næsset, Liknes et al. (2019) also suggest expressing the uncertainties of the map estimates as Mean Square Error (MSE), or Mean Square Difference (MSD), which, unlike MSE, does not assume unbiased reference data rather than variance because the map-based estimators of the means are not necessarily unbiased.

Box 6.1

Case study showing the impact of the harmonization of the reference data on the validation results.

We assessed the global biomass map for the year 2010 produced by the GlobBiomass project (Santoro et al., 2018) for the area of Europe using national statistics. We performed the validation four times, using different levels of harmonization of the reference data. First, we computed the total biomass stock per country from the map and compared the results with the most recent statistics provided by the National Forest Inventory (NFI) for 21 countries, using the relative RMSE (defined as RMSE divided by the mean value of the reference data) as measure of accuracy. The NFI reference data use country-specific definitions and parameters, and are not harmonized among themselves and with the map to validate in terms of forest definition, biomass definition, and reference period. Hence, in a second step we harmonized the NFI reference statistics to the same biomass definition employed by the biomass map using country-specific biomass correction factors (see Avitabile & Camia, 2018). Then, in a third step we further harmonized the statistics to the reference year of the biomass map (2010) using the Carbon Budget Model (CBM-CFS3) developed by the Canadian Forest Service and adapted to the specific European conditions (Pilli et al., 2018). Lastly, as the fourth step, we repeated the previous validations accounting also for the differences among forest definitions. This was done by applying a forest mask to the biomass map, in order to quantify the biomass stock of forest areas only. Among the numerous forest maps available for Europe, we selected the Copernicus Forest Map for the year 2012 because it was the map best matching the NFI statistics on forest area at European level. Nonetheless, in some countries, the difference in forest area between the NFI statistics and the map was still relevant, and thus we selected only the countries (16) where such differences were smaller than 15%, to reduce the impact of different forest areas in the validation results.

The validation results show that the agreement between the reference statistics and the biomass map increased at each step of the harmonization process, with the relative RMSE decreasing steadily from 24% to 15% when the reference statistics were harmonized for biomass definition, temporal resolution and forest definition (Figure 6.4). In other words, a substantial part of the difference initially found between the reference statistics and the biomass map was only due to the lack of harmonization and not to "errors" in the biomass map. This analysis shows the importance of the harmonization of the reference data for a fair and accurate assessment of the biomass maps.

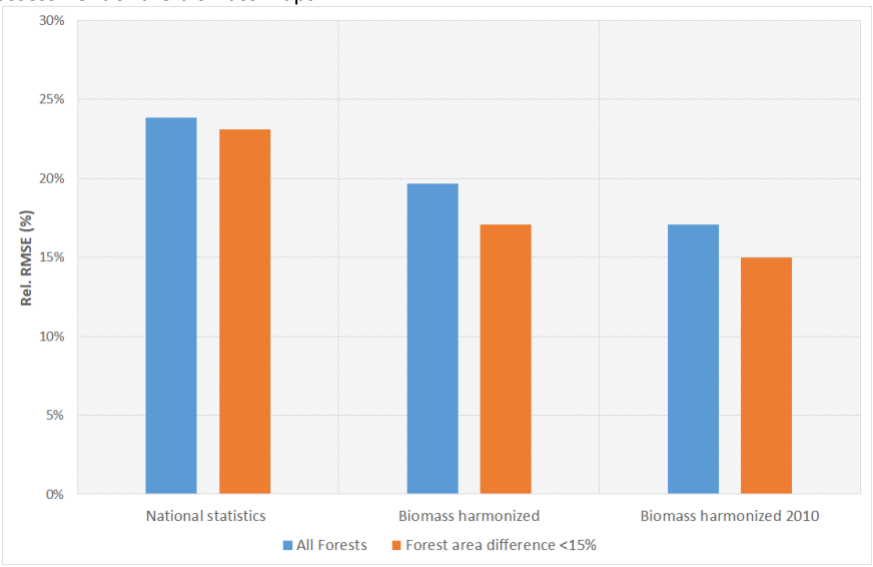

Figure 6.4 Validation results of the GlobBiomass map using the original national statistics ("National statistics", left), the statistics harmonized for biomass definition ("Biomass harmonized", centre), temporal resolution ("Biomass harmonized 2010", right) and forest area (orange bars)

6.5 Summary and current knowledge gaps

This chapter provides guidance for the use of existing biomass reference data to validate biomass maps. Validation data should be independent of the map production process and should ideally be coming from a design-based reference sample. However, the scarcity of biomass reference data often makes it necessary to use existing and available biomass data produced for various purposes, which may be based on a statistical design tailored to a different objective or study area. The collation of existing reference data is likely to produce an "opportunistic" dataset that may not fully represent the biomass variability within an area of interest. Hence, it is important to determine which ecoregions, forest types and biomass ranges are most under-sampled and targeted for additional data acquisition exercises. The use of lidar (from aircraft, UAVs and terrestrial) is an important evolving source of reference data for both calibration and validation purposes.

This chapter barely touches on the consideration of the uncertainty of the reference data itself, but this aspect is treated in more detail in Chapter 4. In fact, assessing the quality of the reference data is very important to obtain reliable validation results, and requires further attention and more examples and approaches to make it an essential part of any upcoming map validation exercise.

Chapter 7: Knowledge and Data Gaps

This chapter seeks to summarise the gaps and pressing areas for new data and research identified in previous chapters. The authors who have compiled the overviews in each case were asked to consider where there might be a need for new measurements, models and methods in order to improve our ability to both derive and assess AGB from EO data across scales and biomes. The gaps identified inevitably reflect the various different requirements when AGB is used for different purposes and applications.

7.1 Generation of reference datasets

A consistent framework is needed to define, quantify and propagate errors from the field data that underpins all reference data sets, through to the generation of AGB estimates from these data, as well as their use in assessing the quality of EO-derived estimates.

To achieve this requires that a number of issues are addressed. Firstly, accurate plot location (precise geolocation of EO estimates), agreement on how edge effects are to be dealt with, and co-location of spatial / EO-derived estimates. As outlined by Clark and Kellner (2012), a global database of biomass harvest plots will require a significant research commitment to improving our ability to quantify measurement errors in ground-based estimates of biomass, and how these may be minimised.

Key gaps here include:

Plot location

- Common consensus and guidance on how to accurately locate ground-based plots, their orientation, boundaries and 'edge effects' i.e. missing in/out trees. This is in addition to considering whether there is a 'desirable' size and/or shape of plots in different ecosystems and what those should be. The issue of edge effects will be particularly important for EO estimates derived from observations made at oblique angles (e.g. radar). The impact of edge effects is likely to decline with increasing plot area.
- How to deal with uncertainty in co-location of EO footprints is of particular importance. The location accuracy of EO footprints varies from product to product and so this needs to be considered in the definition of ground plot location.
- Consistency in measurement protocols for seemingly commonly-measured properties (H and wood density for example).
- Clear consideration of the trade-off between individual and stand-scale accuracy: at the stand-scale, sampling intensity has a much greater impact on precision of basal area (and hence, stand biomass estimates) than other factors such as the height of the stem, size and form of the individual and tool used.

Allometric models

- Agreement over the form and type of allometric regression models to be used, their definition, and clear statement of assumptions that are made in employing a particular model and how the estimates of uncertainty in the model parameters and predictions have been quantified. Tools exist to do this but they should be adopted as a matter of course.
- Consideration of the issue of local vs. general models (aggregated regionally or across species or biome type. etc.): locally-calibrated models may be better than general models in some cases, but not always or by much, and locally-calibrated models are often less robust, because based on small sample sizes.
- Wood density is used to convert tree volume into biomass, and is a critical information in forest biomass estimation. Wood density can be inferred indirectly from tree species identity combined with global species-level wood density averages. If possible, wood density is measured from wood samples.

TLS and UAVs

TLS, and more recently, UAV lidar, are acknowledged as being potentially transformative technologies to derive tree- and plot-scale estimates of tree volume and hence AGB. This opens the possibility for direct estimates of wood volume and AGB at the same scale as EO estimates; improved allometric models with many more tree samples; improved characterisation of the uncertainty in allometry-derived estimates. Calders et al. (Section 2.3) identify the following potential gaps and requirements for TLS and UAV data collection for AGB estimation:

- A need to systematically incorporate TLS as local calibration data to refit current allometric biomass equations, i.e., as more trees are sampled with TLS, allometric models should be revised with improved uncertainty estimates, NB this doesn't necessarily imply **reduced** uncertainty, but better-characterised.
- Systematic, ideally automated, location of stems in TLS and census data. This could be automated via markers, quick response (QR) codes etc., rather than manual (Disney, 2019), but at the very least, every effort needs to be made to match the stems in both samples to enable species and wood density attribution in the lidar data.
- Improved estimates of wood density to reduce uncertainty related to volume-to-mass conversions. Improved (semi-)automated tree species recognition would also help in this attribution.
- Tool development: full automation of structural retrieval, i.e., extraction of single trees from larger point clouds, and then volume fitting (or even volume fitting to larger plot-scale point clouds). Can we develop plot-level volume-fitting algorithms as a "plug & play" approach?
- Tool development: interface development to make volume-fitting algorithms more user-friendly and accessible, as well as automated over large numbers of trees.
- Tool development: a quality control (QC) framework that includes quantification of uncertainty of volume estimates at tree level, as well as uncertainty of plot-level AGB estimates.
- Tool development: routine integration of methods can bridge the (data) gap between ALS and TLS. UAV-LS should not be seen as a replacement for ALS or TLS (or both) but as a

new measurement somewhere between the two. Extracting information on AGB from UAV-LS will use elements of processing methods from both ALS and TLS, but will likely also require new dedicated methods to take into account the particular nature of the measurements.

7.2 Linking reference plots to satellite data

Réjou-Méchain et al. (Chapter 3) discuss the requirements for linking reference plots to EO estimates, particularly the impact and possible mitigation of mismatches in space and time. They make the point that a key requirement is a clear definition of context for the use of EO data, i.e., the purpose for which they will be used needs to be very clearly defined, as this will determine the minimum plot size, as well as the acceptable time gap between ground and EO data collection.

Réjou-Méchain et al. use the example of ESA's BIOMASS, with a 50 m footprint and a 30° view/illumination incidence angle (often used in radar to minimise impacts from soil roughness), can cause a large spatial mismatch with smaller validation plots. There is therefore a need for making co-registration accurate and reducing edge effects by increasing plot size. Larger plots will also tend to reduce temporal differences. Thus, the size of the plots and the spatial scale at which the RS product is calibrated or validated should be large enough to minimize these errors.

7.3 Linking field plots to spaceborne data using airborne lidar

Large and small footprint ALS have long been established as tools for near direct estimation of canopy height and cover, and subsequently through empirical modeling AGBD. Both capture landscape scale variation in structure, providing a bridge between EO and local scale observations (UAV-LS, TLS, field). Large footprint lidar systems (e.g. NASA LVIS) enable rapid and large area wall-to-wall or transect acquisitions (e.g. entire countries in some cases) but are of limited availability, however small footprint ALS systems are now capable of this scale of acquisition as well.

Key issues for global validation of EO products are the practical and logistical constraints on ALS data acquisition. Government restrictions on ALS data availability (e.g. China, India) and flight authorization (e.g. Brazil) can make access to existing data or acquisition of new data difficult. Depending on transit distance and areal coverage required, ALS data acquisitions can be expensive, with estimates on the order of \$250 – \$600 per km^2 (in US dollars) for small footprint ALS (Réjou-Méchain et al., 2019). Non-commercial instruments such as NASA facility instruments can acquire much larger areas than current commercial small footprint ALS systems for the same cost, however they are not as readily accessible and may involve large transit costs.

The following are recommended when producing a reference ALS AGBD map for EO biomass map validation. These are in addition to recommendations on linking field plot data with ALS data that are aimed at minimizing the impact of geolocation error and plot size (Section 3.2).

- Use recently-collected, high-quality, and well-vetted lidar instruments and survey configurations
- Acquire wall-to-wall data that are spatially and temporally coincident with field plots
- Develop local AGBD models using an area-based approach
- Adopt a statistically rigorous framework for reference map generation

We have identified the following gaps in methods development and the tools to enable a systematic validation with a global set of spatially and temporally coincident field plots and ALS AGBD reference maps:

- A quality control framework for AGBD map production
- Improved uncertainty estimates at the pixel level that account for multiple sources of error (see Chapter 4)
- Methods to aggregate AGBD uncertainties from ALS pixel level to the spatial resolution of EO map products
- Evaluation of ALS metric repeatability and biomass model performances when applied to ALS datasets from new instruments or acquisitions that are not coincident with field plots.
- A tool for propagation of *in situ* and ALS measurement and model errors (see Chapters 2 and 4) to the resolution of EO AGBD products that is accessible and standardized
- A tool for comparison of local maps derived from ALS, to global maps derived from EO, at multiple spatial resolutions and by user-defined strata.

7.4 Characterization and propagation of error

Roxburgh and McRoberts (Chapter 4) provide systematic definitions and assessment of the nature of error and uncertainty in estimates of AGB. They make the point that adherence to the IPCC good practice guidelines should be a key consideration of any measurement of AGB, namely: (i) neither over- nor underestimates so far as can be judged and (ii) uncertainties are reduced as far as practicable (GFOI, 2016, p. 15).

- Therefore, when developing new or using existing biomass maps, estimates based on those maps are required to comply with these guidelines. In general terms, this means ensuring that estimates are either obtained or evaluated using unbiased estimators, that is, satisfying criteria (i), and that all of the key error sources are recognised and accounted for, and that the methods applied are rigorous, appropriate, and quantifiable (satisfying criteria (ii).

Roxburgh and McRoberts make the distinction between design-based and model-based inference and point out that,on its own, traditional design-based inference is of little utility for creating maps of biomass (Section 4.3).

A total of thirteen possible sources of error are identified, spanning the breadth of allometric model development through to the calculation of final map biomass and its uncertainty. A

generalised Monte Carlo approach is given to show how these errors can be propagated through to AGB estimation. A key requirement therefore is the specification of these errors in any attempt to estimate AGB, and then how these are propagated to the final maps.

7.5 Considerations for utility of the CEOS biomass protocol: user communities

7.5.1 Utility for the land surface and dynamic vegetation modelling communities

Williams, MacBean and De Kauwe (Chapter 5) consider the requirements of AGB estimates for the various C-cycle modelling (land surface model and DVM) communities. They note the range of modelling scales (time and space) to which AGB is relevant e.g. to address issues of within-plant allocation of carbon and turnover rates of these issues, as well as the individual- to community-scale rate of productivity i.e. the difference between photosynthesis and carbon losses (respiration and turnover of plant tissues). They make several recommendations for the validation requirements of biomass products for these modelling efforts:

- That biomass products are integrated within community land model benchmarking schemes such as the International Land Model Benchmarking (ILAMB) project.
- To maximise value for DVM validation, clarity on the relationship between height and biomass estimates should be provided i.e. specification of whether these measurements are correlated or independent.
- Uncertainty metrics should be provided with/linked to all biomass and height data products. Metadata should explain the nature of the uncertainty estimate and its derivation.

7.5.2 Policy Relevance

Herold et al. (2019) discuss the policy requirements and evolving needs for biomass map data being properly and independently validated in a way that is technically robust (as described in this protocol) and useful for different policy applications and users. They make the point that a one-size-fits-all approach will not work, thus key requirements are:

- Development and assessment of biomass products should be performed by both producers and users in concert.
- National level biomass estimation is most relevant for integration of space-based information into country GHG inventories.
- The final biomass estimate needed by users varies considerably by application. It can be the total or average biomass per country or forest type or a local/pixel-specific value.
- Specification of whether a biomass estimate is needed for one time (i.e. one year) or for multiple years and whether it includes the estimation of biomass changes.
- uncertainty: compliant with IPCC good practice guidelines: (1) “neither over- nor underestimates so far as can be judged,” and (2) “uncertainties are reduced as far as is practicable” (IPCC, 2006).

For IPCC, GCOS, and for AGB ECV global maps

- Optimal: regular global forest biomass monitoring at 50-100 m resolution at annual intervals.
- Focus on long-term (i.e. decadal) and temporal consistency as input to global and regional climate and vegetation models
- Uncertainty and stability of ECV products are to be assessed and reported for relative and absolute systematic deviation and confidence interval or RMSE, overall and by biomass class/range estimated using reference data of better quality.
- Coordinated international activities for uncertainty characterization and reference data collection are required.

Herold et al. note that the reporting will be required at the national level for some schemes and policy frameworks, but regionally for others. They also note that the requirements for uncertainty, validation and reporting will likely differ substantially between the scientific and policy requirements. For example, in the context of REDD+ efforts:

- Uncertainty requirements vary by performance scheme.
- Consistency, accuracy and transparency should be sufficiently convincing for technical assessments (i.e. by the UNFCCC)
- Discounts in payments due to high uncertainty or conservative adjustments possible.

7.5.3 Utility of protocol by non-forest communities

Savanna ecosystems

Section 5.3 notes the importance in terms of AGB of regions not considered 'forest': *non-forests contribute less than 20% of global biomass, they represent half of terrestrial productivity and cover ca. 70% of the Earth's land surface* (Pan et al., 2013). The point is that regions classified as savanna are very often structurally dominated by trees, with, on average, at least 50% tree cover coexisting with C4 and C3 tropical grasses (Lloyd et al., 2008). As a result, there is a need for:

- Bespoke allometric relationships for non-forest trees should be generated via a combination of TLS assessment and destructive harvesting to enable AGB assessment of these trees.
- Consistency checks between satellite-derived aboveground biomass estimates and data on soil properties and belowground carbon storage.
- Definitions of "forest" and "non-forest" boundaries that allow for biomass dynamics to be tracked across these boundaries in a consistent and as unbiased way as possible.

Woodland ecosystems

- Woodland biomass products would provide area and biomass estimates that would then be used in conjunction with models and other data to estimate total C stock changes in biomass, dead organic matter (woodland floor, coarse woody debris, etc), and soil C.
- Uncertainty would be provided as variances or covariances in total biomass C for the spatiotemporal product (covariances in time and space).

7.6 User-led validation with pre-existing reference data

Avitabile et al. discuss the requirements and considerations for facilitating integration of the large amount of pre-existing reference data with new AGB estimates (Chapter 6). Reference data for the validation of large-scale biomass maps includes field plots, local high-quality maps and regional statistics. These pre-existing data are not necessarily produced or derived for the purpose of calibrating and validating biomass maps and so may not be compliant with the recommendations made in this protocol. Errors in the reference data and mismatches (e.g., spatial, temporal, definition) between the reference data and the biomass map can introduce substantial uncertainty in the validation results. However, with appropriate screening and harmonization procedures (described in the chapter 6), they can provide useful reference data for the EO-derived AGB estimates, as well as help to build trust and support the uptake of maps for certain users and applications. The importance of this latter aspect should not be overlooked. Avitabile et al. note in particular the following requirements:

- Screen the existing reference data according to requirements and quality/criteria for inclusion to create “opportunistic” reference datasets appropriate for the validation objectives.
- Harmonize the reference data with EO-derived maps by matching, as much as possible, the biomass definition, the spatial scale and the temporal domain of the two datasets.
- Stratify the existing reference data by ecoregions, forest types and biomass ranges to identify the strata that are most under-sampled and hence in most need of additional reference data.

Chapter 8: Implementation Considerations

This protocol has included specific recommendations for both biomass product developers and users. For the former, recommendations include how to report uncertainties when producing a map, including uncertainties from field measurements, allometric models, geolocation, and propagation to remote sensed products (both airborne as reference maps and spaceborne as global biomass products). For users aiming to interpret the uncertainties reported in products, the protocol should inform the process by which these uncertainties were estimated. For users wishing to conduct their own independent validation, the protocol makes recommendations for how to collect new reference data and use it to estimate uncertainties in products. There is also a chapter focused on recommendations for using existing datasets for product evaluation, with caution for common pitfalls in this process.

For all of the above protocol uses, the fundamental limitation of both product generation and validation is the availability of high-quality, globally-distributed reference datasets. While many such datasets already exist, there are considerable spatial data gaps that limit the ability to estimate uncertainties in new and forthcoming biomass products over certain geographic domains. Even in well sampled areas, reference datasets require updating to be contemporaneous with satellite data used in biomass map production.

Further, it is apparent that product validation activities (both by developers and users) require consistency and transparency if the reported uncertainties are to be trusted. To mobilize the use of biomass products for many applications (e.g. as inputs to carbon cycle models or use in national contributions to policy commitments) we require transparent, open source tools.

To implement the recommendations of this protocol, we require new and updated reference data (including field plots, TLS and airborne lidar datasets), and the development and refinement of tools to use these new measurements for product validation.

8.1 Collection of new reference data

It is recommended to collect new reference data at sites following the recommendations in Chapters 2, 3 and 4. Sites that follow these recommendations are termed ‘Biomass Reference Measurement (BRM) sites’, existing field stations selected based on the following requirements: (1) Availability of at least 10 already established 1 ha permanent sampling plots, according to the best forestry standards. Within the plot, each stem is mapped, its diameter measured, and its species is identified. The plots must have been inventoried in the past (multiple censuses) and be accurately geolocated; (2) Potential for airborne lidar scanning (ALS) coverage over at least 1000 ha, flown over the permanent plots; capacity to conduct new airborne lidar scanning coverage on a regular basis; (3) Potential for terrestrial lidar scanning (TLS) of a subset of the sampling plots; (4) Availability of a weather station and automated soil moisture monitoring (ideally encompassing the landscape-scale variation of soil moisture). These requirements are expanded upon below.

(1) Permanent forest sampling plots provide the most accurate estimates of aboveground biomass, which is inferred from highly precise field measurement through allometric models. A minimal forest sampling coverage is necessary to avoid bias (e.g., selecting only mature forests as representative of a diverse landscape). Many potential BRM sites have on the order of ten 1 ha plots (either in the form of several independent 1 ha plots, or one large plot), as this sampling intensity is manageable. Larger sampling intensities do exist but they are rare. In a forest, only trees greater than 10 cm in stem diameter are commonly measured, representing between 300 and 1000 stems per hectare. Accurate tree positioning, stem diameter measurement, and tree identification are all time-consuming and require skilled personnel on-site. In regrowing forests, a large fraction of biomass is held in stems smaller than 10 cm diameter, and the reference minimal diameter is often reduced to 5 cm.

(2) Airborne lidar scanning (ALS) has been intensively used for the upscaling of plot-based tropical forest biomass measurements. Calibration of the ALS data locally, with permanent sampling plots, results in accurate and precise aboveground biomass maps over areas of typically 1000 ha (10 km^2). This area is also typically within walking distance of the field station of the BRM site, and the biomass map can be thoroughly ground-truthed. Quality requirements for ALS data include high-quality global positioning of the acquisition, and a sufficiently high density of returns (at least 4 returns per m^2, but preferably at least 10 returns per m^2).

(3) Terrestrial laser scanning (TLS) surveys are an essential complement to sample plot data acquisition by providing: (i) an unbiased measure of wood volume, (ii) a reliable measure of total tree height; (iii) an accurate correction of stem geolocation (relative, at stand scale). This considerably increases the quality of the key plot data on which all of the other estimates rely.

(4) Ancillary data from weather stations or soil moisture probes are often needed in EO validation plans, and are easily obtained from established sites where human revisit is frequent. Continuous measurement of these environmental observables is needed during the overflight period of the missions. Data gaps should be minimized and gap-filling methods in place to mitigate instrumentation problems.

8.2 Potential Biomass Reference Measurement Sites

To minimize the cost, the selected sites should preferably belong to existing networks of field observation sites (Chave et al., 2019), because these sites ensure local leadership, established partnerships, data quality, local knowledge on environment and botany, and historical information. Capitalising on existing initiatives where previous expertise, capacity and infrastructure have been built also minimizes cost. This principle will also benefit the data quality and value, as many of these plots have a longstanding history.

Figures 8.1-8.3 depict sites that have been chosen carefully from global ecological networks and initiatives representing different forest types and that have the potential to be extended to become BRM sites. Appendix D provides the site details.

We propose to select 100 of these candidate BRM sites, based on their representativeness of environmental conditions, forest types and continents. Representativeness of a BRM site is evaluated by its contribution to major environmental conditions and global forest types. Based on the sites in Figures 8.1-8.3, it is thus possible to select the 100 sites that maximize representativeness, but also comply with the measurement and logistical criteria.

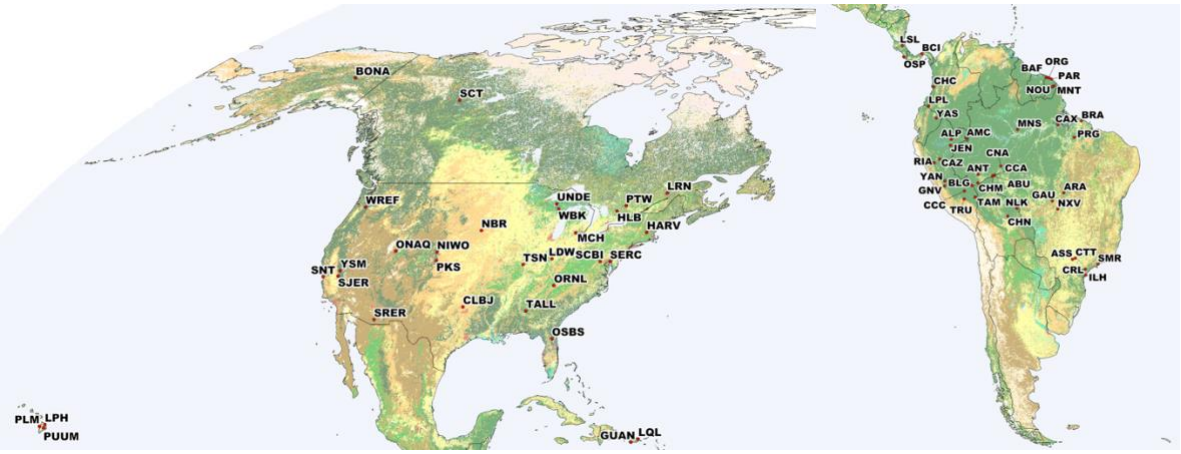

Figure 8.1. Location of the potential Biomass Reference Measurement sites for inclusion in the proposed Forest Biomass Reference System in the Americas (left:North America; right: Central and South America, as of March 2021. The background is the land cover CCI map for year 2015 (cf. https://www.esa-landcover-cci.org).

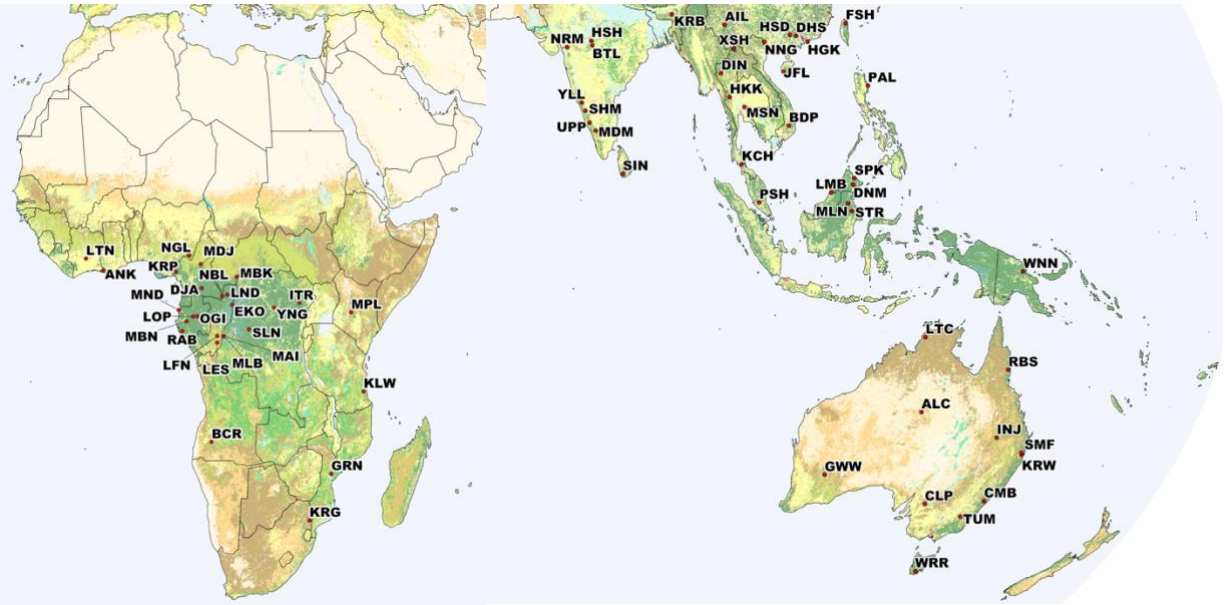

Figure 8.2. Location of the potential Biomass Reference Measurement sites for inclusion in the proposed Forest Biomass Reference System in Africa (left), tropical Asia and Oceania (right), as of March 2021.

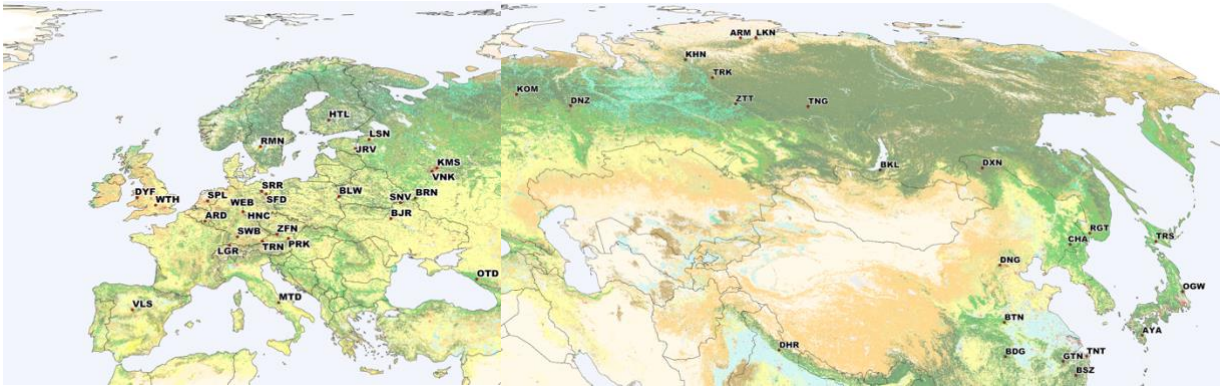

Figure 8.3. Location of the potential Biomass Reference Measurement sites for inclusion in the proposed Forest Biomass Reference System in Europe (left) and temperate Asia (right), as of March 2021.

How many BRM sites are needed? CEOS LPV has defined four stages of validation maturity for land products (https://lpvs.gsfc.nasa.gov). Currently, biomass is at stage 1, with an ultimate goal to achieve stage 4, but our immediate goal in establishing a CEOS Biomass Reference System is to achieve stage 3. We do not know at this time how many biomass reference sites will be needed to achieve validation stage 3. We can, however, state the level of coverage that is achieved for a given effort in relation to the key target variable (global coverage of forest AGB carbon). The Earth has a total forest cover of nearly 40 million km^2, and the key variables that affect forest AGB are all extremely variable, spatially. These include temperature, moisture, soil nutrition, soil depth, slope, gross and net productivity, as well as species diversity.

Furthermore, forests in near-identical climates on different continents have very different biomass, canopy structure, and tree form due to different biogeographical history and species composition. African, South American, Asian and Australian tropical forests share almost no species, have different biomass and growth rates, and different sensitivities to climate change (e.g., Hubau et al. 2020; Sullivan et al. 2020). Finally, forests are also subject to multiple human processes which alter structure and biomass, often creating very heterogeneous and dynamic landscapes.

Hundreds of sites are therefore desirable, but in practice, we have to start from the current situation with sites that already meet or come close to meeting our criteria. There are currently about 170 of these. This suggests that at least 100 'high-intensity' BRM sites (following protocol recommendations including field, TLS and airborne lidar collection) are an achievable target to implement the core of the proposed long-term Forest Biomass Reference System.

To first order, this ground effort should be distributed to represent the uneven distribution of biomass productivity and biodiversity across the Earth's forests (Table 8.1). This shows the overwhelming importance of adequate representation of tropical forests, which have high biomass and productivity and a very high diversity of tree species, structure and composition.

Accounting for the distribution of the key biological variables suggests at least 70% of the effort needs to be dedicated to tropical forests (wet forests, moist forests, dry forests and tropical montane forests). This is a minimum given the complexity of tropical forests.

On average, 100 BRM sites provide coverage of one site per 385,000 km^2 of forest world-wide (Table 8.1). This sample intensity will not satisfy all the needs of the EO community. The 100 sites should therefore be considered a minimum that CEOS should build on with an ambition to go further.

Table 8.1 Site Sample Intensity for each Global Forest Zone after allocating 100 BRM sites as 70:15:15 to Tropical: Temperate: Boreal. In each Zone, one site is allocated on average for the area, biomass, productivity, and diversity indicated here.

	Area (km^2)	**AGB (Pg carbon)**	**GPP (Pg C yr^{-1})**	**NPP (Pg C yr^{-1})**	**Tree Species**
Tropical	278,429	3.74	0.58	0.31	671.4
Temperate	511,333	3.11	0.66	0.54	91.6
Boreal	756,667	3.59	0.55	0.17	9.9
Earth's Forests	385,100	3.63	0.59	0.33	485.2

The level of sampling described above nevertheless leaves very large geographical gaps, especially in the tropics. For example, within high-biomass Amazonia, the distance between many proposed BRM sites is > 1,000 km. We know from previous work that biomass, canopy height, structure, species, carbon dynamics and human impacts all vary at much smaller scales than this. To fill these gaps, the addition of suitable sites will be assessed after the first round of product validation to get to CEOS validation stage 3.

8.3 Low-intensity sampling

The creation of a global Forest Biomass Reference System would be a critical advance for the calibration and validation of ongoing and planned biomass missions. However, as seen above, the spatial and environmental coverage of 100 sites is necessarily limited. Moreover, there is already considerable demand from the EO user community for a much higher plot sampling intensity to validate biomass products. This demand is explicitly recognised for example by the ESA-funded Forest Observation System which seeks to acquire high-quality tree-by-tree data from hundreds of tropical plots.

We therefore propose to define additional low-cost but highly distributed BRM plots, henceforth 'distributed BRM' sites. These sites comprise just single or a few long-term (multi-census) 1 ha plots, with no upscaling using ALS. They do not require the long-term infrastructure of BRM sites, nor do they need a high local density of permanent plots. If during CEOS stage 3 validation it is discovered that some of these distributed BRM sites are critical to validation, they could be promoted to BRM sites.

The key rationale for including distributed BRM sites is to provide much better strategic gap-filling than is otherwise possible.

The number of distributed BRM sites is still limited by what is available. Currently, there are at least 1,100 high-quality, networked, revisited tropical forest plots (including some >1 ha). Of these, up to 500 (700 ha) could be included in the 70 recommended core tropical BRM sites. But the remaining 600 plots are more widely distributed, alone or in small clusters, within more than 300 sites. Discounting those that are close to BRMs still provides a sufficient number of sites from which to easily select more than 200 additional distributed BRMs with an average of two plots in each.

It is therefore recommended to include 210 additional, distributed BRM sites within the proposed Forest Biomass Reference System. This will allow for optimal filling of large gaps between tropical BRMs, and achieve an overall tripling of the tropical site sampling intensity compared to that proposed in Table 8.1. Including these additional sites in the CEOS Forest Biomass Reference System, through periodic re-visit at each site, contributes to the global integration of the ground-based and remotely-sensed forest observation communities, while providing an 'easy win' to ensure sampling captures the complex variation in biomass and biomass change in the tropics. By designating these as distributed BRM sites, the Earth Observation community also efficiently leverages a large historic investment in monitoring some of the more remote forests on Earth.

8.4 Validation tools

To mobilize the ingestion of these new and updated reference measurements for validation of biomass products, we require open source, transparent validation tools. An example of the type of software that enables uncertainty estimation is the BIOMASS package in R (https://cran.r-project.org/web/packages/BIOMASS/index.html, Rejou-Mechain et al. 2017). The extension of these tools to include the measurement and allometric modeling uncertainties, outlined in Chapter 2, and propagated to airborne lidar biomass maps (Chapter 4) is highly desirable. We envision the creation and publication of other community-developed tools in the coming years that will enable transparent validation. Leveraging such tools, and creating novel approaches to product intercomparison and validation **are recommended**.

One such example is the new ESA NASA Multi Mission Algorithm and Analysis Platform (MAAP) (Albinet et al., 2020). The MAAP is a virtual, open and collaborative environment for the

processing, analysis and sharing of data and development and sharing of algorithms. The MAAP will provide a common platform with computing capabilities co-located with the data, as well as a set of tools and algorithms developed to support this specific field of research. The goal for this platform is to establish a collaboration framework between ESA and NASA to share data, science algorithms and compute resources in order to foster and accelerate scientific research conducted by ESA and NASA EO data users.

The MAAP is a virtual research environment that (i) will provide access to all BIOMASS, NISAR, ICESat-2 and GEDI campaign data in a unified format; (ii) includes software tools that allows users to implement and run algorithms in common open programming languages (Python, C, Fortran, R); (iii) makes available resources for processing and interacting with large volumes of data; and (iv) provides the tools for algorithm development (Eclipse Che, Jupyter Notebooks) and sharing (GitLab).

Merging the capabilities of the MAAP, or a similarly transparent platform, with both CEOS mission datasets and biomass products and new reference measurements will enable full traceability in validation activities for biomass. Users will be able to reproduce the validation results of products or other published product inter-comparisons, and thus trust their reported uncertainties. Further, if the uncertainty statistics do not suit a given application, they can be updated by a product user. This would therefore mobilize the integration of biomass products into the suite of applications outlined in Chapter 6.

8.5 Reference Measurement Recommendations summary

The table below summarizes the recommendations for new biomass estimation acquisitions with respect to field plots, airborne lidar (ALS), and terrestrial laser scanning (TLS). Table 8.3 highlights where these recommendations differ from established protocols.

Table 8.2 Summary of recommended specifications for new reference data acquisitions

<table>
<tr><td colspan="2">Recommendations for all data collections
Data should be free and open within at most 1 year after data collection
Data should be acquired in collaboration with long term field plot networks and local partners wherever possible</td></tr>
<tr><td>Field plot recommendations
Square plots
Easier to link to gridded products
Large plots (minimum 0.25 ha in tropics, ideally 1 ha plots with 25 or 50 m subplots)
Minimizes edge effects and geolocation uncertainties
Smaller plots (<0.25 ha) are acceptable outside of tropics provided airborne lidar available
Stem-mapped where possible
Geolocated with high accuracy, and reported uncertainties</td><td>Airborne Lidar Recommendations
Minimum 4 shots /m^2
Preferably acquired same season as field plots
Acquired within 2 years of field data acquisition
Repeated every ~5 years or when disturbance is detected
Covering at least 3 km^2
Cover both the plots and local environmental and forest structure gradient
Smaller area of coverage acceptable is only UAV lidar available</td></tr>
</table>

• Trained botanist should be employed for species identification	

Spatial Distribution of Field Plots
• Plots cover environmental gradients under airborne lidar collection that are locally or regionally correlated to biomass (e.g. topographic gradients)
• Sufficient number of plots collected to train a lidar model (min approximately 30, depending on complexity of system)
	Terrestrial lidar Recommendations
	• Data collection in new or existing long-term plots
	  ○ Data augments field measurements, does not replace them
	• 1 ha plots preferable
	• Data acquired in a grid pattern
	• Spacing 10 m in dense forests, 20 m in open areas
	  ○ Can be changed to ensure consistent sampling and minimize occlusion
	• Instrument must have ability to range tallest trees in 1 ha plots (150 m range)
	• Repeated ~ every 5 years or when disturbance is detected
	• Multiple scans need to be coregistered (either through use of targets or with sensor that has automatic coregistration)

Table 8.3 Biomass acquisition recommendations compared to existing network field protocols. Protocols have been rated as "Meets" or "Exceeds" for specifications that satisfy the CEOS recommendation. For specifications that do not match, the difference is noted. If a protocol does not include a specification, the cell has been left blank.

Protocol /Measurement	CEOS Recommendation	ForestGEO	RAINFOR-GEM- AfriTRON	TERN	NEON
Field plots					
Plot shape	Square	Square or rectangle, but large size also minimizes edge effect	Square or rectangle	Meets	Meets
Plot size	1 ha (0.25 ha minimum in dense forests)	Exceeds	Meets	Meets	0.16 ha, woody plants not sampled across entire plot
Subplot size	25x25 m	20x20 m	20x20 m	20x20 m	20x20 m, subplots not contiguous for aggregation
Stem map	Yes, where possible	Meets	Meets, but coordinates should always be measured precisely, not estimated	Meets	Meets
# of plots	> 10 - 30	Single large plot	Several to 20		Meets
Cover local gradients (topography, biomass range, etc.)	Yes	Meets, but less so than multiple plots across landscape	Usually	Meets	Meets
ALS sampling					

Shot density	4 shots/m^2 minimum			Exceeds	Meets
Area	~3x3 km minimum.			Exceeds	Exceeds
Return interval	~5 years				Exceeds
Time from field acquisition	<2 years				Meets
TLS sampling					
Plot size	1 ha			Meets	
Sampling pattern	Grid			Meets	
Spacing	10 m (dense veg.) 20 m (open veg.)			Meets	
Return interval	~5 years				
Instrument	Range > 150 m			Meets	

Appendix A: Field plot survey guide

The goal of this guide is to provide recommendations for field plot measurements so that they are as useful as possible for validation of aboveground biomass estimates from remote sensing. These guidelines have been synthesized from recommendations in this document, and survey protocols from ForestGEO (Condit, 1998), the National Ecological Observation Network (NEON; Meier, 2014, 2017, 2018; Thorpe et al., 2016), the Amazon Forest Inventory Network (RAINFOR; O. Phillips et al., 2018), the Terrestrial Ecosystem Research Network (TERN; A. White et al., 2012; Wood et al., n.d.), and the Socio-Ecological Observatory for Southern African Woodlands (SEOSAW; Ryan & Berry, 2020; the SEOSAW partnership, 2020). The guidelines will make general recommendations for plot establishment and data collection, then expand into specific recommendations for different biomes. It is important to note that these recommendations are extremely general and focused on collection from an EO perspective, and we anticipate collections will often differ from the specifics of these recommendations based on local and regional considerations and constraints. However, where existing local protocols are not available, this appendix provides a guide for collection of field data.

A.1 Plot establishment

The first consideration when establishing a field plot for validation of Earth Observing remote sensing missions is the nature of measurements taken by the mission. The plot size should be determined by resolution of the biomass product. Also, aboveground biomass may be defined as forest biomass, woody biomass, or vegetation biomass by various measurements, and field measurements should reflect this definition.

A plot could be measured only once, but would ideally be established as a long-term monitoring site. Either situation can be useful for validation as long as there is a temporal match between the field and remote sensing measurements. Especially in cases where a monitoring site is being established, stakeholder involvement should be considered:

- Local scientific participation
- Long term protection from human disturbance
- Long term institutional support

Where plots are established for the purposes of training airborne lidar biomass maps, a sufficient number of plots should be established to train a lidar model. It is important to ensure the range of sample values in the plot data and associated airborne lidar metrics are representative, i.e. they are similar in the sample as in the population. The minimum number plots necessary to develop a local wall-to-wall biomass map is dependent on the heterogeneity of the region, and is often between 10-30 plots, but may be greater.

A.1.1 Site Selection

The plot location should be determined before going to the field using maps or satellite imagery. An ideal plot location would be representative of the biogeography of the region. Plots should represent the range of environmental and structure gradients sampled by airborne lidar (e.g. topography, height, biomass, etc.), if available. Plots should be placed randomly within strata, such as those defined by regional ecosystem extent maps. This is important to avoid any bias towards large trees or aesthetically pleasing locations while in the field. To minimize the impacts of spatial and temporal mismatches between field plot and map product (see Section 3.1), placement of plots on or immediately adjacent to landscape boundaries such as roads or a forest clear-cut edges should be avoided. In addition, selection of field plot locations needs to consider land tenure to ensure plot access is possible. Ideally site selection will support long-term monitoring so that plot revisits are feasible.

A.1.1.1 Change plots

The guidance, provided in A.1.1, on plot selection makes sense when trying to create a reasonably representative set of training plots for remote sensing data. However, when trying to validate biomass *change* maps, different priorities may exist, as most plots will change only a small amount between censuses. Instead, there may be a desire to bias plot locations actively to areas where they will have more likelihood of being degraded (e.g. close to settlements, within logging concessions with loggers committing to logging them). Or the field teams can go further and actively select trees within their concessions to be logged, so varying fixed quantities of biomass are removed (as in the FODEX project for example, https://mitchardgroup.wordpress.com/fodex/).

A.1.2 Plot geometry

After choosing a location, plot geometry (size, shape, and orientation) should also be determined before visiting the field.

A.1.2.1 Size

- Minimum recommended plot size of 0.25 ha in tropical forests
- Ideal plot size is 1 ha with 25 x 25 m subplots
    - Large plots (≥ 1 ha) help reduce edge effects and mismatches between field and remotely sensed data (Réjou-Méchain et al., 2019)
    - 1 ha plots with 20 x 20 m subplots match the RAINFOR and TERN networks, but 20-m subplots are less useful for matching resolution of current and planned spaceborne biomass campaigns.
- Plots may be as large as 50 ha to match the size of ForestGEO plots
- Acceptable minimum plot size varies by biome (Table A.2)
    - See A.3 for biome-specific considerations
- Plot size should always match or exceed the resolution of the remote sensing product

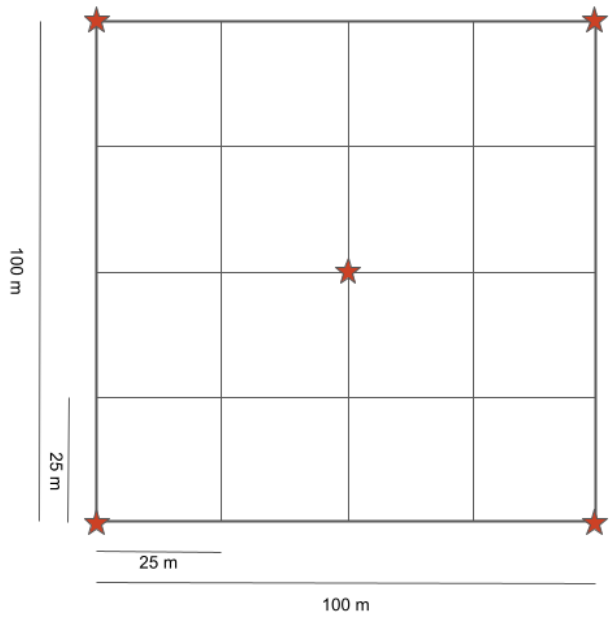

Figure A.1 Example layout of 1 ha survey plot. GPS should be taken from starred locations. (adapted from Wood et al., n.d.)

Table A.2 Minimum recommended plot size for different biomes. 1-ha plots are preferable when logistically possible, but smaller plots can be acceptable based on biome-specific considerations (see A.3).

Biome	Minimum plot size
Tropical forest	0.25 ha
Temperate forest	0.01 ha
Boreal forest	0.01 ha
Mangrove	0.01 ha
Dryland	0.25 ha
Savanna/Woodland	0.25 ha

A.1.2.2 Shape

- Large, square plots are generally recommended over circular or rectangular plots:
    - Square plots can be easier to set up, especially in the case of large plots
    - It can be easier to locate subplots and trees in square plots using a coordinate system
    - Square plots match the shape of remote sensing pixels

        - Reduced edge effects in comparison to rectangular plots
            - Less uncertainty about which trees are in or out of plot
- Circular plots are also acceptable, and reduce edge effects further than square plots. However, these are less preferable to square plots based on the considerations listed above.
- Over slopes, plots may require adjustment for matching to remote sensing pixels. At minimum, the angle and direction of the slope should be recorded.

Nested plots are quite common, particularly in a forestry or ecology context. These are often but not always circular, with different tree diameter thresholds for different sections. These should be avoided where possible in a carbon and remote sensing context, as the multiplication up of the small parcels can lead to high errors in the overall biomass estimates, and they are highly prone to errors in the field or in the data analysis stage.

A.1.2.3 Orientation

- N/S orientation preferable when possible, accounting for magnetic deviation
- Alternative orientations are acceptable if stand or topographic features do not allow for a N/S orientation

A.1.3 Pairing field and remote sensing data

A.1.3.1 Geolocation

Proper geolocation is essential for matching field and remote sensing data. Once a plot's location and geometry have been determined, the plot can be set up in the field (see Condit, 1998; Wood et al., n.d. for survey techniques). The latitude and longitude of the plot center and all four corners should be if possible geolocated using a differential GPS, though note that our experience suggests that differential corrections often do not work under a dense forest canopy, suggesting a high specification standard GPS is necessary (Chapter 3.1). To minimize geolocation errors caused by taking GPS measurements beneath the canopy, no less than 20 GPS acquisitions should be averaged at each point (Réjou-Méchain et al., 2019), ideally with averaging done by revising the points at several times in different parts of the day/over different days (see Chapter 3.1).

A.1.3.2 Timing

The temporal difference between field and remote sensing measurements should be minimized. For plots that are to be measured only once, this means taking field measurements as close as possible to a remote sensing acquisition. For long term monitoring plots, a sampling schedule should be set and adhered to, with field measurements occurring at the same time of year for each census. The season of data acquisition should be noted to facilitate linking to remote sensing data as phenology will have important impacts on vegetation structure.

Airborne lidar or UAV-based acquisitions **are recommended** to be repeated every 5 years or when a disturbance has been detected. Field measurements should be collected at least as frequently as airborne lidar data, and preferably in the same year as airborne lidar data. Field data acquisitions may be more frequent than lidar acquisitions. Common sampling schedules in established networks are:

- ~1-5 year sampling interval (1 ha-plot, RAINFOR)
- 1-3 year sampling interval (0.16-ha plot, NEON)
- 5 year sampling interval (10-50-ha plot, ForestGEO)

As plot size increases toward the size of the remote sensing pixel, the error associated with a temporal mismatch between field and remote sensing data decreases (Réjou-Méchain et al., 2019), but the mismatch should always be as small as possible.

A.2 Tree measurements

Field plots should be stem-mapped whenever possible, as this allows partitioning larger plots to the specific shape and size of remote sensing observations. Individual tree measurements, which underpin all field-based biomass estimates, should of course be as accurate and unbiased as possible. Within a plot, all stems ≥ 10 cm in diameter should be measured. These stems should be mapped by either:

- X and Y coordinates from SW plot corner (RAINFOR)
- Distance and azimuth from designated points within the plot (NEON)

If multiple censuses will be taken, stems should also be tagged with a unique identifier. With each new census, it is also important to include any stems that have newly crossed the 10 cm threshold ('new recruits').

A.2.1 Diameter measurements

The default height for diameter measurement is 1.3 m above the ground (breast height). There are a number of situations where this point of measurement may be altered because of unusual tree shape (see Condit, 1998 for details) or for savanna trees where sometimes 0.3 m may be more appropriate (Ryan & Berry, 2020).

A diameter tape is the preferred tool for measuring diameter, but a Stepped Diameter Gauge may be more precise for irregularly shaped stems. Errors in diameter measurements can be detected through repeated censuses (Chave et al., 2004).

A.2.2 Height measurements

Measuring tree heights within a plot can be used to develop local height-diameter models. Plot-specific height-diameter models reduce error in over- or underestimation of height caused by reliance on regional or global height-diameter. Refer to Chapter 2.1 for uncertainty considerations regarding height estimation, as well as (Sullivan et al., 2018).

Errors in height measurements may be reduced by incorporating lidar-derived height metrics. This is especially true if tree height has been measured with a laser rangefinder using the "sine" method, when lidar coverage can be used to quantify the bias in height (Réjou-Méchain et al., 2019).

A.2.3 Data recording

Below are the minimum fields that should be included in field measurements. Sample data sheets can be found at ForestPlots.net.

- Trees
    - Plot ID
    - Subplot ID
    - X and Y coordinates in m
    - Tree ID
    - Family and species
    - Diameter in mm
    - Point of measurement
    - Health status
- Plot
    - Latitude and Longitude
    - Elevation
    - Bearing of plot axis

Errors in recording tree measurements can lead to downstream over- or underestimation of biomass. To reduce errors, double entry and proofreading of field measurements is recommended (Condit, 1998).

A.3 Biome-specific considerations

A.3.1 Tropical forests

A.3.1.1 Diameter measurements

In tropical forests, trees with buttress roots and lianas growing against stems are common, and can result in overestimation of diameter measurements. To reduce this bias, a consistent protocol should be followed:

- Lianas: when measuring a tree, diameter tape should be slid underneath lianas
- Buttress roots: the diameter point of measurement (POM) should be 50 cm above the top buttress
    - Carefully record POM
    - If buttress is within 30 cm of POM
        - Measure original POM
        - Establish and measure new POM 50 cm above original POM

Figure A.2 Trees with buttress roots (left) and lianas growing against the stem (right) at Barro Colorado Island, Panama.

A.3.1.2 Timing

To reduce errors in diameter measurements caused by variations in stem water content, it is best to measure tropical forests during the wet season, when water availability is most consistent. At minimum, remeasurements of seasonal forests should be done at the same time of year as the previous censuses. The error in not doing so is not negligible: reliable plot measurements in Lope National Park, Gabon, show large trees shrinking by as much as 6 cm over a few months from wet to dry season.

A.3.2 Temperate forests

A.3.2.1 Sampling design

In temperate forests, plots smaller than 0.25-ha are acceptable, as long as they are accompanied by airborne lidar. As always, matching the remote sensing measurements should be considered when deciding plot size and sampling interval.

- Plot size should match or exceed the resolution of the remote sensing product

- Minimum recommended plot size is 10 m radius, or 20 x 20 m square plots, but larger plots are highly desirable where practicable and especially in high biomass forests
- High-quality geolocation is necessary for small plots
- Temporal proximity to remote sensing acquisition is especially important for small plots

A.3.2.2 Plot type

The use of variable radius plots should be avoided if the prime purpose of the plots is for comparison or cal/val of remote sensing products. This plot type is historically common in temperate forests, but because it does not have defined plot boundaries, it is difficult to pair with remote sensing data.

A.3.2.3 Timing

For consistency between field censuses, diameter measurements should occur after the end of the growing season.

A.3.3 Boreal forests

A.3.3.1 Sampling design

In boreal forests, because of greater homogeneity of stand structure and composition, and slow growth rates, plots can be smaller and sampled less often. However, because of landscape patchiness from disturbance and recovery and extreme micro-topographic variation, these plots must represent the range of stem densities and tree cover that are controlled by this variation. Furthermore, the measurements should include woody shrubs. Plots smaller than 0.25-ha are acceptable, as long as they are accompanied by airborne lidar or part of a large extensive sample that accounts for landscape variation in woody structure. As always, matching the remote sensing measurements should be considered when deciding plot size and sampling interval.

- Plot size should match or exceed the resolution of the remote sensing product
- Minimum recommended plot size is 10 x 10 m
- High-quality geolocation is necessary for small plots
- Individuals with DBH ≥ 3 cm should be measured
- Temporal proximity to remote sensing acquisition is especially important for small plots
- In NEON boreal sites, sampling interval is increased from 3 years to 6 years
- Plot locations and bounds should be placed carefully to account for micro-site variation in tree stem densities.

A.3.3.2 Bryophyte sampling

Bryophytes such as *Sphagnum* spp. form dense mats that significantly impact above ground productivity in boreal systems. Sampling bryophyte biomass should be considered for sites with high bryophyte cover (see Meier, 2014 for sampling design). However, the applicability of bryophyte biomass measurements to validation of remote sensing products will vary (i.e. more applicable to spectral data than lidar)

A.3.4 Mangroves

In mangrove forests, plots smaller than 0.25-ha are acceptable, due to the difficulty of making the measurements (moving and standing water, aboveground root systems, muddy soils), however larger plots are recommended. Plots should be co-located with airborne Lidar measurements and the number of plots should be stratified by height classes. Plot size should match the remote sensing measurements and forest structure - with bigger plots for taller trees and smaller plots for denser, shorter forests. For detailed description of how to measure biomass in mangroves, see (Howard et al, 2014)

- Minimum recommended plot size is 10 m radius or 0.03 ha m, but larger plots strongly recommended where possible
- High-quality geolocation is necessary for small plots
- Individuals with DBH ≥ 5 cm should be measured
- Prop and Buttress roots: the diameter point of measurement (POM) should be measured at the point directly above the buttress, for some individuals with prop roots extending well into the canopy, it is not necessary, practical or accurate to measure above the highest prop root
- For shrub and low stature mangroves under 130 cm, total height and stem diameter should be measured
- Measurements should be taken at both high and low tide, where possible, to minimize effects of tide and assist in image interpretation

A.3.5 Drylands

Low tree cover and greater dominance of shrubs is often typical in dryland systems. Special considerations for when shrubs make up the majority of the overstory and a plot lacks trees include:

- Individuals with DBH ≥ 5 cm should be mapped and measured
- Shrub stem diameter should be measured at the base of the stem (or the POM specified by an applicable shrub allometric model; see Section 2.1.4) instead of breast height. Record POM.
- Cover should be quantified using the point-intercept method along transects in a 20-m grid (see A. White et al., 2012 for details)

The use of variable radius plots should be avoided. This plot type is sometimes used to extend dryland plots with low tree cover, but because it does not have defined plot boundaries, it is difficult to pair with remote sensing data.

A.3.6 Savannas/Woodlands

Savannas and woodlands differ from Drylands (A3.5) in that shrub cover tends to be low, with vegetation dominated by a mixture of grasses and trees. In many savanna landscapes, a large proportion of total biomass is found in small trees (<10 cm), meaning a cut off of diameter measurements at 10 cm, as is typical in many forest sampling protocols, is not appropriate. Large trees are often found however: indeed, the impact of fire, which can only be survived by large trees, can lead to a bimodal distribution of tree diameters, with few trees of intermediate sizes.

Finding these large, rare, trees, and patches of smaller trees (clumping again offers some protection against fire) often necessitates large plots to produce robust estimates of biomass. Nested plots are not normally recommended for use with remote sensing, but are common in savanna landscapes in order to allow robust assessments of the biomass of more homogeneous, small-size vegetation (grass, seedlings, shrubs) which need distributed small sub-plots, within the large plots needed to estimate biomass of distributed trees. For example, in the Bateke savanna landscape of the Republic of Congo, typified by very low biomass values (6.5 Mg C/ha including above and below-ground biomass of trees, shrubs and grass), four 25 ha plots were set up and further analysis of the data found that a plot size of at least 10 hectares were needed to provide a good estimate of the mean tree biomass, due to the clumpiness of the trees (Nieto-Quintano et al., 2018).

Guidance on collecting good savanna and woodland biomass information is provided in the protocols from the Socio-Ecological Observatory for Southern African Woodlands (SEOSAW) project, https://seosaw.github.io/manuals.html, which are based on the RAINFOR protocol for tropical forests where appropriate, but adapted based on the experience of scientists in a broad consortium with expertise on the savannas and woodlands of sub-Saharan Africa (including scientists from 12 African countries), though also applicable throughout the dry tropics.

A.3.6.1 Sampling design

- Plots larger than 1 ha may be necessary in order to capture a sufficient number of stems, particularly in the drier end of the savanna spectrum.
    - >200 stems is desirable in order to be able to estimate mortality over time
    - if the trees are noticeably organised into 'clumps', then ensuring that the plot size is sufficient to capture several such clumps ensures that the exact placement of the plot does not change the estimate of local biomass significantly. Plot sizes well over 1 ha are often desirable
- Careful consideration is needed of a minimum DBH threshold. An initial rule of thumb is that individuals with DBH ≥5 cm should be mapped and measured, but in some landscapes a smaller threshold is appropriate.
- If few trees reach 1.3 m, the default POM for the plot should be lowered to 0.3 m (The SEOSAW partnership, 2020).
- If woody biomass is the focus of the study, then only trees should be measured. But in some landscapes shrub and especially grass biomass may be significant: grass represented over half the biomass in the Bateke landscape studied by Nieto-Quintano (2018). For shrubs, see the recommendations for Drylands (A.3.5). Estimation of grass biomass is outside the scope of this protocol, however it can be measured in small subplots distributed throughout the larger plot, with destructive sampling used to calibrate a Disk Pasture Meter, which, once calibrated, can be used to collect quick estimates of grass biomass at high accuracy.

A.3.6.2 Timing

It is best if measurements are taken at the same time each year to minimize errors caused by stem water content. In savannas, measuring the dry season is easiest, when there is less

understory vegetation. However, site visits during the wet season may be necessary for species identification.

A.4 Accommodating terrestrial laser scanner measurements

Ideally, plots should be established to accommodate terrestrial laser scanning (TLS) measurements. Scanning positions are distributed throughout a plot, with each scan generating a point cloud. Neighboring scans are combined to create a scan of the entire plot.

Sampling may also occur in existing long-term plots. A plot size of 1 ha is preferable. However, to provide consistent point cloud quality throughout the plot, sampling over a larger area than the plot size is often required (Chapter 2.3).

A.4.1 Sampling pattern

The primary concerns for acquiring TLS data over a large area are to have a uniform point density sampled across the whole plot, and sufficient retro-reflective targets to register the location of each scan in relation to others. Dense vegetation blocks TLS sampling, requiring a denser sampling pattern. A sampling pattern should be chosen which yields consistent sampling and minimizes occlusion.

- Dense understory: 10 m scan position grid
- Open understory: 20 m scan position grid
- TLS measurement points and retroreflector targets must continue outside of the plot
- Sample in a chain pattern (Wilkes et al., 2017)
- Time campaigns when forest conditions and climate are consistent
- Instrument must have ability to range to tallest trees (e.g. RIEGL VZ-400(i) or greater)

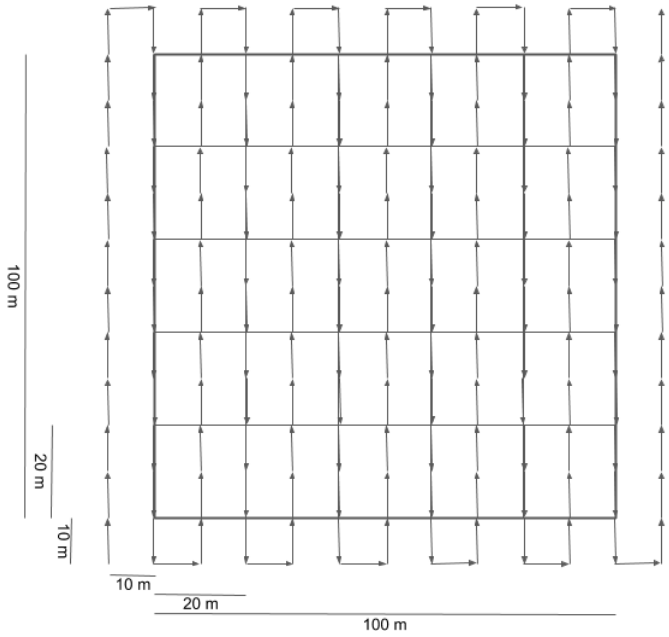

Figure A.3 Example TLS sampling design in a chain pattern. A 100 x 100 m plot with a 10-m sampling grid. A TLS scan location would be placed at the ends of each arrow.

A.4.2 Linking to census measurements

For long term monitoring plots, steps should be taken so that repeated tree censuses and TLS measurements can be linked over time.

- Use same local origin, measurement units and reference coordinate system for TLS and field measurements (Wilkes et al., 2017)
- Repeated scanning acquisitions should be <1 m from original location (Wilkes et al., 2017)
- Permanent retroreflector targets at subplot corners for co-registration (Wilkes et al., 2017)

Acquisitions of TLS should be repeated every 5 years or when a disturbance has been detected. Preferably, TLS and field data acquisitions should occur in the same season. In ecosystems that have leaf-off seasons, scanning in leaf-off conditions is preferred in the context of deriving AGB from point clouds.

A.5 Airborne lidar measurements

Ideally, airborne lidar measurements will be acquired in the same season as field plot measurements. Airborne lidar acquisition should occur no more than 2 years (ideally < 1 year) from field data acquisition. Airborne lidar measurements should be repeated approximately every 5 years or following a detected disturbance.

A minimum area of 3 x 3 km should be sampled with airborne lidar. This area should cover all field plots, and any local environmental and forest structure gradients. A smaller sampling area is acceptable when using a drone. A minimum shot density of 4 pulses/m^2 and instruments capable of recording a minimum of 4 returns per pulse is recommended for small footprint discrete return lidar. Large footprint waveform lidar would have value added for many forest structure researchers, but either high to medium resolution waveform lidar (e.g. NASA LVIS), discrete return lidar (e.g. Optech, NASA GLiHT) or many UAV-LS lidar systems provide sufficient accuracy for collection of reference data.

A.6 Data availability

Ideally, all field validation data should be made open and public to maximize its usefulness across projects. Yet, unlike satellite measurements the actual acquisition of ground data is rarely funded via government space agencies. Field-work requires specific skills and for many field-leaders and PIs the prime motivation may not be Cal-Val, but one or more of a host of other interesting scientific questions - most of which require long-term, multi-site, and painstaking effort to address. It follows that the incentives to make data open and public are maximised if the Earth Observation community actually contributes to the cost. Calls for ‘public’ and ‘open’ data therefore need to be tied to efforts to contribute to the long-term sustainability of plots, and the valorisation of the scientists and assistants who make measurements.

Once funded and acquired, field data can be contributed to a repository so that the data are well-documented and easy to access for others. There are several available options, assuming continued funding and support for these data curation efforts, including:

- ForestPlots.net
- The Joint ESA-NASA Multi-mission Algorithm and Analysis Platform
- The Forest Observation System (FOS; https://forest-observation-system.net/)

Appendix B: Quantifying GNSS geolocation errors (protocol used in Chapter 3)

To quantify the error associated with GNSS measurements in a dense forest, we took advantage of a grid laid out by a professional topography consultant using a total station over a 25-ha forest plot located in Gabon, the Rabi CTFS plot (Memiaghe et al., 2016). A terrestrial lidar acquisition (unpublished data) suggested that the position of the grid markers yielded an error between 10 and 30 cm. We took 208 reference points over the materialized plot markers using a survey grade GNSS device (Trimble L1/L2 GEOXT 7000 rover) with the Floodlight technology option activated (to filter out multipath signals). Differential correction was performed through post-processing using base data acquired by a SXBlue L1/L2 GNSS device with full sky view, and located 1.3 km away from the plot center. Post-processing was applied using Trimble Pathfinder Office software. **Only code** (not phase) signal correction proved possible, as is customary under dense vegetation. Terrestrial lidar data were acquired over a total of 1.5 ha, using a Leica C10 scan station, with high resolution acquisitions (i.e. 0.05 m between point at 100 m) taken every 10 m along the NE (X axis) direction, and every 20 m along the NW (Y axis) direction. Scans and targets were positioned on plot marker positions, using a closed traverse survey approach (Figure B1). 2D rigid (translation, rotation) transforms (angle and scale preserving) were estimated from the TLS global coordinate system to the plot grid referencing system (relative X and Y) to the geographic coordinate systems on the basis of the positions of the plot markers in these different systems (Torr & Zisserman, 2000). The reference (best) transform was obtained using all 208 measurements. A sensitivity analysis was then performed by deriving new transforms in a Monte Carlo permutation approach (1000 iterations) by progressively removing points. The accuracy of the new transforms was assessed using the positions estimated using each transformation matrix on the full set of points. X and Y differences were averaged and then converted to Euclidean distance.

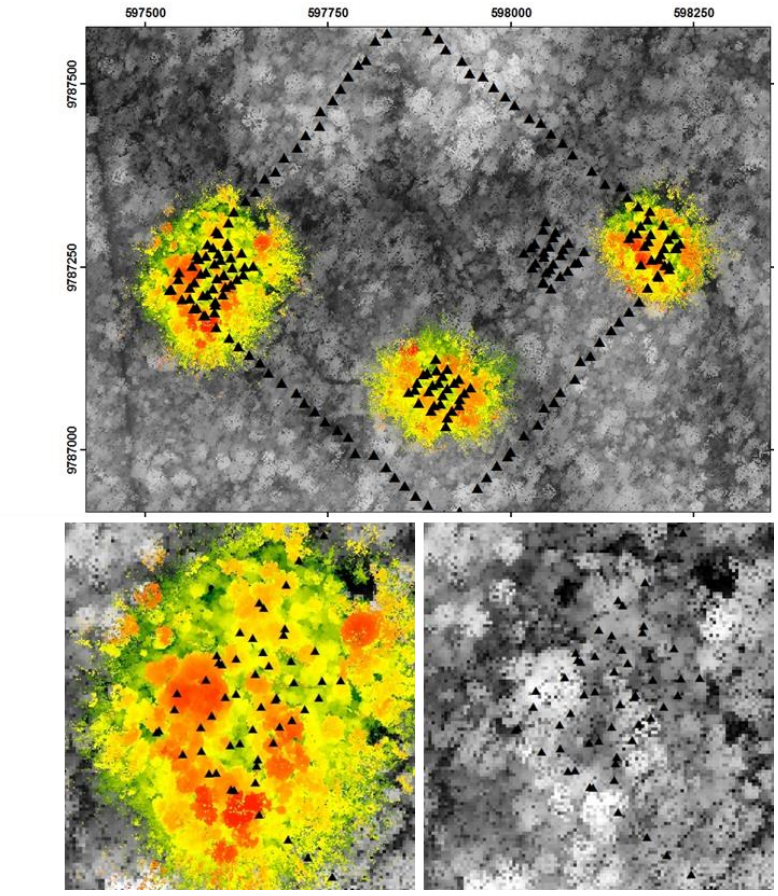

Figure B.1 Sampling setup for assessing the quality of the reference projection. Green to red height maps illustrate the canopy height model (CHM) derived from terrestrial lidar and grey levels colors depict an airborne lidar CHM (Labrière et al. in press). Black triangles indicate the DGNSS positions of measured plot markers. The two panels at the bottom illustrate the co-registration difference between terrestrial and airborne lidar data. Metric reference system (UTM 32S). We acknowledge Drs. Alfonso Alonso, Hervé Memiaghe and David Kenfack for giving access to the Rabi plot and to the data acquired by the professional topography consultant. We also acknowledge S. Momo Takoudjou for giving access to the TLS acquisitions.

Appendix C: Quantifying errors associated with incidence angle, plot size and geolocation (protocol used in Chapter 3)

We used a simplistic conceptual sensitivity analysis, using a set of ellipsoid crowns, produced using realistic dendrometric curves and allometric models, as described in Barbier et al. (2012). We did not account for radiative transfer. For the simulation, representative 200 x 200 m forest stands were generated. To facilitate the numerical analysis, crown ellipsoids were converted to voxels (cubic cells) of 0.5 m sides. We considered that a square field plot with finite limits allows for an exact accounting of the volume of tree crowns within these limits. Similarly, a square remote sensing footprint provides an exact measurement of the intercepted crown volumes down to ground level. We assessed the effect of plot/footprint size (from 20 to 100 m), geolocation precision (from 0 to 30 m) and satellite incidence angle (from 0 to 50°) on total crown volume estimation obtained from the two viewpoints. To obtain a distribution of errors, each combination of the tested parameters was applied on three independently generated forest stands, to which ten different random horizontal rotations were applied. The discrepancy between the crown volume obtained for a given combination of shift, length and tilt and that measured in the field plot can be assessed in two ways. The first one is to only consider the difference in the actual measured volume, allowing for compensations from the surrounding forest. The other measure only considers the proportion of the crown volume initially measured in the plot that is also measured by the RS signal (no-compensation).

Appendix D: List of potential biomass reference measurement sites

Table D.1 List of potential BRM sites as of 09th March 2021. This list is regularly being updated and currently includes 189 sites. Note not all of these BRM sites fulfill all the inclusion criteria listed in Section 8.2.

Realm	Country	Site	Code	Coordinator
Afrotropical	Angola	Bicuar	BCR	SEOSAW + ForestPlots.net
Afrotropical	Cameroon	Dja	DJA	AfriTRON + ForestPlots.net
Afrotropical	Cameroon	Korup	KRP	ForestGEO
Afrotropical	Cameroon	Mbam Djerem	MDJ	AfriTRON + ForestPlots.net
Afrotropical	Central African Republic	M'Baiki	MBK	TmFO
Afrotropical	Cote d'Ivoire	La Tene	LTN	CIRAD
Afrotropical	DRC	Ituri	ITR	ForestGEO
Afrotropical	DRC	Mai-Ndombe	MAI	NASA
Afrotropical	DRC	Malebo	MLB	WWF
Afrotropical	DRC	Salonga	SLN	AfriTRON + ForestPlots.net
Afrotropical	DRC	Yangambi	YNG	AfriTRON + ForestPlots.net
Afrotropical	Gabon	Lope	LOP	AfriTRON + ForestPlots.net
Afrotropical	Gabon	Mabounie	MBN	AMAP
Afrotropical	Gabon	Mondah	MND	NASA + AfriTRON + ForestPlots.net
Afrotropical	Gabon	Ogooue-Ivindo	OGI	FODEX
Afrotropical	Gabon	Rabi	RAB	ForestGEO
Afrotropical	Ghana	Ankasa	ANK	AfriTRON + ForestPlots.net
Afrotropical	Kenya	Mpala	MPL	ForestGEO
Afrotropical	Mozambique	Gorongosa	GRN	SEOSAW + ForestPlots.net
Afrotropical	Nigeria	Ngel Nyaki	NGL	ForestGEO

Afrotropical	Republic of the Congo	Ekolongouma	EKO	AfriTRON + ForestPlots.net
Afrotropical	Republic of the Congo	Lefini	LFN	SEOSAW + ForestPlots.net
Afrotropical	Republic of the Congo	Lesio Louna	LES	SEOSAW + ForestPlots.net
Afrotropical	Republic of the Congo	Loundoungou	LND	CIRAD
Afrotropical	Republic of the Congo	Nouabale-Ndoki	NBL	AfriTRON + ForestPlots.net
Afrotropical	South Africa	Kruger Lowveld	KRG	CSIR
Afrotropical	Tanzania	Kilwa	KLW	SEOSAW + ForestPlots.net
Australasian	Australia	Alice Mulga	ALC	TERN
Australasian	Australia	Calperum Mallee	CLP	TERN
Australasian	Australia	Cumberland Plain	CMB	TERN
Australasian	Australia	Great Western Woodlands	GWW	TERN
Australasian	Australia	Injune	INJ	TERN
Australasian	Australia	Karawatha	KRW	TERN
Australasian	Australia	Litchfield Savanna	LTC	TERN
Australasian	Australia	Robson Creek	RBS	TERN
Australasian	Australia	Samford	SMF	TERN
Australasian	Australia	Tumbarumba Wet Eucalypt	TUM	TERN
Australasian	Australia	Warra Tall Eucalypt	WRR	TERN
Australasian	Papua New Guinea	Wanang	WNN	ForestGEO
Indo-Malayan	China	Ailaoshan	AIL	ForestGEO
Indo-Malayan	China	Baishanzu	BSZ	ForestGEO
Indo-Malayan	China	Dinghushan	DHS	ForestGEO
Indo-Malayan	China	Heishiding	HSD	ForestGEO
Indo-Malayan	China	Hong Kong	HGK	ForestGEO
Indo-Malayan	China	Jianfengling	JFL	ForestGEO
Indo-Malayan	China	Nonggang	NNG	ForestGEO
Indo-Malayan	China	Tiantongshan	TNT	ForestGEO
Indo-Malayan	China	Xishuangbanna	XSH	ForestGEO
Indo-Malayan	India	Betul	BTL	ISRO

Indo-Malayan	India	Dehradun	DHR	ISRO
Indo-Malayan	India	Hoshangabad	HSH	ISRO
Indo-Malayan	India	Karbi Anglong	KRB	ISRO
Indo-Malayan	India	Narmada	NRM	ISRO
Indo-Malayan	India	Shimoga	SHM	ISRO
Indo-Malayan	India	Mudumalai	MDM	ForestGEO
Indo-Malayan	India	Uppangala	UPP	IFP
Indo-Malayan	India	Yellapur	YLL	AMAP
Indo-Malayan	Indonesia	Malinau	MLN	TmFO + ForestPlots.net
Indo-Malayan	Indonesia	STREK	STR	TmFO
Indo-Malayan	Malaysia_Borneo	Danum Valley	DNM	ForestGEO + ForestPlots.net
Indo-Malayan	Malaysia_Borneo	Lambir	LMB	ForestGEO
Indo-Malayan	Malaysia_Borneo	Sepilok	SPK	Aberdeen, Leeds, Cambridge + ForestPlots.net
Indo-Malayan	Malaysia_Penisular	Pasoh	PSH	ForestGEO + TmFO
Indo-Malayan	Philippines	Palanan	PAL	ForestGEO
Indo-Malayan	Sri Lanka	Sinharaja	SIN	ForestGEO
Indo-Malayan	Taiwan	Fushan	FSH	ForestGEO
Indo-Malayan	Thailand	Khao Chong	KCH	ForestGEO
Indo-Malayan	Thailand	Doi Inthanon	DIN	ForestGEO
Indo-Malayan	Thailand	Huai Kha Khaeng	HKK	ForestGEO
Indo-Malayan	Thailand	Mo Singto	MSN	ForestGEO + AMAP
Indo-Malayan	Vietnam	Bidoup	BDP	ForestGEO
Neartic	Canada	Haliburton	HLB	ForestGEO
Neartic	Canada	Laurentides	LRN	NASA
Neartic	Canada	Petawawa	PTW	Canadian Forest Service
Neartic	Canada	Scotty Creek	SCT	ForestGEO
Neartic	Mexico	Sorona	SRN	ForestGEO
Neartic	USA	Caribou-Poker Creeks Res Watershed	BONA	NEON
Neartic	USA	Harvard	HARV	ForestGEO + NEON

Neartic	USA	LBJ Natl Grassland	CLBJ	NEON
Neartic	USA	Lilly Dickey Woods	LDW	ForestGEO
Neartic	USA	Michigan Big Woods	MCH	ForestGEO
Neartic	USA	Niobrara	NBR	ForestGEO
Neartic	USA	Niwot Ridge Mt Res Stn	NIWO	NEON
Neartic	USA	Oak Ridge	ORNL	NEON
Neartic	USA	Onaqui	ONAQ	NEON
Neartic	USA	Ordway-Swisher Biol Stn	OSBS	ForestGEO + NEON
Neartic	USA	Pikes Peak	PKS	Colorado State Univ
Neartic	USA	San Joaquin Exp Range	SJER	NEON
Neartic	USA	Santa Cruz	SNT	ForestGEO
Neartic	USA	Santa Rita Exp Range	SRER	NEON
Neartic	USA	Smithsonian Conserv Biol Inst	SCBI	ForestGEO + NEON
Neartic	USA	Smithsonian Environ Res Cent	SERC	ForestGEO + NEON
Neartic	USA	Talladega Natl For	TALL	NEON
Neartic	USA	Tyson	TSN	ForestGEO
Neartic	USA	Univ Notre Dame Environ Res Cent	UNDE	NEON
Neartic	USA	Wabikon	WBK	ForestGEO
Neartic	USA	Wind River	WREF	ForestGEO + NEON
Neartic	USA	Yosemite	YSM	ForestGEO
Neotropical	Bolivia	La Chonta	CHN	TmFO
Neotropical	Bolivia	Noel Kempff	NLK	RAINFOR + ForestPlots.net
Neotropical	Brazil	Abuna	ABU	PPBio + ForestPlots.net
Neotropical	Brazil	Antimary	ANT	EMBRAPA
Neotropical	Brazil	Araguaia	ARA	RAINFOR + ForestPlots.net
Neotropical	Brazil	Assis	ASS	Instituto Florestal
Neotropical	Brazil	Braganca	BRA	EMBRAPA + ForestPlots.net
Neotropical	Brazil	Caetetus	CTT	Instituto Florestal

Neotropical	Brazil	Caicara	CCA	PPBio + ForestPlots.net
Neotropical	Brazil	Carlos Botelho	CRL	Instituto Florestal
Neotropical	Brazil	Caxiuana	CAX	RAINFOR + TEAM + ForestPlots.net
Neotropical	Brazil	Chico Mendes	CHM	RAINFOR + ForestPlots.net
Neotropical	Brazil	Cunia	CNA	PPBio + ForestPlots.net
Neotropical	Brazil	Gaucha do Norte	GAU	RAINFOR + ForestPlots.net
Neotropical	Brazil	Ilha do Cardoso	ILH	Instituto Florestal + ForestGEO
Neotropical	Brazil	Manaus	MNS	ForestGEO + RAINFOR
Neotropical	Brazil	Nova Xavantina	NXV	RAINFOR + ForestPlots.net
Neotropical	Brazil	Paragominas	PRG	TmFO
Neotropical	Brazil	Serra do Mar	SMR	Instituto Florestal + ForestPlots.net
Neotropical	Colombia	Amacayacu	AMC	ForestGEO
Neotropical	Colombia	Choco	CHC	NASA
Neotropical	Colombia	La Planada	LPL	ForestGEO
Neotropical	Costa Rica	La Selva	LSL	OTS
Neotropical	Costa Rica	Osa Peninsula	OSP	Wien Univ
Neotropical	Ecuador	Yasuni	YAS	ForestGEO
Neotropical	French Guiana	BAFOG	BAF	Guyafor
Neotropical	French Guiana	Montagne Tortue	MNT	Guyafor
Neotropical	French Guiana	Nouragues	NOU	RAINFOR + ForestPlots.net
Neotropical	French Guiana	Organabo	ORG	Guyafor
Neotropical	French Guiana	Paracou	PAR	TmFO
Neotropical	Panama	Barro Colorado Island	BCI	ForestGEO
Neotropical	Peru	Allpahuayo	ALP	RAINFOR + ForestPlots.net
Neotropical	Peru	Belgica	BLG	FODEX
Neotropical	Peru	Cocha Cashu	CCC	RAINFOR + ForestPlots.net
Neotropical	Peru	Cordillera Azul	CAZ	RAINFOR + ForestPlots.net
Neotropical	Peru	Genova	GNV	Univ La Molina, RAINFOR + ForestPlots.net

Neotropical	Peru	Jenaro Herrera	JEN	RAINFOR + ForestPlots.net
Neotropical	Peru	Rio Abiseo	RIA	RAINFOR + ForestPlots.net
Neotropical	Peru	Tambopata	TAM	RAINFOR + ForestPlots.net
Neotropical	Peru	Trocha Union	TRU	ABERG, RAINFOR + ForestPlots.net
Neotropical	Peru	Yanachaga	YAN	RAINFOR + ForestPlots.net
Neotropical	USA	Guanica Forest	GUAN	NEON
Neotropical	USA	Luquillo	LQL	ForestGEO
Oceanian	USA	Laupahoehoe	LPH	ForestGEO
Oceanian	USA	Palamanui	PLM	ForestGEO
Oceanian	USA	Puu Makaala Nat Area Reserve	PUUM	NEON
Paleartic	Austria	Purkersdorf	PRK	IIASA
Paleartic	Belgium	Ardennes	ARD	Liege Univ
Paleartic	China	Badagongshan	BDG	ForestGEO
Paleartic	China	Baotianman	BTN	ForestGEO
Paleartic	China	Changbaishan	CHA	ForestGEO
Paleartic	China	Daxing'anling	DXN	ForestGEO
Paleartic	China	Donglingshan	DNG	ForestGEO
Paleartic	China	Gutianshan	GTN	ForestGEO
Paleartic	Czech Republic	Zofin	ZFN	ForestGEO
Paleartic	Estonia	Jarvselja	JRV	Tartu Observatory
Paleartic	Finland	Hyytiala	HTL	Univ Helsinki
Paleartic	Germany	Hainich-Dün	HNC	DFG Biodiversity Exploratories
Paleartic	Germany	Schorfheide-Chorin	SFD	DFG Biodiversity Exploratories
Paleartic	Germany	Serrahn	SRR	Luneburg Univ
Paleartic	Germany	Schwabische Alb	SWB	DFG Biodiversity Exploratories
Paleartic	Germany	Traunstein	TRN	ForestGEO
Paleartic	Germany	Weberstedter Holz	WEB	Max-Planck Inst Biogeochem, Freiburg Univ and Hainich Natl Park
Paleartic	Italy	Montedimezzo	MTD	Univ Molise

Paleartic	Japan	Aya	AYA	Forestry For Prod Res Inst
Paleartic	Japan	Ogawa	OGW	Forestry For Prod Res Inst & Kyoto Univ
Paleartic	Japan	Tomakomai Res Site	TRS	Forestry For Prod Res Inst
Paleartic	Netherlands	Speulderbos	SPL	ForestGEO
Paleartic	Poland	Bialowieski	BLW	IIASA
Paleartic	Russia	Baikal	BKL	IIASA
Paleartic	Russia	Bryansk	BRN	IIASA
Paleartic	Russia	Denezhkin	DNZ	IIASA
Paleartic	Russia	Kamshilovka	KMS	IIASA
Paleartic	Russia	Khantaisk	KHN	IIASA
Paleartic	Russia	Komi	KOM	IIASA
Paleartic	Russia	Lisino	LSN	IIASA
Paleartic	Russia	Otdalennyy	OTD	IIASA
Paleartic	Russia	Tunguska	TNG	IIASA
Paleartic	Russia	Turukhansk	TRK	IIASA
Paleartic	Russia	Vnukovo	VNK	IIASA
Paleartic	Russia	Zotino	ZTT	IIASA
Paleartic	Russia	Ary-Mas	ARM	IIASA
Paleartic	Russia	Lukunskoe	LKN	IIASA
Paleartic	Russia	Gorno-Tayezhnoe	RGT	IIASA
Paleartic	Spain	Valsain	VLS	Univ Politec Madrid
Paleartic	Sweden	Remningstorp	RMN	Swedish Univ Agric Sci
Paleartic	Switzerland	Laegern	LGR	Univ Zurich
Paleartic	UK	Dyfi Valley	DYF	Aberystwyth Univ
Paleartic	UK	Wytham Woods	WTH	ForestGEO
Paleartic	Ukraine	Bojarka	BJR	IIASA
Paleartic	Ukraine	Snovsk	SNV	IIASA

References

Abshire, J. B., Sun, X., Riris, H., Marcos Sirota, J., McGarry, J. F., Palm, S., Yi, D., & Liiva, P. (2005). Geoscience Laser Altimeter System (GLAS) on the ICESat Mission: On-orbit measurement performance. In *Geophysical Research Letters* (Vol. 32, Issue 21). https://doi.org/10.1029/2005gl024028

Albinet, C., Nouvellon, S., Frommknecht, B., Rutakaza, R., Daniel, S., & Saüt, C. (2020). *MAAP: The Mission Algorithm and Analysis Platform: A New Virtual and Collaborative Environment for the Scientific Community*. https://doi.org/10.5194/egusphere-egu2020-19989

Alvarez, E., Duque, A., Saldarriaga, J., Cabrera, K., de las Salas, G., del Valle, I., Lema, A., Moreno, F., Orrego, S., & Rodríguez, L. (2012). Tree above-ground biomass allometries for carbon stocks estimation in the natural forests of Colombia. In *Forest Ecology and Management* (Vol. 267, pp. 297–308). https://doi.org/10.1016/j.foreco.2011.12.013

Andersen, H.-E., Reutebuch, S. E., & McGaughey, R. J. (2006). A rigorous assessment of tree height measurements obtained using airborne lidar and conventional field methods. In *Canadian Journal of Remote Sensing* (Vol. 32, Issue 5, pp. 355–366). https://doi.org/10.5589/m06-030

Anderson, K., Hancock, S., Disney, M., & Gaston, K. J. (2016). Is waveform worth it? A comparison of LiDAR approaches for vegetation and landscape characterization. In *Remote Sensing in Ecology and Conservation* (Vol. 2, Issue 1, pp. 5–15). https://doi.org/10.1002/rse2.8

Archibald, S., & Bond, W. J. (2003). Growing tall vs growing wide: tree architecture and allometry of Acacia karroo in forest, savanna, and arid environments. In *Oikos* (Vol. 102, Issue 1, pp. 3–14). https://doi.org/10.1034/j.1600-0706.2003.12181.x

Armston, J., Bunting, P., Flood, N., & Gillingham, S. (2020). *PyLidar*. http://www.pylidar.org/

Armston, J., Disney, M., Lewis, P., Scarth, P., Phinn, S., Lucas, R., Bunting, P., & Goodwin, N. (2013). Direct retrieval of canopy gap probability using airborne waveform lidar. *Remote Sensing of Environment*, *134*, 24–38.

Armston, J., Tang, H., Marselis, S., Duncanson, L. I., Hofton, M., Blair, J. B., Fatoyinbo, T., & Dubayah, R. O. (2020). *AfriSAR: Gridded Forest Biomass and Canopy Metrics Derived from LVIS, Gabon, 2016* (ORNL DAAC, Oak Ridge, Tennessee, USA) [Data set]. https://doi.org/10.3334/ORNLDAAC/1775

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.

Asner, G. P., Mascaro, J., Anderson, C., Knapp, D. E., Martin, R. E., Kennedy-Bowdoin, T., van Breugel, M., Davies, S., Hall, J. S., Muller-Landau, H. C., Potvin, C., Sousa, W., Wright, J., & Bermingham, E. (2013). High-fidelity national carbon mapping for resource management and REDD+. *Carbon Balance and Management*, *8*(1), 7.

Asner, G. P., Mascaro, J., Muller-Landau, H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., Hall, J. S., & van Breugel, M. (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia*, *168*(4), 1147–1160.

Asner, G. P., Powell, G. V. N., Mascaro, J., Knapp, D. E., Clark, J. K., Jacobson, J., Kennedy-Bowdoin, T., Balaji, A., Paez-Acosta, G., Victoria, E., Secada, L., Valqui, M., & Hughes, R. F. (2010). High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(38), 16738–16742.

Aubry-Kientz, M., Dutrieux, R., Ferraz, A., Saatchi, S., Hamraz, H., Williams, J., Coomes, D., Piboule, A., & Vincent, G. (2019). A Comparative Assessment of the Performance of Individual Tree Crowns Delineation Algorithms from ALS Data in Tropical Forests. In *Remote Sensing* (Vol. 11, Issue 9, p. 1086). https://doi.org/10.3390/rs11091086

Auclair, D. (1986). Measurement errors in forest biomass estimation. In H. Wharton & T. Cunia (Eds.), *Estimating tree biomass regressions and their error* (pp. 133–138). Broomall, USDA.

Averill, C., Turner, B. L., & Finzi, A. C. (2014). Mycorrhiza-mediated competition between plants and decomposers drives soil carbon storage. *Nature*, *505*(7484), 543–545.

Avitabile, V., & Camia, A. (2018). An assessment of forest biomass maps in Europe using harmonized national statistics and inventory plots. *Forest Ecology and Management*, *409*, 489–498.

Avitabile, V., Herold, M., Heuvelink, G. B. M., Lewis, S. L., Phillips, O. L., Asner, G. P., Armston, J., Ashton, P. S., Banin, L., Bayol, N., Berry, N. J., Boeckx, P., de Jong, B. H. J., DeVries, B., Girardin, C. A. J., Kearsley, E., Lindsell, J. A., Lopez-Gonzalez, G., Lucas, R., … Willcock, S. (2016). An integrated pan-tropical biomass map using multiple reference datasets. *Global Change Biology*, *22*(4), 1406–1420.

Babcock, C., Finley, A. O., Andersen, H.-E., Pattison, R., Cook, B. D., Morton, D. C., Alonzo, M., Nelson, R., Gregoire, T. G., Ene, L., Gobakken, T., & Næsset, E. (2018). Geostatistical estimation of forest biomass in interior Alaska combining Landsat-derived tree cover, sampled airborne lidar and field observations. In *Remote Sensing of Environment* (Vol. 212, pp. 212–230). https://doi.org/10.1016/j.rse.2018.04.044

Babcock, C., Finley, A. O., Cook, B. D., Weiskittel, A., & Woodall, C. W. (2016). Modeling forest biomass and growth: Coupling long-term inventory and LiDAR data. In *Remote Sensing of Environment* (Vol. 182, pp. 1–12). https://doi.org/10.1016/j.rse.2016.04.014

Baccini, A., & Asner, G. P. (2013). Improving pantropical forest carbon maps with airborne LiDAR sampling. In *Carbon Management* (Vol. 4, Issue 6, pp. 591–600). https://doi.org/10.4155/cmt.13.66

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R. O., Friedl, M. A., Samanta, S., & Houghton, R. A. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. In *Nature Climate Change* (Vol. 2, Issue 3, pp. 182–185). https://doi.org/10.1038/nclimate1354

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. In *Environmental Research Letters* (Vol. 3, Issue 4, p. 045011). https://doi.org/10.1088/1748-9326/3/4/045011

Baker, T. R., Pennington, R. T., Dexter, K. G., Fine, P. V. A., Fortune-Hopkins, H., Honorio, E. N., Huamantupa-Chuquimaco, I., Klitgård, B. B., Lewis, G. P., de Lima, H. C., Ashton, P., Baraloto, C., Davies, S., Donoghue, M. J., Kaye, M., Kress, W. J., Lehmann, C. E. R., Monteagudo, A., Phillips, O. L., & Vasquez, R. (2017). Maximising Synergy among Tropical Plant Systematists, Ecologists, and Evolutionary Biologists. *Trends in Ecology & Evolution*, *32*(4), 258–267.

Baker, T. R., Phillips, O. L., Malhi, Y., Almeida, S., Arroyo, L., Di Fiore, A., Erwin, T., Killeen, T. J., Laurance, S. G., Laurance, W. F., Lewis, S. L., Lloyd, J., Monteagudo, A., Neill, D. A., Patino, S., Pitman, N. C. A., Silva, J. N. M., & Martinez, R. V. (2004). Variation in wood density determines spatial patterns in Amazonian forest biomass. In *Global Change Biology* (Vol. 10, Issue 5, pp. 545–562). https://doi.org/10.1111/j.1365-2486.2004.00751.x

Barbier, N., Couteron, P., Gastelly-Etchegorry, J.-P., & Proisy, C. (2012). Linking canopy images to forest structural parameters: potential of a modeling framework. In *Annals of Forest Science* (Vol. 69, Issue 2, pp. 305–311). https://doi.org/10.1007/s13595-011-0116-9

Barker, J. R., Bollman, M., Ringold, P. L., Sackinger, J., & Cline, S. P. (2002). Evaluation of metric precision for a riparian forest survey. *Environmental Monitoring and Assessment*, *75*(1), 51–72.

Barlow, J., Gardner, T. A., Araujo, I. S., Avila-Pires, T. C., Bonaldo, A. B., Costa, J. E., Esposito, M. C., Ferreira, L. V., Hawes, J., Hernandez, M. I. M., Hoogmoed, M. S., Leite, R. N., Lo-Man-Hung, N. F., Malcolm, J. R., Martins, M. B., Mestre, L. A. M., Miranda-Santos, R., Nunes-Gutjahr, A. L., Overal, W. L., … Peres, C. A. (2007). Quantifying the biodiversity value of tropical primary, secondary, and plantation forests. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(47), 18555–18560.

Baskerville, G. L. (1972). Use of Logarithmic Regression in the Estimation of Plant Biomass. In *Canadian Journal of Forest Research* (Vol. 2, Issue 1, pp. 49–53). https://doi.org/10.1139/x72-009

Bastin, J.-F., Barbier, N., Réjou-Méchain, M., Fayolle, A., Gourlet-Fleury, S., Maniatis, D., de Haulleville, T., Baya, F., Beeckman, H., Beina, D., Couteron, P., Chuyong, G., Dauby, G., Doucet, J.-L., Droissart, V., Dufrêne, M., Ewango, C., Gillet, J. F., Gonmadje, C. H., … Bogaert, J. (2015). Seeing Central African forests through their largest trees. *Scientific Reports*, *5*, 13156.

Bastin, J.-F., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., Patriarca, C., Picard, N., Sparrow, B., Abraham, E. M., Aloui, K., Atesoglu, A., Attore, F., Bassüllü, Ç., Bey, A., Garzuglia, M., García-Montero, L. G., Groot, N., Guerin, G., … Castro, R. (2017). The extent of forest in dryland biomes. *Science*, *356*(6338), 635–638.

Batjes, N. H. (1996). Total carbon and nitrogen in the soils of the world. In *European Journal of Soil Science* (Vol. 47, Issue 2, pp. 151–163). https://doi.org/10.1111/j.1365-2389.1996.tb01386.x

Behre, E. C. (1926). Comparison of diameter tape and caliper measurements in second-growth spruce. *Journal of Forestry*, *24*(2), 178–182.

Berger, A., Gschwantner, T., McRoberts, R. E., & Schadauer, K. (2014). Effects of Measurement Errors on Individual Tree Stem Volume Estimates for the Austrian National Forest Inventory. In *Forest Science* (Vol. 60, Issue 1, pp. 14–24). https://doi.org/10.5849/forsci.12-164

Bienert, A., Georgi, L., Kunz, M., Maas, H.-G., & Von Oheimb, G. (2018). Comparison and Combination of Mobile and Terrestrial Laser Scanning for Natural Forest Inventories. *Forests*, *9*(7), 395.

Biging, G. S., & Wensel, L. C. (1988). The effect of eccentricity on the estimation of basal area and basal area increment of coniferous trees. *Forest Science*, *34*(3), 621–633.

Blair, J. B., Bryan Blair, J., Rabine, D. L., & Hofton, M. A. (1999). The Laser Vegetation Imaging Sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 54, Issues 2-3, pp. 115–122). https://doi.org/10.1016/s0924-2716(99)00002-7

Bloom, A. A., Exbrayat, J.-F., van der Velde, I. R., Feng, L., & Williams, M. (2016). The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(5), 1285–1290.

Bolton, D. K., Coops, N. C., & Wulder, M. A. (2013). Investigating the agreement between global canopy height maps and airborne Lidar derived height estimates over Canada. In *Canadian Journal of Remote Sensing* (Vol. 39, Issue sup1, pp. S139–S151). https://doi.org/10.5589/m13-036

Boni Vicari, M., Disney, M., Wilkes, P., Burt, A., Calders, K., & Woodgate, W. (2019). Leaf and wood classification framework for terrestrial LiDAR point clouds. *Methods in Ecology and Evolution / British Ecological Society*, *10*(5), 680–694.

Bonser, S. P., & Aarssen, L. W. (1994). Plastic allometry in young sugar maple (Acer saccharum): adaptive responses to light availability. In *American Journal of Botany* (Vol. 81, Issue 4, pp. 400–406). https://doi.org/10.1002/j.1537-2197.1994.tb15463.x

Bortolot, Z. J., & Wynne, R. H. (2005). Estimating forest biomass using small footprint LiDAR data: An individual tree-based approach that incorporates training data. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 59, Issue 6, pp. 342–360). https://doi.org/10.1016/j.isprsjprs.2005.07.001

Boshuizen, C., Mason, J., Klupar, P., & Spanhake, S. (2014). Results from the planet labs flock constellation. *Technical Session I: Private Endeavors*. Small Satellite Conference.

Boudreau, J., Nelson, R., Margolis, H., Beaudoin, A., Guindon, L., & Kimes, D. (2008). Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec. *Remote Sensing of Environment*, *112*(10), 3876–3890.

Brancalion, P. H. S., Bello, C., Chazdon, R. L., Galetti, M., Jordano, P., Lima, R. A. F., Medina, A., Pizo, M. A., & Leighton Reid, J. (2018). Maximizing biodiversity conservation and carbon stocking in restored

tropical forests. In *Conservation Letters* (Vol. 11, Issue 4, p. e12454). https://doi.org/10.1111/conl.12454

Brancalion, P. H. S., Lamb, D., Ceccon, E., Boucher, D., Herbohn, J., Strassburg, B., & Edwards, D. P. (2017). Using markets to leverage investment in forest and landscape restoration in the tropics. In *Forest Policy and Economics* (Vol. 85, pp. 103–113). https://doi.org/10.1016/j.forpol.2017.08.009

Brede, B., Calders, K., Lau, A., Raumonen, P., Bartholomeus, H. M., Herold, M., & Kooistra, L. (2019). Non-destructive tree volume estimation through quantitative structure modelling: Comparing UAV laser scanning with terrestrial LIDAR. *Remote Sensing of Environment*, *223*, 111355.

Brede, B., Lau, A., Bartholomeus, H. M., & Kooistra, L. (2017). Comparing RIEGL RiCOPTER UAV LiDAR Derived Canopy Height and DBH with Terrestrial LiDAR. *Sensors* , *17*(10). https://doi.org/10.3390/s17102371

Breidenbach, J., Antón-Fernández, C., Petersson, H., McRoberts, R. E., & Astrup, R. (2014). Quantifying the Model-Related Variability of Biomass Stock and Change Estimates in the Norwegian National Forest Inventory. In *Forest Science* (Vol. 60, Issue 1, pp. 25–33). https://doi.org/10.5849/forsci.12-137

Brown, I. F., Foster Brown, I., Martinelli, L. A., Wayt Thomas, W., Moreira, M. Z., Cid Ferreira, C. A., & Victoria, R. A. (1995). Uncertainty in the biomass of Amazonian forests: An example from Rondônia, Brazil. In *Forest Ecology and Management* (Vol. 75, Issues 1-3, pp. 175–189). https://doi.org/10.1016/0378-1127(94)03512-u

Brown, S. (1997). *Estimating Biomass and Biomass Change of Tropical Forests: A Primer* (Food and Agriculture Organization of the United Nations (ed.)). Food & Agriculture Org.

Brown, S., Gillespie, A. J., & Lugo, A. E. (1989). Biomass estimation methods for tropical forests with applications to forest inventory data. *Forest Science*, *35*(4), 881–902.

Bunting, P., Armston, J., Clewley, D., & Lucas, R. M. (2013). Sorted pulse data (SPD) library—Part II: A processing framework for LiDAR data from pulsed laser systems in terrestrial environments. *Computers & Geosciences*, *56*, 207–215.

Burt, A., Calders, K., Cuni-Sanchez, A., Gómez-Dans, J., Lewis, P., Lewis, S. L., Malhi, Y., Phillips, O. L., & Disney, M. (2020). Assessment of Bias in Pan-Tropical Biomass Predictions. *Frontiers in Forests and Global Change*, *3*, 297.

Burt, A., Disney, M., & Calders, K. (2019). Extracting individual trees from lidar point clouds using treeseg. In *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210x.13121

Burt, A., Vicari, M. B., da Costa, A. C. L., Coughlin, I., Meir, P., Rowland, L., & Disney, M. (2021). New insights into large tropical tree mass and structure from direct harvest and terrestrial lidar. *Royal Society Open Science*, *8*(2), 2054–5703.

Bustamante, M. M. C., Roitman, I., Aide, T. M., Alencar, A., Anderson, L. O., Aragão, L., Asner, G. P., Barlow, J., Berenguer, E., Chambers, J., Costa, M. H., Fanin, T., Ferreira, L. G., Ferreira, J., Keller, M., Magnusson, W. E., Morales-Barquero, L., Morton, D., Ometto, J. P. H. B., … Vieira, I. C. G. (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*(1), 92–109.

Cairns, M. A., Brown, S., Helmer, E. H., & Baumgardner, G. A. (1997). Root biomass allocation in the world's upland forests. In *Oecologia* (Vol. 111, Issue 1, pp. 1–11). https://doi.org/10.1007/s004420050201

Calders, K., Adams, J., Armston, J., Bartholomeus, H., Bauwens, S., Bentley, L. P., Chave, J., Danson, F. M., Demol, M., Disney, M., Gaulton, R., Krishna Moorthy, S. M., Levick, S. R., Saarinen, N., Schaaf, C., Stovall, A., Terryn, L., Wilkes, P., & Verbeeck, H. (2020). Terrestrial laser scanning in forest ecology: Expanding the horizon. *Remote Sensing of Environment*, *251*, 112102.

Calders, K., Disney, M., Armston, J., Burt, A., Brede, B., Origo, N., Muir, J., & Nightingale, J. (2017). Evaluation of the Range Accuracy and the Radiometric Calibration of Multiple Terrestrial Laser

Scanning Instruments for Data Interoperability. In *IEEE Transactions on Geoscience and Remote Sensing* (Vol. 55, Issue 5, pp. 2716–2724). https://doi.org/10.1109/tgrs.2017.2652721

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., & Kaasalainen, M. (2015). Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods in Ecology and Evolution*, *6*(2), 198–208.

Calders, K., Origo, N., Burt, A., Disney, M., Nightingale, J., Raumonen, P., Åkerblom, M., Malhi, Y., & Lewis, P. (2018). Realistic Forest Stand Reconstruction from Terrestrial LiDAR for Radiative Transfer Modelling. In *Remote Sensing* (Vol. 10, Issue 6, p. 933). https://doi.org/10.3390/rs10060933

Calders, K., Schenkels, T., Bartholomeus, H., Armston, J., Verbesselt, J., & Herold, M. (2015). Monitoring spring phenology with high temporal resolution terrestrial LiDAR measurements. *Agricultural and Forest Meteorology*, *203*, 158–168.

Calder, W. A. (1984). *Size, Function, and Life History*. Courier Corporation.

Cao, L., Coops, N., Innes, J., Dai, J., & She, G. (2014). Mapping Above- and Below-Ground Biomass Components in Subtropical Forests Using Small-Footprint LiDAR. In *Forests* (Vol. 5, Issue 6, pp. 1356–1373). https://doi.org/10.3390/f5061356

Case, B. S., & Hall, R. J. (2008). Assessing prediction errors of generalized tree biomass and volume equations for the boreal forest region of west-central Canada. In *Canadian Journal of Forest Research* (Vol. 38, Issue 4, pp. 878–889). https://doi.org/10.1139/x07-212

Cassel, C.-M., Cassel, C., Claes-Magnus, Särndal, C.-E., & Wretman, J. H. (1977). *Foundations of Inference in Survey Sampling*. Krieger Publishing Company.

Chambers, J. Q., Negron-Juarez, R. I., Marra, D. M., Di Vittorio, A., Tews, J., Roberts, D., Ribeiro, G. H. P. M., Trumbore, S. E., & Higuchi, N. (2013). The steady-state mosaic of disturbance and succession across an old-growth Central Amazon forest landscape. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(10), 3949–3954.

Chao, K.-J., -J. Chao, K., Phillips, O. L., Baker, T. R., Peacock, J., Lopez-Gonzalez, G., Vásquez Martínez, R., Monteagudo, A., & Torres-Lezama, A. (2009). After trees die: quantities and determinants of necromass across Amazonia. In *Biogeosciences* (Vol. 6, Issue 8, pp. 1615–1626). https://doi.org/10.5194/bg-6-1615-2009

Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.-P., Nelson, B. W., Ogawa, H., Puig, H., Riéra, B., & Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, *145*(1), 87–99.

Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., & Perez, R. (2004). Error propagation and scaling for tropical forest biomass estimates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *359*(1443), 409–420.

Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G., & Zanne, A. E. (2009). Towards a worldwide wood economics spectrum. *Ecology Letters*, *12*(4), 351–366.

Chave, J., Davies, S. J., Phillips, O. L., Lewis, S. L., Sist, P., Schepaschenko, D., Armston, J., Baker, T. R., Coomes, D., Disney, M., Duncanson, L. I., Hérault, B., Labrière, N., Meyer, V., Réjou-Méchain, M., Scipal, K., & Saatchi, S. S. (2019). Ground Data are Essential for Biomass Remote Sensing Missions. In *Surveys in Geophysics* (Vol. 40, Issue 4, pp. 863–880). https://doi.org/10.1007/s10712-019-09528-w

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B. C., Duque, A., Eid, T., Fearnside, P. M., Goodman, R. C., Henry, M., Martínez-Yrízar, A., Mugasha, W. A., Muller-Landau, H. C., Mencuccini, M., Nelson, B. W., Ngomanda, A., Nogueira, E. M., Ortiz-Malavassi, E., … Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*, *20*(10), 3177–3190.

Chazdon, R. L., Broadbent, E. N., Rozendaal, D. M. A., Bongers, F., Zambrano, A. M. A., Aide, T. M., Balvanera, P., Becknell, J. M., Boukili, V., Brancalion, P. H. S., Craven, D., Almeida-Cortez, J. S., Cabral,

G. A. L., de Jong, B., Denslow, J. S., Dent, D. H., DeWalt, S. J., Dupuy, J. M., Durán, S. M., … Poorter, L. (2016). Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. *Science Advances*, *2*(5), e1501639.

Chen, Q. (2015). Modeling aboveground tree woody biomass using national-scale allometric methods and airborne lidar. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 106, pp. 95–106). https://doi.org/10.1016/j.isprsjprs.2015.05.007

Chen, Q., Kutser, T., Collin, A., & Warner, T. A. (2018). Fine resolution remote sensing of species in terrestrial and coastal ecosystems. *International Journal of Remote Sensing*, *39*(17), 5597–5599.

Chen, Q., Laurin, G. V., & Valentini, R. (2015). Uncertainty of remotely sensed aboveground biomass over an African tropical forest: Propagating errors from trees to plots to pixels. In *Remote Sensing of Environment* (Vol. 160, pp. 134–143). https://doi.org/10.1016/j.rse.2015.01.009

Chen, Q., McRoberts, R. E., Wang, C., & Radtke, P. J. (2016). Forest aboveground biomass mapping and estimation across multiple spatial scales using model-based inference. In *Remote Sensing of Environment* (Vol. 184, pp. 350–360). https://doi.org/10.1016/j.rse.2016.07.023

Chen, Q., Vaglio Laurin, G., & Valentini, R. (2015). Uncertainty of remotely sensed aboveground biomass over an African tropical forest: Propagating errors from trees to plots to pixels. *Remote Sensing of Environment*, *160*, 134–143.

Chisholm, R. A., Muller-Landau, H. C., Abdul Rahman, K., Bebber, D. P., Bin, Y., Bohlman, S. A., Bourg, N. A., Brinks, J., Bunyavejchewin, S., Butt, N., Cao, H., Cao, M., Cárdenas, D., Chang, L.-W., Chiang, J.-M., Chuyong, G., Condit, R., Dattaraja, H. S., Davies, S., … Zimmerman, J. K. (2013). Scale-dependent relationships between tree species richness and ecosystem function in forests. *The Journal of Ecology*, *101*(5), 1214–1224.

Chojnacky, D. C., Heath, L. S., & Jenkins, J. C. (2014). Updated generalized biomass equations for North American tree species. *Forestry*, *87*(1), 129–151.

Clark, D. A. (2002). Are tropical forests an important carbon sink? Reanalysis of the long-term plot data. In *Ecological Applications* (Vol. 12, Issue 1, pp. 3–7). https://doi.org/10.1890/1051-0761(2002)012[0003:atfaic]2.0.co;2

Clark, D. B., & Kellner, J. R. (2012). Tropical forest biomass estimation and the fallacy of misplaced concreteness. In *Journal of Vegetation Science* (Vol. 23, Issue 6, pp. 1191–1196). https://doi.org/10.1111/j.1654-1103.2012.01471.x

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C., & Cox, P. M. (2011). The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics. In *Geoscientific Model Development* (Vol. 4, Issue 3, pp. 701–722). https://doi.org/10.5194/gmd-4-701-2011

Cleverly, J., Eamus, D., Edwards, W., Grant, M., Grundy, M. J., Held, A., Karan, M., Lowe, A. J., Prober, S. M., Sparrow, B., & Morris, B. (2019). TERN, Australia's land observatory: addressing the global challenge of forecasting ecosystem responses to climate variability and change. In *Environmental Research Letters* (Vol. 14, Issue 9, p. 095004). https://doi.org/10.1088/1748-9326/ab33cb

Clifford, D., Cressie, N., England, J. R., Roxburgh, S. H., & Paul, K. I. (2013). Correction factors for unbiased, efficient estimation and prediction of biomass from log–log allometric models. In *Forest Ecology and Management* (Vol. 310, pp. 375–381). https://doi.org/10.1016/j.foreco.2013.08.041

Clough, B. J., Russell, M. B., Domke, G. M., & Woodall, C. W. (2016). Quantifying allometric model uncertainty for plot-level live tree biomass stocks with a data-driven, hierarchical framework. In *Forest Ecology and Management* (Vol. 372, pp. 175–188). https://doi.org/10.1016/j.foreco.2016.04.001

Cochran, W. G. (1977). *Sampling techniques*. John Wiley & Sons.

Cohen, R., Kaino, J., Okello, J. A., Bosire, J. O., Kairo, J. G., Huxham, M., & Mencuccini, M. (2013).

Propagating uncertainty to estimates of above-ground biomass for Kenyan mangroves: A scaling procedure from tree to landscape level. In *Forest Ecology and Management* (Vol. 310, pp. 968–982). https://doi.org/10.1016/j.foreco.2013.09.047

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M., & Randerson, J. T. (2018). The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation. In *Journal of Advances in Modeling Earth Systems* (Vol. 10, Issue 11, pp. 2731–2754). https://doi.org/10.1029/2018ms001354

Comas, X., Terry, N., Hribljan, J. A., Lilleskov, E. A., Suarez, E., Chimner, R. A., & Kolka, R. K. (2017). Estimating belowground carbon stocks in peatlands of the Ecuadorian páramo using ground-penetrating radar (GPR). In *Journal of Geophysical Research: Biogeosciences* (Vol. 122, Issue 2, pp. 370–386). https://doi.org/10.1002/2016jg003550

Condit, R. (1998). *Tropical Forest Census Plots: Methods and Results from Barro Colorado Island, Panama and a Comparison with Other Plots*. Springer-Verlag Berlin Heidelberg and R.G. Landes Company.

Cook, B., Corp, L., Nelson, R., Middleton, E., Morton, D., McCorkel, J., Masek, J., Ranson, K., Ly, V., & Montesano, P. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing*, *5*(8), 4045–4066.

Coomes, D. A., Dalponte, M., Jucker, T., Asner, G. P., Banin, L. F., David F R, Lewis, S. L., Nilus, R., Phillips, O. L., Phua, M.-H., & Qie, L. (2017). Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data. In *Remote Sensing of Environment* (Vol. 194, pp. 77–88). https://doi.org/10.1016/j.rse.2017.03.017

Corona, P., Fattorini, L., Franceschi, S., Scrinzi, G., & Torresan, C. (2014). Estimation of standing wood volume in forest compartments by exploiting airborne laser scanning information: model-based, design-based, and hybrid perspectives. In *Canadian Journal of Forest Research* (Vol. 44, Issue 11, pp. 1303–1311). https://doi.org/10.1139/cjfr-2014-0203

Côté, J.-F., Fournier, R. A., & Egli, R. (2011). An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR. In *Environmental Modelling & Software* (Vol. 26, Issue 6, pp. 761–777). https://doi.org/10.1016/j.envsoft.2010.12.008

Côté, J.-F., Widlowski, J.-L., Fournier, R. A., & Verstraete, M. M. (2009). The structural and radiative consistency of three-dimensional tree reconstructions from terrestrial lidar. In *Remote Sensing of Environment* (Vol. 113, Issue 5, pp. 1067–1081). https://doi.org/10.1016/j.rse.2009.01.017

Cunia, T. (1986). Error of forest inventory estimates: its main components. In H. Wharton & T. Cunia (Eds.), *Estimating tree biomass regressions and their error* (pp. 1–13). Broomall, USDA.

Dalponte, M., Jucker, T., David F R, Lewis, S. L., Nilus, R., Phillips, O., Qie, L., & Coomes, D. A. (2016). Aboveground biomass estimation in tropical forests at single tree level with ALS data. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. https://doi.org/10.1109/igarss.2016.7730390

Danson, F. M., Gaulton, R., Armitage, R. P., Disney, M., Gunawan, O., Lewis, P., Pearson, G., & Ramirez, A. F. (2014). Developing a dual-wavelength full-waveform terrestrial laser scanner to characterize forest canopy structure. In *Agricultural and Forest Meteorology* (Vols. 198–199, pp. 7–14). https://doi.org/10.1016/j.agrformet.2014.07.007

Danson, F. M., Sasse, F., & Schofield, L. A. (2018). Spectral and spatial information from a novel dual-wavelength full-waveform terrestrial laser scanner for forest ecology. *Interface Focus*, *8*(2), 20170049.

da Páscoa, K. J. V., Gomide, L. R., Tng, D. Y. P., Scolforo, J. R. S., Filho, A. C. F., & de Mello, J. M. (2020). How many trees and samples are adequate for estimating wood-specific gravity across different tropical forests? *Trees*, *574*, 654.

Dargie, G. C., Lewis, S. L., Lawson, I. T., Mitchard, E. T. A., Page, S. E., Bocko, Y. E., & Ifo, S. A. (2017). Age, extent and carbon storage of the central Congo Basin peatland complex. *Nature*, *542*(7639), 86–90.

Dassot, M., Colin, A., Santenoise, P., Fournier, M., & Constant, T. (2012). Terrestrial laser scanning for measuring the solid wood volume, including branches, of adult standing trees in the forest environment. In *Computers and Electronics in Agriculture* (Vol. 89, pp. 86–93). https://doi.org/10.1016/j.compag.2012.08.005

Deckert, C., & Bolstad, P. V. (1996). Forest Canopy, Terrain, and Distance Effects on Global Positioning System Point Accuracy. *Photogrammetric Engineering & Remote Sensing*, *62*(3), 317–321.

Di Marco, M., Watson, J. E. M., Currie, D. J., Possingham, H. P., & Venter, O. (2018). The extent and predictability of the biodiversity-carbon correlation. *Ecology Letters*, *21*(3), 365–375.

Disney, M. (2019). *New 3D measurements and models of canopy structure, and the insights this can provide*. ESA Living Planet Symposium, Milan, Italy.

Disney, M., 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*(2), 20170048.

Disney, M., Burt, A., Calders, K., Schaaf, C., & Stovall, A. (2019). Innovations in ground and airborne technologies as reference and for training and validation: terrestrial laser scanning (TLS). *Surveys in Geophysics*, 1–22.

Disney, M., Kalogirou, V., Lewis, P., Prieto-Blanco, A., Hancock, S., & Pfeifer, M. (2010). Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests. *Remote Sensing of Environment*, *114*(7), 1546–1560.

Disney, M., Lewis, P., & Raumonen, P. (2012). Testing a new vegetation structure retrieval algorithm from terrestrial lidar scanner data using 3d models. *SilviLaser 2012*.

Dorigo, W., Hollaus, M., Wagner, W., & Schadauer, K. (2010). An application-oriented automated approach for co-registration of forest inventory and airborne laser scanning data. In *International Journal of Remote Sensing* (Vol. 31, Issue 5, pp. 1133–1153). https://doi.org/10.1080/01431160903380581

Dot, E. E. (2016). *Guidelines for field sampling of biomass* (Version 1.0). The Department of the Environment, Canberra.

Drake, J. B., Knox, R. G., Dubayah, R. O., Clark, D. B., Condit, R., Blair, J. B., & Hofton, M. (2003). Above-ground biomass estimation in closed canopy neotropical forests using lidar remote sensing: Factors affecting the generality of relationships. *Global Ecology and Biogeography: A Journal of Macroecology*, *12*(2), 147–159.

Dubayah, R. O., Blair, J. B., Goetz, S. J., 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. *Egyptian Journal of Remote Sensing and Space Sciences*, *1*, 100002.

Dubayah, R. O., Knox, R., Hofton, M., Blair, J. B., & Drake, J. (2000). Land Surface Characterization Using Lidar Remote Sensing. In M. J. H. Hill & R. J. Apinall (Eds.), *Spatial Information for Land Use Management* (pp. 53–66).

Dubayah, R. O., Sheldon, S. L., Clark, D. B., Hofton, M. A., Blair, J. B., Hurtt, G. C., & Chazdon, R. L. (2010). Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. In *Journal of Geophysical Research: Biogeosciences* (Vol. 115, Issue G2). https://doi.org/10.1029/2009jg000933

Dubayah, R. O., Swatantran, A., Huang, W., Duncanson, L., Tang, H., Johnson, K., Dunne, J. O., & Hurtt, G. C. (2017). *CMS: LiDAR-derived Biomass, Canopy Height and Cover, Sonoma County, California, 2013* (ORNL DAAC, Oak Ridge, Tennessee, USA) [Data set]. https://doi.org/10.3334/ORNLDAAC/1523

Duffy, J. E., Godwin, C. M., & Cardinale, B. J. (2017). Biodiversity effects in the wild are common and as strong as key drivers of productivity. *Nature*, *549*(7671), 261–264.

Duncanson, L. I., Armston, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., Carter, S., Chave, J., Herold,

M., Crowther, T. W., Falkowski, M., Kellner, J. R., Labrière, N., Lucas, R., MacBean, N., McRoberts, R. E., Meyer, V., Næsset, E., Nickeson, J. E., … Williams, M. (2019). The Importance of Consistent Global Forest Aboveground Biomass Product Validation. *Surveys in Geophysics*, *40*(4), 979–999.

Duncanson, L. I., Cook, B. D., Rourke, O., Hurtt, G. C., & Dubayah, R. O. (2013). *Evaluating the Importance of Local Environment on Tree Structural Allometries*. *2013*, B43C – 0513.

Duncanson, L. I., & Dubayah, R. O. (2018). Monitoring individual tree-based change with airborne lidar. *Ecology and Evolution*, *8*(10), 5079–5089.

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

Duncanson, L. I., 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. O. (2020). Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. *Remote Sensing of Environment*, *242*, 111779.

Duncanson, L. I., Niemann, K. O., & Wulder, M. A. (2010). Integration of GLAS and Landsat TM data for aboveground biomass estimation. In *Canadian Journal of Remote Sensing* (Vol. 36, Issue 2, pp. 129–141). https://doi.org/10.5589/m10-037

Duncanson, L. I., Rourke, O., & Dubayah, R. O. (2015). Small Sample Sizes Yield Biased Allometric Equations in Temperate Forests. *Scientific Reports*, *5*, 17153.

Elzinga, C., Shearer, R. C., & Elzinga, G. (2005). Observer Variation in Tree Diameter Measurements. In *Western Journal of Applied Forestry* (Vol. 20, Issue 2, pp. 134–137). https://doi.org/10.1093/wjaf/20.2.134

Ene, L. T., Gobakken, T., Andersen, H.-E., Næsset, E., Cook, B. D., Morton, D. C., Babcock, C., & Nelson, R. (2018). Large-area hybrid estimation of aboveground biomass in interior Alaska using airborne laser scanning data. *Remote Sensing of Environment*, *204*, 741–755.

Ene, L. T., Næsset, E., Gobakken, T., Gregoire, T. G., Ståhl, G., & Holm, S. (2013). A simulation approach for accuracy assessment of two-phase post-stratified estimation in large-area LiDAR biomass surveys. *Remote Sensing of Environment*, *133*, 210–224.

Ene, L. T., Næsset, E., Gobakken, T., Mauya, E. W., Bollandsås, O. M., Gregoire, T. G., Ståhl, G., & Zahabu, E. (2016). Large-scale estimation of aboveground biomass in miombo woodlands using airborne laser scanning and national forest inventory data. *Remote Sensing of Environment*, *186*, 626–636.

Enquist, B. J., & Niklas, K. J. (2002). Global allocation rules for patterns of biomass partitioning in seed plants. *Science*, *295*(5559), 1517–1520.

Espírito-Santo, F. D. B., Gloor, M., Keller, M., Malhi, Y., Saatchi, S. S., Nelson, B., Junior, R. C. O., Pereira, C., Lloyd, J., Frolking, S., Palace, M., Shimabukuro, Y. E., Duarte, V., Mendoza, A. M., López-González, G., Baker, T. R., Feldpausch, T. R., Brienen, R. J. W., Asner, G. P., … Phillips, O. L. (2014). Size and frequency of natural forest disturbances and the Amazon forest carbon balance. *Nature Communications*, *5*, 3434.

Esteban, J., McRoberts, R., Fernández-Landa, A., Tomé, J., & Næsset, E. (2019). Estimating Forest Volume and Biomass and Their Changes Using Random Forests and Remotely Sensed Data. In *Remote Sensing* (Vol. 11, Issue 16, p. 1944). https://doi.org/10.3390/rs11161944

Exbrayat, J.-F., Anthony Bloom, A., Carvalhais, N., Fischer, R., Huth, A., MacBean, N., & Williams, M. (2019). Understanding the Land Carbon Cycle with Space Data: Current Status and Prospects. In *Surveys in Geophysics* (Vol. 40, Issue 4, pp. 735–755). https://doi.org/10.1007/s10712-019-09506-2

FAO and ITPS. (2018). *Global Soil Organic Carbon Map - GSOCmap* (Version 1.2.0). FAO.

Fatoyinbo, L. E., Armston, J., Simard, M., Saatchi, S., Debina, M., Lavalle, M., Hofton, M., Tang, H., Marselis, S. M., Pinto, N., Hancock, S., Hawkins, B., Duncanson, L. I., Blair, J. B., Hansen, C., Lou, Y., Dubayah, R. O., Hensley, S., Silva, C. A., … Hibbard, K. (2021). The NASA AfriSAR Campaign: Airborne

SAR and Lidar Measurements of Tropical Forest Structure and Biomass in Support of Future Space Missions. In *Preprint on https://essoar.org*. https://doi.org/10.1002/essoar.10506317.1

Fattorini, L. (2012). Design-based or model-based inference? The role of hybrid approaches in environmental surveys. In L. Fattorini (Ed.), *Studies in Honor of Claudio Scala* (pp. 173–214). Department of Economics and Statistics, University of Siena.

Fayolle, A., Doucet, J.-L., Gillet, J.-F., Bourland, N., & Lejeune, P. (2013). Tree allometry in Central Africa: Testing the validity of pantropical multi-species allometric equations for estimating biomass and carbon stocks. In *Forest Ecology and Management* (Vol. 305, pp. 29–37). https://doi.org/10.1016/j.foreco.2013.05.036

Fayolle, A., Ngomanda, A., Mbasi, M., Barbier, N., Bocko, Y., Boyemba, F., Couteron, P., Fonton, N., Kamdem, N., Katembo, J., Kondaoule, H. J., Loumeto, J., Maïdou, H. M., Mankou, G., Mengui, T., Mofack, G. I. I., Moundounga, C., Moundounga, Q., Nguimbous, L., … Medjibe, V. P. (2018). A regional allometry for the Congo basin forests based on the largest ever destructive sampling. In *Forest Ecology and Management* (Vol. 430, pp. 228–240). https://doi.org/10.1016/j.foreco.2018.07.030

Feldpausch, T. R., Phillips, O. L., Brienen, R. J. W., Gloor, E., Lloyd, J., Lopez-Gonzalez, G., Monteagudo-Mendoza, A., Malhi, Y., Alarcón, A., Álvarez Dávila, E., Alvarez-Loayza, P., Andrade, A., Aragao, L. E. O. C., Arroyo, L., Aymard C., G. A., Baker, T. R., Baraloto, C., Barroso, J., Bonal, D., … Vos, V. A. (2016). Amazon forest response to repeated droughts: Amazon forest response to droughts. *Global Biogeochemical Cycles*, *30*(7), 964–982.

Ferraz, A., Saatchi, S., Mallet, C., Jacquemoud, S., Gonçalves, G., Silva, C., Soares, P., Tomé, M., & Pereira, L. (2016). Airborne Lidar Estimation of Aboveground Forest Biomass in the Absence of Field Inventory. In *Remote Sensing* (Vol. 8, Issue 8, p. 653). https://doi.org/10.3390/rs8080653

Ferraz, A., Saatchi, S., Mallet, C., & Meyer, V. (2016). Lidar detection of individual tree size in tropical forests. In *Remote Sensing of Environment* (Vol. 183, pp. 318–333). https://doi.org/10.1016/j.rse.2016.05.028

Ferreira, J., Lennox, G. D., Gardner, T. A., Thomson, J. R., Berenguer, E., Lees, A. C., Nally, R. M., Luiz E O, Ferraz, S. F. B., Louzada, J., Moura, N. G., Oliveira, V. H. F., Pardini, R., Solar, R. R. C., Vieira, I. C. G., & Barlow, J. (2018). Carbon-focused conservation may fail to protect the most biodiverse tropical forests. In *Nature Climate Change* (Vol. 8, Issue 8, pp. 744–749). https://doi.org/10.1038/s41558-018-0225-7

Finer, M., Novoa, S., Weisse, M. J., Petersen, R., Mascaro, J., Souto, T., Stearns, F., & Martinez, R. G. (2018). Combating deforestation: From satellite to intervention. *Science*, *360*(6395), 1303–1305.

Fischer, G., Nachtergaele, F., Prieler, S., Van Velthuizen, H. T., Verelst, L., & Wiberg, D. (2008). *Global agro-ecological zones assessment for agriculture (GAEZ 2008)*. https://books.google.com/books/about/Global_Agro_ecological_Assessment_for_Ag.html?hl=&id=E611AQAACAAJ

Fisher, A., Armston, J., Goodwin, N., & Scarth, P. (2020). Modelling canopy gap probability, foliage projective cover and crown projective cover from airborne lidar metrics in Australian forests and woodlands. *Remote Sensing of Environment*, *237*, 111520.

Fisher, J. B., Sweeney, S., Brzostek, E. R., Evans, T. P., Johnson, D. J., Myers, J. A., Bourg, N. A., Wolf, A. T., Howe, R. W., & Phillips, R. P. (2016). Tree-mycorrhizal associations detected remotely from canopy spectral properties. *Global Change Biology*, *22*(7), 2596–2607.

Food and Agriculture Organization of the United Nations. (2018). *Global Forest Resources Assessment 2015: Desk reference*. Food & Agriculture Org.

Forkel, M., Andela, N., Harrison, S. P., Lasslop, G., van Marle, M., Chuvieco, E., Dorigo, W., Forrest, M., Hantson, S., Heil, A., Li, F., Melton, J., Sitch, S., Yue, C., & Arneth, A. (2019). Emergent relationships with respect to burned area in global satellite observations and fire-enabled vegetation models.

Biogeosciences, *16*(1), 57–76.
Forrester, D. I., Tachauer, I. H. H., Annighoefer, P., Barbeito, I., Pretzsch, H., Ruiz-Peinado, R., Stark, H., Vacchiano, G., Zlatanov, T., Chakraborty, T., Saha, S., & Sileshi, G. W. (2017). Generalized biomass and leaf area allometric equations for European tree species incorporating stand structure, tree age and climate. In *Forest Ecology and Management* (Vol. 396, pp. 160–175). https://doi.org/10.1016/j.foreco.2017.04.011
Frazer, G. W., Magnussen, S., Wulder, M. A., & Niemann, K. O. (2011). Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. In *Remote Sensing of Environment* (Vol. 115, Issue 2, pp. 636–649). https://doi.org/10.1016/j.rse.2010.10.008
Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers, R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin, P., Schaphoff, S., … Woodward, F. I. (2014). Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(9), 3280–3285.
Gaston, K. J. (2000). Global patterns in biodiversity. In *Nature* (Vol. 405, Issue 6783, pp. 220–227). https://doi.org/10.1038/35012228
Genova, E., & Barton, C. C. (2003). *Global positioning system accuracy and precision at Hubbard Brook Experimental Forest, Grafton County, New Hampshire; a guide to the limits of handheld GPS receivers* (Open-File Report 2003–316). US Geological Survey. https://doi.org/10.3133/ofr2003316
Gentry, A. H. (1988). Tree species richness of upper Amazonian forests. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(1), 156–159.
Gertner, G. Z. (1990). The sensitivity of measurement error in stand volume estimation. In *Canadian Journal of Forest Research* (Vol. 20, Issue 6, pp. 800–804). https://doi.org/10.1139/x90-105
Gertner, G. Z., & Dzialowy, P. J. (1984). Effects of measurement errors on an individual tree-based growth projection system. In *Canadian Journal of Forest Research* (Vol. 14, Issue 3, pp. 311–316). https://doi.org/10.1139/x84-057
Gertner, G. Z., & Köhl, M. (1992). An assessment of some nonsampling errors in a national survey using an error budget. *Forest Science*, *38*(3), 525–538.
Gilroy, J. J., Woodcock, P., Edwards, F. A., Wheeler, C., Baptiste, B. L. G., Medina Uribe, C. A., Haugaasen, T., & Edwards, D. P. (2014). Cheap carbon and biodiversity co-benefits from forest regeneration in a hotspot of endemism. In *Nature Climate Change* (Vol. 4, Issue 6, pp. 503–507). https://doi.org/10.1038/nclimate2200
Global Forest Observations Initiative. (2016). *Integration of remote-sensing and ground-based observations for estimation of emissions and removals of greenhouse gases in forests: Methods and guidance from the Global Forest Observations Initiative, edition 2.0*. UN Food and Agriculture Organization.
Gobakken, T., & Næsset, E. (2008). Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. In *Canadian Journal of Forest Research* (Vol. 38, Issue 5, pp. 1095–1109). https://doi.org/10.1139/x07-219
Gobakken, T., & Næsset, E. (2009). Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. In *Canadian Journal of Forest Research* (Vol. 39, Issue 5, pp. 1036–1052). https://doi.org/10.1139/x09-025
Gobakken, T., Næsset, E., Nelson, R., Bollandsås, O. M., Gregoire, T. G., Ståhl, G., Holm, S., Ørka, H. O., & Astrup, R. (2012). Estimating biomass in Hedmark County, Norway using national forest inventory field plots and airborne laser scanning. In *Remote Sensing of Environment* (Vol. 123, pp. 443–456). https://doi.org/10.1016/j.rse.2012.01.025

Goetz, S. J., Baccini, A., Laporte, N. T., Johns, T., Walker, W., Kellndorfer, J., Houghton, R. A., & Sun, M. (2009). Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. *Carbon Balance and Management*, *4*, 2.

Goetz, S. J., & Dubayah, R. O. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. In *Carbon Management* (Vol. 2, Issue 3, pp. 231–244). https://doi.org/10.4155/cmt.11.18

Goetz, S. J., Hansen, M., Houghton, R. A., Walker, W., Laporte, N., & Busch, J. (2015). Measurement and monitoring needs, capabilities and potential for addressing reduced emissions from deforestation and forest degradation under REDD. In *Environmental Research Letters* (Vol. 10, Issue 12, p. 123001). https://doi.org/10.1088/1748-9326/10/12/123001

Gonçalves, F., Treuhaft, R., Law, B., Almeida, A., Walker, W., Baccini, A., dos Santos, J., & Graça, P. (2017). Estimating Aboveground Biomass in Tropical Forests: Field Methods and Error Analysis for the Calibration of Remote Sensing Observations. *Remote Sensing*, *9*(1), 47.

Gonzalez de Tanago, J., Lau, A., Bartholomeus, H., Herold, M., Avitabile, V., Raumonen, P., Martius, C., Goodman, R. C., Disney, M., Manuri, S., Burt, A., & Calders, K. (2018). Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR. In *Methods in Ecology and Evolution* (Vol. 9, Issue 2, pp. 223–234). https://doi.org/10.1111/2041-210x.12904

Goodman, L. A. (1960). On the Exact Variance of Products. *Journal of the American Statistical Association*, *55*(292), 708–713.

Goodman, R. C., Phillips, O. L., & Baker, T. R. (2014). The importance of crown dimensions to improve tropical tree biomass estimates. In *Ecological Applications* (Vol. 24, Issue 4, pp. 680–698). https://doi.org/10.1890/13-0070.1

Goodwin, N. R., Coops, N. C., & Culvenor, D. S. (2007). Development of a simulation model to predict LiDAR interception in forested environments. *Remote Sensing of Environment*, *111*(4), 481–492.

Goulden, T., & Scholl, V. (2019). *NEON Algorithm Theoretical Basis Document (ATBD): Ecosystem Structure* (NEON Doc. #: NEON.DOC.002387; Revision A).

Gregoire, T. G., Næsset, E., McRoberts, R. E., Ståhl, G., Andersen, H.-E., Gobakken, T., Ene, L., & Nelson, R. (2016). Statistical rigor in LiDAR-assisted estimation of aboveground forest biomass. *Remote Sensing of Environment*, *173*, 98–108.

Gregoire, T. G., Zedaker, S. M., & Nicholas, N. S. (1990). Modeling relative error in stem basal area estimates. In *Canadian Journal of Forest Research* (Vol. 20, Issue 5, pp. 496–502). https://doi.org/10.1139/x90-065

Griffith, D. M., Lehmann, C. E. R., Strömberg, C. A. E., Parr, C. L., Pennington, R. T., Sankaran, M., Ratnam, J., Still, C. J., Powell, R. L., Hanan, N. P., Nippert, J. B., Osborne, C. P., Good, S. P., Anderson, T. M., Holdo, R. M., Veldman, J. W., Durigan, G., Tomlinson, K. W., Hoffmann, W. A., … Bond, W. J. (2017). Comment on "The extent of forest in dryland biomes" [Review of *Comment on "The extent of forest in dryland biomes"*]. *Science*, *358*(6365). https://doi.org/10.1126/science.aao1309

Hackenberg, J., Spiecker, H., Calders, K., Disney, M., & Raumonen, P. (2015). SimpleTree —An Efficient Open Source Tool to Build Tree Models from TLS Clouds. In *Forests* (Vol. 6, Issue 12, pp. 4245–4294). https://doi.org/10.3390/f6114245

Hall, F. G., Bergen, K., Blair, J. B., Dubayah, R., Houghton, R., Hurtt, G., Kellndorfer, J., Lefsky, M., Ranson, J., Saatchi, S., Shugart, H. H., & Wickland, D. (2011). Characterizing 3D vegetation structure from space: Mission requirements. *Remote Sensing of Environment*, *115*(11), 2753–2775.

Hall, S. A., Burke, I. C., Box, D. O., Kaufmann, M. R., & Stoker, J. M. (2005). Estimating stand structure using discrete-return lidar: an example from low density, fire prone ponderosa pine forests. *Forest Ecology and Management*, *208*(1-3), 189–209.

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. In *Earth and Space Science*. https://doi.org/10.1029/2018ea000506
Hancock, S., Gaulton, R., & Danson, F. M. (2017). Angular Reflectance of Leaves With a Dual-Wavelength Terrestrial Lidar and Its Implications for Leaf-Bark Separation and Leaf Moisture Estimation. In *IEEE Transactions on Geoscience and Remote Sensing* (Vol. 55, Issue 6, pp. 3084–3090). https://doi.org/10.1109/tgrs.2017.2652140
Hansen, E., Ene, L., Mauya, E., Patočka, Z., Mikita, T., Gobakken, T., & Næsset, E. (2017). Comparing Empirical and Semi-Empirical Approaches to Forest Biomass Modelling in Different Biomes Using Airborne Laser Scanner Data. In *Forests* (Vol. 8, Issue 5, p. 170). https://doi.org/10.3390/f8050170
Hansen, E., Gobakken, T., Bollandsås, O., Zahabu, E., & Næsset, E. (2015). Modeling Aboveground Biomass in Dense Tropical Submontane Rainforest Using Airborne Laser Scanner Data. In *Remote Sensing* (Vol. 7, Issue 1, pp. 788–807). https://doi.org/10.3390/rs70100788
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*(6160), 850–853.
Hauglin, M., Lien, V., Næsset, E., & Gobakken, T. (2014). Geo-referencing forest field plots by co-registration of terrestrial and airborne laser scanning data. In *International Journal of Remote Sensing* (Vol. 35, Issue 9, pp. 3135–3149). https://doi.org/10.1080/01431161.2014.903440
Hegarty, E. E., & Caballé, G. (1991). Community ecology of vines. In *The Biology of Vines* (pp. 313–335). Cambridge University Press.
Heidemann, H. K. (2018). Lidar base specification. In *U.S. Geological Survey Techniques and Methods* (ver. 1.3, February 2018, Vols. 11–B4, p. 101). USGS.
Held, A., Phinn, S., Soto-Berelov, M., & Jones, S. (Eds.). (2018). *A TERN Landscape Assessment Initiative: Effective field calibration and validation practices* (1.3 ed.). TERN Australia.
Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J. G. B., Walsh, M. G., & Gonzalez, M. R. (2014). SoilGrids1km — Global Soil Information Based on Automated Mapping. In *PLoS ONE* (Vol. 9, Issue 8, p. e105992). https://doi.org/10.1371/journal.pone.0105992
Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PloS One*, *12*(2), e0169748.
Henry, M., Besnard, A., Asante, W. A., Eshun, J., Adu-Bredu, S., Valentini, R., Bernoux, M., & Saint-André, L. (2010). Wood density, phytomass variations within and among trees, and allometric equations in a tropical rainforest of Africa. In *Forest Ecology and Management* (Vol. 260, Issue 8, pp. 1375–1388). https://doi.org/10.1016/j.foreco.2010.07.040
Henry, M., Bombelli, A., Trotta, C., Alessandrini, A., Birigazzi, L., Sola, G., Vieilledent, G., Santenoise, P., Longuetaud, F., Valentini, R., Picard, N., & Saint-André, L. (2013). GlobAllomeTree: international platform for tree allometric equations to support volume, biomass and carbon assessment. *iForest*, *6*(6), 326–330.
Hernández-Clemente, R., Navarro-Cerrillo, R., Ramírez, F., Hornero, A., & Zarco-Tejada, P. (2014). A Novel Methodology to Estimate Single-Tree Biophysical Parameters from 3D Digital Imagery Compared to Aerial Laser Scanner Data. In *Remote Sensing* (Vol. 6, Issue 11, pp. 11627–11648). https://doi.org/10.3390/rs61111627
Herold, M., Carter, S., Avitabile, V., Espejo, A. B., Jonckheere, I., Lucas, R., McRoberts, R. E., Næsset, E., Nightingale, J., Petersen, R., Reiche, J., Romijn, E., Rosenqvist, A., Rozendaal, D. M. A., Seifert, F. M., Sanz, M. 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*(4), 757–778.
Heuvelink, G., & Snepvangers, J. (2005). Space–Time Geostatistics. In *Books in Soils, Plants, and the Environment* (pp. 437–451). https://doi.org/10.1201/9781420028188.ch16
Holdaway, R. J., McNeill, S. J., Mason, N. W. H., & Carswell, F. E. (2014). Propagating Uncertainty in Plot-based Estimates of Forest Carbon Stock and Carbon Stock Change. In *Ecosystems* (Vol. 17, Issue 4, pp. 627–640). https://doi.org/10.1007/s10021-014-9749-5
Holmgren, J., & Persson, Å. (2004). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment*, *90*(4), 415–423.
Hopkinson, C., Chasmer, L., Gynan, C., Mahoney, C., & Sitar, M. (2016). Multisensor and Multispectral LiDAR Characterization and Classification of a Forest Environment. In *Canadian Journal of Remote Sensing* (Vol. 42, Issue 5, pp. 501–520). https://doi.org/10.1080/07038992.2016.1196584
Hosoi, F., Nakai, Y., & Omasa, K. (2013). 3-D voxel-based solid modeling of a broad-leaved tree for accurate volume estimation using portable scanning lidar. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 82, pp. 41–48). https://doi.org/10.1016/j.isprsjprs.2013.04.011
Hruska, J., Cermák, J., & Sustek, S. (1999). Mapping tree root systems with ground-penetrating radar. *Tree Physiology*, *19*(2), 125–130.
Huang, W., Sun, G., Ni, W., Zhang, Z., & Dubayah, R. O. (2015). Sensitivity of Multi-Source SAR Backscatter to Changes in Forest Aboveground Biomass. In *Remote Sensing* (Vol. 7, Issue 8, pp. 9587–9609). https://doi.org/10.3390/rs70809587
Hubau, W., Lewis, S. L., Phillips, O. L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., Daniels, A. K., Ewango, C. E. N., Fauset, S., Mukinzi, J. M., Sheil, D., Sonké, B., Sullivan, M. J. P., Sunderland, T. C. H., Taedoumg, H., Thomas, S. C., White, L. J. T., Abernethy, K. A., Adu-Bredu, S., … Zemagho, L. (2020). Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature*, *579*(7797), 80–87.
Hudak, A. T., Fekety, P. A., Kane, V. R., Kennedy, R. E., Filippelli, S. K., Falkowski, M. J., Tinkham, W. T., Smith, A. M. S., Crookston, N. L., Domke, G. M., Corrao, M. V., Bright, B. C., Churchill, D. J., Gould, P. J., McGaughey, R. J., Kane, J. T., & Dong, J. (2020). A carbon monitoring system for mapping regional, annual aboveground biomass across the northwestern USA. In *Environmental Research Letters* (Vol. 15, Issue 9, p. 095003). https://doi.org/10.1088/1748-9326/ab93f9
Hunter, M. O., Keller, M., Victoria, D., & Morton, D. C. (2013). Tree height and tropical forest biomass estimation. In *Biogeosciences* (Vol. 10, Issue 12, pp. 8385–8399). https://doi.org/10.5194/bg-10-8385-2013
Huxley, J. S., & Teissier, G. (1936). Terminology of Relative Growth. In *Nature* (Vol. 137, Issue 3471, pp. 780–781). https://doi.org/10.1038/137780b0
Hyde, P., Dubayah, R., Peterson, B., Blair, J., Hofton, M., Hunsaker, C., Knox, R., & Walker, W. (2005). Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sensing of Environment*, *96*(3-4), 427–437.
Hyyppa, J., & Inkinen, M. (1999). Detecting and estimating attributes for single trees using laser scanner. *Photogramm J Finland*, *16*, 27–42.
Hyyppa, J., Kelle, O., Lehikoinen, M., & Inkinen, M. (2001). A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. In *IEEE Transactions on Geoscience and Remote Sensing* (Vol. 39, Issue 5, pp. 969–975). https://doi.org/10.1109/36.921414
Iglhaut, J., Cabo, C., Puliti, S., Piermattei, L., O'Connor, J., & Rosette, J. (2019). Structure from Motion Photogrammetry in Forestry: a Review. *Current Forestry Reports*, *5*(3), 155–168.
Intergovernmental Panel on Climate Change. (2000). *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*. IPCC/OECD/IEA/IGES.
Intergovernmental Panel on Climate Change. (2003). *Good Practice Guidance for Land Use, Land-Use*

Change and Forestry (J. Penman, M. Gytarsky, T. Hiraishi, T. Krug, D. Kruger, R. Pipatti, L. Buendia, K. Miwa, T. Ngara, K. Tanabe, & F. Wagner (Eds.)). Institute for Global Environmental Strategies (IGES) for the IPCC.

Intergovernmental Panel on Climate Change. (2006). *2006 IPCC Guidelines for National Greenhouse Gas Inventories* (S. Eggleston, L. Buendia, K. Miwa, T. Ngara, & K. Tanabe (Eds.)). Institute for Global Environmental Strategies (IGES), Hayama, Japan on behalf of the IPCC.

Intergovernmental Panel on Climate Change. (2019). *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories* (E. C. Buendia, K. Kiyoto Tanabe, A. Andrej Kranjc, B. Jamsranjav, M. Fukuda, S. Ngarize, A. Osako, Y. Pyrozhenko, P. Shermanau, & S. Federici (Eds.)). Intergovernmental Panel on Climate Change.

Jaakkola, A., Hyyppä, J., Yu, X., Kukko, A., Kaartinen, H., Liang, X., Hyyppä, H., & Wang, Y. (2017). Autonomous Collection of Forest Field Reference—The Outlook and a First Step with UAV Laser Scanning. In *Remote Sensing* (Vol. 9, Issue 8, p. 785). https://doi.org/10.3390/rs9080785

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2003). National-scale biomass estimators for United States tree species. *Forest Science*, *49*(1), 12–35.

Jobbágy, E. G., & Jackson, R. B. (2000). The vertical distribution of soil organic carbon and its relation to climate and vegetation. In *Ecological Applications* (Vol. 10, Issue 2, pp. 423–436). https://doi.org/10.1890/1051-0761(2000)010[0423:tvdoso]2.0.co;2

Johnson, C. E., & Barton, C. C. (2004). Where in the world are my field plots? Using GPS effectively in environmental field studies. In *Frontiers in Ecology and the Environment* (Vol. 2, Issue 9, pp. 475–482). https://doi.org/10.1890/1540-9295(2004)002[0475:witwam]2.0.co;2

Johnson, K. D., Domke, G. M., Russell, M. B., Walters, B., Hom, J., Peduzzi, A., Birdsey, R., Dolan, K., & Huang, W. (2017). Estimating aboveground live understory vegetation carbon in the United States. In *Environmental Research Letters* (Vol. 12, Issue 12, p. 125010). https://doi.org/10.1088/1748-9326/aa8fdb

Johnson, M. O., Galbraith, D., Gloor, M., De Deurwaerder, H., Guimberteau, M., Rammig, A., Thonicke, K., Verbeeck, H., von Randow, C., Monteagudo, A., Phillips, O. L., Brienen, R. J. W., Feldpausch, T. R., Lopez Gonzalez, G., Fauset, S., Quesada, C. A., Christoffersen, B., Ciais, P., Sampaio, G., … Baker, T. R. (2016). Variation in stem mortality rates determines patterns of above-ground biomass in Amazonian forests: implications for dynamic global vegetation models. *Global Change Biology*, *22*(12), 3996–4013.

Joint Committee for Guides in Metrology. (2008). *JCGM 100:2008: Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. JCGM.

Jones, C. D., Arora, V., Friedlingstein, P., Bopp, L., Brovkin, V., Dunne, J., Graven, H., Hoffman, F., Ilyina, T., John, J. G., Jung, M., Kawamiya, M., Koven, C., Pongratz, J., Raddatz, T., Randerson, J. T., & Zaehle, S. (2016). C4MIP – The Coupled Climate–Carbon Cycle Model Intercomparison Project: experimental protocol for CMIP6. *Geoscientific Model Development*, *9*(8), 2853–2880.

Jucker, T., Asner, G. P., Dalponte, M., Brodrick, P. G., Philipson, C. D., Vaughn, N. R., Teh, Y. A., Brelsford, C., Burslem, D. F. R. P., Deere, N. J., Ewers, R. M., Kvasnica, J., Lewis, S. L., Malhi, Y., Milne, S., Nilus, R., Pfeifer, M., Phillips, O. L., Qie, L., … Coomes, D. A. (2018). Estimating aboveground carbon density and its uncertainty in Borneo's structurally complex tropical forests using airborne laser scanning. *Biogeosciences*, *15*(12), 3811–3830.

Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., Dalponte, M., van Ewijk, K. Y., Forrester, D. I., Haeni, M., Higgins, S. I., Holdaway, R. J., Iida, Y., Lorimer, C., Marshall, P. L., Momo, S., Moncrieff, G. R., Ploton, P., Poorter, L., … Coomes, D. A. (2017). Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global Change Biology*, *23*(1), 177–190.

Jung, J., Kim, S., Hong, S., Kim, K., Kim, E., Im, J., & Heo, J. (2013). Effects of national forest inventory plot

location error on forest carbon stock estimation using k-nearest neighbor algorithm. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 81, pp. 82–92). https://doi.org/10.1016/j.isprsjprs.2013.04.008

Justice, C. O., Giglio, L., Korontzi, S., Owens, J., Morisette, J. T., Roy, D., Descloitres, J., Alleaume, S., Petitcolin, F., & Kaufman, Y. (2002). The MODIS fire products. In *Remote Sensing of Environment* (Vol. 83, Issues 1-2, pp. 244–262). https://doi.org/10.1016/s0034-4257(02)00076-7

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B. M., & Wu, J.-C. (2012). An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. In *Remote Sensing* (Vol. 4, Issue 4, pp. 950–974). https://doi.org/10.3390/rs4040950

Kalliovirta, J., Laasasenaho, J., & Kangas, A. (2005). Evaluation of the Laser-relascope. In *Forest Ecology and Management* (Vol. 204, Issues 2-3, pp. 181–194). https://doi.org/10.1016/j.foreco.2004.09.020

Kampe, T. U., Johnson, B. R., Kuester, M. A., & Keller, M. (2010). NEON: the first continental-scale ecological observatory with airborne remote sensing of vegetation canopy biochemistry and structure. In *Journal of Applied Remote Sensing* (Vol. 4, Issue 1, p. 043510). https://doi.org/10.1117/1.3361375

Kelbe, D., van Aardt, J., Romanczyk, P., van Leeuwen, M., & Cawse-Nicholson, K. (2016). Marker-Free Registration of Forest Terrestrial Laser Scanner Data Pairs With Embedded Confidence Metrics. In *IEEE Transactions on Geoscience and Remote Sensing* (Vol. 54, Issue 7, pp. 4314–4330). https://doi.org/10.1109/tgrs.2016.2539219

Keller, M., Palace, M., Asner, G. P., Pereira, R., & Silva, J. N. M. (2004). Coarse woody debris in undisturbed and logged forests in the eastern Brazilian Amazon. In *Global Change Biology* (Vol. 10, Issue 5, pp. 784–795). https://doi.org/10.1111/j.1529-8817.2003.00770.x

Keller, M., Palace, M., & Hurtt, G. (2001). Biomass estimation in the Tapajos National Forest, Brazil. In *Forest Ecology and Management* (Vol. 154, Issue 3, pp. 371–382). https://doi.org/10.1016/s0378-1127(01)00509-6

Kellner, J. R., Armston, J., Birrer, M., Cushman, K. C., Duncanson, L. I., Eck, C., Falleger, C., Imbach, B., Král, K., Krůček, M., Trochta, J., Vrška, T., & Zgraggen, C. (2019). New Opportunities for Forest Remote Sensing Through Ultra-High-Density Drone Lidar. *Surveys in Geophysics*, *40*(4), 959–977.

Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. In *Remote Sensing of Environment* (Vol. 114, Issue 12, pp. 2897–2910). https://doi.org/10.1016/j.rse.2010.07.008

Kitahara, F., Mizoue, N., & Yoshida, S. (2009). Evaluation of data quality in Japanese National Forest Inventory. *Environmental Monitoring and Assessment*, *159*(1-4), 331–340.

Kitahara, F., Mizoue, N., & Yoshida, S. (2010). Effects of training for inexperienced surveyors on data quality of tree diameter and height measurements. In *Silva Fennica* (Vol. 44, Issue 4). https://doi.org/10.14214/sf.133

Knapp, N., Fischer, R., Cazcarra-Bes, V., & Huth, A. (2020). Structure metrics to generalize biomass estimation from lidar across forest types from different continents. *Remote Sensing of Environment*, *237*, 111597.

Knapp, N., Fischer, R., & Huth, A. (2018). Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states. *Remote Sensing of Environment*, *205*, 199–209.

Koch, B., Heyder, U., & Weinacker, H. (2006). Detection of Individual Tree Crowns in Airborne Lidar Data. In *Photogrammetric Engineering & Remote Sensing* (Vol. 72, Issue 4, pp. 357–363). https://doi.org/10.14358/pers.72.4.357

Kodani, E., & Awaya, Y. (2013). Estimating stand parameters in manmade coniferous forest stands using low-density LiDAR. *Journal of the Japan Society of Photogrammetry and Remote Sensing*, *52*(2),

44–55.
Krause, K., & Goulden, T. (2015). *NEON L0-to-L1 Discrete Return LiDAR Algorithm Theoretical Basis Document (ATBD)* (NEON Doc. #: NEON.DOC.001292vA). NSF NEON.
Krishna Moorthy, S. M., Calders, K., Boni Vicari, M., & Verbeeck, H. (2019). Improved Supervised Learning-Based Approach for Leaf and Wood Classification From LiDAR Point Clouds of Forests. *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society*, 1–14.
Kükenbrink, D., Schneider, F. D., Leiterer, R., Schaepman, M. E., & Morsdorf, F. (2017). Quantification of hidden canopy volume of airborne laser scanning data using a voxel traversal algorithm. *Remote Sensing of Environment*, *194*, 424–436.
Labriere, N., Tao, S., Chave, J., Scipal, K., Le Toan, T., Abernethy, K., Alonso, A., Barbier, N., Bissiengou, P., Casal, T., Davies, S. J., Ferraz, A., Herault, B., Jaouen, G., Jeffery, K. J., Kenfack, D., Korte, L., Lewis, S. L., Malhi, Y., … Saatchi, S. S. (2018). In Situ Reference Datasets From the TropiSAR and AfriSAR Campaigns in Support of Upcoming Spaceborne Biomass Missions. In *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (Vol. 11, Issue 10, pp. 3617–3627). https://doi.org/10.1109/jstars.2018.2851606
Langner, A., Achard, F., & Grassi, G. (2014). Can recent pan-tropical biomass maps be used to derive alternative Tier 1 values for reporting REDD activities under UNFCCC? In *Environmental Research Letters* (Vol. 9, Issue 12, p. 124008). https://doi.org/10.1088/1748-9326/9/12/124008
Larjavaara, M., & Muller-Landau, H. C. (2013). Measuring tree height: a quantitative comparison of two common field methods in a moist tropical forest. *Methods in Ecology and Evolution / British Ecological Society*, *4*(9), 793–801.
Lasky, J. R., Uriarte, M., Boukili, V. K., Erickson, D. L., John Kress, W., & Chazdon, R. L. (2014). The relationship between tree biodiversity and biomass dynamics changes with tropical forest succession. *Ecology Letters*, *17*(9), 1158–1167.
Lau, A., Calders, K., Bartholomeus, H., Martius, C., Raumonen, P., Herold, M., Vicari, M., Sukhdeo, H., Singh, J., & Goodman, R. (2019). Tree Biomass Equations from Terrestrial LiDAR: A Case Study in Guyana. In *Forests* (Vol. 10, Issue 6, p. 527). https://doi.org/10.3390/f10060527
Laurin, G. V., Ding, J., Disney, M., Bartholomeus, H., Herold, M., Papale, D., & Valentini, R. (2019). Tree height in tropical forest as measured by different ground, proximal, and remote sensing instruments, and impacts on above ground biomass estimates. *International Journal of Applied Earth Observation and Geoinformation*, *82*, 101899.
Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong, L., & Shreenan, R. (2003). Combined high-density lidar and multispectral imagery for individual tree crown analysis. In *Canadian Journal of Remote Sensing* (Vol. 29, Issue 5, pp. 633–649). https://doi.org/10.5589/m03-024
Ledo, A., Paul, K. I., Burslem, D. F. R. P., Ewel, J. J., Barton, C., Battaglia, M., Brooksbank, K., Carter, J., Eid, T. H., England, J. R., Fitzgerald, A., Jonson, J., Mencuccini, M., Montagu, K. D., Montero, G., Mugasha, W. A., Pinkard, E., Roxburgh, S., Ryan, C. M., … Chave, J. (2018). Tree size and climatic water deficit control root to shoot ratio in individual trees globally. *The New Phytologist*, *217*(1), 8–11.
Lefsky, M. A. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters*, *37*(15). https://doi.org/10.1029/2010GL043622
Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A., & Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sensing of Environment*, *70*(3), 339–361.
Lefsky, M. A., Cohen, W. B., Harding, D. J., Parker, G. G., Acker, S. A., & Gower, S. T. (2002). Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology and Biogeography: A Journal of*

Macroecology, *11*(5), 393–399.
Lehtonen, A., Cienciala, E., Tatarinov, F., & Mäkipää, R. (2007). Uncertainty estimation of biomass expansion factors for Norway spruce in the Czech Republic. In *Annals of Forest Science* (Vol. 64, Issue 2, pp. 133–140). https://doi.org/10.1051/forest:2006097
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 and Management*, *10*(1), 3.
Letcher, S. G., & Chazdon, R. L. (2009). Rapid Recovery of Biomass, Species Richness, and Species Composition in a Forest Chronosequence in Northeastern Costa Rica. In *Biotropica* (Vol. 41, Issue 5, pp. 608–617). https://doi.org/10.1111/j.1744-7429.2009.00517.x
Liang, J., Crowther, T. W., Picard, N., Wiser, S., Zhou, M., Alberti, G., Schulze, E.-D., McGuire, A. D., Bozzato, F., Pretzsch, H., de-Miguel, S., Paquette, A., Hérault, B., Scherer-Lorenzen, M., Barrett, C. B., Glick, H. B., Hengeveld, G. M., Nabuurs, G.-J., Pfautsch, S., … Reich, P. B. (2016). Positive biodiversity-productivity relationship predominant in global forests. *Science*, *354*(6309). https://doi.org/10.1126/science.aaf8957
Liang, X., Wang, Y., Jaakkola, A., Kukko, A., Kaartinen, H., Hyyppä, J., Honkavaara, E., & Liu, J. (2015). Forest Data Collection Using Terrestrial Image-Based Point Clouds From a Handheld Camera Compared to Terrestrial and Personal Laser Scanning. *IEEE Transactions on Geoscience and Remote Sensing: A Publication of the IEEE Geoscience and Remote Sensing Society*, *53*(9), 5117–5132.
Lim, K., Treitz, P., Wulder, M., St-Onge, B., & Flood, M. (2003). LiDAR remote sensing of forest structure. *Progress in Physical Geography: Earth and Environment*, *27*(1), 88–106.
Lisein, J., Pierrot-Deseilligny, M., Bonnet, S., & Lejeune, P. (2013). A Photogrammetric Workflow for the Creation of a Forest Canopy Height Model from Small Unmanned Aerial System Imagery. In *Forests* (Vol. 4, Issue 4, pp. 922–944). https://doi.org/10.3390/f4040922
Li, Z., Schaefer, M., Strahler, A., Schaaf, C., & Jupp, D. (2018). On the utilization of novel spectral laser scanning for three-dimensional classification of vegetation elements. *Interface Focus*, *8*(2), 20170039.
Lloyd, J., Bird, M. I., Vellen, L., Miranda, A. C., Veenendaal, E. M., Djagbletey, G., Miranda, H. S., Cook, G., & Farquhar, G. D. (2008). Contributions of woody and herbaceous vegetation to tropical savanna ecosystem productivity: a quasi-global estimate. In *Tree Physiology* (Vol. 28, Issue 3, pp. 451–468). https://doi.org/10.1093/treephys/28.3.451
Longo, M., Keller, M., dos-Santos, M. N., Leitold, V., Pinagé, E. R., Baccini, A., Saatchi, S., Nogueira, E. M., Batistella, M., & Morton, D. C. (2016). Aboveground biomass variability across intact and degraded forests in the Brazilian Amazon: AMAZON INTACT AND DEGRADED FOREST BIOMASS. *Global Biogeochemical Cycles*, *30*(11), 1639–1660.
Lopez-Gonzalez, G., Lewis, S. L., Burkitt, M., & Phillips, O. L. (2011). ForestPlots.net: a web application and research tool to manage and analyse tropical forest plot data. In *Journal of Vegetation Science* (Vol. 22, Issue 4, pp. 610–613). https://doi.org/10.1111/j.1654-1103.2011.01312.x
Luck, L., Hutley, L. B., Calders, K., & Levick, S. R. (2020). Exploring the Variability of Tropical Savanna Tree Structural Allometry with Terrestrial Laser Scanning. *Remote Sensing*, *12*(23), 3893.
Magalhães, T. M., & Seifert, T. (2015a). Estimation of Tree Biomass, Carbon Stocks, and Error Propagation in Mecrusse Woodlands. In *Open Journal of Forestry* (Vol. 05, Issue 04, pp. 471–488). https://doi.org/10.4236/ojf.2015.54041
Magalhães, T. M., & Seifert, T. (2015b). Tree component biomass expansion factors and root-to-shoot ratio of Lebombo ironwood: measurement uncertainty. *Carbon Balance and Management*, *10*, 9.
Magnussen, S., Köhl, M., & Olschofsky, K. (2014). Error propagation in stock-difference and gain–loss estimates of a forest biomass carbon balance. In *European Journal of Forest Research* (Vol. 133, Issue 6, pp. 1137–1155). https://doi.org/10.1007/s10342-014-0828-0

Malhi, Y., Wood, D., Baker, T. R., Wright, J., Phillips, O. L., Cochrane, T., Meir, P., Chave, J., Almeida, S., Arroyo, L., Higuchi, N., Killeen, T. J., Laurance, S. G., Laurance, W. F., Lewis, S. L., Monteagudo, A., Neill, D. A., Vargas, P. N., Pitman, N. C. A., … Vinceti, B. (2006). The regional variation of aboveground live biomass in old-growth Amazonian forests. In *Global Change Biology* (Vol. 12, Issue 7, pp. 1107–1138). https://doi.org/10.1111/j.1365-2486.2006.01120.x

Maltamo, M., Peuhkurinen, J., Malinen, J., Vauhkonen, J., Packalén, P., & Tokola, T. (2009). Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. In *Silva Fennica* (Vol. 43, Issue 3). https://doi.org/10.14214/sf.203

Mandallaz, D. (2007). *Sampling Techniques for Forest Inventories*. https://doi.org/10.1201/9781584889779

Margolis, H. A., Nelson, R. F., Montesano, P. M., Beaudoin, A., Sun, G., Andersen, H.-E., & Wulder, M. A. (2015a). Combining satellite lidar, airborne lidar, and ground plots to estimate the amount and distribution of aboveground biomass in the boreal forest of North America. In *Canadian Journal of Forest Research* (Vol. 45, Issue 7, pp. 838–855). https://doi.org/10.1139/cjfr-2015-0006

Margolis, H. A., Nelson, R. F., Montesano, P. M., Beaudoin, A., Sun, G., Andersen, H.-E., & Wulder, M. A. (2015b). Combining satellite lidar, airborne lidar, and ground plots to estimate the amount and distribution of aboveground biomass in the boreal forest of North America. *Canadian Journal of Forest Research. Journal Canadien de La Recherche Forestiere*, *45*(7), 838–855.

Martin, A. R., Doraisami, M., & Thomas, S. C. (2018). Global patterns in wood carbon concentration across the world's trees and forests. In *Nature Geoscience* (Vol. 11, Issue 12, pp. 915–920). https://doi.org/10.1038/s41561-018-0246-x

Martin, A. R., & Thomas, S. C. (2011). A reassessment of carbon content in tropical trees. *PloS One*, *6*(8), e23533.

Mascaro, J., Detto, M., Asner, G. P., & Muller-Landau, H. C. (2011). Evaluating uncertainty in mapping forest carbon with airborne LiDAR. In *Remote Sensing of Environment* (Vol. 115, Issue 12, pp. 3770–3774). https://doi.org/10.1016/j.rse.2011.07.019

Mavouroulou, Q. M., Ngomanda, A., Obiang, N. L. E., Lebamba, J., Gomat, H., Mankou, G. S., Loumeto, J., Iponga, D. M., Ditsouga, F. K., Koumba, R. Z., Bobé, K. H. B., Lépengué, N., Mbatchi, B., & Picard, N. (2014). How to improve allometric equations to estimate forest biomass stocks? Some hints from a central African forest. In *Canadian Journal of Forest Research* (Vol. 44, Issue 7, pp. 685–691). https://doi.org/10.1139/cjfr-2013-0520

McDowell, N. G., Fisher, R. A., Xu, C., Domec, J. C., Hölttä, T., Mackay, D. S., Sperry, J. S., Boutz, A., Dickman, L., Gehres, N., Limousin, J. M., Macalady, A., Martínez-Vilalta, J., Mencuccini, M., Plaut, J. A., Ogée, J., Pangle, R. E., Rasse, D. P., Ryan, M. G., … Pockman, W. T. (2013). Evaluating theories of drought-induced vegetation mortality using a multimodel-experiment framework. *The New Phytologist*, *200*(2), 304–321.

McGaughey, R. (2020). *FUSION/LDV LIDAR analysis and visualization software*.

McRoberts, R. E. (2010). Probability- and model-based approaches to inference for proportion forest using satellite imagery as ancillary data. In *Remote Sensing of Environment* (Vol. 114, Issue 5, pp. 1017–1025). https://doi.org/10.1016/j.rse.2009.12.013

McRoberts, R. E., Chen, Q., Domke, G. M., Ståhl, G., Saarela, S., & Westfall, J. A. (2016). Hybrid estimators for mean aboveground carbon per unit area. In *Forest Ecology and Management* (Vol. 378, pp. 44–56). https://doi.org/10.1016/j.foreco.2016.07.007

McRoberts, R. E., Hahn, J. T., Hefty, G. J., & Van Cleve, J. R. (1994). Variation in forest inventory field measurements. In *Canadian Journal of Forest Research* (Vol. 24, Issue 9, pp. 1766–1770). https://doi.org/10.1139/x94-228

McRoberts, R. E., Moser, P., Oliveira, L. Z., & Vibrans, A. C. (2015). A general method for assessing the effects of uncertainty in individual-tree volume model predictions on large-area volume estimates

with a subtropical forest illustration. In *Canadian Journal of Forest Research* (Vol. 45, Issue 1, pp. 44–51). https://doi.org/10.1139/cjfr-2014-0266

McRoberts, R. E., Næsset, E., & Gobakken, T. (2013). Inference for lidar-assisted estimation of forest growing stock volume. In *Remote Sensing of Environment* (Vol. 128, pp. 268–275). https://doi.org/10.1016/j.rse.2012.10.007

McRoberts, R. E., Næsset, E., Gobakken, T., Chirici, G., Condés, S., Hou, Z., Saarela, S., Chen, Q., Ståhl, G., & Walters, B. F. (2018). Assessing components of the model-based mean square error estimator for remote sensing assisted forest applications. *Canadian Journal of Forest Research. Journal Canadien de La Recherche Forestiere*, *48*(6), 642–649.

McRoberts, R. E., Næsset, E., Liknes, G. C., Chen, Q., Walters, B. F., Saatchi, S., & Herold, M. (2019). Using a Finer Resolution Biomass Map to Assess the Accuracy of a Regional, Map-Based Estimate of Forest Biomass. In *Surveys in Geophysics* (Vol. 40, Issue 4, pp. 1001–1015). https://doi.org/10.1007/s10712-019-09507-1

McRoberts, R. E., Næsset, E., Saatchi, S., Liknes, G. C., Walters, B. F., & Chen, Q. (2019). Local validation of global biomass maps. *International Journal of Applied Earth Observation and Geoinformation*, *83*, 101931.

McRoberts, R. E., & Westfall, J. A. (2016). Propagating uncertainty through individual tree volume model predictions to large-area volume estimates. In *Annals of Forest Science* (Vol. 73, Issue 3, pp. 625–633). https://doi.org/10.1007/s13595-015-0473-x

Means, J. E., Acker, S. A., Harding, D. J., Blair, J. B., Lefsky, M. A., Cohen, W. B., Harmon, M. E., & McKee, W. A. (1999). Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the Western Cascades of Oregon. *Remote Sensing of Environment*, *67*(3), 298–308.

Meier, C. (2014). *TOS science design for plant biomass, productivity, and leaf area index* (Version NEON.DOC.000914). National Ecological Observatory Network. https://data.neonscience.org/documents

Meier, C. (2017). *TOS protocol and procedure: Measurement of vegetation structure* (Version NEON.DOC.000987). National Ecological Observatory Network. https://data.neonscience.org/documents

Meier, C. (2018). *NEON user guide to woody plant vegetation structure (NEON.DP1.10098) and non-herbaceous perennial vegetation structure (NEON.DP1.10045)*. National Ecological Observatory Network. https://data.neonscience.org/documents

Melson, S. L., Harmon, M. E., Fried, J. S., & Domingo, J. B. (2011). Estimates of live-tree carbon stores in the Pacific Northwest are sensitive to model selection. *Carbon Balance and Management*, *6*, 2.

Memiaghe, H. R., Lutz, J. A., Korte, L., Alonso, A., & Kenfack, D. (2016). Ecological Importance of Small-Diameter Trees to the Structure, Diversity and Biomass of a Tropical Evergreen Forest at Rabi, Gabon. *PloS One*, *11*(5), e0154988.

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. In *Remote Sensing of Environment* (Vol. 159, pp. 307–317). https://doi.org/10.1016/j.rse.2014.12.019

Methley, J. (2001). Getting the measure of your timber: let's start with diameter. *Forestry & British Timber*, *30*, 26–28.

Meyer, V., Saatchi, S. S., Chave, J., Dalling, J. W., Bohlman, S., Fricker, G. A., Robinson, C., Neumann, M., & Hubbell, S. (2013). Detecting tropical forest biomass dynamics from repeated airborne lidar measurements. In *Biogeosciences* (Vol. 10, Issue 8, pp. 5421–5438). https://doi.org/10.5194/bg-10-5421-2013

Miller, C. E., Griffith, P. C., Goetz, S. J., Hoy, E. E., Pinto, N., McCubbin, I. B., Thorpe, A. K., Hofton, M., Hodkinson, D., Hansen, C., Woods, J., Larson, E., Kasischke, E. S., & Margolis, H. A. (2019). An overview of ABoVE airborne campaign data acquisitions and science opportunities. *Environmental*

Research Letters: ERL [Web Site], *14*(8), 080201.
Mitchard, E. T. A., Feldpausch, T. R., Brienen, R. J. W., Lopez-Gonzalez, G., Monteagudo, A., Baker, T. R., Lewis, S. L., Lloyd, J., Quesada, C. A., Gloor, M., Ter Steege, H., Meir, P., Alvarez, E., Araujo-Murakami, A., Aragão, L. E. O. C., Arroyo, L., Aymard, G., Banki, O., Bonal, D., … Phillips, O. L. (2014). Markedly divergent estimates of Amazon forest carbon density from ground plots and satellites. *Global Ecology and Biogeography: A Journal of Macroecology*, *23*(8), 935–946.
Mitchard, E. T. A., & Flintrop, C. M. (2013). Woody encroachment and forest degradation in sub-Saharan Africa's woodlands and savannas 1982-2006. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1625), 20120406.
Mitchard, E. T. A., Saatchi, S. S., Baccini, A., Asner, G. P., Goetz, S. J., Harris, N. L., & Brown, S. (2013). Uncertainty in the spatial distribution of tropical forest biomass: a comparison of pan-tropical maps. *Carbon Balance and Management*, *8*(1), 10.
Mokany, K., John Raison, R., & Prokushkin, A. S. (2006). Critical analysis of root : shoot ratios in terrestrial biomes. In *Global Change Biology* (Vol. 12, Issue 1, pp. 84–96). https://doi.org/10.1111/j.1365-2486.2005.001043.x
Molto, Q., Rossi, V., & Blanc, L. (2013). Error propagation in biomass estimation in tropical forests. In *Methods in Ecology and Evolution* (Vol. 4, Issue 2, pp. 175–183). https://doi.org/10.1111/j.2041-210x.2012.00266.x
Momo, S. T., Ploton, P., Martin-Ducup, O., Lehnebach, R., Fortunel, C., Sagang, L. B. T., Boyemba, F., Couteron, P., Fayolle, A., Libalah, M., Loumeto, J., Medjibe, V., Ngomanda, A., Obiang, D., Pélissier, R., Rossi, V., Yongo, O., Sonké, B., & Barbier, N. (2020). Leveraging Signatures of Plant Functional Strategies in Wood Density Profiles of African Trees to Correct Mass Estimations From Terrestrial Laser Data. In *Scientific Reports* (Vol. 10, Issue 1). https://doi.org/10.1038/s41598-020-58733-w
Momo Takoudjou, S., Ploton, P., Sonké, B., Hackenberg, J., Griffon, S., Coligny, F., Kamdem, N. G., Libalah, M., Mofack, G. I. I., Le Moguédec, G., Pélissier, R., & Barbier, N. (2018). Using terrestrial laser scanning data to estimate large tropical trees biomass and calibrate allometric models: A comparison with traditional destructive approach. In *Methods in Ecology and Evolution* (Vol. 9, Issue 4, pp. 905–916). https://doi.org/10.1111/2041-210x.12933
Moncrieff, G. R., Chamaillé-Jammes, S., Higgins, S. I., O'Hara, R. B., & Bond, W. J. (2011). Tree allometries reflect a lifetime of herbivory in an African savanna. In *Ecology* (Vol. 92, Issue 12, pp. 2310–2315). https://doi.org/10.1890/11-0230.1
Moorcroft, P. R., Hurtt, G. C., & Pacala, S. W. (2001). A method for scaling vegetation dynamics: The ecosystem demography model (ED). In *Ecological Monographs* (Vol. 71, Issue 4, pp. 557–586). https://doi.org/10.1890/0012-9615(2001)071[0557:amfsvd]2.0.co;2
Moran, L. A., & Williams, R. A. (2002). Field Note—Comparison of Three Dendrometers in Measuring Diameter at Breast Height Field Note. In *Northern Journal of Applied Forestry* (Vol. 19, Issue 1, pp. 28–33). https://doi.org/10.1093/njaf/19.1.28
Morsdorf, F., Eck, C., Zgraggen, C., Imbach, B., Schneider, F. D., & Kükenbrink, D. (2017). UAV-based LiDAR acquisition for the derivation of high-resolution forest and ground information. In *The Leading Edge* (Vol. 36, Issue 7, pp. 566–570). https://doi.org/10.1190/tle36070566.1
Motohka, T., Isoguchi, O., Sakashita, M., & Shimada, M. (2018). Results of ALOS-2 PALSAR-2 Calibration and Validation After 3 Years of Operation. In *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*. https://doi.org/10.1109/igarss.2018.8519118
Moulatlet, G. M., Zuquim, G., Figueiredo, F. O. G., Lehtonen, S., Emilio, T., Ruokolainen, K., & Tuomisto, H. (2017). Using digital soil maps to infer edaphic affinities of plant species in Amazonia: Problems and prospects. *Ecology and Evolution*, *7*(20), 8463–8477.
Mulatu, K., Mora, B., Kooistra, L., & Herold, M. (2017). Biodiversity Monitoring in Changing Tropical Forests: A Review of Approaches and New Opportunities. In *Remote Sensing* (Vol. 9, Issue 10, p.

1059). https://doi.org/10.3390/rs9101059
Myers, C. A. (1961). *Variation in Measuring Diameter at Breast Height of Mature Ponderosa Pine*.
Næsset, E. (2001). Effects of Differential Single- and Dual-Frequency GPS and GLONASS Observations on Point Accuracy under Forest Canopies. *Photogrammetric Engineering & Remote Sensing*, *67*(9), 1021–1026.
Næsset, E. (2003). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research / Issued Bimonthly by the Nordic Forest Research Cooperation Committee* , *19*(2), 164–179.
Næsset, E. (2009). Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. In *Remote Sensing of Environment* (Vol. 113, Issue 1, pp. 148–159). https://doi.org/10.1016/j.rse.2008.09.001
Næsset, E., Bollandsås, O. M., Gobakken, T., Solberg, S., & McRoberts, R. E. (2015). The effects of field plot size on model-assisted estimation of aboveground biomass change using multitemporal interferometric SAR and airborne laser scanning data. *Remote Sensing of Environment*, *168*, 252–264.
Næsset, E., Gobakken, T., Bollandsås, Ole Martin, Gregoire, Timothy G, Nelson, Ross, & Ståhl, G. (2013). Comparison of precision of biomass estimates in regional field sample surveys and airborne LiDAR-assisted surveys in Hedmark County, Norway. *Remote Sensing of Environment*, *130*, 108–120.
Narine, L. L., Popescu, S., Neuenschwander, A., Zhou, T., Srinivasan, S., & Harbeck, K. (2019). Estimating aboveground biomass and forest canopy cover with simulated ICESat-2 data. *Remote Sensing of Environment*, *224*, 1–11.
Neigh, C. S. R., Nelson, R. F., Ranson, K. J., Margolis, H. A., Montesano, P. M., Sun, G., Kharuk, V., Næsset, E., Wulder, M. A., & Andersen, H.-E. (2013). Taking stock of circumboreal forest carbon with ground measurements, airborne and spaceborne LiDAR. *Remote Sensing of Environment*, *137*, 274–287.
Nelson, R. F., Hyde, P., Johnson, P., Emessiene, B., Imhoff, M. L., Campbell, R., & Edwards, W. (2007). Investigating RaDAR–LiDAR synergy in a North Carolina pine forest. *Remote Sensing of Environment*, *110*(1), 98–108.
Nelson, R., Gobakken, T., Næsset, E., Gregoire, T. G., Ståhl, G., Holm, S., & Flewelling, J. (2012). Lidar sampling — Using an airborne profiler to estimate forest biomass in Hedmark County, Norway. In *Remote Sensing of Environment* (Vol. 123, pp. 563–578). https://doi.org/10.1016/j.rse.2011.10.036
Nelson, R., Krabill, W., & MacLean, G. (1984). Determining forest canopy characteristics using airborne laser data. In *Remote Sensing of Environment* (Vol. 15, Issue 3, pp. 201–212). https://doi.org/10.1016/0034-4257(84)90031-2
Nelson, R., Krabill, W., & Tonelli, J. (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment*, *24*(2), 247–267.
Nelson, R., Margolis, H., Montesano, P., Sun, G., Cook, B., Corp, Larry, Andersen, H.-E., deJong, B., Pellat, F. P., Fickel, T., Kauffman, J., & Prisley, S. (2017). Lidar-based estimates of aboveground biomass in the continental US and Mexico using ground, airborne, and satellite observations. *Remote Sensing of Environment*, *188*, 127–140.
Nester, M. R. (1981). *Assessment and measurement errors in slash pine research plots*.
Neumann, M., Moreno, A., Mues, V., Härkönen, S., Mura, M., Bouriaud, O., Lang, M., Achten, W. M. J., Thivolle-Cazat, A., Bronisz, K., Merganič, J., Decuyper, M., Alberdi, I., Astrup, R., Mohren, F., & Hasenauer, H. (2016). Comparison of carbon estimation methods for European forests. In *Forest Ecology and Management* (Vol. 361, pp. 397–420). https://doi.org/10.1016/j.foreco.2015.11.016
Newnham, G., Armston, J., Muir, J., Goodwin, N., Tindall, D., Culvenor, D., Püschel, P., Nyström, M., & Johansen, K. (2012). *Evaluation of terrestrial laser scanners for measuring vegetation structure* (Manuscript ID: EP124571). CSIRO Sustainable Agriculture Flagship.

Newnham, G. J., Armston, J. D., Calders, K., Disney, M., Lovell, J. L., Schaaf, C. B., Strahler, A. H., & Mark Danson, F. M. (2015). Terrestrial Laser Scanning for Plot-Scale Forest Measurement. In *Current Forestry Reports* (Vol. 1, Issue 4, pp. 239–251). https://doi.org/10.1007/s40725-015-0025-5

Ngomanda, A., Obiang, N. L. E., Lebamba, J., Mavouroulou, Q. M., Gomat, H., Mankou, G. S., Loumeto, J., Iponga, D. M., Ditsouga, F. K., Koumba, R. Z., Bobé, K. H. B., Okouyi, C. M., Nyangadouma, R., Lépengué, N., Mbatchi, B., & Picard, N. (2014). Site-specific versus pantropical allometric equations: Which option to estimate the biomass of a moist central African forest? In *Forest Ecology and Management* (Vol. 312, pp. 1–9). https://doi.org/10.1016/j.foreco.2013.10.029

Nieto-Quintano, P., Mitchard, E. T. A., Odende, R., Mouwembe, M. A. B., Rayden, T., & Ryan, C. M. (2018). The mesic savannas of the Bateke Plateau: carbon stocks and floristic composition. *Biotropica*, *50*(6), 868–880.

Niklas, K. J. (1994). *Plant Allometry: The Scaling of Form and Process*. University of Chicago Press.

Ni-Meister, W., Lee, S., Strahler, A. H., Woodcock, C. E., Schaaf, C., Yao, T., Jon Ranson, K., Sun, G., & Bryan Blair, J. (2010). Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from lidar remote sensing. In *Journal of Geophysical Research: Biogeosciences* (Vol. 115, Issue G2). https://doi.org/10.1029/2009jg000936

Nogueira, E. M., Nelson, B. W., & Fearnside, P. M. (2006). Volume and biomass of trees in central Amazonia: influence of irregularly shaped and hollow trunks. In *Forest Ecology and Management* (Vol. 227, Issues 1-2, pp. 14–21). https://doi.org/10.1016/j.foreco.2006.02.004

Ogle, S. (n.d.). [Personal communication].

Omasa, K., Qiu, G. Y., Watanuki, K., Yoshimi, K., & Akiyama, Y. (2003). Accurate estimation of forest carbon stocks by 3-D remote sensing of individual trees. *Environmental Science & Technology*, *37*(6), 1198–1201.

Omule, S. A. Y. (1980). Personal Bias in Forest Measurements. In *The Forestry Chronicle* (Vol. 56, Issue 5, pp. 222–224). https://doi.org/10.5558/tfc56222-5

Pan, Y., Birdsey, R. A., Phillips, O. L., & Jackson, R. B. (2013). The Structure, Distribution, and Biomass of the World's Forests. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 593–622.

Patterson, P. L., Healey, S. P., Ståhl, G., Saarela, S., Holm, S., Andersen, H.-E., Dubayah, R. O., Duncanson, L., Hancock, S., Armston, J., Kellner, J. R., Cohen, W. B., & Yang, Z. (2019). Statistical properties of hybrid estimators proposed for GEDI—NASA's global ecosystem dynamics investigation. In *Environmental Research Letters* (Vol. 14, Issue 6, p. 065007). https://doi.org/10.1088/1748-9326/ab18df

Paul, K. I. (2019). [Personal communication].

Paul, K. I., Larmour, J., Specht, A., Zerihun, A., Ritson, P., Roxburgh, S. H., Sochacki, S., Lewis, T., Barton, C. V. M., England, J. R., Battaglia, M., O'Grady, A., Pinkard, E., Applegate, G., Jonson, J., Brooksbank, K., Sudmeyer, R., Wildy, D., Montagu, K. D., … Hobbs, T. (2019). Testing the generality of below-ground biomass allometry across plant functional types. In *Forest Ecology and Management* (Vol. 432, pp. 102–114). https://doi.org/10.1016/j.foreco.2018.08.043

Paul, K. I., Larmour, J. S., Roxburgh, S. H., England, J. R., Davies, M. J., & Luck, H. D. (2017). Measurements of stem diameter: implications for individual- and stand-level errors. *Environmental Monitoring and Assessment*, *189*(8), 416.

Paul, K. I., Radtke, P. J., Roxburgh, S. H., Larmour, J., Waterworth, R., Butler, D., Brooksbank, K., & Ximenes, F. (2018). Validation of allometric biomass models: How to have confidence in the application of existing models. *Forest Ecology and Management*, *412*, 70–79.

Paul, K. I., Roxburgh, S. H., Chave, J., England, J. R., Zerihun, A., Specht, A., Lewis, T., Bennett, L. T., Baker, T. G., Adams, M. A., Huxtable, D., Montagu, K. D., Falster, D. S., Feller, M., Sochacki, S., Ritson, P., Bastin, G., Bartle, J., Wildy, D., … Sinclair, J. (2016). Testing the generality of above-ground biomass allometry across plant functional types at the continent scale. *Global Change Biology*, *22*(6),

2106–2124.
Paul, K. I., Roxburgh, S. H., & Larmour, J. S. (2017). Moisture content correction: Implications of measurement errors on tree- and site-based estimates of biomass. In *Forest Ecology and Management* (Vol. 392, pp. 164–175). https://doi.org/10.1016/j.foreco.2017.02.007
Paynter, I., Genest, D., Peri, F., & Schaaf, C. (2018). Bounding uncertainty in volumetric geometric models for terrestrial lidar observations of ecosystems. *Interface Focus*, *8*(2), 20170043.
Paynter, I., Genest, D., Saenz, E., Peri, F., Li, Z., Strahler, A., & Schaaf, C. (2018). Quality Assessment of Terrestrial Laser Scanner Ecosystem Observations Using Pulse Trajectories. In *IEEE Transactions on Geoscience and Remote Sensing* (Vol. 56, Issue 11, pp. 6324–6333). https://doi.org/10.1109/tgrs.2018.2836947
Paynter, I., Saenz, E., Genest, D., Peri, F., Erb, A., Li, Z., Wiggin, K., Muir, J., Raumonen, P., Schaaf, E. S., Strahler, A., & Schaaf, C. (2016). Observing ecosystems with lightweight, rapid-scanning terrestrial lidar scanners. *Remote Sensing in Ecology and Conservation*, *2*(4), 174–189.
Pérez-Cruzado, C., Fehrmann, L., Magdon, P., Cañellas, I., Sixto, H., & Kleinn, C. (2015). On the site-level suitability of biomass models. In *Environmental Modelling & Software* (Vol. 73, pp. 14–26). https://doi.org/10.1016/j.envsoft.2015.07.019
Phillips, D. L., Brown, S. L., Schroeder, P. E., & Birdsey, R. A. (2000). Toward error analysis of large-scale forest carbon budgets. In *Global Ecology and Biogeography* (Vol. 9, Issue 4, pp. 305–313). https://doi.org/10.1046/j.1365-2699.2000.00197.x
Phillips, O., Baker, T., Feldpausch, T., & Brienen, R. (2018). *RAINFOR: Field manual for plot establishment and remeasurement*. https://forestplots.net
Phillips, O. L., Aragão, L. E. O. C., Lewis, S. L., Fisher, J. B., Lloyd, J., López-González, G., Malhi, Y., Monteagudo, A., Peacock, J., Quesada, C. A., van der Heijden, G., Almeida, S., Amaral, I., Arroyo, L., Aymard, G., Baker, T. R., Bánki, O., Blanc, L., Bonal, D., … Torres-Lezama, A. (2009). Drought sensitivity of the Amazon rainforest. *Science*, *323*(5919), 1344–1347.
Phillips, O. L., Malhi, Y., Vinceti, B., Baker, T., Lewis, S. L., Higuchi, N., Laurance, W. F., Núñez Vargas, P., Vásquez Martinez, R., Laurance, S., Ferreira, L. V., Stern, M., Brown, S., & Grace, J. (2002). Changes in growth of tropical forests: Evaluating potential biases. In *Ecological Applications* (Vol. 12, Issue 2, pp. 576–587). https://doi.org/10.1890/1051-0761(2002)012[0576:cigotf]2.0.co;2
Phillips, O. L., Sullivan, M. J. P., Baker, T. R., Monteagudo Mendoza, A., Vargas, P. N., & Vásquez, R. (2019). Species Matter: Wood Density Influences Tropical Forest Biomass at Multiple Scales. *Surveys in Geophysics*, *40*(4), 913–935.
Picard, N., Bosela, F. B., & Rossi, V. (2015). Reducing the error in biomass estimates strongly depends on model selection. In *Annals of Forest Science* (Vol. 72, Issue 6, pp. 811–823). https://doi.org/10.1007/s13595-014-0434-9
Picard, N., Henry, M., Fonton, N. H., Kondaoulé, J., Fayolle, A., Birigazzi, L., Sola, G., Poultouchidou, A., Trotta, C., & Maïdou, H. (2016). Error in the estimation of emission factors for forest degradation in central Africa. In *Journal of Forest Research* (Vol. 21, Issue 1, pp. 23–30). https://doi.org/10.1007/s10310-015-0510-5
Picard, N., Saint-Andre, L., & Henry, M. (2012). *Manual for building tree volume and biomass allometric equations: from field measurement to prediction*. Food and Agriculture Organization of the United Nations, Centre de Cooperation Internationale en Recherche Agronomique pour le Developpement.
Pilli, R., Kull, S. J., Blujdea, V. N. B., & Grassi, G. (2018). The Carbon Budget Model of the Canadian Forest Sector (CBM-CFS3): customization of the Archive Index Database for European Union countries. In *Annals of Forest Science* (Vol. 75, Issue 3). https://doi.org/10.1007/s13595-018-0743-5
Ploton, P., Mortier, F., Barbier, N., Cornu, G., Réjou-Méchain, M., Rossi, V., Alonso, A., Bastin, J.-F., Bayol, N., Bénédet, F., Bissiengou, P., Chuyong, G., Demarquez, B., Doucet, J.-L., Droissart, V., Kamdem, N. G., Kenfack, D., Memiaghe, H., Moses, L., … Gourlet-Fleury, S. (2020). A map of African humid

tropical forest aboveground biomass derived from management inventories. *Scientific Data*, *7*(1), 221.

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., Bayol, N., Lyapustin, A., Gourlet-Fleury, S., & Pélissier, R. (2020). Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature Communications*, *11*(1), 1–11.

Poorter, H., Jagodzinski, A. M., Ruiz-Peinado, R., Kuyah, S., Luo, Y., Oleksyn, J., Usoltsev, V. A., Buckley, T. N., Reich, P. B., & Sack, L. (2015). How does biomass distribution change with size and differ among species? An analysis for 1200 plant species from five continents. *The New Phytologist*, *208*(3), 736–749.

Popescu, S. C. (2007). Estimating biomass of individual pine trees using airborne lidar. In *Biomass and Bioenergy* (Vol. 31, Issue 9, pp. 646–655). https://doi.org/10.1016/j.biombioe.2007.06.022

Popescu, S. C., Wynne, R. H., & Nelson, R. F. (2003). Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. In *Canadian Journal of Remote Sensing* (Vol. 29, Issue 5, pp. 564–577). https://doi.org/10.5589/m03-027

Popescu, S. C., Zhao, K., Neuenschwander, A., & Lin, C. (2011). Satellite lidar vs. small footprint airborne lidar: Comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level. *Remote Sensing of Environment*, *115*(11), 2786–2797.

Puliti, S., Breidenbach, J., & Astrup, R. (2020). Estimation of Forest Growing Stock Volume with UAV Laser Scanning Data: Can It Be Done without Field Data? *Remote Sensing*, *12*(8), 1245.

Puliti, S., Ørka, H., Gobakken, T., & Næsset, E. (2015). Inventory of Small Forest Areas Using an Unmanned Aerial System. In *Remote Sensing* (Vol. 7, Issue 8, pp. 9632–9654). https://doi.org/10.3390/rs70809632

Puliti, S., Solberg, S., & Granhus, A. (2019). Use of UAV Photogrammetric Data for Estimation of Biophysical Properties in Forest Stands Under Regeneration. *Remote Sensing*, *11*(3), 233.

Pyörälä, J., Liang, X., Vastaranta, M., Saarinen, N., Kankare, V., Wang, Y., Holopainen, M., & Hyyppä, J. (2018). Quantitative Assessment of Scots Pine (Pinus Sylvestris L.) Whorl Structure in a Forest Environment Using Terrestrial Laser Scanning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, *11*(10), 3598–3607.

Quadros, N., & Keysers, J. (2018). Airborne LiDAR Acquisition and Validation. In A. Held, S. Phinn, M. Soto-Berelov, & S. Jones (Eds.), *A TERN Landscape Assessment Initiative: Effective field calibration and validation practices* (1.3 ed., pp. 244–275). TERN Australia.

Quantum systems. (2018). *Airborne LiDAR*. https://www.quantum-systems.com/download/31626/

Radtke, P. J., Walker, D. M., Weiskittel, A. R., Frank, J., Coulston, J. W., & Westfall, J. A. (2015). Legacy tree data: a national database of detailed tree measurements for volume, weight, and physical properties. In S. M. Stanton, G. A. Christensen, & comps. (Eds.), *Pushing boundaries: new directions in inventory techniques and applications* (Vol. 931, pp. 25–30). Gen. Tech. Rep. PNW-GTR-931. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., & Lewis, P. (2013). Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. In *Remote Sensing* (Vol. 5, Issue 2, pp. 491–520). https://doi.org/10.3390/rs5020491

Réjou-Méchain, M., Barbier, N., Couteron, P., Ploton, P., Vincent, G., Herold, M., Mermoz, S., Saatchi, S., Chave, J., de Boissieu, F., Féret, J.-B., Takoudjou, S. M., & Pélissier, R. (2019). Upscaling Forest Biomass from Field to Satellite Measurements: Sources of Errors and Ways to Reduce Them. In *Surveys in Geophysics* (Vol. 40, Issue 4, pp. 881–911). https://doi.org/10.1007/s10712-019-09532-0

Réjou-Méchain, M., Muller-Landau, H. C., Detto, M., Thomas, S. C., Le Toan, T., Saatchi, S. S., Barreto-Silva, J. S., Bourg, N. A., Bunyavejchewin, S., Butt, N., Brockelman, W. Y., Cao, M., Cárdenas, D., Chiang, J.-M., Chuyong, G. B., Clay, K., Condit, R., Dattaraja, H. S., Davies, S. J., … Chave, J. (2014).

Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks. *Biogeosciences Discussions*, *11*, 5711.

Réjou-Méchain, M., Tanguy, A., Piponiot, C., Chave, J., & Hérault, B. (2017). biomass : an r package for estimating above-ground biomass and its uncertainty in tropical forests. In *Methods in Ecology and Evolution* (Vol. 8, Issue 9, pp. 1163–1167). https://doi.org/10.1111/2041-210x.12753

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. In *Remote Sensing of Environment* (Vol. 169, pp. 93–101). https://doi.org/10.1016/j.rse.2015.08.001

Requena Suarez, D., Rozendaal, D. M. A., De Sy, V., Phillips, O. L., Alvarez-Dávila, E., Anderson-Teixeira, K., Araujo-Murakami, A., Arroyo, L., Baker, T. R., Bongers, F., Brienen, R. J. W., Carter, S., Cook-Patton, S. C., Feldpausch, T. R., Griscom, B. W., Harris, N., Hérault, B., Honorio Coronado, E. N., Leavitt, S. M., … Herold, M. (2019). Estimating aboveground net biomass change for tropical and subtropical forests: Refinement of IPCC default rates using forest plot data. *Global Change Biology*, *25*(11), 3609–3624.

Reutebuch, S. E., McGaughey, R. J., Andersen, H.-E., & Carson, W. W. (2003). Accuracy of a high-resolution lidar terrain model under a conifer forest canopy. In *Canadian Journal of Remote Sensing* (Vol. 29, Issue 5, pp. 527–535). https://doi.org/10.5589/m03-022

Richardson, A. D., Carbone, M. S., Keenan, T. F., Czimczik, C. I., Hollinger, D. Y., Murakami, P., Schaberg, P. G., & Xu, X. (2013). Seasonal dynamics and age of stemwood nonstructural carbohydrates in temperate forest trees. *The New Phytologist*, *197*(3), 850–861.

Riegl. (2018). *Riegl VUX-1UAV*. http://www.riegl.com/uploads/tx_pxpriegldownloads/RIEGL_VUX-1UAV_Datasheet_2017-09-01.pdf

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. In *Ecography* (Vol. 40, Issue 8, pp. 913–929). https://doi.org/10.1111/ecog.02881

Roberts, S. D., Dean, T. J., Evans, D. L., McCombs, J. W., Harrington, R. L., & Glass, P. A. (2005). Estimating individual tree leaf area in loblolly pine plantations using LiDAR-derived measurements of height and crown dimensions. In *Forest Ecology and Management* (Vol. 213, Issues 1-3, pp. 54–70). https://doi.org/10.1016/j.foreco.2005.03.025

Robinson, C., Saatchi, S. S., Neumann, M., & Gillespie, T. (2013). Impacts of Spatial Variability on Aboveground Biomass Estimation from L-Band Radar in a Temperate Forest. In *Remote Sensing* (Vol. 5, Issue 3, pp. 1001–1023). https://doi.org/10.3390/rs5031001

Romijn, E., De Sy, V., Herold, M., Böttcher, H., Roman-Cuesta, R. M., Fritz, S., Schepaschenko, D., Avitabile, V., Gaveau, D., Verchot, L., & Martius, C. (2018). Independent data for transparent monitoring of greenhouse gas emissions from the land use sector – What do stakeholders think and need? In *Environmental Science & Policy* (Vol. 85, pp. 101–112). https://doi.org/10.1016/j.envsci.2018.03.016

Roşca, S., Suomalainen, J., Bartholomeus, H., & Herold, M. (2018). Comparing terrestrial laser scanning and unmanned aerial vehicle structure from motion to assess top of canopy structure in tropical forests. *Interface Focus*, *8*(2), 20170038.

Roussel, J.-R., Auty, D., De Boissieu, F., & Meador, A. S. (2018). *lidR: Airborne LiDAR data manipulation and visualization for forestry applications* (R package version 1(1)) [Computer software].

Roussel, J.-R., Caspersen, J., Béland, M., Thomas, S., & Achim, A. (2017). Removing bias from LiDAR-based estimates of canopy height: Accounting for the effects of pulse density and footprint size. *Remote Sensing of Environment*, *198*, 1–16.

Rowland, L., Hill, T. C., Stahl, C., Siebicke, L., Burban, B., Zaragoza-Castells, J., Ponton, S., Bonal, D., Meir, P., & Williams, M. (2014). Evidence for strong seasonality in the carbon storage and carbon use efficiency of an Amazonian forest. *Global Change Biology*, *20*(3), 979–991.

Roxburgh, S. H., Karunaratne, S. B., Paul, K. I., Lucas, R. M., Armston, J. D., & Sun, J. (2019). A revised above-ground maximum biomass layer for the Australian continent. In *Forest Ecology and Management* (Vol. 432, pp. 264–275). https://doi.org/10.1016/j.foreco.2018.09.011

Roxburgh, S. H., & Paul, K. I. (2019). Comprehensive propagation of errors for the prediction of woody biomass. *MS*.

Roxburgh, S. H., Paul, K. I., Clifford, D., England, J. R., & Raison, R. J. (2015). Guidelines for constructing allometric models for the prediction of woody biomass: How many individuals to harvest? In *Ecosphere* (Vol. 6, Issue 3, p. art38). https://doi.org/10.1890/es14-00251.1

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons.

Ryan, C. M., & Berry, N. (2020). *SEOSAW Field Manual* (Version V3.2).

Ryan, C. M., & Williams, M. (2011). How does fire intensity and frequency affect miombo woodland tree populations and biomass? *Ecological Applications: A Publication of the Ecological Society of America*, *21*(1), 48–60.

Saarela, S., Holm, S., Grafström, A., Schnell, S., Næsset, E., Gregoire, T. G., Nelson, R. F., & Ståhl, G. (2016). Hierarchical model-based inference for forest inventory utilizing three sources of information. *Annals of Forest Science*, *73*(4), 895–910.

Saarela, S., Holm, S., Healey, S., Andersen, H.-E., Petersson, H., Prentius, W., Patterson, P., Næsset, E., Gregoire, T. G., & Ståhl, G. (2018). Generalized Hierarchical Model-Based Estimation for Aboveground Biomass Assessment Using GEDI and Landsat Data. In *Remote Sensing* (Vol. 10, Issue 11, p. 1832). https://doi.org/10.3390/rs10111832

Saarela, S., Wästlund, A., Holmström, E., Mensah, A. A., Holm, S., Nilsson, M., Fridman, J., & Ståhl, G. (2020). Mapping aboveground biomass and its prediction uncertainty using LiDAR and field data, accounting for tree-level allometric and LiDAR model errors. In *Forest Ecosystems* (Vol. 7, Issue 1). https://doi.org/10.1186/s40663-020-00245-0

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M., & Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(24), 9899–9904.

Saatchi, S. 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. In *Remote Sensing of Environment* (Vol. 115, Issue 11, pp. 2836–2849). https://doi.org/10.1016/j.rse.2010.07.015

Sagang, L. B. T., Momo, S. T., Libalah, M. B., Rossi, V., Fonton, N., Mofack, G. I. I., Kamdem, N. G., Nguetsop, V. F., Sonké, B., Ploton, P., & Barbier, N. (2018). Using volume-weighted average wood specific gravity of trees reduces bias in aboveground biomass predictions from forest volume data. In *Forest Ecology and Management* (Vol. 424, pp. 519–528). https://doi.org/10.1016/j.foreco.2018.04.054

Santoro, M., Beaudoin, A., Beer, C., Cartus, O., Fransson, J. E. S., Hall, R. J., Pathe, C., Schmullius, C., Schepaschenko, D., Shvidenko, A., Thurner, M., & Wegmüller, U. (2015). Forest growing stock volume of the northern hemisphere: Spatially explicit estimates for 2010 derived from Envisat ASAR. In *Remote Sensing of Environment* (Vol. 168, pp. 316–334). https://doi.org/10.1016/j.rse.2015.07.005

Santoro, M., Cartus, O., Mermoz, S., Bouvet, A., Le Toan, T., Carvalhais, N., Rozendaal, D., Herold, M., Avitabile, V., Quegan, S., Carreiras, J., Rauste, Y., Balzter, H., Schmullius, C., & Seifert, F. M. (2018). *GlobBiomass global above-ground biomass and growing stock volume datasets* [Data set]. http://globbiomass.org/products/global-mapping

Scharlemann, J. P. W., Tanner, E. V. J., Hiederer, R., & Kapos, V. (2014). Global soil carbon: understanding and managing the largest terrestrial carbon pool. In *Carbon Management* (Vol. 5, Issue 1, pp. 81–91). https://doi.org/10.4155/cmt.13.77

Schepaschenko, D., Moltchanova, E., Shvidenko, A., Blyshchyk, V., Dmitriev, E., Martynenko, O., See, L., & Kraxner, F. (2018). Improved Estimates of Biomass Expansion Factors for Russian Forests. *Forests, Trees and Livelihoods*, *9*(6), 312.

Schepaschenko, D., Shvidenko, A., Usoltsev, V., Lakyda, P., Luo, Y., Vasylyshyn, R., Lakyda, I., Myklush, Y., See, L., McCallum, I., Fritz, S., Kraxner, F., & Obersteiner, M. (2017). A dataset of forest biomass structure for Eurasia. *Scientific Data*, *4*, 170070.

Schimel, D., Stephens, B. B., & Fisher, J. B. (2015). Effect of increasing CO2 on the terrestrial carbon cycle. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(2), 436–441.

Schofield, L. A., Danson, F. M., Entwistle, N. S., Gaulton, R., & Hancock, S. (2016). Radiometric calibration of a dual-wavelength terrestrial laser scanner using neural networks. *Remote Sensing Letters* , *7*(4), 299–308.

Schreuder, H. T., Gregoire, T. G., & Wood, G. B. (1993). *Sampling methods for multiresource forest inventory*. John Wiley & Sons.

Schroeder, P., Brown, S., Mo, J., Birdsey, R., & Cieszewski, C. (1997). Biomass estimation for temperate broadleaf forests of the United States using inventory data. *Forest Science*, *43*(3), 424–434.

Shettles, M., Temesgen, H., Gray, A. N., & Hilker, T. (2015). Comparison of uncertainty in per unit area estimates of aboveground biomass for two selected model sets. In *Forest Ecology and Management* (Vol. 354, pp. 18–25). https://doi.org/10.1016/j.foreco.2015.07.002

Shinozaki, K., Yoda, K., K., Hozumi, & Kira, T. (1964). A quantitative analysis of plant form-the pipe model theory: II. Further evidence of the theory and its application in forest ecology. *Japanese Journal of Ecology*, *14*(4), 133–139.

Sigrist, P., Coppin, P., & Hermy, M. (1999). Impact of forest canopy on quality and accuracy of GPS measurements. In *International Journal of Remote Sensing* (Vol. 20, Issue 18, pp. 3595–3610). https://doi.org/10.1080/014311699211228

Sileshi, G. W. (2014). A critical review of forest biomass estimation models, common mistakes and corrective measures. *Forest Ecology and Management*, *329*, 237–254.

Sillett, S. C., Van Pelt, R., Kramer, R. D., Carroll, A. L., & Koch, G. W. (2015). Biomass and growth potential of Eucalyptus regnans up to 100m tall. In *Forest Ecology and Management* (Vol. 348, pp. 78–91). https://doi.org/10.1016/j.foreco.2015.03.046

Silva, C. A., Saatchi, S., Garcia, M., Labrière, 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*, 1–15.

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. In *Journal of Geophysical Research* (Vol. 116, Issue G4). https://doi.org/10.1029/2011jg001708

Slik, J. W. F., Paoli, G., McGuire, K., Amaral, I., Barroso, J., Bastian, M., Blanc, L., Bongers, F., Boundja, P., Clark, C., Collins, M., Dauby, G., Ding, Y., Doucet, J.-L., Eler, E., Ferreira, L., Forshed, O., Fredriksson, G., Gillet, J.-F., … Zweifel, N. (2013). Large trees drive forest aboveground biomass variation in moist lowland forests across the tropics: Large trees and tropical forest biomass. *Global Ecology and Biogeography: A Journal of Macroecology*, *22*(12), 1261–1271.

Smallman, T. L., Exbrayat, J.-F., Mencuccini, M., Bloom, A. A., & Williams, M. (2017). Assimilation of repeated woody biomass observations constrains decadal ecosystem carbon cycle uncertainty in aggrading forests. In *Journal of Geophysical Research: Biogeosciences* (Vol. 122, Issue 3, pp.

528–545). https://doi.org/10.1002/2016jg003520
Snedecor, G. W., & Cochran, W. G. (1967). *Statistical methods*. Iowa State University Press.
Ståhl, G., Heikkinen, J., Petersson, H., Repola, J., & Holm, S. (2014). Sample-Based Estimation of Greenhouse Gas Emissions From Forests—A New Approach to Account for Both Sampling and Model Errors. In *Forest Science* (Vol. 60, Issue 1, pp. 3–13). https://doi.org/10.5849/forsci.13-005
Ståhl, G., Saarela, S., Schnell, S., Holm, S., Breidenbach, J., Healey, S. P., Patterson, P. L., Magnussen, S., Næsset, E., McRoberts, R. E., & Gregoire, T. G. (2016). Use of models in large-area forest surveys: comparing model-assisted, model-based and hybrid estimation. In *Forest Ecosystems* (Vol. 3, Issue 1). https://doi.org/10.1186/s40663-016-0064-9
Stehman, S. V. (2009). Sampling designs for accuracy assessment of land cover. In *International Journal of Remote Sensing* (Vol. 30, Issue 20, pp. 5243–5272). https://doi.org/10.1080/01431160903131000
Stephens, P. R., Watt, P. J., Loubser, D., Haywood, A., & Kimberley, M. O. (2007). Estimation of carbon stocks in New Zealand planted forests using airborne scanning LiDAR. *International Archives of Photogrammetry. Remote Sensing and Spatial Information Sciences*, *36*, 389–394.
Stoker, J. M., Abdullah, Q. A., Nayegandhi, A., & Winehouse, J. (2016). Evaluation of Single Photon and Geiger Mode Lidar for the 3D Elevation Program. *Remote Sensing*, *8*, 767.
Stoker, J. M., Brock, J. C., Soulard, C. E., Ries, K. G., Sugarbaker, L. J., Newton, W. E., Haggerty, P. K., Lee, K. E., & Young, J. A. (2016). *USGS lidar science strategy—Mapping the technology to the science* (Open-File Report 2015-1209). U.S. Geological Survey. https://doi.org/10.3133/ofr20151209
St-Onge, B., Treitz, P., Wulder, M., Kurz, W., & Gillis, M. (2004). Retrospective mapping of structural and biomass changes in forest ecosystems using photogrammetry and laser altimetry. *AGU Spring Meeting Abstracts*.
Storey, J., Choate, M., & Lee, K. (2014). Landsat 8 Operational Land Imager On-Orbit Geometric Calibration and Performance. In *Remote Sensing* (Vol. 6, Issue 11, pp. 11127–11152). https://doi.org/10.3390/rs61111127
Stovall, A. E. L., & Shugart, H. H. (2018). Improved Biomass Calibration and Validation With Terrestrial LiDAR: Implications for Future LiDAR and SAR Missions. In *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (Vol. 11, Issue 10, pp. 3527–3537). https://doi.org/10.1109/jstars.2018.2803110
Stovall, A. E. L., Vorster, A. G., Anderson, R. S., Evangelista, P. H., & Shugart, H. H. (2017). Non-destructive aboveground biomass estimation of coniferous trees using terrestrial LiDAR. In *Remote Sensing of Environment* (Vol. 200, pp. 31–42). https://doi.org/10.1016/j.rse.2017.08.013
Strahler, A. H., Jupp, D. L. B., Woodcock, C. E., Schaaf, C. B., Yao, T., Zhao, F., Yang, X., Lovell, J., Culvenor, D., Newnham, G., Ni-Miester, W., & Boykin-Morris, W. (2008). Retrieval of forest structural parameters using a ground-based lidar instrument (Echidna®). In *Canadian Journal of Remote Sensing* (Vol. 34, Issue sup2, pp. S426–S440). https://doi.org/10.5589/m08-046
Sullivan, M. J. P., Lewis, S. L., Affum-Baffoe, K., Castilho, C., Costa, F., Sanchez, A. C., Ewango, C. E. N., Hubau, W., Marimon, B., Monteagudo-Mendoza, A., Qie, L., Sonké, B., Martinez, R. V., Baker, T. R., Brienen, R. J. W., Feldpausch, T. R., Galbraith, D., Gloor, M., Malhi, Y., … Phillips, O. L. (2020). Long-term thermal sensitivity of Earth's tropical forests. *Science*, *368*(6493), 869–874.
Sullivan, M. J. P., Lewis, S. L., Hubau, W., Qie, L., Baker, T. R., Banin, L. F., Chave, J., Cuni-Sanchez, A., Feldpausch, T. R., Lopez-Gonzalez, G., Arets, E., Ashton, P., Bastin, J.-F., Berry, N. J., Bogaert, J., Boot, R., Brearley, F. Q., Brienen, R., Burslem, D. F. R. P., … Phillips, O. L. (2018). Field methods for sampling tree height for tropical forest biomass estimation. *Methods in Ecology and Evolution / British Ecological Society*, *9*(5), 1179–1189.
Sullivan, M. J. P., Talbot, J., Lewis, S. L., Phillips, O. L., Qie, L., Begne, S. K., Chave, J., Cuni-Sanchez, A., Hubau, W., Lopez-Gonzalez, G., Miles, L., Monteagudo-Mendoza, A., Sonké, B., Sunderland, T., Ter Steege, H., White, L. J. T., Affum-Baffoe, K., Aiba, S.-I., de Almeida, E. C., … Zemagho, L. (2017).

Diversity and carbon storage across the tropical forest biome. *Scientific Reports*, *7*, 39102.
Suomalainen, J., Anders, N., Iqbal, S., Roerink, G., Franke, J., Wenting, P., Hünniger, D., Bartholomeus, H., Becker, R., & Kooistra, L. (2014). A Lightweight Hyperspectral Mapping System and Photogrammetric Processing Chain for Unmanned Aerial Vehicles. In *Remote Sensing* (Vol. 6, Issue 11, pp. 11013–11030). https://doi.org/10.3390/rs61111013
Su, Y., Guo, Q., Xue, B., Hu, T., Alvarez, O., Tao, S., & Fang, J. (2016). Spatial distribution of forest aboveground biomass in China: Estimation through combination of spaceborne lidar, optical imagery, and forest inventory data. In *Remote Sensing of Environment* (Vol. 173, pp. 187–199). https://doi.org/10.1016/j.rse.2015.12.002
Swatantran, A., Tang, H., Barrett, T., DeCola, P., & Dubayah, R. (2016). Rapid, High-Resolution Forest Structure and Terrain Mapping over Large Areas using Single Photon Lidar. *Scientific Reports*, *6*, 28277.
Tansey, K., Selmes, N., Anstee, A., Tate, N. J., & Denniss, A. (2009). Estimating tree and stand variables in a Corsican Pine woodland from terrestrial laser scanner data. In *International Journal of Remote Sensing* (Vol. 30, Issue 19, pp. 5195–5209). https://doi.org/10.1080/01431160902882587
The SEOSAW partnership. (2020). A network to understand the changing socio-ecology of the southern African woodlands (SEOSAW): Challenges, benefits, and methods. *Plants People Planet*, *00*, 0–19.
Thiel, C., & Schmullius, C. (2017). Comparison of UAV photograph-based and airborne lidar-based point clouds over forest from a forestry application perspective. In *International Journal of Remote Sensing* (Vol. 38, Issues 8-10, pp. 2411–2426). https://doi.org/10.1080/01431161.2016.1225181
Thomas, R. Q., Williams, M., Cavaleri, M. A., -F. Exbrayat, J., Smallman, T. L., & Street, L. E. (2019). Alternate Trait-Based Leaf Respiration Schemes Evaluated at Ecosystem-Scale Through Carbon Optimization Modeling and Canopy Property Data. In *Journal of Advances in Modeling Earth Systems* (Vol. 11, Issue 12, pp. 4629–4644). https://doi.org/10.1029/2019ms001679
Thomas, V., Treitz, P., McCaughey, J. H., & Morrison, I. (2006). Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: an examination of scanning density. *Canadian Journal of Forest Research. Journal Canadien de La Recherche Forestiere*, *36*(1), 34–47.
Thorpe, A. S., Barnett, D. T., Elmendorf, S. C., Hinckley, E. S., Hoekman, D., Jones, K. D., LeVan, K. E., Meier, C. L., Stanish, L. F., & Thibault, K. M. (2016). Introduction to the sampling designs of the National Ecological Observatory Network Terrestrial Observation System. *Ecosphere* , *7*(12). https://doi.org/10.1002/ecs2.1627
Thurner, M., Beer, C., Santoro, M., Carvalhais, N., Wutzler, T., Schepaschenko, D., Shvidenko, A., Kompter, E., Ahrens, B., Levick, S. R., & Schmullius, C. (2014). Carbon stock and density of northern boreal and temperate forests. In *Global Ecology and Biogeography* (Vol. 23, Issue 3, pp. 297–310). https://doi.org/10.1111/geb.12125
Tian, J., Dai, T., Li, H., Liao, C., Teng, W., Hu, Q., Ma, W., & Xu, Y. (2019). A Novel Tree Height Extraction Approach for Individual Trees by Combining TLS and UAV Image-Based Point Cloud Integration. *Forests*, *10*(7), 537.
Tittmann, P., Saatchi, S., & Sharma, B. D. (2015). *VT0005: Tool for measuring aboveground live forest biomass using remote sensing* (Version Vesion 1.0). VCS.
Torello-Raventos, M., Feldpausch, T. R., Veenendaal, E., Schrodt, F., Saiz, G., Domingues, T. F., Djagbletey, G., Ford, A., Kemp, J., Marimon, B. S., Hur Marimon Junior, B., Lenza, E., Ratter, J. A., Maracahipes, L., Sasaki, D., Sonké, B., Zapfack, L., Taedoumg, H., Villarroel, D., … Lloyd, J. (2013). On the delineation of tropical vegetation types with an emphasis on forest/savanna transitions. *Plant Ecology & Diversity*, *6*(1), 101–137.
Torr, P. H. S., & Zisserman, A. (2000). MLESAC: A New Robust Estimator with Application to Estimating Image Geometry. In *Computer Vision and Image Understanding* (Vol. 78, Issue 1, pp. 138–156). https://doi.org/10.1006/cviu.1999.0832

Trochta, J., Krůček, M., Vrška, T., & Král, K. (2017). 3D Forest: An application for descriptions of three-dimensional forest structures using terrestrial LiDAR. *PloS One*, *12*(5), e0176871.

Ustin, S. L., & Gamon, J. A. (2010). Remote sensing of plant functional types. *The New Phytologist*, *186*(4), 795–816.

van Breugel, M., Ransijn, J., Craven, D., Bongers, F., & Hall, J. S. (2011). Estimating carbon stock in secondary forests: Decisions and uncertainties associated with allometric biomass models. In *Forest Ecology and Management* (Vol. 262, Issue 8, pp. 1648–1657). https://doi.org/10.1016/j.foreco.2011.07.018

Van de Perre, F., Willig, M. R., Presley, S. J., Andemwana, F. B., Beeckman, H., Boeckx, P., Cooleman, S., de Haan, M., De Kesel, A., Dessein, S., Grootaert, P., Huygens, D., Janssens, S. B., Kearsley, E., Kabeya, P. M., Leponce, M., Van den Broeck, D., Verbeeck, H., Würsten, B., … Verheyen, E. (2018). Reconciling biodiversity and carbon stock conservation in an Afrotropical forest landscape. In *Science Advances* (Vol. 4, Issue 3, p. eaar6603). https://doi.org/10.1126/sciadv.aar6603

van der Sande, M. T., Poorter, L., Kooistra, L., Balvanera, P., Thonicke, K., Thompson, J., Eric J M, Alaniz, N. G., Jones, L., Mora, F., Mwampamba, T. H., Parr, T., & Peña-Claros, M. (2017). Biodiversity in species, traits, and structure determines carbon stocks and uptake in tropical forests. In *Biotropica* (Vol. 49, Issue 5, pp. 593–603). https://doi.org/10.1111/btp.12453

van Leeuwen, M., & Nieuwenhuis, M. (2010). Retrieval of forest structural parameters using LiDAR remote sensing. In *European Journal of Forest Research* (Vol. 129, Issue 4, pp. 749–770). https://doi.org/10.1007/s10342-010-0381-4

Vastaranta, M., Melkas, T., Holopainen, M., Kaartinen, H., Hyyppä, J., & Hyyppä, H. (2008). Comparison of different laser-based methods to measure stem diameter. *Proceedings of the SilviLaser*, 17–19.

Vauhkonen, J., Tokola, T., Packalén, P., & Maltamo, M. (2009). Identification of Scandinavian Commercial Species of Individual Trees from Airborne Laser Scanning Data Using Alpha Shape Metrics. *Forest Science*, *55*(1), 37–47.

Vega, C. (n.d.). [Personal communication].

Venter, O., Laurance, W. F., Iwamura, T., Wilson, K. A., Fuller, R. A., & Possingham, H. P. (2009). Harnessing carbon payments to protect biodiversity. *Science*, *326*(5958), 1368.

Vidal, C., Alberdi, I. A., Mateo, L. H., & Redmond, J. J. (2016). *National Forest Inventories: Assessment of Wood Availability and Use*. Springer.

Villard, L., & Le Toan, T. (2015). Relating P-Band SAR Intensity to Biomass for Tropical Dense Forests in Hilly Terrain: $\gamma^0 or t^0$? In *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (Vol. 8, Issue 1, pp. 214–223). https://doi.org/10.1109/jstars.2014.2359231

Wallace, L., Lucieer, A., Watson, C., & Turner, D. (2012). Development of a UAV-LiDAR System with Application to Forest Inventory. In *Remote Sensing* (Vol. 4, Issue 6, pp. 1519–1543). https://doi.org/10.3390/rs4061519

Wang, D., Momo Takoudjou, S., & Casella, E. (2020). LeWoS: A universal leaf-wood classification method to facilitate the 3D modelling of large tropical trees using terrestrial LiDAR. *Methods in Ecology and Evolution / British Ecological Society*, *11*(3), 376–389.

Wang, Y., Lehtomäki, M., Liang, X., Pyörälä, J., Kukko, A., Jaakkola, A., Liu, J., Feng, Z., Chen, R., & Hyyppä, J. (2019). Is field-measured tree height as reliable as believed – A comparison study of tree height estimates from field measurement, airborne laser scanning and terrestrial laser scanning in a boreal forest. In *ISPRS Journal of Photogrammetry and Remote Sensing* (Vol. 147, pp. 132–145). https://doi.org/10.1016/j.isprsjprs.2018.11.008

Watson, J. E. M., Evans, T., Venter, O., Williams, B., Tulloch, A., Stewart, C., Thompson, I., Ray, J. C., Murray, K., Salazar, A., McAlpine, C., Potapov, P., Walston, J., Robinson, J. G., Painter, M., Wilkie, D., Filardi, C., Laurance, W. F., Houghton, R. A., … Lindenmayer, D. (2018). The exceptional value of intact forest ecosystems. *Nature Ecology & Evolution*, *2*(4), 599–610.

Watson, J. E. M., Shanahan, D. F., Di Marco, M., Allan, J., Laurance, W. F., Sanderson, E. W., Mackey, B., & Venter, O. (2016). Catastrophic Declines in Wilderness Areas Undermine Global Environment Targets. *Current Biology: CB*, *26*(21), 2929–2934.

Weaver, S. A., Ucar, Z., Bettinger, P., Merry, K., Faw, K., & Cieszewski, C. J. (2015). Assessing the accuracy of tree diameter measurements collected at a distance. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, *36*(1), 73–83.

West, P. W. (2009). *Tree and Forest Measurement*. https://doi.org/10.1007/978-3-540-95966-3

White, A., Sparrow, B., Leitch, E., Foulkes, J., Flitton, R., Lowe, A. J., & Caddy-Retalic, S. (2012). *AusPlots Rangelands Survey Protocols Manual*. The University of Adelaide Press.

White, J. C., Woods, M., Krahn, T., Papasodoro, C., Bélanger, D., Onafrychuk, C., & Sinclair, I. (2021). Evaluating the capacity of single photon lidar for terrain characterization under a range of forest conditions. *Remote Sensing of Environment*, *252*, 112169.

White, J. C., Wulder, M. A., Varhola, A., Vastaranta, M., Coops, N. C., Cook, B. D., Pitt, D., & Woods, M. (2013). A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. In *The Forestry Chronicle* (Vol. 89, Issue 06, pp. 722–723). https://doi.org/10.5558/tfc2013-132

Wilkes, P., Lau, A., Disney, M., Calders, K., Burt, A., de Tanago, J. G., Bartholomeus, H., Brede, B., & Herold, M. (2017). Data acquisition considerations for Terrestrial Laser Scanning of forest plots. In *Remote Sensing of Environment* (Vol. 196, pp. 140–153). https://doi.org/10.1016/j.rse.2017.04.030

Williamson, G. B., & Wiemann, M. C. (2010). Measuring wood specific gravity…Correctly. In *American Journal of Botany* (Vol. 97, Issue 3, pp. 519–524). https://doi.org/10.3732/ajb.0900243

Wilson, E. R., Murray, J., Ryding, I., & Mont, E. C. (2007). Comparison of three tools for measuring tree diameter in stands of different age and stem size. *Quarterly Journal of Forestry*, *101*(4), 267.

Wing, M. G., Solmie, D., & Kellogg, L. (2004). Comparing digital range finders for forestry applications. *Journal of Forestry*, *102*(4), 16–20.

Wollschläger, U., Gerhards, H., Yu, Q., & Roth, K. (2010). Multi-channel ground-penetrating radar to explore spatial variations in thaw depth and moisture content in the active layer of a permafrost site. In *The Cryosphere* (Vol. 4, Issue 3, pp. 269–283). https://doi.org/10.5194/tc-4-269-2010

Wood, S., Stephens, H., Foulkes, J., Ebsworth, E., & Bowman, D. (n.d.). *AusPlots Forests Survey Protocols Manual* (Version 1.6). University of Tasmania. https://static1.squarespace.com/static/54c18c59e4b04884b35c7843/t/54ebc32fe4b03ce768b68c40/1424737071050/Ausplots+Forests+Field+Manual_v1.6.pdf

Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., Hilker, T., Bater, C. W., & 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., Bastin, J.-F., Banks, N., Boeckx, P., Verbeeck, H., Lewis, S. L., Muanza, E. T., Bongwele, E., Kayembe, F., Mbenza, D., Kalau, L., Mukendi, F., Ilunga, F., & Ebuta, D. (2017). Spatial Distribution of Carbon Stored in Forests of the Democratic Republic of Congo. *Scientific Reports*, *7*(1), 15030.

Yanai, R. D., Battles, J. J., Richardson, A. D., Blodgett, C. A., Wood, D. M., & Rastetter, E. B. (2010). Estimating Uncertainty in Ecosystem Budget Calculations. In *Ecosystems* (Vol. 13, Issue 2, pp. 239–248). https://doi.org/10.1007/s10021-010-9315-8

Yu, X., Hyyppä, J., Kaartinen, H., & Maltamo, M. (2004). Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. In *Remote Sensing of Environment* (Vol. 90, Issue 4, pp. 451–462). https://doi.org/10.1016/j.rse.2004.02.001

Zarco-Tejada, P. J., Diaz-Varela, R., Angileri, V., & Loudjani, P. (2014). Tree height quantification using very high resolution imagery acquired from an unmanned aerial vehicle (UAV) and automatic 3D photo-reconstruction methods. In *European Journal of Agronomy* (Vol. 55, pp. 89–99).

https://doi.org/10.1016/j.eja.2014.01.004
Zhao, K., Popescu, S., & Nelson, R. (2009). Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. In *Remote Sensing of Environment* (Vol. 113, Issue 1, pp. 182–196). https://doi.org/10.1016/j.rse.2008.09.009
Zhao, K., Suarez, J. C., Garcia, M., Hu, T., Wang, C., & Londo, A. (2018). Utility of multitemporal lidar for forest and carbon monitoring: Tree growth, biomass dynamics, and carbon flux. In *Remote Sensing of Environment* (Vol. 204, pp. 883–897). https://doi.org/10.1016/j.rse.2017.09.007
Zhao, P., Lu, D., Wang, G., Wu, C., Huang, Y., & Yu, S. (2016). Examining Spectral Reflectance Saturation in Landsat Imagery and Corresponding Solutions to Improve Forest Aboveground Biomass Estimation. *Remote Sensing*, *8*(6), 469.
Zhong, L., Cheng, L., Xu, H., Wu, Y., Chen, Y., & Li, M. (2017). Segmentation of Individual Trees From TLS and MLS Data. In *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (Vol. 10, Issue 2, pp. 774–787). https://doi.org/10.1109/jstars.2016.2565519
Zianis, D., & Mencuccini, M. (2004). On simplifying allometric analyses of forest biomass. In *Forest Ecology and Management* (Vol. 187, Issues 2-3, pp. 311–332). https://doi.org/10.1016/j.foreco.2003.07.007
Zolkos, S. G., Goetz, S. J., & Dubayah, R. (2013). A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. In *Remote Sensing of Environment* (Vol. 128, pp. 289–298). https://doi.org/10.1016/j.rse.2012.10.017