Biomass estimation models and methods for miombo woodlands of Malawi using field and remotely sensed data

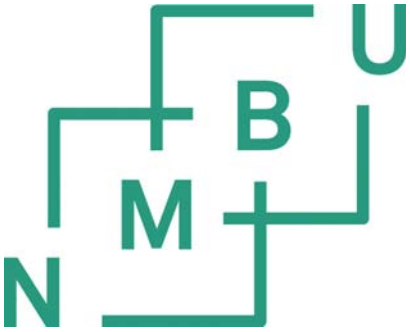
Modeller og metoder for å estimere biomasse i miomboskog i Malawi ved hjelp av feltinventering og fjernmåling

Philosophiae Doctor (PhD) Thesis

Daud Jones Kachamba

Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences

Ås 2016



Thesis number 2016:92
ISSN 1894-6402
ISBN 978-82-575-1406-8

Supervisors

Professor Tron Eid
Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, NO-1432, Ås, Norway

Professor Terje Gobakken
Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, NO-1432, Ås, Norway

Dr Hans Ole Ørka
Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, NO-1432, Ås, Norway

Professor Weston Mwase
Department of Forestry
Lilongwe University of Agriculture and Natural Resources
P.O. Box 219, Lilongwe, Malawi

Evaluation committee

Professor Gherardo Chirici, PhD
Department of Agricultural, Food and Forestry Systems
University of Florence
Via San Bonaventura, 13
50145 Florence, Italy

Dr Casey Ryan
School of GeoSciences
University of Edinburgh
Crew Building
The King's Buildings
Alexander Crum Brown Road
Edinburgh EH9 3FF, UK

Dr Ole Martin Bollandsås
Department of Ecology and Natural Resource Management
Norwegian University of Life Sciences
P.O. Box 5003, NO-1432, Ås, Norway

Acknowledgements

I feel honoured to have Professors Tron Eid, Terje Gobakken and Weston Mwase and Dr Hans Ole Ørka as supervisors for my PhD study. Tron (main supervisor) worked tirelessly in securing research funds; designing the study; as well as guiding fieldwork, data analysis and manuscript development. I would also like to thank Terje and Hans for introducing me to remote sensing and unmanned aerial vehicles (UAVs). Weston has been my mentor since my undergraduate studies and ensured that my fieldwork went smoothly. My gratitude also extends to Dr Wilson Mugasha, Dr Endre Hofstad Hansen, Dr Liviu Theodor Ene, Dr Ernest Mauya, Dr Deo Shirima and Stefano Puliti for their support during data analysis.

I am indebted to the Norwegian Government through the Quota Scholarship Scheme for funding my studies and stay in Norway, the Norwegian University of Life Sciences (NMBU), the World Wildlife Fund (WWF) under the Russel E Train Education for Nature Fellowship and the programme on Capacity Building for Managing Climate Change (CABMACC) in Malawi for funding the field work.

The Forest Research Institute of Malawi deserves a special mention for facilitating the granting of research permit from the National Research Council of Malawi. Happy Kachamba, Devey Lekapo, Humphrey Chapama, Boyd Zulu, Steven Mphamba, Herbert Jenya, Martin Nyoni, Kola Daitoni and Steve Hara; thank you for your jokes and team spirit during data collection.

Fellow students and administrative staff at INA and NMBU also played a very critical role by creating a conducive working and social environment. Special mention goes to Mengesteab, Samson, Moses, Trust, Hambulo, Linette, Pakwanja, Cecelia, Sarah, Fundi, Greyson, Erik, Fredrick, Dawit, Mekdes, Walid, Victor and Rafal, just to mention a few.

My lovely wife Lexa, cute daughter Nicole and my twin babies Tiyamike and Takondwa, deserve medals for enduring while I was away pursuing this PhD study. I dedicate this work to my parents, Mr and Mrs Kachamba, for instilling in me the spirit of hard work.

Ås, 10.10.2016

Daud Jones Kachamba

Contents

Abstract

Dry tropical forests, such as the miombo woodlands, play an important role in the global carbon budget as well as in contributing towards the sustainable development of countries such as Malawi. To ensure sustainability of these forests, availability of models and methods for assisting forest managers in quantifying volume and biomass are indispensable. This thesis therefore sought to develop volume and biomass prediction models as well as to test the potential of applying unmanned aerial vehicles (UAVs) in biomass prediction and estimation in miombo woodlands. In Paper 1 and 2, we developed models for predicting tree sectional (twigs, merchantable stem and branches) volume and biomass, total tree volume as well as tree above-and belowground biomass. The performances and evaluations suggested that the models can be used over a wide range of geographical and ecological conditions in Malawi with an appropriate accuracy in predictions. Application of UAVs for biomass prediction and estimation were tested and the results are presented in Papers 3 and 4. In Paper 3, we tested methods to derive digital terrain models (DTMs) while Paper 4 focused on the assessment of the efficiency of UAV-assisted inventories as well as the influence of sample plot sizes and number of sample plots on the precision of biomass estimates. The results, presented in Paper 3, show that among the tested DTMs, the model developed from unsupervised ground filtering based on a grid search approach performed best. Furthermore, the observed prediction errors for the biomass predictions are similar to results from previous studies using airborne laser scanning (ALS) data, thus showing the potential of applying this technology in miombo woodlands. Finally, Paper 4 demonstrated that UAV-assisted inventories produce more precise estimates compared to those based on purely field-based inventories. The results also indicated that large sample plot sizes and sample sizes favour UAV-assisted inventories and that UAV-assisted inventories are more efficient than purely field-based inventories. The developed models and the results from the tested methods presented in the thesis have taken us some steps forward that are expected to support and improve forest management decision-making in general as well as the implementation of a REDD+ MRV system covering the miombo woodlands of Malawi.

List of papers

Paper 1: Kachamba, D.J. and Eid, T. (2016). Total tree, merchantable stem and branch volume models for miombo woodlands of Malawi. *Southern Forests, 78, 41-51.*

Paper 2: Kachamba, D.J., Eid, T. and Gobakken, T. (2016). Above- and belowground biomass models for trees in miombo woodlands of Malawi. *Forests 7(2):38*

Paper 3: Kachamba, D.J., Ørka, H.O., Gobakken, T., Eid, T. and Mwase W. (2016). Biomass prediction using an unmanned aerial vehicle in a tropical woodland. Under revision.

Paper 4: Kachamba, D.J., Ørka, H.O., Gobakken, T., Eid, T. and Næsset, E. (2016). Influence of plot size and sample size on efficiency of biomass estimates in inventories of dry tropical forests assisted by photogrammetric data from an unmanned aerial vehicle. Manuscript.

1.0 Introduction

Dry tropical forests cover central and south America, Africa, India, south-east Asia and northern Australia (Miles et al. 2006). In southern Africa, dry forests are mainly dominated by miombo woodlands. These woodlands were estimated to cover an area of approximately 2.7 million km^2 (Frost 1996), but this area is most likely lower today due to deforestation and forest degradation. Miombo woodlands are presently spanning 11 countries in Africa, including Malawi (Chidumayo & Gumbo 2010; Ryan et al. 2011). The miombo ecoregion occurs in a climate with a dry season of three months or more and has mean annual precipitations and temperatures of 710 – 1365 mm and 18.0 – 23.1°C, respectively (Frost 1996). Unlike other African savannas and woodlands, miombo woodlands are dominated by three key deciduous tree species belonging to the family Fabaceae, subfamily Caesalpinioideae in the genera *Brachystegia, Julbernadia and Isoberlinia* (Frost 1996; Ryan et al. 2011). A similar tree species composition is found in the miombo woodlands of Malawi (Mwase et al. 2007).

Miombo woodlands are multi-species and multi-layered and are regenerated through coppicing as well as seed dispersal. Structurally, the canopy of miombo woodlands is dominated by trees that are umbrella-shaped whose heights usually range from 14 to 18 m. The sub-canopy is composed of a highly variable scattered layer of shrubs, suppressed saplings of canopy layer trees, grasses and sedges (Abbot et al. 1997; Frost 1996). Tree forms in these woodlands vary from small, multi-stemmed trees to tall single-stemmed trees with straight boles (Abbot et al. 1997). Fires occur frequently in miombo woodlands both in time and space (e.g. Tarimo et al. 2015). Fires are regarded essential to the structure and stability of miombo woodlands (Frost 1996), and the biomass may be reduced substantially if the fire frequency is high (Ryan & Williams 2011). Some tree species have a thick bark to protect them from fires (Frost 1996).

In miombo woodlands, tree species richness and densities vary widely with location, i.e. ranging between 70 and 300 species, and up to 4100 stems ha^{-1} depending on rainfall and anthropogenic factors (Abbot et al. 1997; Dewees et al. 2011; Frost 1996; Furley et al. 2008; Giliba et al. 2011; Malimbwi et al. 2016; Williams et al. 2008). In Malawi, the number of tree species is estimated to exceed 130 with tree densities ranging from about 260 to 1640 stems ha^{-1} (Government of Malawi 2012).

Miombo woodlands provide a wide variety of food and ecosystem services to millions of people in the region including fruits, bush meat, edible insects, beeswax, honey, traditional medicines, biodiversity and watershed conservation (Abbot & Homewood 1999; Blackie et al. 2014; Chidumayo & Gumbo 2010; Kajembe et al. 2015; Luoga et al. 2005; Mwase et al. 2007; Ryan et al. 2016). In Malawi, the woodlands constitute 92% of the country's total forest area (Government of Malawi 2010; Government of Malawi 2012). The Malawi government recognizes the role the woodlands play towards achieving sustainable development for the country. However, increases in population growth has led to high demand for firewood, charcoal and timber products leading to deforestation, currently estimated at 1% per annum (Government of Malawi 2001; Government of Malawi 2010).

Dry tropical forests, including miombo woodlands, are currently the least studied compared to wet tropical forests despite their significant contribution to the global carbon budget and to livelihoods of a lot of people (e.g. Dirzo et al. 2011). In recognition of the importance of forests, including the dry tropical forests, the global community, through the United Nations Framework Convention on Climate Change (UNFCCC), established the Reducing Emissions from Deforestation and Forest Degradation, plus forest conservation, sustainable management of forest and enhancement of carbon stocks (REDD+) mechanism (Barquín et al. 2014; Goetz et al. 2015; UNFCCC 2014). This mechanism has given a financial incentive to developing countries in their efforts to reduce deforestation and forest degradation through increased forest conservation and implementation of sustainable forest management. The payment scheme for REDD+ is based on reported national level carbon stock estimates to the UNFCCC (Goetz et al. 2015). To implement REDD+, each participating country is therefore expected to have a credible forest monitoring system that supports the functions of monitoring, reporting and verification (MRV) of forest carbon stocks at a national scale (Gizachew & Duguma 2016). The system is thus expected to establish a national baseline carbon stock estimate as well as changes of carbon stocks over time (Goetz et al. 2015).

Currently, Malawi is in the preparatory phase of implementing REDD+. The first step in this phase has involved establishing legal and institutional frameworks. In 2012, the government released a draft version of the national climate change policy to support all climate change related programs in the country. In the same year, the government also launched the Malawi REDD+ program aiming for streamlining the process of operationalizing REDD+

(Government of Malawi 2015). Through the program, the Forestry Institute of Malawi (FRIM) was mandated with the development of a national forest carbon MRV system.

A fully functional national forest carbon MRV system for REDD+ implementation in Malawi shall require establishment of a data collection and management system comprising three pillars, namely: a) a remote sensing based land monitoring system for collecting and assessing activity data related to forest cover changes over time (Goetz et al. 2015; UNFCCC 2014), b) conducting national forest inventories (NFIs) for quantifying carbon stock changes and c) a data analysis and reporting system for production of reports to the UNFCCC. Among these pillars, NFIs form a critical component since they are directly linked with carbon stock changes, which is a key component of the REDD+ payment system (Gizachew & Duguma 2016; Goetz et al. 2015). NFIs rely on the utilisation of both reliable biomass/volume models and state of the art remote sensing techniques. Currently, reliable biomass and volume models are lacking and modern remote sensing techniques for volume or biomass prediction and estimation are yet to be tested for application in REDD+ in Malawi. So far, through funding from a number of multi-and bilateral donors, including Food and Agriculture Organization (FAO), United States Agency for International Development (USAID), Japan International Cooperation Agency (JICA) and World Bank, national land use and land cover maps have been developed for benchmarking land use and land use change in the forests of Malawi. However, a fully functional national forest carbon MRV system is yet to be established.

Apart from the anticipated financial benefits from carbon credits, the Malawi government considers REDD+ as an opportunity for instituting sustainable forest management in the country. Currently, the management of miombo woodlands is suffering from lack of reliable models and methods that may support forest managers in decision-making. The existence of such models and methods is instrumental in the efforts to accomplish a sustainable management of these resources.

2.0 Objectives

The main objective of this thesis was to develop models and methods for predicting and estimating volume and biomass of miombo woodlands in Malawi. The models and methods developed are based on both field and remotely sensed data and are expected to support forest management decision-making in general as well as the implementation of a REDD+ MRV system in the country. The following specific sub-objectives were addressed in four different papers;

a) Develop general (multiple tree species from several sites) models for volume prediction in miombo woodlands (Paper 1);
b) Develop general (multiple tree species from several sites) models for biomass prediction in miombo woodlands (Paper 2);
c) Explore the possibility of using UAVs in biomass prediction in miombo woodlands (Paper 3);
d) Assess the efficiency of UAV-assisted inventories as well as the influence of sample plot size and sample size on error estimates in biomass estimation in miombo woodlands (Paper 4).

3.0 Conceptual framework

Forest volume and biomass estimates are basic information needed generally for forest management decision-making as well as when implementing a REDD+ MRV system. A conceptual framework for estimation of volume and above- and belowground biomass for forest areas is presented in Figure 1.

When employing field-based methods for volume or biomass estimation, sample plot inventories are first conducted, and subsequently, individual tree volume or above- and belowground biomass models, if readily available, are applied. In cases where reliable individual tree models are lacking, they can be developed. The process of developing the models involves conducting sample plot inventories to guide the selection of representative trees for destructive sampling. The destructive sampled tree data are then used to develop individual tree volume or above-and belowground biomass models, which can finally be used for forest area volume or above-and belowground biomass estimation.

In cases where remote sensing is the main method applied for estimating forest area volume or biomass, remotely sensed data can be collected using different sensors mounted on different platforms. Application of imagery captured from UAVs is an example of a method that has recently gained ground in forestry. In addition to the remote sensing based data (processed UAV images), sample plot inventory data are also needed. The processed UAV images, sample plot inventory data and individual tree aboveground biomass models (alternatively individual tree volume models) are then used to develop area-based models that can finally be used to estimate forest area aboveground biomass (or forest area volume).

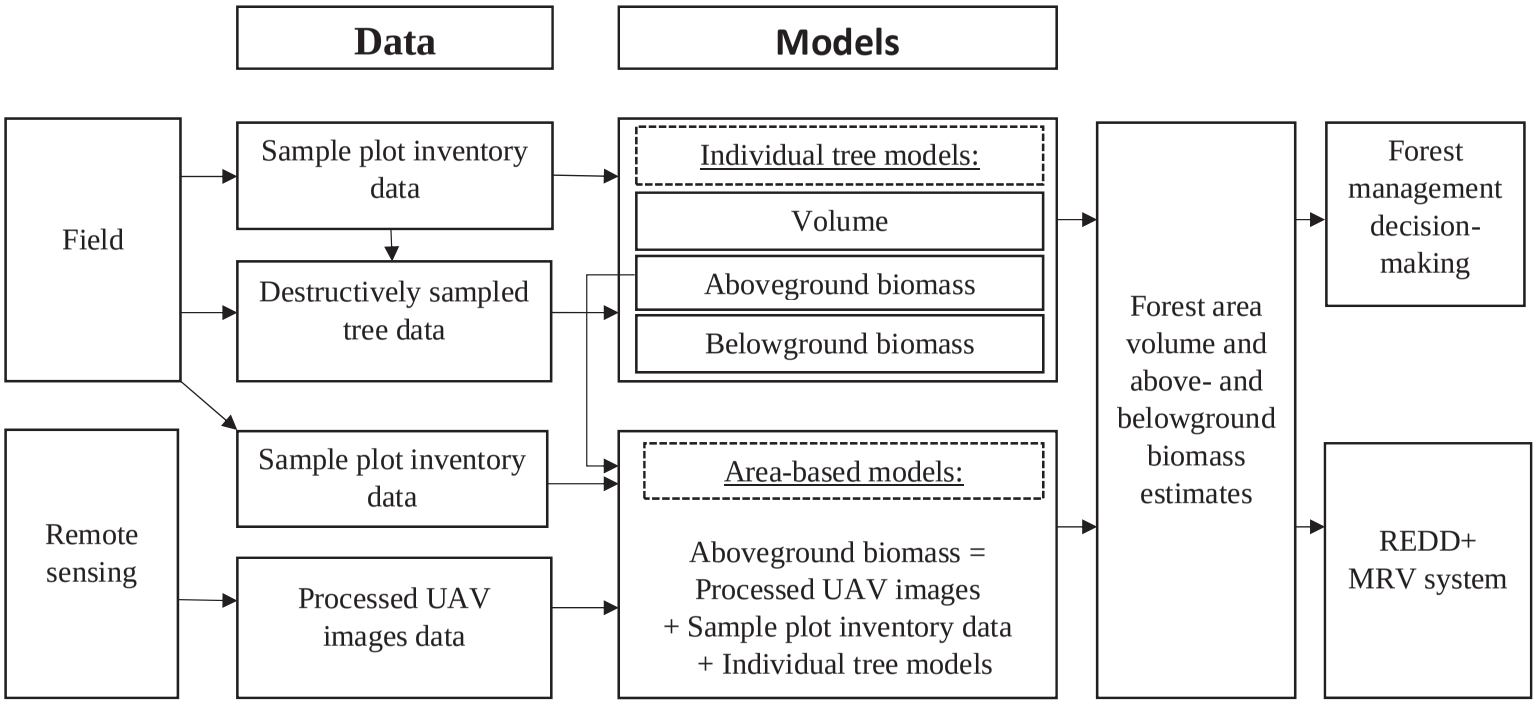


Figure 1. A conceptual framework for estimation of forest area volume and above- and belowground biomass.

4.0 Background

4.1 Volume models

Availability of volume models is regarded as a basic prerequisite for implementation of sustainable forest management. Volume models are important for establishing current growing stock of forests, timber valuation, selection of forest areas in harvest scheduling, growth and yield studies and as a basis for estimation of biomass and carbon stocks. Furthermore, the government of Malawi uses a licensing system that permits the issuance of permits to individuals for accessing timber in public forests. In this context, merchantable stem volume models are required. Merchantable stem volume models may also be useful in cases where compensation payments to tree/forest owners are required when trees are being cleared for infrastructure development, such as roads, railways and buildings. Branch volume models can be used as tools for assessing wood quantities related to brick burning as well as in the production of domestic fuelwood, charcoal and construction poles.

A review by Henry et al. (2011) showed that many models for predicting tree volume in miombo woodlands have been developed previously. Most of these models were developed in miombo woodlands located in neighbouring countries like Tanzania (Chamshama et al. 2004; Malimbwi et al. 1994; Mauya et al. 2014; Mwakalukwa et al. 2014), Zambia (Chidumayo 1988) and Mozambique (Mate 2014). Due to high biogeographical variability in the miombo ecoregion, there is a need for developing models that can be applied locally. The only existing tree volume models for miombo woodlands in Malawi were developed by Abbot et al. (1997). However, application of these models is limited due several reasons regarding the data used for model calibration: a) narrow geographical ranges of study sites, (b) relatively small ranges of diameter at breast height and (c) relatively few tree species.

The main objective of Paper 1 was therefore to develop general total tree volume models, as well as general tree sectional models for branches and merchantable stems, applicable across the entire distribution of miombo woodlands in Malawi.

4.2 Biomass models

Estimation of biomass is the first step towards calculation of carbon stocks in forest ecosystems. Due to the natural capacity of trees to sequester carbon dioxide, miombo woodlands are considered an important element in global climate change mitigation programs such as REDD+. Establishment of credible national MRV systems for REDD+

implementation requires estimation of biomass using either field or remote sensing-based methods. Both these methods rely on the availability of reliable biomass models.

Biomass can be estimated using either direct or indirect methods. Direct methods involve harvesting all trees in a known area and measuring the oven dry weight of the different components of the harvested trees such as the stem, leaves, roots and branches. Although this method determines biomass accurately for a particular area, it is time and resource consuming, strenuous, destructive, expensive and not feasible for large scale analysis (Vashum & Jayakumar 2012). On the other hand, indirect methods involve applying individual tree models for predicting biomass, or expansion factors and/or root to shoot ratios (Brown 2002). Application of individual tree models is now the most widely used method in forest biomass estimation.

By 2011 there were approximately 370 models for predicting tree biomass in sub-Saharan Africa (Henry et al. 2011). The majority of these models were developed for tropical rainforests in western Africa. Among the reviewed models, and those developed after the review in south-eastern Africa, only a few were developed for miombo woodlands (Chamshama et al. 2004; Chidumayo 2014; Kuyah et al. 2016; Malimbwi et al. 1994; Mate et al. 2014; Mugasha et al. 2013; Mwakalukwa et al. 2014; Ryan et al. 2011). Among these, the models developed by Kuyah et al. (2016) are the only ones based on data from Malawi. However, these models are also limited for the same reasons limiting existing volume models (narrow geographical ranges, relatively small ranges of diameters and relatively few tree species). The models developed by Kuyah et al. (2016) were also developed using miombo trees from outside forests, hence limiting their applicability in the REDD+ mechanism which is currently targeting trees in forest reserves.

Furthermore, most of the described models for miombo woodlands focused on aboveground biomass only. However, estimation of belowground biomass is also vital. Existing belowground biomass models for miombo woodlands in neighbouring countries were developed by Mugasha et al. (2013), Chidumayo (2014) and Ryan et al. (2011). For Malawi, however, no belowground biomass models exist.

The main objective of Paper 2 was therefore to develop general above- and belowground biomass models applicable across the entire distribution of miombo woodlands in Malawi.

The models were also accompanied with information on their covariance structure to enable quantification of model-related uncertainties in biomass and carbon estimation.

4.3 Application of UAVs in biomass prediction

Remote sensing methods can be used to collect data for estimating forest volume or aboveground biomass. Prediction of these attributes using this approach involve conducting sample plot forest inventories based on a relatively small number of sample plots to determine field reference biomass. The field reference biomass is then regressed with metrics derived from remotely sensed data for the respective sample plots. The developed models are finally used to predict volume or biomass for the entire study area. For forestry applications, remotely sensed data is mainly sourced from three main systems, namely, airborne laser scanning, radio detection and ranging (e.g. synthetic aperture radar) and optical (e.g. satellite and aerial images) (Kumar et al. 2015). Currently, application of UAVs for predicting volume or aboveground biomass is slowly gaining ground due to UAVs ability to acquire high quality 3D data on forests at relatively low costs (Dandois & Ellis 2013; Getzin et al. 2012; Puliti et al. 2015; Tang & Shao 2015). Furthermore, the availability of user-friendly image processing software has made the application of the technology attractive (Dandois & Ellis 2013; Puliti et al. 2015). Application of this technology to potential REDD+ projects in Malawi could be an attractive option since the sizes of approximately 50% of potential project areas are ideal for efficient application of UAVs in biomass prediction (see Puliti et al. 2015). However, the application of this newly developed technology for biomass prediction in the miombo woodlands of Malawi still needs to be tested.

Successful prediction of forest attributes using remotely sensed data is dependent on the availability of a reliable digital terrain model for correct estimation of ground elevation for the study area. Images collected by UAVs may not be suitable for generating reliable digital terrain models since it is mainly concentrated in the top of the forest canopy. Reliable digital terrain models are usually generated from airborne laser scanning data. However, due to the high costs associated with acquiring such data, it is imperative for researchers utilizing UAVs in developing countries to strive to search for relatively accurate, but also cost efficient digital terrain model generating approaches.

The main objective of Paper 3 was therefore to evaluate the application of photogrammetric point cloud data generated from UAV acquired images in aboveground biomass prediction

for miombo woodlands. Digital terrain models generated from the photogrammetric point cloud based on different methods and parameter settings were also compared.

4.4 Influence of plot and sample size on UAV-assisted biomass estimates

Data from field-based probability sample plot inventories are important for estimating forest biomass/volume during UAV-assisted inventories as they reportedly improve the estimates (Næsset et al. 2011). Determination of field plot size is an important design decision when planning field-based probability sample inventories. In estimation based on field-based probability sample data combined with auxiliary data from remote sensing i.e. design-based and model-assisted inferential framework, an appropriate geographical correspondence between plots on the ground and the remotely sensed data is paramount. An increased sample plot size can reduce the effects of errors arising from co-registration problems (Frazer et al. 2011). Larger plots will also tend to reduce the plot boundary effects (McRoberts et al. 2014).

Several authors have studied the effect of sample plot size on biomass estimates and other forest attributes in inventories assisted by remotely sensed data in tropical wet forests (Asner et al. 2009; Hansen et al. 2015; Keller et al. 2001; Mascaro et al. 2011; Mauya et al. 2015b; Saatchi et al. 2011), temperate forests (Frazer et al. 2011; Levick et al. 2016), boreal forests (Gobakken & Næsset 2008; Næsset et al. 2015) among others. Apart from sample plot size, sample size, i.e. the number of sample plots employed during an inventory, will also have a large effect on the efficiency of biomass estimates and the associated total inventory costs (Eid et al. 2004; Gobakken & Næsset 2008; Strunk et al. 2012).

To the best of our knowledge, no studies on the influence of sample plot size and sample size on efficiency of biomass estimates (or other forest attributes) have been done in UAV-assisted sample plot inventories, i.e. using design-based and model-assisted inferential framework in miombo woodlands. The main objective of Paper 4 was therefore to assess the efficiency of using a UAV-assisted estimation of biomass in a case study in miombo woodlands of Malawi based on different sample sizes and sample plot sizes.

5.0 Materials and methods

5.1 Study sites

Figure 2 presents the location of the study sites. The sample trees for the development of volume and biomass models in Papers 1 and 2 were selected from four forest reserves, namely Mtangatanga (northern Malawi), Kongwe (central Malawi), Mua-livulezi (central Malawi) and Tsamba (southern Malawi). The selection of sites was based on geographical location and climatic conditions to capture a wide range of factors influencing tree growth. Data for Papers 3 and 4 was collected from Muyobe community forest reserve in the northern Malawi.

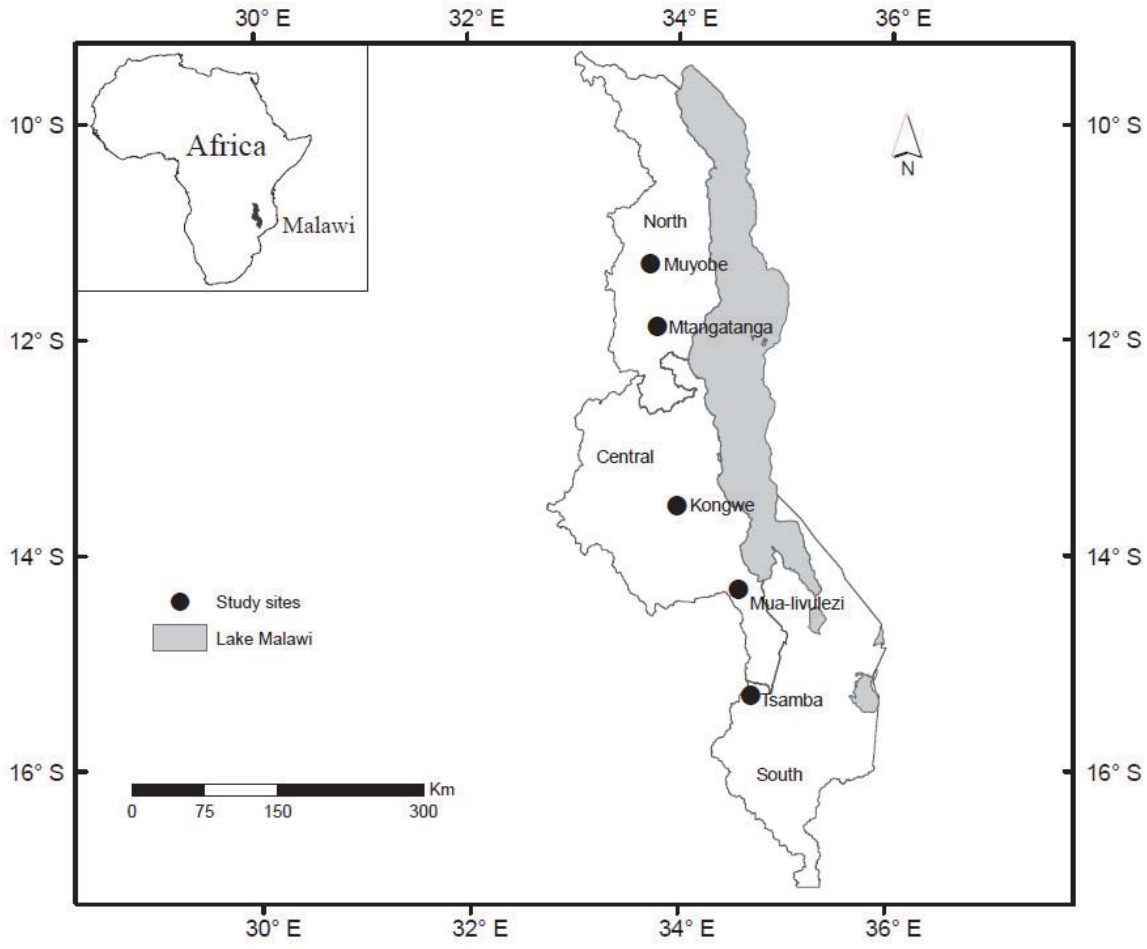


Figure 2. Map of Malawi showing the location of the study sites.

5.2 Data collection

a) Sample plot inventory data

Sample plot field inventory data was required for all the four papers. For Papers 1 and 2 the inventories were conducted on systematically distributed 0.04 ha circular plots. The inventories covered a total of 221 plots with 70, 30, 71 and 50 plots for Mtangatanga, Kongwe, Mua-livulezi and Tsamba, respectively. On each plot, all trees with diameters at breast height > 4 cm were identified and had their diameters at breast height measured. In addition, we sampled three trees within each plot (with the smallest, medium and largest diameters at breast height), and measured their total height using a Vertex hypsometer. In total, for all the study sites, we identified 139 tree species. The sample plot inventory data was then used for selection of sample trees that were destructively sampled.

For Papers 3 and 4, the inventory was conducted on 107 systematically distributed probability sample plots which were circular (radius = 17.84 m, 0.1 ha each). On each plot, the following tree variables were recorded: Total horizontal distances from the plot centres to each tree (using a Haglöf vertex hypsometer), diameter at breast height (using a caliper or a diameter tape) and scientific name of all trees ≥ 5 cm. The total horizontal distances from the plot centres to each tree were calculated as the sum of the horizontal distance to the front of each tree and half of the tree's diameter at breast height. These distances were subsequently used to subset the sample plot data into different sizes, i.e. 250, 500, 750 and 1000 m^2, for further analysis.

In order to assess the effect of sample size on precision of biomass estimates we considered three different systematic samples of different sizes, i.e., the full sample of 107 plots, one sample with half the size (54 plots) in which every second plot was excluded, and finally one sample of one third of the full size (36 plots) in which every third plot was retained. In total 12 datasets (i.e. four sample plot sizes × three sets of sample sizes) were created and used for the analyses.

Furthermore, total tree height of up to 10 randomly selected sample trees within each plot were measured using a Haglöf vertex hypsometer. Precise registration of the positions of centres for sample plots is very important in remote sensing-assisted forest inventories. In this study, positions of the plot centres were measured with a differential Global Navigation Satellite Systems (dGNSS) unit. The dGNSS unit is comprised of two Topcon legacy- E +40

dual frequency receivers. One of the receivers was used as a base station unit and the other as a rover field unit. The receivers observe pseudo-range and carrier phase of both the Global Positioning System (GPS) and the Global Navigation Satellite System (GLONASS). During the study, the baseline between the base station and rover units was approximately 25 km. The position of the base station was determined using Precise Point Positioning (PPP) with GPS and GLONASS data collected continuously for 24 hours as suggested by Kouba (2015) before commencement of the forest inventory. The rover field unit was placed at the centre of each sample plot on a 2.98 m rod for an average of 33±20 minutes using a one-second logging rate. The recorded plot centre coordinates were post-processed using the RTKLIB software (Takasu 2009) and the results revealed that the maximum deviations for northing, easting and height were 1.16 cm, 3.02 cm and 3.06 cm, respectively.

b) Destructively sampled tree data

For development of above- and belowground biomass models, as well as volume models in Papers 1 and 2, a total of 74 trees were selected based on the observed diameters at breast height and tree species frequency within the sites. We ensured that the trees were selected from all diameter at breast height classes observed in the sample plot inventories. In addition, we selected a total of eight trees with larger diameter at breast height than those observed in the sample plot inventories to reduce uncertainty when predicting biomass of very large trees. We also selected at least one tree among the eight most frequently observed species in each site. The remaining sample trees were selected randomly among all species. In total, 33 tree species were selected, comprising 10, 10, 12 and 10 different tree species in Mtangatanga, Kongwe, Mua-livulezi and Tsamba, respectively. Before felling the selected trees, we measured their diameters at breast height and total tree height, and also determined their species names. Out of the 74 trees, 41 trees were excavated for determination of belowground biomass.

For determination of aboveground biomass, the aboveground portion of each of the 74 trees was separated into the following components: merchantable stem (from the stump at 30 cm above ground to the point where the first branches start), branches (all parts of the tree above the defined merchantable stem and up to a minimum diameter of 2.5 cm) and twigs (all branches with a diameter less than 2.5 cm). For small trees not considered suitable for timber production (diameter at breast height < 15 cm, in total 14 trees), merchantable stem volume/biomass were allocated to branches. To facilitate measurements, the stems and

branches were crosscut into manageable logs of approximately 1–2 m in length. We measured the lengths and the mid-diameters over bark of each of the logs and then weighed their fresh weight using a mechanical hanging spring balance (0 – 200 kg). Twigs from each tree were separately bundled and weighed to determine their fresh weight.

For determination of belowground biomass, our strategy involved root sampling at two levels (Mugasha et al. 2013), namely main roots (roots branching directly from the root crown) and side roots (roots branching from the main roots). The first step in excavation involved clearing the topsoil around the tree base to expose the points at which the roots were branching. We then selected three main roots, i.e. the main roots with the largest, medium and smallest diameters and recorded their diameters at the points where they joined the root crown. The diameters of all main roots not excavated were recorded at the point where they joined the root crown. From each of the selected main roots, we selected up to three side roots, i.e. the side roots with the largest, medium and smallest diameters. For each of the selected side roots, we recorded the diameter where they joined the main root. For the remaining side roots, we also recorded the diameters at the branching point from the mainroot. The selected side and main roots were then fully excavated up to a minimum diameter of 1 cm and then weighed.

In cases where the full roots could not be excavated due to obstacles such as rocks, the diameter of the last bit of the root was recorded and we treated the remaining unexcavated part as a side root. An effort was made to ensure that all the taproots were fully excavated up to a diameter of 1 cm. In total, 38 out of the 41 trees had taproots. Out of these 38 trees, we were not able to fully excavate the taproots of 16 trees. In such cases, the diameter at the breaking point of the unexcavated taproot was recorded and treated as a side root. On average, tap roots were dug down to 2.5 m depth. Lastly, we recorded the fresh weight of the root crown for each tree. For all sample trees, three small sub-samples, varying in weight between 0.1 and 1.0 kg, were taken from each main and side root, and one was taken from the root crown. We obtained the fresh weight of the sub-samples using an electronic balance and brought them to the laboratory for oven drying.

Photo 1. Miombo woodlands during dry season (a), weighing a log during destructive sampling (b), Sensefly eBee Unmanned Aerial Vehicle (c), preparing to fly the Sensefly eBee Unmanned Aerial Vehicle (d).

c) Processed UAV images data

The images used in Papers 3 and 4 were acquired using a SenseFly eBee fixed-wing UAV (Sensefly 2015). The UAV was made from flexible foam weighing 537 g without camera. The UAV was equipped with a Canon IXUS 127 HS Digital camera. The dimensions and weight of camera with battery and memory card were 93.2 × 57.0 × 20.0 mm and 135 g, respectively. The camera produces 16.1 megapixel images in the red, green and blue spectral bands. The UAV is also equipped with an inertial measurement unit as well as an on-board Global Navigation Satellite Systems (GNSS) to control the flight and to provide rough positioning (Sensefly 2015). Prior to taking images, positions of ground control points (GCPs) as well as landing and take-off points, e.g. on open areas with no trees within the forest and agricultural fields near the forest, were identified and measured. The GCPs were

made of a set of 1 × 1 m cross-shaped timber planks painted white and some black and white 50 × 50 cm checkerboards. The position of the centre of each GCP was fixed using the same procedure as used when locating plot centres for the sample plot inventory described above. The data were collected for an average of 13±6 minutes for each GCP with a 1-second logging rate. The recorded coordinates for each GCP were post-processed similarly as the sample plots. The results revealed that maximum deviations for northing, easting and height were 2.24 cm, 4.50 cm and 4.46 cm, respectively.

Acquisition of images was controlled from a laptop computer with a mission control software eMotion 2 version 2.4 (Sensefly 2015). All the flights were planned in the mission control software prior to flying. For navigation purposes, a georeferenced base map from Microsoft Bing maps covering the study area. For this study we applied percentage end and side image overlaps of 80 and 90% respectively, as well as a fixed flight height above the ground of 325 m. In total 20 flights were carried out to cover the forest.

5.3 Data analyses

For development of volume models (Paper 1), volumes of individual logs were calculated by multiplying the basal area of the mid-section of each log by its length. Subsequently, the stem and branch volumes for each tree were determined by summing all individual log volumes for the respective sections. Total tree volumes were determined by summing the merchantable stem and branches volumes for individual trees.

Development of biomass models (Paper 2) started by first drying all sub-samples from both above- and belowground portions of each tree in an oven at a temperature of 80°C until a constant weight was achieved (constant weight was observed in 2–3 days) and subsequently recording their dry weights. The sub-sample dry and fresh weights were then used to determine the tree- and section specific dry to fresh weight ratios (DF-ratios) which were then used to calculate the dry weight of each section as a product of tree- and section specific DF-ratios and the fresh weights of the respective trees and tree sections. Subsequently, we computed the total aboveground dry weight each tree by summing the dry weights of the merchantable stem, branches and twigs.

To determine the total belowground dry weights of the excavated parts of the trees we first converted all the fresh weights from the different sections to dry weight biomass by

multiplying the tree- and section specific DF-ratios and their respective fresh weights. We then developed a general (combining data from all sites) side root model by regressing the dry weight biomass of the fully excavated side roots and their diameters (cm). The side root model was used to predict the dry weight biomass of all the side roots that were not excavated for the main sample root. The total dry weight of all side roots for each main sample root was then determined by summing dry weights of the excavated side roots and predicted dry weights of unexcavated side roots. Finally the complete dry weight of the sample main root was determined by summing the total dry weights of all side roots and the excavated parts of the main root. A main root model was then developed and applied to predict the dry weights of main roots not excavated. To determine the dry weight of unexcavated parts of the taproots (16 trees), we applied the general side root model. Total belowground dry weight biomass for each tree was finally determined by adding the dry weights of all excavated and unexcavated main roots, dry weight of the taproot and the dry weight of the root crown.

Using the respective datasets, general and site specific volume, aboveground and belowground models for total tree, merchantable stem and branch were developed utilizing diameter at breast height, total tree height and species-specific mean wood specific gravity as independent variables. The species-specific mean wood specific gravity values were extracted from the global wood density database (Chave et al. 2009; Zanne et al. 2009). Since the data demonstrated heteroscedasticity for volume, above-and belowground biomass, we applied generalized methods of moments (GMM) estimation method for volume models and weighted nonlinear regression for above- and belowground biomass models. The analysis was implemented using SAS Institute (2012) software. For all models, pseudo-R^2, root mean square error and mean prediction error values were reported. However, model efficiency and performance were based on root mean square error values calculated using leave-one-out cross validation procedure. Previously developed models were also tested and compared with the models developed in the current study.

Both Papers 3 and 4 required calculating aboveground biomass of each tree in respective sample plots. Before calculating aboveground biomass, total heights of trees whose height was not measured were predicted using a height-diameter model developed (Paper 3) using the measured heights of sample trees from all the sample plots. We then calculated aboveground biomass for each tree in the sample plots by using a model developed in Paper

2, with diameter at breast height and total tree height as independent variables. Per hectare values for aboveground biomass of the respective plots were calculated by first summing up the individual tree aboveground biomass values within a given plot and scaling them to per hectare values.

For both Papers 3 and 4, Agisoft Photoscan Professional version 1.1 (AgiSoft 2015) was used to generate a 3D dense point cloud from the acquired UAV images. To normalize the point clouds and subsequently extract metrics describing canopy height, canopy density and canopy spectral information in both Paper3 and 4, we developed, tested, and selected the best digital terrain models in Paper 3 using different approaches, and compared their performance to determine a suitable digital terrain model since the study area did not have an existing one. The tested approaches included a) supervised ground filtering based on visual classification, b) supervised ground filtering based on logistic regression, c) supervised ground filtering based on quantile regression and d) Shuttle Radar Topography Mission with quantile regression. In Paper 4, the metrics were extracted for each of the datasets for respective plot sizes (i.e. 250, 500, 750 and 1000 m^2).

To compare the performance of the different DTMs in Paper 3 as well as to estimate aboveground biomass for the study area in Paper 4 models relating reference aboveground biomass and the generated metrics were fitted on square root transformed dependent variables using multiple linear regression in R software (R Core Team 2016).

For both Papers 3 and 4, the developed models were evaluated using the squared Pearson correlation coefficient, root mean square error, relative root mean square error, mean prediction error and relative mean prediction error. Model selection was however based on the root mean square error values.

To assess the efficiency of UAV-assisted as well as the effect of sample plot and sample sizes on error estimates in biomass estimation in Paper 4, field-based biomass estimates and corresponding variances were based on the simple random sampling estimator. On the other hand, a model-assisted regression estimator described by Särndal et al. (1992), and its corresponding variance estimator, were applied for UAV-assisted biomass estimation. The relative efficiency (RE) of UAV-assisted inventory was assessed by a ratio of the variance

estimates for the biomass based on purely field-based inventory data to that based on UAV-assisted inventory data.

Furthermore, to assess the cost efficiency of UAV-assisted over pure field-based inventories in Paper 4, during field work we randomly selected 16 sample plots and for each plot recorded three categories of time consumption, i.e. fixed time (time spent when recording sample plot attributes such as plot number, date, etc.), variable time (time spent on measuring trees) and walking time (time spent during walking from one plot to another). The average recorded time consumption was 7.5, 25.0 and 7.0 minutes for each of the aspects, respectively. We then set the relative cost of a sample plot inventory of 107 sample plots (1000 m^2 each) in a 220 × 220 m grid to 100% based on the recorded information. We then used the cost information from the current inventory (4 persons working for 15 days with a daily salary of USD 25.13 each) to calculate the variable costs for each plot scaled according to plot size and walking distance.

The costs for the UAV data acquisition were fixed for all sample plot sizes and sample sizes because the need for auxiliary remotely sensed information would be the same regardless of plot size and sample size. The cost was computed based on the experience from the current study. The cost included pre-flight preparations and the actual flying where a two-man crew was required. Each person worked five days with a salary similar to the field crew. Post-processing of the acquired images required four days.

6.0 Main findings and discussion

6.1 Volume and biomass models

The developed volume, above- and belowground biomass models (Papers 1 and 2) offer options for forest inventory scenarios in which data on diameter at breast height only or both diameter at breast height and total tree height are available. For both volume and biomass, the root mean square error and mean prediction error values of the models with both diameter at breast height and total tree heights as independent variables were better than those of the models with diameter at breast height as the only independent variable. This result also conforms to previous studies (Abbot et al. 1997; Mauya et al. 2014; Mwakalukwa et al. 2014) for volume models and Mugasha et al. (2013) for aboveground biomass models. On the other hand, for belowground biomass, the only viable model had diameter at breast height as the only independent variable. The fit of this model is similar to that of the models developed by Mugasha et al. (2013), Chidumayo (2014) and Ryan et al. (2011).

If diameters at breast height and total heights of all trees are measured in an inventory, the model including both variables should, of course, be applied. Otherwise, models with diameter at breast height alone are still reliable since much of the variation in volume and aboveground biomass was explained by diameter at breast height, while the addition of total tree height resulted in only small improvements. Since total tree height measurements are time consuming, they are usually estimated from height-diameter models developed from a few sample trees. If all tree forms in the forest are not represented among the sample trees, additional uncertainties in predictions are introduced. With appropriate sample trees and small measurement errors in tree heights, the accuracy of predictions will probably be improved by including total tree height as an independent variable, in spite of the uncertainty added by using a height-diameter model. For aboveground biomass models, inclusion of species-specific mean wood specific gravity values in place of total tree height did not improve the performance of the model. This could be attributed to the fact that the species-specific mean wood specific gravity values were not obtained directly from the sampled trees, but from the global wood density database (Chave et al. 2009; Zanne et al. 2009).

Tree component volume and aboveground biomass models, i.e. for twigs, branches and merchantable stem, may be useful when planning commercial extraction of timber or quantification of volume or aboveground biomass for domestic fuelwood or charcoal production. All tree component models with significant parameter estimates produced mean

prediction error values not significantly different from zero, an indication of appropriate model performance.

When the selected volume and aboveground biomass models were tested on our dataset over different sites, none of the mean prediction error values were significantly different from zero, except for the volume model with both diameter at breast height and total tree height as independent variables in Tsamba, where volume was over-estimated. Furthermore, when previously developed models were tested on our dataset, the results showed that these models either over- or underestimated biomass (Table 1) or tree volume (Figure 3). These results demonstrate the importance of developing local models and also highlight the dangers of applying models beyond their geographical ranges because a change in geographical site in most cases also mean changes in ecological, climatic and edaphic conditions.

Table 1. Performance of previously developed biomass models tested on our dataset.

Component	Model	Independent variable(s)	No. of trees	Observed	Predicted	MPE	
				(kg)	(kg)	(kg)	(%)
Above-ground	Mugasha et al. (2013)	dbh	74	1239.7	1135.7	104.0	8.4
	Mugasha et al. (2013)	dbh, ht	74	1239.7	1076.7	163.0	13.2 **
	Ryan et al. (2011)	dbh	74	1239.7	1068.8	170.9	13.8 *
	Chidumayo (2014)	dbh	74	1239.7	1205.6	34.1	2.8
	Chave et al. (2014)	dbh, ρ, ht	74	1239.7	953.7	286.1	23.1 ***
Below-ground	Mugasha et al. (2013)	dbh	41	527.2	377.5	149.7	28.4 ***
	Mugasha et al. (2013)	dbh, ht	41	527.2	364.8	162.4	30.8 ***
	Ryan et al. (2011)	dbh	41	527.2	426.9	100.3	19.0 ***
	Chidumayo (2014)	dbh	41	527.2	551.9	−24.7	−4.7

***** MPE is significantly different from zero at ($p < 0.05$); ** MPE is significantly different from zero at ($p < 0.01$) and *** MPE is significantly different from zero at ($p < 0.001$), dbh = diameter at breast height, ht = total tree height, ρ = species-specific mean wood specific gravity.

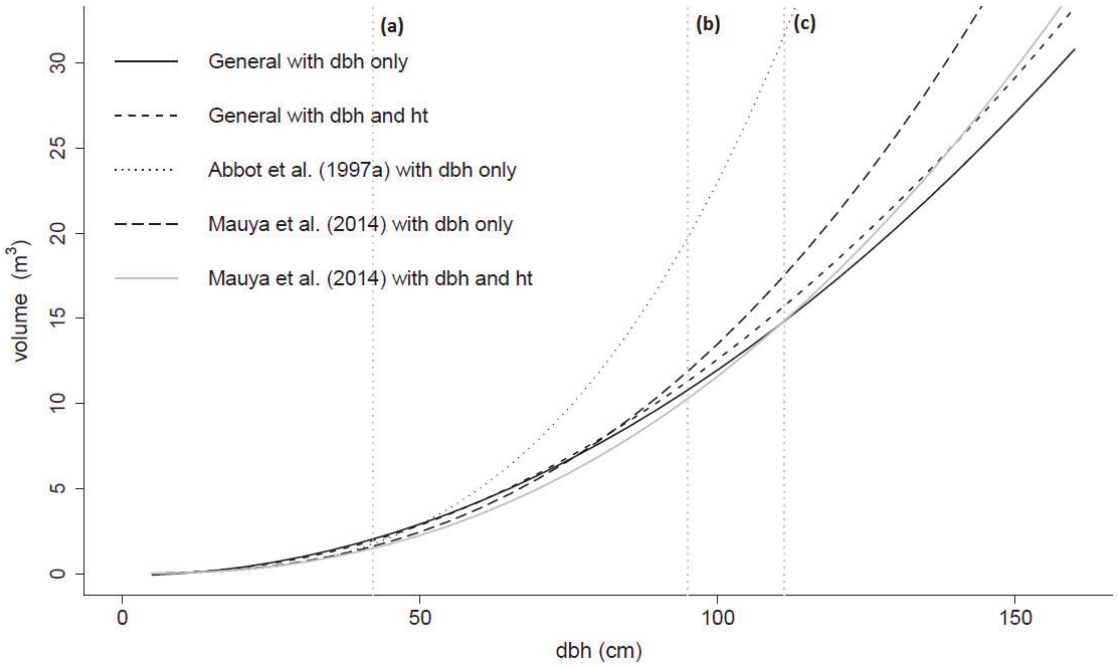


Figure 3. Display of total tree volume over diameter at breast height (dbh) for models developed in this study and previously. For the models with total tree height included as an independent variable, a height–diameter model developed from our sample trees was applied. Vertical dotted lines are the maximum diameter at breast height of the modelling datasets used by Abbot et al. (1997) (a), Mauya et al. (2014) (b) and in this study (c), respectively.

Recently, Kuyah et al. (2016) also developed aboveground biomass models for miombo woodlands in Malawi. These models were based on miombo trees outside forests collected from three sites in the central and southern region of Malawi. These models are thus suitable for biomass estimation for miombo trees in agroforestry systems during national forest inventories when biomass of trees outside forests is also considered (Schnell et al. 2014). Unlike the models developed in this study, application of models developed by Kuyah et al. (2016) for the REDD+ mechanism in Malawi is limited, since the potential project areas are forest reserves scattered across the country.

6.2 Application of UAVs in biomass prediction

Reliable biomass estimates from remotely sensed 3D data are heavily reliant on the availability of a good digital terrain model. In paper 3 we first tested different methods of generating digital terrain models. The comparisons of plot centre height predictions from different digital terrain models showed that predictions from the digital terrain model

generated using Shuttle Radar Topography Mission data are unreliable as compared to those derived from the other methods. This indicates that when digital terrain models based on Shuttle Radar Topography Mission data are used in biomass estimation, the estimates can hardly be trusted.

Biomass predictions from the digital terrain models developed based on the tested approaches show that the digital terrain model developed using unsupervised ground filtering based on a grid search approach performed slightly better than others. This performance demonstrated that with some effort, it is possible to find good parameter settings in the AgiSoft Photoscan software (AgiSoft 2015). Furthermore, despite performing slightly less than the digital terrain model developed using unsupervised ground filtering based on a grid search approach, the digital terrain model based on supervised ground filtering using visual classification, was equally good. However, since unsupervised ground filtering is relatively easier to implement, future studies should consider application of this approach. On the other hand, the relatively poor performance of the digital terrain model developed from unsupervised ground filtering based on Shuttle Radar Topography Mission could be attributed to the inherent random errors in heights associated with shuttle radar topography mission data (Hofton et al. 2006; Karwel & Ewiak 2008; Rodríguez et al. 2006).

The root mean square error value for the best model from our study is similar to that reported in a study conducted in miombo woodlands of Tanzania by Mauya et al. (2015a) when using ALS data. On the other hand, in a study by Puliti et al. (2015), where data acquired from UAV was applied in boreal forests, a smaller root mean square error value compared to our study was observed when estimating forest stand volume. This might be attributed to the differences in forest structures between miombo woodlands and boreal forests. It should also be noted that Puliti et al. (2015) utilized ALS data for digital terrain model determination, which are superior in describing forest ground surface compared to optical sensors such as those applied in the current study (Baltsavias 1999). It is worth noting that the observed root mean square error in Puliti et al. (2015) is comparable to that observed in a study by Gobakken et al. (2015) also conducted in boreal forests. However, Gobakken et al. (2015) used exclusively ALS data. This demonstrates the efficiency of UAV data in forest inventories.

The findings from our study have also demonstrated that data generated by the UAV system have potential of being successfully used in estimating forest biomass in dry tropical forests such as miombo woodlands. Similar studies in other dry tropical forests are however recommended to validate the results of the current study because of the wide range of variations in structure, weather and terrain conditions seen in these forests.

6.3 Influence of plot and sample size on UAV-assisted biomass estimates

The results from Paper 4 have demonstrated that incorporation of UAV derived photogrammetric data in a forest inventory can improve forest biomass estimates beyond what can be achieved by purely field-based sample plot inventories (see Table 2). The relatively smaller mean biomass standard error values for the UAV-based estimates indicate that inclusion of remotely sensed data from UAV-imagery can improve the precision of biomass estimates. Thus the application of UAV-assisted inventories for REDD+ implementation in Malawi could potentially result in improved biomass estimates compared to pure field-based inventories.

Table 2. Relative efficiency (RE), estimated mean biomass and associated standard error (SE) estimates based on field-based and UAV-assisted estimation for different sample plot sizes and sample sizes.

Plot size (m^2)	Sample size *(n)*	Field-based (Mg ha^{-1})		UAV-assisted (Mg ha^{-1})		Relative efficiency
		$\hat{B}_{field}$	SE	$\hat{B}_{uav}$	SE	
250	107	36.86	3.29	44.12	2.75	1.44
250	54	36.23	4.58	43.63	3.79	1.47
250	36	36.37	6.21	49.69	5.20	1.43
500	107	37.38	2.96	42.49	2.22	1.77
500	54	39.87	4.57	45.60	3.59	1.62
500	36	34.21	4.68	42.42	3.52	1.76
750	107	38.12	2.79	42.16	1.86	2.26
750	54	39.50	4.13	43.39	3.07	1.81
750	36	32.63	4.15	43.81	2.56	2.63
1000	107	38.99	2.85	43.30	1.72	2.75
1000	54	39.59	4.09	43.11	2.30	3.16
1000	36	33.12	4.39	40.96	2.36	3.46

$\hat{B}_{field}$ = Estimated mean biomass from ground-based sample, $\hat{B}_{uav}$ = Estimated mean biomass from UAV-assisted data, SE = Estimated standard error of mean biomass, RE = ratio of the variance estimates for the biomass based on purely field-based inventory data to that based on UAV-assisted inventory data.

Furthermore, correct choice of sample plot sizes is critical to the precision and accuracy of biomass estimates in remote sensing based forest inventories (Frazer et al. 2011). This is demonstrated by the increase in the magnitude of relative efficiency values with increasing sample plot size and the decrease in RMSE values with increasing sample plot sizes. The

same trend was observed by Frazer et al. (2011) and Mauya et al. (2015b). The improvement of biomass estimates with increasing sample plot sizes shows that large sample plot sizes favour UAV-assisted inventories. This could be attributed to reduction in plot boundary effects as the sample plot sizes increases as suggested by Goetz and Dubayah (2011). Thus for small sample plots, canopies of trees with wide crowns such as those in miombo woodlands (Frost 1996) tend to be partially included and thus under predicting sample plot biomass. On the other hand, as the sample plot sizes increase, this effect tends to decrease substantially since these variations are averaged out at larger sample plot sizes (Saatchi et al. 2011).

The fact that the results in the current study indicate that larger plots and larger sample sizes favour UAV-assisted forest inventories does not imply that larger sample plots and sample sizes should always be applied during the UAV-assisted inventory because of the associated costs. The results on cost efficiency analysis indicate that there is a trade-off between costs and required precision. On one hand, acquiring UAV data and field reference data from many large plots is expensive but produces more precise results. On the other, acquiring the data from many small plots is less expensive but produces less precise results. Based on the observed trends, if a standard error estimate of less than approximately 3 Mg ha^{-1} was targeted during a forest inventory, then a UAV-assisted forest inventory should be applied to ensure cost efficient and precise estimates. This demonstrates the need for carrying out a cost analysis during UAV-assisted inventory in order to determine the optimal sample plot size and sample size to apply.

Finally, it should be noted that careful planning is needed for application of UAV-assisted inventories under the REDD+ mechanism in Malawi to be accomplished. For example, if the inventory is intended for smaller forest reserves, wall-to-wall coverage using a UAV is possible. On the other hand, in cases where inventories are conducted in larger forest reserves, the UAV can be applied as a sampling tool because wall-to-wall operations maybe found economically and logistically infeasible. Furthermore, this study was conducted on a single site and thus represents a forest inventory scenario at specific location. Although this case study has provided evidence of great efficiency of UAV-assisted inventory, similar studies should be conducted in other reserves across the country in order to be able to generalize and provide guidance for future operational inventories.

7.0 Concluding remarks and future studies

The main objective of the thesis was to develop models and methods for estimating volume and biomass of miombo woodlands in Malawi (Figure 1). The results from this thesis have taken us some steps forward that are expected to support and improve forest management decision-making in general as well as the implementation of a REDD+ MRV system in the country. Still, however, much work and research are needed. In the following, we point at main achievements as well as some weaknesses and corresponding suggestions on more research directly linked to the individual papers. We have also tried to go beyond the scope of the thesis, and have identified a few interesting and relevant topics for future studies that potentially could provide valuable inputs for further improvements in forest management decision-making and REDD+ MRV implementation in Malawi.

The performances and the evaluations of the models developed in Papers 1 and 2 suggest that they can be used over a wide range of geographical and ecological conditions in Malawi with an appropriate accuracy in predictions. The appropriateness of the models, and the importance of using local models in biomass estimation, was also supported by the fact that their mean prediction errors were much lower than some previously developed models tested on our data. In addition to the models for facilitating carbon assessments, we have also developed section models that can be applied when quantifying fuelwood and for timber valuation in compensation payments.

It should, however, be noted that the number of tree species included in the modelling datasets were relatively low when considering the total number of tree species found in miombo woodlands. Future studies should therefore aim at updating the current datasets (displayed in full in Papers 1 and 2) with additional species to improve the robustness of the models (e.g Chave et al. 2014).

The leaves were excluded from the biomass modelling dataset because most of the trees had started to shed leaves when the destructive sampling was carried out. Future studies may therefore attempt to collect the data when all the trees have leaves on them. Furthermore, inclusion of wood specific gravity as an independent variable, in addition to diameter at breast height did not improve the biomass predictions probably because the wood specific gravity values were obtained from the global wood density database. According to Baker et al. (2004), inclusion of wood specific gravity values from sample trees in biomass modelling

is important as it helps in explaining variation in aspects of forest structure that vary significantly at regional scales (Baker et al. 2004; Chave et al. 2014; Ramananantoandro et al. 2015). So future studies should aim at utilizing wood specific gravity values from the sample trees.

The results from the biomass predictions based on a combination of remotely sensed data captured using UAV and field-based inventory data (Paper 3), show that the observed prediction errors are similar to those from previous studies using ALS data in miombo woodlands, thus showing the potential of applying this technology in miombo woodlands. Furthermore, the study highlighted that digital terrain models developed using unsupervised ground filtering based on a grid search approach can produce reliable results in miombo woodlands. Additional studies, however, are recommended to validate these results under other conditions using different flight settings, i.e. flying altitude and image overlaps, to search for the optimum settings. According to Bohlin et al. (2012), both flight altitude and degree of image overlaps influence the accuracy of the 3D data produced.

The results presented in Paper 4 demonstrated that UAV-assisted inventories produced more precise biomass estimates compared to those utilizing exclusively field-based methods. Furthermore, larger plot and sample sizes favour UAV-assisted estimates. The results on cost analysis of UAV-assisted inventory has shown that if a standard error estimate of mean biomass of less than approximately 3 Mg ha^{-1} is targeted during a forest inventory, then a UAV-assisted forest inventory should be applied to ensure cost efficient and precise estimates. However, similar studies should be conducted in other forest reserves across the country in order to be able to generalize and provide guidance for future operational inventories.

If we go beyond the scope of this thesis, an exercise where questions related to error propagations in biomass estimation are approached, would be very important, since the Intergovernmental panel on climate change (IPCC) requires biomass and carbon estimates reporting under the REDD+ mechanism to be accompanied by appropriate measures of uncertainty. Such uncertainties occur when applying the sampling design (sample plot size and shape, sample size), during tree measurements and when applying biomass models (Chave et al. 2004; Clark & Kellner 2012; Magnussen & Carillo 2015; McRoberts & Westfall 2016; Molto et al. 2013). All datasets described in this theses could be applied for

error propagation in volume or biomass estimation. This could be done by using different biomass models (with corresponding covariance matrices as displayed in Paper 2), exclusively field-based methods and in combination with UAVs.

Another step would to be study methods and uncertainty related to determination of biomass changes over time. This is also important in the context of IPCC requirements on biomass and carbon reporting under the REDD+ mechanism. In particular, uncertainties related to different biomass change detection procedures (e.g. Magnussen et al. 2015; McRoberts et al. 2015) would be important. A study directly relevant for the miombo woodlands of Malawi could be done for the same study area as used in Papers 3 and 4, where, after some years, the sample plot inventory in combination with the UAV acquisition is repeated for estimating biomass.

A third step to ensure a sustained reduction in emissions from deforestation and forest degradation should be to conduct further research to understand the drivers of deforestation and forest degradation (Gizachew & Duguma 2016; Kissinger et al. 2012). Further studies on the establishment of sustainable forest management regimes capable of enhancing forest conservation and carbon stocks are also necessary (Edwards et al. 2010). In order to facilitate a better planning environment, a decision-making tool based on growth, mortality and recruitment models, like the one developed from miombo woodlands in Tanzania, is required (see Mugasha et al. 2016a; Mugasha et al. 2016b).

Finally, participation of local communities is critical for the sustainability of REDD+ in Malawi. Studies on assessing the feasibility of incorporating aspects of participatory MRV in the current preparatory phase of REDD+ implementation would therefore be important to check the interest of local communities surrounding the forest reserves (Danielsen et al. 2011; Hawthorne & Boissière 2014; Zahabu 2008).

8.0 References

Abbot, J. I. O. & Homewood, K. (1999). A history of change: causes of miombo woodland decline in a protected area in Malawi. *Journal of Applied Ecology*, 36 (3): 422-433.

Abbot, P., Lowore, J. & Werren, M. (1997). Models for the estimation of single tree volume in four Miombo woodland types. *Forest Ecology and Management*, 97 (1): 25-37.

AgiSoft. (2015). Agisoft PhotoScan User Manual: Professional Edition. Version 1.1: 85. Available at: downloads.agisoft.ru/pdf/photoscan-pro_1_0_0_en.pdf (accessed: 04/09/2015).

Asner, G. P., Flint Hughes, R., Varga, T. A., Knapp, D. E. & Kennedy-Bowdoin, T. (2009). Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. *Ecosystems*, 12 (2): 261-278.

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., et al. (2004). Variation in wood density determines spatial patterns in Amazonian forest biomass. *Global Change Biology*, 10 (5): 545-562.

Baltsavias, E. P. (1999). A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (2–3): 83-94.

Barquín, L., Chacón, M., Panfil, S., Adeleke, A., Florian, E. & Triraganon, R. (2014). The knowledge and skills needed to engage in REDD+: a competencies framework. Arlington, Virginia, USA. : International Union for the Conservation of Nature. 171 pp.

Blackie, R., Baldauf, C., Gautier, D., Gumbo, D., Kassa, H., Parthasarathy, N., Paumgarten, F., Sola, P., Pulla, S., Waeber, P., et al. (2014). *Tropical dry forests: the state of global knowledge and recommendetations for future research.* Discussion Paper. Bogor, Indonesia: CIFOR. 38 pp.

Bohlin, J., Wallerman, J. & Fransson, J. E. S. (2012). Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM. *Scandinavian Journal of Forest Research*, 27 (7): 692-699.

Brown, S. (2002). Measuring carbon in forests: current status and future challenges. *Environmental Pollution*, 116 (3): 363-372.

Chamshama, S. A. O., Mugasha, A. G. & Zahabu, E. (2004). Stand biomass and volume estimation for miombo woodlands at Kitulangalo, Morogoro, Tanzania. *The Southern African Forestry Journal*, 200 (1): 59-70.

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 B: Biological Sciences*, 359 (1443): 409-420.

Chave, J., Coomes, D. A., 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., Réjou-Méchain, M., Búrquez, A., Chidumayo, E. N., Colgan, M. S., Delitti, W. B. C., Duque, A., Eid, T., Fearnside, P. M., Goodman, R. C., et al. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*, 20 (10): 3177-3190.

Chidumayo, E. N. (1988). Estimating fuelwood production and yield in regrowth dry miombo woodland in Zambia. *Forest Ecology and Management*, 24 (1): 59-66.

Chidumayo, E. N. & Gumbo, D. J. (2010). *The dry forests and woodlands of Africa: managing for products and services*. London, UK: Earthscan. 304 pp.

Chidumayo, E. N. (2014). Estimating tree biomass and changes in root biomass following clear-cutting of Brachystegia-Julbernardia (miombo) woodland in central Zambia. *Environmental Conservation*, 41 (1): 54-63.

Clark, D. B. & Kellner, J. R. (2012). Tropical forest biomass estimation and the fallacy of misplaced concreteness. *Journal of Vegetation Science*, 23 (6): 1191-1196.

Dandois, J. P. & Ellis, E. C. (2013). High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. *Remote sensing of environment*, 136: 259-276.

Danielsen, F., Skutsch, M., Burgess, N. D., Jensen, P. M., Andrianandrasana, H., Karky, B., Lewis, R., Lovett, J. C., Massao, J., Ngaga, Y., et al. (2011). At the heart of REDD+: a role for local people in monitoring forests? *Conservation Letters*, 4 (2): 158-167.

Dewees, P. A., Campbell, B. M., Katerere, Y., Sitoe, A., Cunningham, A. B., Angelsen, A. & Wunder, S. (2011). *Managing the miombo woodlands of southern Africa: policies, incentives, and options for the rural poor*. Washington, D.C., USA: Program on Forests (PROFOR).

Dirzo, R., Young, H. S., Mooney, H. A. & Ceballos, G. (2011). *Seasonally dry tropical forests ecology and conservation*. 1718 Connecticut Avenue NW, Suite 300, Washington, DC 20009, USA: Island Press. 407 pp.

Edwards, D. P., Fisher, B. & Boyd, E. (2010). Protecting degraded rainforests: enhancement of forest carbon stocks under REDD+. *Conservation Letters*, 3 (5): 313-316.

Eid, T., Gobakken, T. & Næsset, E. (2004). Comparing stand inventories for large areas based on photo-interpretation and laser scanning by means of cost-plus-loss analyses. *Scandinavian Journal of Forest Research*, 19 (6): 512-523.

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. *Remote Sensing of Environment*, 115 (2): 636-649.

Frost, P. (1996). The ecology of miombo woodlands. In Campbell, B. (ed.) *The miombo in transition: woodlands and welfare in Africa*, pp. 11-57. Bogor, Indonesia: Centre for International Forestry Research.

Furley, P. A., Rees, R. M., Ryan, C. M. & Saiz, G. (2008). Savanna burning and the assessment of long-term fire experiments with particular reference to Zimbabwe. *Progress in Physical Geography*, 32 (6): 611-634.

Getzin, S., Wiegand, K. & Schöning, I. (2012). Assessing biodiversity in forests using very high-resolution images and unmanned aerial vehicles. *Methods in ecology and evolution*, 3 (2): 397-404.

Giliba, R. A., E.K., B., Kayombo, C. J., Musamba, E. B., Kashindye, A. M. & Shayo, P. F. (2011). Species composition, richness and diversity in miombo woodlands of Bereku forest reserve, Tanzania. *Journal of Biodiversity*, 2 (1): 1-7.

Gizachew, B. & Duguma, L. A. (2016). Forest carbon monitoring and reporting for REDD+: what future for Africa? *Environmental Management*: 1-9.

Gobakken, T. & Næsset, E. (2008). Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, 38 (5): 1095-1109.

Gobakken, T., Bollandsås, O. M. & Næsset, E. (2015). Comparing biophysical forest characteristics estimated from photogrammetric matching of aerial images and airborne laser scanning data. *Scandinavian Journal of Forest Research*, 30 (1): 73-86.

Goetz, S. & Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Management*, 2 (3): 231-244.

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+. *Environmental Research Letters*, 10 (12): 123001.

Government of Malawi. (2001). Malawi's national forestry programme: priorities for improving forestry and livelihoods. Lilongwe: Department of Forestry, NFP co-ordinating unit. 68 pp.

Government of Malawi. (2010). Malawi state of environment and outlook report. Lilongwe, Malawi: Ministry of Natural Resources, Energy and Environment. 302 pp.

Government of Malawi. (2012). Forest resource mapping project under the Japanese grant for the forest preservation programme to the Republic of Malawi: Ministry of Environment and Climate Change Management. 483 pp.

Government of Malawi. (2015). Government of Malawi REDD+ Action Plan 2014-2019. Lilongwe, Malawi: Department of Forestry, Ministry of Environment and Climate Change Management. 23 pp.

Hansen, E., Gobakken, T., Bollandsås, O., Zahabu, E. & Næsset, E. (2015). Modeling aboveground biomass in dense tropical submontane rainforest using airborne laser scanner data. *Remote Sensing*, 7: 788-807.

Hawthorne, S. & Boissière, M. (2014). *Literature review of participatory measurement, reporting and verification (PMRV),* vol. 152. Bogor, Indonesia: CIFOR.

Henry, M., Picard, N., Trotta, C., Manlay, R., Valentini, R., Bernoux, M. & Saint-André, L. (2011). Estimating tree biomass of sub-Saharan African forests: a review of available allometric equations. *Silva Fennica*, 45 (3B): 477–569.

Hofton, M., Dubayah, R., Blair, J. & Rabine, D. (2006). Validation of SRTM elevations over vegetated and non-vegetated terrain using medium footprint Lidar. *Photogrammetric Engineering & Remote Sensing*, 72 (3): 279-285.

Kajembe, G., Silayo, D. & Vatn, A. (2015). The adaptation of REDD+ initiatives in forest management regimes in two pilot projects in Kondoa and Kilosa districts, Tanzania. Ås, Norway: Norwegian University of Life Sciences (NMBU). 40 pp.

Karwel, A. & Ewiak, I. (2008). *Silk Road for Information from Imagery*. Estimation of the accuracy of the SRTM terrain model on the area of Poland, Beijing, China: International Society for Photogrammetry and Remote Sensing. 169-172 pp.

Keller, M., Palace, M. & Hurtt, G. (2001). Biomass estimation in the Tapajos National Forest, Brazil: Examination of sampling and allometric uncertainties. *Forest Ecology and Management*, 154 (3): 371-382.

Kissinger, G., Herold, M. & Sy, V. D. (2012). Drivers of deforestation and forest degradation: a synthesis report for REDD+ policymakers. Lexeme Consulting, Vancouver Canada.

Kouba, J. (2015). A guide to using international GNSS service (IGS) products. In *Geodetic Survey Division*. Available at: ftp://igscb.jpl.nasa.gov/pub/resource/pubs/UsingIGSProductsVer21.pdf (accessed: 06/10/2015).

Kumar, L., Sinha, P., Taylor, S. & Alqurashi, A. (2015). Review of the use of remote sensing for biomass estimation to support renewable energy generation. *Journal of Applied Remote Sensing*, 9 (1): 1-29.

Kuyah, S., Sileshi, G. & Rosenstock, T. (2016). Allometric models based on bayesian frameworks give better estimates of aboveground biomass in the miombo Woodlands. *Forests*, 7 (2): 1-13.

Levick, S. R., Hessenmöller, D. & Schulze, E. D. (2016). Scaling wood volume estimates from inventory plots to landscapes with airborne LiDAR in temperate deciduous forest. *Carbon Balance and Management*, 11 (7): 1-14.

Luoga, E., Kajembe, G., Shemweta, D., Zahabu, E., Mwaipopo, C. & Kweka, D. (2005). Assessment of tree stocking and diversity for joint forest management (JFM) in Nkweshoo village forest management area, Kilimanjaro, Tanzania. *Forests, Trees and Livelihoods*, 15 (3): 259-273.

Magnussen, S. & Carillo, N. (2015). Model errors in tree biomass estimates computed with an approximation to a missing covariance matrix. *Carbon Balance and Management*, 10 (21): 1-14.

Magnussen, S., Næsset, E. & Gobakken, T. (2015). LiDAR-supported estimation of change in forest biomass with time-invariant regression models. *Canadian Journal of Forest Research*, 45 (11): 1514-1523.

Malimbwi, R., Solberg, B. & Luoga, E. (1994). Estimation of biomass and volume in miombo woodland at Kitulangalo forest reserve, Tanzania. *Journal of Tropical Forest Science*, 7 (2): 230-242.

Malimbwi, R., Eid, T. & Chamshama, S. (2016). *Allometric tree biomass and volume models in Tanzania*. Morogoro, Tanzania: Department of Forest Mensuration and Management, Sokoine University of Agriculture. 129 pp.

Mascaro, J., Detto, M., Asner, G. P. & Muller-Landau, H. C. (2011). Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment*, 115 (12): 3770-3774.

Mate, R. (2014). *Biomass and volume estimates for valuable timber species in Mozambique*. PhD Thesis. Uppsala: Swedish University of Agricultural Sciences, Department of Energy and Technology. 58 pp.

Mate, R., Johansson, T. & Sitoe, A. (2014). Biomass equations for tropical forest tree species in Mozambique. *Forests*, 5 (3): 535-553.

Mauya, E., Ene, L., Bollandsås, O., Gobakken, T., Næsset, E., Malimbwi, R. & Zahabu, E. (2015a). Modelling aboveground forest biomass using airborne laser scanner data in the miombo woodlands of Tanzania. *Carbon Balance and Management*, 10 (28): 1-16.

Mauya, E., Hansen, E., Gobakken, T., Bollandsås, O., Malimbwi, R. & Næsset, E. (2015b). Effects of field plot size on prediction accuracy of aboveground biomass in airborne laser scanning-assisted inventories in tropical rain forests of Tanzania. *Carbon Balance and Management*, 10 (10): 1-14.

Mauya, E. W., Mugasha, W. A., Zahabu, E., Bollandsås, O. M. & Eid, T. (2014). Models for estimation of tree volume in the miombo woodlands of Tanzania. *Southern Forests: a Journal of Forest Science*, 76 (4): 209-219.

McRoberts, R., Andersen, H.-E. & Naesset, E. (2014). Using airborne laser scanning data to support forest sample surveys. Forestry applications of airborne laser scanning. In *Forestry applications of airborne laser scanning--concepts and case studies*, pp. 269 - 292. Dordrecht: Springer.

McRoberts, R. E., Næsset, E., Gobakken, T. & Bollandsås, O. M. (2015). Indirect and direct estimation of forest biomass change using forest inventory and airborne laser scanning data. *Remote Sensing of Environment*, 164: 36-42.

McRoberts, R. E. & Westfall, J. A. (2016). Propagating uncertainty through individual tree volume model predictions to large-area volume estimates. *Annals of Forest Science*, 73 (3): 625-633.

Miles, L., Newton, A. C., DeFries, R. S., Ravilious, C., May, I., Blyth, S., Kapos, V. & Gordon, J. E. (2006). A global overview of the conservation status of tropical dry forests. *Journal of Biogeography*, 33 (3): 491-505.

Molto, Q., Rossi, V. & Blanc, L. (2013). Error propagation in biomass estimation in tropical forests. *Methods in Ecology and Evolution*, 4 (2): 175-183.

Mugasha, W., Eid, T., Bollandsås, O., Malimbwi, R., Chamshama, S., Zahabu, E. & Katani, J. (2013). Allometric models for prediction of above- and belowground biomass of

trees in the miombo woodlands of Tanzania. *Forest Ecology and Management*, 310: 87-101.

Mugasha, W., Bollandsås, O., Eid, T., Gobakken, T., Zahabu, E. & Katani, J. (2016a). Decision-support tool for management of miombo woodlands - a matrix model approach. *Southern Forests*, Accepted.

Mugasha, W., Eid, T., Bollandsås, O. & Mbwambo, L. (2016b). Modelling diameter growth, mortality and recruitment of trees in miombo woodlands in Tanzania. *Southern Forests*, Accepted.

Mwakalukwa, E. E., Meilby, H. & Treue, T. (2014). Volume and aboveground biomass models for dry miombo woodland in Tanzania. *International Journal of Forestry Research*, 2014: 11.

Mwase, W. F., Bjørnstad, Å., Bokosi, J. M., Kwapata, M. B. & Stedje, B. (2007). The role of land tenure in conservation of tree and shrub species diversity in miombo woodlands of southern Malawi. *New Forests*, 33 (3): 297-307.

Næsset, E., Gobakken, T., Solberg, S., Gregoire, T. G., Nelson, R., Ståhl, G. & Weydahl, D. (2011). Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: A case study from a boreal forest area. *Remote Sensing of Environment*, 115 (12): 3599–3614.

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.

Puliti, S., Ørka, H., Gobakken, T. & Næsset, E. (2015). Inventory of small forest areas using an unmanned aerial system. *Remote Sensing*, 7 (8): 9632-9654.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ramananantoandro, T., Rafidimanantsoa, H. P. & Ramanakoto, M. F. (2015). Forest aboveground biomass estimates in a tropical rainforest in Madagascar: new insights from the use of wood specific gravity data. *Journal of Forestry Research*, 26 (1): 47-55.

Rodríguez, E., Morris, C. & Belz, J. (2006). A Global Assessment of the SRTM Performance. *Photogrammetric Engineering & Remote Sensing*, 72 (3): 249–260.

Ryan, C. M. & Williams, M. (2011). How does fire intensity and frequency affect miombo woodland tree populations and biomass? *Ecological Applications*, 21 (1): 48-60.

Ryan, C. M., Williams, M. & Grace, J. (2011). Above- and belowground carbon stocks in a miombo woodland landscape of Mozambique. *Biotropica*, 43 (4): 423-432.

Ryan, C. M., Pritchard, R., McNicol, I., Owen, M., Fisher, J. A. & Lehmann, C. (2016). Ecosystem services from southern African woodlands and their future under global change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371 (1703): 1-16.

Saatchi, S., Marlier, M., Chazdon, R. L., Clark, D. B. & Russell, A. E. (2011). Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. *Remote Sensing of Environment*, 115 (11): 2836-2849.

SAS Institute. (2012). *SAS 9.4*. Level 1. Cary: SAS Institute.

Schnell, S., Altrell, D., Ståhl, G. & Kleinn, C. (2014). The contribution of trees outside forests to national tree biomass and carbon stocks—a comparative study across three continents. *Environmental Monitoring and Assessment*, 187 (1): 1-18.

Sensefly. (2015). eBee sensefly: Extended user manual. 176. Available at: https://www.sensefly.com/fileadmin/user_upload/sensefly/documents/manuals/Extended_User_Manual_eBee_and_eBee_Ag_v16.pdf. (accessed: 20/08/2015).

Strunk, J., Temesgen, H., Andersen, H.-E., Flewelling, J. P. & Madsen, L. (2012). Effects of lidar pulse density and sample size on a model-assisted approach to estimate forest inventory variables. *Canadian Journal of Remote Sensing*, 38 (5): 644-654.

Särndal, C. E., Swensson, B. & Wretman, J. (1992). *Model assisted survey sampling*. Springer series in statistics. New york: Springer-Verlag, inc. 695 pp.

Takasu, T. (2009). *RTKLIB: Open source program package for RTK-GPS*. Free and open source software for Geospatial Sydney: Open Source Geospatial Foundation (OSGeo).

Tang, L. & Shao, G. (2015). Drone remote sensing for forestry research and practices. *Journal of Forestry Research*, 26 (4): 791-797.

Tarimo, B., Dick, O. B., Gobakken, T. & Totland, O. (2015). Spatial distribution of temporal dynamics in anthropogenic fires in miombo savanna woodlands of Tanzania. *Carbon Balance and Management*, 10 (18): 1-15.

UNFCCC. (2014). Key decisions relevant for reducing emissions from deforestation and forest degradation in developing countries (REDD+): Decision booklet REDD+: UNFCCC secretariat. 44 pp.

Vashum, K. & Jayakumar, S. (2012). Methods to estimate above-ground biomass and carbon stock in natural forests - A review. *Ecosystem and Ecography*, 2 (4): 1-7.

Williams, M., Ryan, C. M., Rees, R. M., Sambane, E., Fernando, J. & Grace, J. (2008). Carbon sequestration and biodiversity of re-growing miombo woodlands in Mozambique. *Forest Ecology and Management*, 254 (2): 145-155.

Zahabu, E. (2008). *Sinks and sources : a strategy to involve forest communities in Tanzania in global climate policy*. PhD Thesis. Enschede: University of Twente. 249 pp.

Zanne, A. E., Lopez-Gonzalez, G., Coomes, D. A., Ilic, J., Jansen, S., Lewis, S. L., Miller, R. B., Swenson, N. G., Wiemann, M. C. & Chave, J. (2009). *Data from: Towards a worldwide wood economics spectrum*: Dryad Data Repository.

Paper I

Southern Forests 2016, 78(1): 41–51
Printed in South Africa — All rights reserved

SOUTHERN FORESTS
ISSN 2070-2620 EISSN 2070-2639
http://dx.doi.org/10.2989/20702620.2015.1108615

Total tree, merchantable stem and branch volume models for miombo woodlands of Malawi

Daud J Kachamba* and Tron Eid

Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway
** Corresponding author, email: dkachamba@gmail.com*

The objective of this study was to develop general (multispecies) models for prediction of total tree, merchantable stem and branch volume including options with diameter at breast height (dbh) only, and with both dbh and total tree height (ht), as independent variables. The modelling data set was based on destructively sampled trees and comprised 74 trees from 33 tree species, collected from four forest reserves located in different ecological zones of Malawi. The dbh and ht ranges for the data set were 5.3–111.2 cm and 3.0–25.0 m, respectively. A number of alternative model forms were tested and the final model selection was based on root mean square error (RMSE) values calculated using a leave-one-out cross-validation procedure. The model performances and the evaluations of the finally selected models (R^2 range 0.72 to 0.92; RMSE range 38% to 71%; mean prediction errors range −1.4% to 1.3%) suggest that all models can be used over a wide range of geographical and ecological conditions in Malawi with an appropriate accuracy in predictions. The appropriateness of the developed models was also supported by the fact that the mean prediction errors of these models were much lower than the mean prediction errors (range −23.6% to 48.9%) of some previously developed models tested on our data.

Keywords: cross-validation, destructive sampling, general volume models

Introduction

Miombo woodlands are dry tropical forests dominated by deciduous trees belonging to the genera *Brachystegia*, *Julbernadia* and *Isoberlinia* that occur in a climate with a dry season of three months or more (Chidumayo and Gumbo 2010). These woodlands cover an area of approximately 2.7 million km^2 that spans 10 countries in East and Central Africa, including Malawi (Abbot et al. 1997; Abdallah and Monela 2007; Ryan et al. 2011). Miombo woodlands provide a wide variety of products, including fruits, fish, bush meat, edible insects, beeswax, honey and traditional medicines, and services such as biodiversity and watershed conservation.

In Malawi, miombo woodlands constitute 92.4% of the country's total forest area (Government of Malawi 2010a). They are mainly located in forest and game reserves, which were exclusively established for water catchment, and soil and biodiversity conservation (Abbot et al. 1997; Government of Malawi 2012). Any sustainable management approach for these woodlands requires the availability of proper methods, models and tools to assist forest managers. A basic prerequisite for this is the availability of appropriate allometric models for predicting tree volumes (Akindele and LeMay 2006; Guendehou et al. 2012; Mauya et al. 2014). Quantification of forest volume is, for example, important for establishing the growing stock of stands and forests, for timber valuation, for selection of forest areas in harvest scheduling, for growth and yield studies, and as a basis for estimation of biomass and carbon stocks (see Chidumayo 1988; Lowore et al. 1994; Chamshama et al. 2004; Hofstad 2005; Adekunle 2007; Picard et al. 2012; Adekunle et al. 2013).

Volume of individual trees is usually determined using models based on diameter at breast height (dbh) and total tree height (ht), and in some cases measures of tree form, as independent variables (see Clutter et al. 1983; Hofstad 2005). Reviews by Henry et al. (2011, 2013) showed that several models for predicting tree volume in miombo woodlands have been developed previously. Based on these reviews and a literature search we have been able to identify tree volume models for miombo woodlands developed by Malimbwi et al. (1994), Chamshama et al. (2004), Mauya et al. (2014) and Mwakalukwa et al. (2014a) in Tanzania, by Mate (2014) in Mozambique and by Chidumayo (1988) in Zambia. The only existing tree volume models for miombo woodlands in Malawi were developed by Abbot et al. (1997).

Many of the previously developed models are species-specific (see Chidumayo 1988; Abbot et al. 1997; Mate 2014; Mwakalukwa et al. 2014a). Recent studies from neighbouring countries have shown that the number of tree species in miombo woodlands may range from 80 to 300 (Dewees et al. 2011; Giliba et al. 2011; Kalaba et al. 2012; Mugasha et al. 2013; Mwakalukwa et al. 2014a). For miombo woodlands in Malawi, Mwase et al. (2007) found 48 tree species and shrubs representing 24 families based on a sample of 6.5 ha, whereas Missanjo et al. (2014) found 22 species on 0.5 ha. Given such high tree species numbers, the applicability of species-specific models is

limited in these woodlands, and general models (covering multiple species) are therefore the most useful alternative in most cases. Only a few general volume estimation models have been developed (Abbot et al. 1997; Mauya et al. 2014; Mwakalukwa et al. 2014a) and the applicability of these is limited by several factors in the data used for modelling: (1) narrow geographical ranges of study sites, (2) relatively small dbh ranges and (3) relatively few tree species. The modelling data of the general total tree volume model developed by Abbot et al. (1997) for Malawi were limited to nine timber tree species, a small dbh range (5–35 cm), and a relatively limited geographical range, i.e. the sites were located in the central and southern parts of Malawi, thus excluding the northern part, which accommodates approximately 25% of miombo woodlands located in forest reserves in the country (Government of Malawi 2010a).

All identified studies on volume models for miombo woodlands, except Mate (2014), include total tree volume models. In addition, all of the studies developed stem volume models while branch volume models were developed by Malimbwi et al. (1994), Chamshama et al. (2004) and Mauya et al. (2014). The stem volume models are usually motivated by timber production and comprise the merchantable parts of stems. Large-scale commercial extraction of timber from miombo woodlands is no longer taking place in Malawi (Abbot et al. 1997; Luhanga 2009) but the Malawian government uses a licensing system that permits the issuance of permits to individuals for accessing timber (Government of Malawi 2010b). In this context general stem volume models would be useful. The same applies to cases where trees are being cleared for infrastructure development, such as roads, railways and buildings, and compensation payments to tree/forest owners are necessary. General branch volume models may be useful tools for assessing quantities related to brick burning, domestic fuelwood, production of charcoal and construction poles.

The main objective of this study was therefore to develop general total tree volume models, as well as general tree sectional models for branches and merchantable stems, which can be applicable across the entire distribution of miombo woodlands in Malawi. The models were based on data from different ecological zones covering all three regions in Malawi. The performance of the models developed in this study was also compared with previously developed models from Malawi and neighbouring countries.

Materials and methods

Site description

Given that tree growth is dependent on species, genetics, climate, soil and management (Chidumayo and Gumbo 2010), the selection of study sites was based on geographical location, management regime and ecological zones. Among the four selected sites, Mtangatanga is located in Mzimba District in northern Malawi, Kongwe and Mua-livulezi are located in Dowa and Dedza districts, respectively, in central Malawi, and Tsamba is located in Neno District in southern Malawi (see Figure 1, Table 1).

Data collection

Information on ranges in tree size and species distribution is important for the selection of sample trees to be used in modelling (Mugasha et al. 2013). We performed systematic sample plot inventories covering each site. Circular plots, with a radius of 11.28 m (400 m^2), were chosen. For each plot, we measured dbh for all trees with dbh ≥ 5 cm and also determined their local and scientific names. We conducted inventories comprising a total of 215 sample plots with 70, 30, 64 and 50 plots for Mtangatanga, Kongwe, Mua-livulezi and Tsamba forest reserves, respectively.

For each study site, we divided the sample plot inventory into two phases. The first phase involved covering approximately half of the plots. This data set enabled the selection of initial sample trees for destructive sampling while the forest inventory was still ongoing. After the completion of

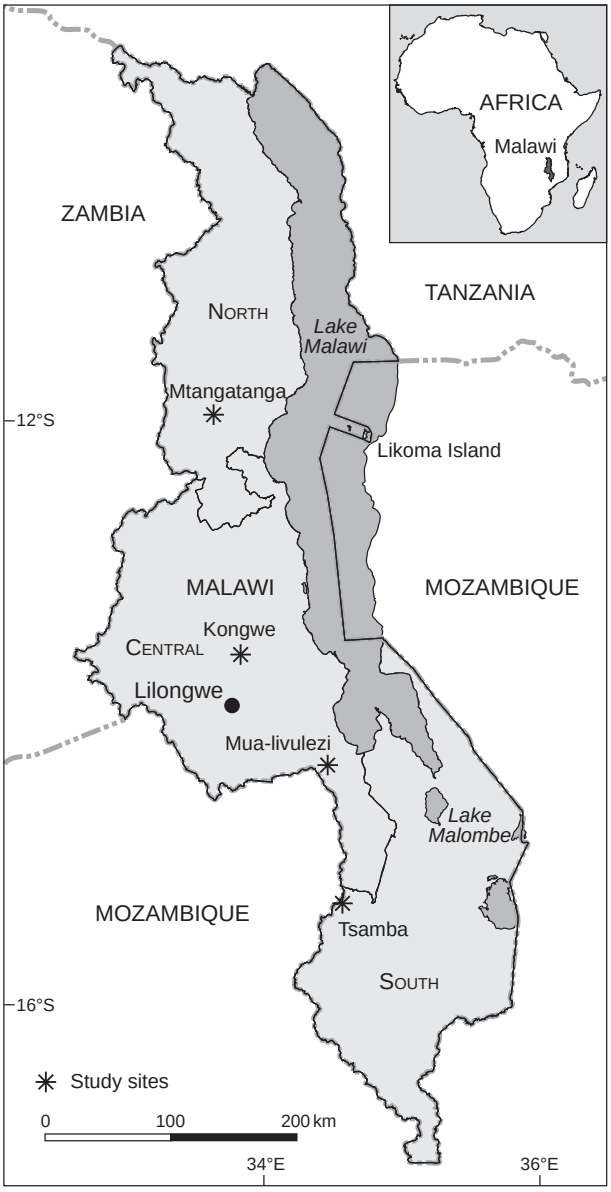


Figure 1: Map of Malawi showing the location of the study sites

the second phase, we selected the remaining sample trees from the complete data set. The maximum dbh based on all sample plots in Mtangatanga, Kongwe, Mua-livulezi and Tsamba was 61, 73, 70 and 56 cm, respectively, and the number of species identified for the respective sites was 66, 45, 75 and 65. In total, for all of the study sites, 139 different species were identified during the inventory. The most frequent species for the respective sites were *Brachystegia spiciformis*, *Diplorhynchus condylocarpon* and *Uapaca kirkiana*.

A total of 74 standing live trees were sampled from the four sites based on the observed dbh and tree species frequency for each site (see Table 2). A deliberate effort was made to ensure that the trees were selected from the entire range of dbh classes observed in the sample plot inventories. We also selected eight trees with larger diameters than those observed in the sample plot inventories to improve predictions of volumes for very large trees. The eight most frequent tree species for each site were always represented according to quantities. The remaining trees were selected randomly within the study site among the remaining tree species. In total, 33 tree species were selected, comprising 10, 10, 12 and 10 different tree species in Mtangatanga, Kongwe, Mua-livulezi and Tsamba, respectively.

Before felling the sampled trees, we measured their dbh (over bark) and ht and also determined their scientific and local names. To identify scientific and local names of tree species, we used an experienced botanist from the Forest Research Institute of Malawi and a local forest guard from each of the study sites. To measure dbh, a caliper or a diameter tape was used depending on tree sizes, and a Suunto hypsometer was used for all ht measurements. A statistical summary of the sample trees is given in Table 2 and information on individual trees is given in the Appendix.

After felling, the trees were separated into two aboveground sections, i.e. merchantable stem and branches. The merchantable stem section comprised the stem from 30 cm aboveground to the point where the first major branches started (see Abbot et al. 1997). The branches comprised all parts of the tree above the defined stem and up to a minimum diameter of 2.5 cm. All branches with a diameter less than 2.5 cm were considered as twigs and therefore not included in the volume. For small trees not considered as suitable for timber production (dbh $\leq$ 15 cm, in total 14 trees), stem volumes were allocated to branches (see Mauya et al. 2014).

We crosscut the stems and branches into manageable pieces (logs), whose lengths ranged from 1 to 3.7 m, to facilitate volume measurements. We then measured the lengths and the mid-diameters over bark of each log. We calculated individual log volumes by multiplying the basal area of the mid-section of each log by its length. Subsequently, we determined the stem and branch volumes for each tree by summing all individual log volumes. We finally determined the total tree volumes by summing merchantable stem and branches volumes. A statistical summary of all volumes is given in Table 3 and total tree volumes over dbh and ht for each site are displayed in Figure 2.

Model development and evaluation

The volume of an individual standing tree may be established as:

Table 1: Description of the study sites. Data sources: rainfall and temperature data (1975–2005) from Ministry of Natural Resources, Energy and Mining, Department of Climate Change and Meteorological Services. Soil and silvicultural classification: Hardcastle (1978)

	Mtangatanga	Kongwe	Mua-livulezi	Tsamba
Region	Northern	Central	Central	Southern
District	Mzimba	Dowa	Dedza	Neno
Location	11°56′ S, 33°42′ E	13°35′ S, 33°55′ E	14°21′ S, 34°37′ E	15°21′ S, 34°36′ E
Area (ha)	8 443	1 813	12 147	3 240
Altitude (m)	1 500–1 700	1 000–1 500	400–900	700–1 500
Dominant soil type	Humic ferrallitic	Ferruginous	Lithosols	Ferrallitic
Management regime	Co-management	Government	Co-management	Government
Silvicultural classification	Moist *Brachystegia*	Moist *Brachystegia*	Dry *Brachystegia*	Moist *Brachystegia*
Mean minimum annual temp (°C)	6	6	13	8
Mean maximum annual temp (°C)	29	29	32	28
Total annual rainfall range (mm)	960–1 050	960–1 050	840–960	1 200–1 600
Rain period	December–April	November–April	November–April	November–April
Dry months	May–November	May–October	May–October	May–October

Table 2: Statistical summary of sample trees. dbh = diameter at breast height, ht = total tree height

Site	dbh (cm)					ht (m)			
	n	Mean	Minimum	Maximum	SD	Mean	Minimum	Maximum	SD
Mtangatanga	20	35.5	6.0	111.2	26.7	10.7	4.0	18.0	4.3
Kongwe	18	34.9	9.0	75.7	19.6	11.7	5.0	22.0	4.7
Mua-livulezi	18	32.8	5.3	81.7	23.0	11.6	3.0	22.0	5.8
Tsamba	18	30.2	8.4	75.0	17.5	12.6	6.5	25.0	5.1
All	74	33.4	5.3	111.2	21.8	11.6	3.0	25.0	4.9

$$V = f(\text{dbh}, \text{ht}, f)$$

where V = total tree volume, dbh = diameter at breast height, ht = total tree height, and f = measure of stem form (Clutter at al. 1983). Since measures of stem form are difficult and expensive to obtain from field measurements (van Laar and Akça 2007), volume is usually predicted from models either with dbh only or with both dbh and ht as independent variables. We decided to develop both types in order to facilitate application of the models also when height measurements are not available from the forest inventory.

Many linear and non-linear allometric model forms were first considered based on a review of models previously developed for miombo woodlands (Chidumayo 1988; Malimbwi et al. 1994; Abbot et al. 1997; Chamshama et al. 2004; Munishi et al. 2010). We finally selected some of the most frequently used model forms for fitting and further evaluations:

$$V = a + b*(\text{dbh})^2$$
$$V = a + b*\text{dbh} + c*(\text{dbh})^2$$
$$V = a*(\text{dbh})^b$$
$$V = a*(\text{dbh}^2*\text{ht})^b$$
$$V = a*(\text{dbh})^b*(\text{ht})^c$$
$$V = a + b*(\text{dbh})^2 + c*\text{ht}$$

where V is total tree, branch or merchantable stem volume over bark (m^3), dbh is the diameter at breast height (cm), ht is total tree height (m), and a, b and c are parameter estimates.

Before fitting the models, we assessed scatter plots of volumes over dbh. As expected, the plots indicated non-linear patterns of the relationship between total volume and dbh (see Figure 2). Non-linear models were initially fitted with PROC MODEL of SAS 9.4 software (SAS Institute, Cary, NC, USA) using non-linear ordinary least-squares estimation. Residual plots indicated that the data required transformation because they violated the assumptions of equal variance (homoscedasticity) and normality (Picard et al. 2012). The most common methods of data transformations in biological studies, e.g. log transformation and weighted regression, were therefore tested (Parresol 1999; Picard et al. 2012; Mauya et al. 2014; Sileshi 2014). However, none of the methods produced satisfactory results. We therefore applied generalised methods of moments (GMM) estimation, a semi-parametric estimation method. According to SAS Institute (2012), GMM estimation can be used if the error variance relationship is unknown to improve the efficiency of the parameter estimation in the presence of the heteroscedastic errors. To achieve an optimal minimisation of the root mean square error (RMSE) of the fitted models, a wide range of initial instrumental values for the models were used.

Model selection was based on the results from a leave-one-out cross-validation procedure (James et al. 2013; Sileshi 2014). The leave-one-out cross-validation procedure involves splitting the data set of n observations into two parts, namely a training data set and validation data set. The validation data set comprises one single observation (x_1, y_1) and the training data set comprises the remaining $\{(x_2, y_2),\ldots(x_n, y_n)\}$ observations. The model is fitted on the $n - 1$ observations in the training data set and a prediction $\hat{y}$

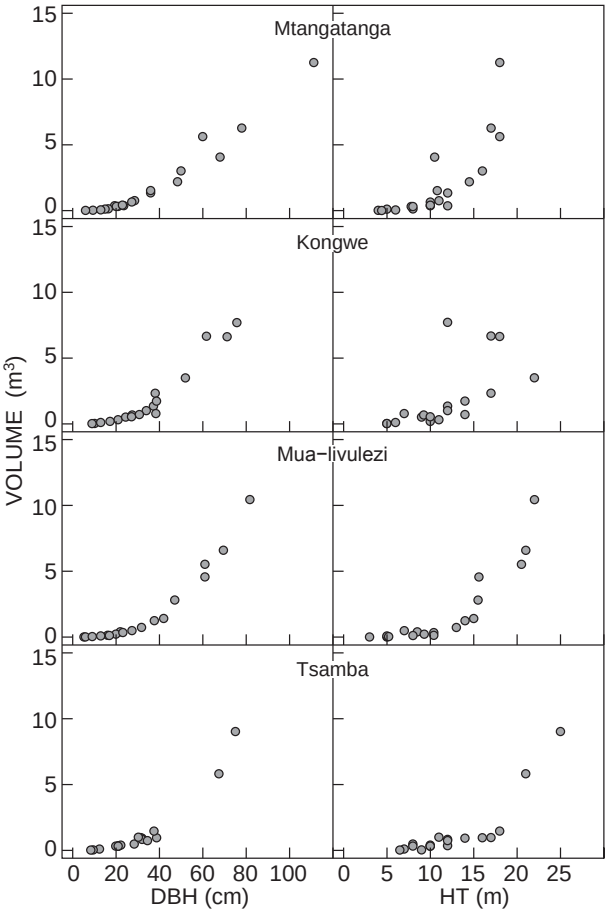


Figure 2: Total tree volume (m^3) distribution over diameter at breast height (DBH) and total tree height (ht) for Mtangatanga, Kongwe, Mua-livulezi and Tsamba forest reserves

Table 3: Statistical summary of total tree, branches and merchantable stem volumes

Site	Total tree volume (m^3)					Branches volume (m^3)					Merchantable stem volume (m^3)				
	n	Mean	Min.	Max.	SD	*n*	Mean	Min.	Max.	SD	*n*	Mean	Min.	Max.	SD
Mtangatanga	20	1.944	0.015	11.254	2.888	20	1.347	0.015	7.437	2.055	16	0.746	0.077	3.817	0.938
Kongwe	18	1.935	0.024	7.698	2.497	18	1.404	0.024	6.238	1.879	15	0.637	0.080	2.209	0.668
Mua-livulezi	18	1.955	0.009	10.442	2.948	18	1.190	0.009	6.571	1.846	14	0.983	0.077	3.871	1.222
Tsamba	18	1.340	0.025	9.027	2.313	18	0.706	0.025	4.581	1.233	15	0.761	0.153	4.446	1.163
Total	74	1.797	0.009	11.254	2.640	74	1.167	0.009	7.437	1.773	60	0.778	0.077	4.446	0.998

is made for the single observation in the validation data set, using its value x_1. Since (x_1, y_1) was not used in the fitting process, square error (SE) = $(y_1 - \hat{y})^2$ provides an estimate for the test error. This procedure is repeated n times and thus producing n test errors, SE_1,…SE_n. The leave-one-out-cross-validation estimate for the test error is the mean of these n test error estimates (MSE).

The cross-validation results were then used to calculate RMSE and mean prediction errors (MPE) for the models. RMSE and MPE were calculated as follows:

$$\text{RMSE} = \sqrt{\text{MSE}}$$

$$\text{RMSE}\,(\%) = \frac{\text{RMSE}}{\bar{y}} \times 100$$

$$\text{MPE} = \sum_{i=1}^{n} \frac{\left(y_i - \hat{y}_i\right)}{n}$$

$$\text{MPE}\,(\%) = \frac{\text{MPE}}{\bar{y}} \times 100$$

where n is the number of sample trees, y_1 is the observed volume of tree i, $\hat{y}_i$ is the predicted volume of tree i, and $\bar{y}$ is the mean observed volume. The RMSE gives an estimate of the predictive power of the model, and the RMSE (%) measures the proportion of the RMSE relative to the mean observed volume. The MPE measures the mean prediction error of the model (and possibly bias if significantly different from zero) and the MPE (%) measures the proportion of the MPE relative to the mean observed volume.

The final model selection was based on RMSE values. Models with insignificant parameter estimates, however, were not considered for selection, irrespective of RMSE values. For all models, we also displayed pseudo-R^2 and MPE values. Student's t-tests were done to determine whether the MPE values were significantly different from zero.

We also tested a number of previously developed general total tree volume models for miombo woodlands on our data set (Table 4). Some of these models (Abbot et al. 1997; Mauya et al. 2014) were also examined by graphical plots that displayed the behaviour of each model over a wide range of diameters (also outside their respective model data ranges). For the model with ht as the independent variable (see Table 4; Mauya et al. 2014), we applied a height–diameter model developed from our sample trees:

$$\text{ht} = 1.3 + \exp(3.787685 - 6.62809 * \text{dbh}^{-0.45222}).$$

Results

Parameter estimates and fit statistics of different candidate general total tree, branches and merchantable stem volume models are presented in Table 5. The main criterion for selecting models was RMSE (lowest). However, the models with non-significant ($p > 0.05$) parameter estimates were not considered for selection. Based on these criteria, model 1 was selected among the total tree volume models with dbh only as the independent variable, whereas model 4 was selected among those with both dbh and ht as independent variables. For branch volume, models 1 and 4 were selected. For the merchantable stem volume, model 1 was selected among those with dbh only, whereas model 5 was selected among those with both dbh and ht as independent variables.

The performance of the selected general total tree volume models was tested over the study sites (see Table 6). For the models with dbh only as an independent variable, no MPE values were significantly different from zero ($p > 0.05$), whereas for the models with both dbh and ht as independent variables, a significant MPE value ($p < 0.05$) was found for Tsamba. The mean standard deviations of the difference between observed and predicted values varied between sites from 18.7% to 47.7%.

We also fitted site-specific total tree models and selected, among these, based on RMSE and significance of parameter estimates (Table 7). For Mtangatanga, Kongwe, Mua-livulezi and Tsamba, we selected models 1, 1, 3 and 3, respectively, among the models with dbh as the only independent variable. Among the models with both dbh and ht, we selected models 4, 5, 4 and 4, for Mtangatanga, Kongwe, Mua-livulezi, and Tsamba, respectively.

Table 8 shows the results when testing the previously developed general total tree volume models on our data set. The model developed by Abbot et al. (1997) on average gave 23.6% higher total volume, whereas that of Mauya et al. (2014), with dbh only as the independent variable, on average gave 8.1% lower total volume than observed in our data set. None of these MPE values were significantly different from zero. All other tested models gave a significant over-estimation ($p < 0.05$) of total volume. Figure 3 displays total volume over dbh for the models developed in this study and for some of the previously developed models.

Table 4: Previously developed models tested on our data. V = total tree volume (m^3), dbh = diameter at breast height (cm), ht = total tree height (m), and a, b and c are parameter estimates

	Model	Authors	n	Number of sites	Number of species	dbh range (cm)	Country
1	$V = 0.0168 + 0.000023*\text{dbh}^3$	Abbot et al. (1997a)	88	4	9	5.0–35.0	Malawi
2	$V = 0.00016*(\text{dbh})^{2.46300}$	Mauya et al. (2014)	158	4	55	1.2–95.0	Tanzania
3	$V = 0.00011*(\text{dbh})^{2.13300}*(\text{ht})^{0.57580}$	Mauya et al. (2014)	158	4	55	1.2–95.0	Tanzania
4	$V = 0.0001(\text{dbh})^{2.032}(\text{ht})^{0.659}$	Malimbwi et al. (1994)	17	1	15	9.3–41.0	Tanzania
5	$V = 0.000048*(\text{dbh})^{1.445}*(\text{ht})^{1.7026}$	Chamshama et al. (2004)	30	1	20	1.0–50.0	Tanzania
6	$V = \exp(-9.5238 + 1.8067*\ln(\text{dbh}) + 1.1940*\ln(\text{ht}))$	Mwakalukwa et al. (2014)	142	1	28	1.4–62.0	Tanzania
7	$V = \exp(-8.4800 + 2.3351*\ln(\text{dbh}))$	Mwakalukwa et al. (2014)	142	1	28	1.4–62.0	Tanzania

Discussion

The data set used for modelling, collected over four sites in different ecological zones spanning all of the three regions of Malawi, captured wide ranges of the natural variability in factors affecting tree growth, such as soil types, temperature and rainfall (see Table 1). The only previously developed volume prediction models for miombo woodlands in Malawi (Abbot et al. 1997) covered different ecological zones well, but lacked data from the northern part of Malawi. The representability and robustness of the developed models is also enhanced by using sample trees selected from an independent systematic sample plot inventory as well as inclusion of a number of large trees to secure appropriate

Table 5: General total tree, branches and merchantable stem volume models. RMSE = root mean square error, MPE = mean prediction error. Selected models are highlighted in bold

Volume section	Model	Parameter estimate			Pseudo-R^2	RMSE		MPE		
		a	*b*	*c*		(m^3)	(%)	(m^3)	(%)	*p*-value
Total tree	**1. $V = a + b*(dbh)^2$**	**−0.111530**	**0.001206**		**0.92**	**0.937**	**52.2**	**−0.020**	**−1.1**	**0.8564**
(*n* = 74)	2. $V = a + b*dbh + c*(dbh)^2$	−0.728670	0.034828*	0.000861	0.92	1.088	60.5	−0.056	−3.1	0.6626
	3. $V = a*(dbh)^b$	0.000487*	2.210493		0.90	1.302	72.5	−0.056	−3.1	0.7142
	4. $V = a*(dbh^2*ht)^b$	**0.000218**	**0.896561**		**0.95**	**0.697**	**38.8**	**−0.008**	**−0.4**	**0.9190**
	5. $V = a*(dbh)^b*(ht)^c$	0.000192	1.909726	0.774122	0.94	1.029	57.3	−0.045	−2.5	0.7095
	6. $V = a + b*(dbh)^2 + c*ht$	−0.956790	0.001051	0.093967	0.93	0.884	49.2	−0.029	−1.6	0.7809
Branches	**1. $V = a + b*(dbh)^2$**	**−0.110130***	**0.000807**		**0.91**	**0.651**	**55.8**	**−0.015**	**−1.3**	**0.8451**
(*n* = 74)	2. $V = a + b*dbh + c*(dbh)^2$	−0.489640*	0.021418*	0.000594	0.91	0.774	66.3	−0.040	−3.4	0.6738
	3. $V = a*(dbh)^b$	0.000235*	2.283238		0.88	0.966	82.8	−0.045	−3.9	0.6944
	4. $V = a*(dbh^2*ht)^b$	**0.000152**	**0.887957**		**0.88**	**0.672**	**57.6**	**0.009**	**0.8**	**0.9051**
	5. $V = a*(dbh)^b*(ht)^c$	0.000148	2.137749	0.379827*	0.89	0.990	84.8	−0.046	−3.9	0.6904
	6. $V = a + b*(dbh)^2 + c*ht$	−0.350150*	0.000763	0.026683*	0.91	0.708	60.7	−0.022	−1.9	0.7879
Merchantable	**1. $V = a + b*(dbh)^2$**	**0.015339***	**0.000396**		**0.75**	**0.553**	**71.1**	**−0.011**	**−1.4**	**0.8791**
stem (*n* = 60)	2. $V = a + b*dbh + c*(dbh)^2$	−0.488650	0.023634	0.000180*	0.76	0.581	74.7	−0.027	−3.5	0.7212
	3. $V = a*(dbh)^b$	0.000361*	2.027542		0.75	0.582	74.8	−0.017	−2.2	0.8254
	4. $V = a*(dbh^2*ht)^b$	0.000121*	0.853676		0.89	0.411	52.8	0.007	0.9	0.8896
	5. $V = a*(dbh)^b*(ht)^c$	**0.000057**	**1.437605**	**1.517811**	**0.93**	**0.295**	**37.9**	**0.010**	**1.3**	**0.7992**
	6. $V = a + b*(dbh)^2 + c*ht$	−0.985480	0.000271	0.095167	0.84	0.427	54.9	−0.001	−0.1	0.9929

* Non-significant parameter estimate at $\alpha = 0.05$ significance level

Table 6: Performance of the selected general total tree models over sites. MPE = mean prediction error

Model	Site	*n*	Observed (m^3)	Predicted (m^3)	MPE			SD	
					(m^3)	(%)	*p*-value	(m^3)	(%)
General with	Mtangatanga	20	1.944	2.226	−0.283	−14.6	0.1865	0.923	47.5
dbh only	Kongwe	18	1.935	1.793	0.142	7.3	0.3711	0.655	33.9
	Mua-livulezi	18	1.955	1.791	0.164	8.4	0.3488	0.720	36.8
	Tsamba	18	1.340	1.332	0.008	0.6	0.9606	0.640	47.7
	All	74	1.797	1.798	−0.000	−0.0	0.9969	0.756	42.1
General with	Mtangatanga	20	1.944	2.006	−0.062	−3.2	0.6659	0.632	32.5
dbh and ht	Kongwe	18	1.935	1.672	0.263	13.6	0.2260	0.887	45.9
	Mua-livulezi	18	1.955	1.938	0.017	0.9	0.8442	0.366	18.7
	Tsamba	18	1.340	1.527	−0.187*	−14.0	0.0062	0.255	19.0
	All	74	1.797	1.791	0.006	0.3	0.9340	0.600	33.4

* Significantly different from zero ($p < 0.05$)

Table 7: Site-specific total tree volume models. Selected models are highlighted in bold. RMSE = root mean square error, MPE = mean prediction error

Model	Site	Parameter estimate			Pseudo-R^2	RMSE		MPE		
		a	*b*	*c*		(m^3)	(%)	(m^3)	(%)	*p*-value
1. $V = a + b*(dbh)^2$	Mtangatanga	0.088649	0.000957		0.96	0.749	38.5	−0.074	−3.8	0.6688
1. $V = a + b*(dbh)^2$	Kongwe	−0.364420	0.001456		0.96	0.570	29.5	−0.011	−0.6	0.9357
3. $V = a*(dbh)^b$	Mua-livulezi	0.000073	2.697713		0.99	0.370	18.9	−0.061	−3.1	0.4999
3. $V = a*(dbh)^b$	Tsamba	0.000085	2.665223		0.99	0.454	33.9	0.038	2.8	0.7314
4. $V = a*(dbh^2*ht)^b$	Mtangatanga	0.000501	0.816614		0.98	0.728	37.4	−0.061	−3.1	0.7191
5. $V = a*(dbh)^b*(ht)^c$	Kongwe	0.000089*	2.371170	0.432938	0.96	0.918	47.4	0.023	1.2	0.9177
4. $V = a*(dbh^2*ht)^b$	Mua-livulezi	0.000058	1.016152		0.99	0.565	28.9	−0.109	−5.6	0.4292
4. $V = a*(dbh^2*ht)^b$	Tsamba	0.000056	1.010726		0.99	0.219	16.3	0.019	1.4	0.7257

* Non-significant parameter estimate at $\alpha = 0.05$ significance level

predictions of volume for larger trees (see also Figure 3). For mature tropical forests this is of particular importance because large trees often account for very large parts of the biomass (Pinard and Putz 1996; Zahabu 2008).

Previous studies have revealed that tree species richness is high in miombo woodlands, with numbers ranging from 80 to 300 (Dewees et al. 2011; Giliba et al. 2011; Kalaba et al. 2012; Mugasha et al. 2013; Mwakalukwa et al. 2014b). The total number of tree species identified during the forest inventories for our sites was 139 and the number included in the modelling data set was 33 (24% of the total). This proportion is, of course, relatively low when considering the presumably high tree species diversity among the species found in these woodlands. There is, however, no reason to believe that our tree selection procedures caused any biases in the models because we first emphasised the inclusion of the most abundant species and then selected randomly among the remaining ones (see Appendix). Such biases, however, are much more likely to appear in the general total tree volume model developed by Abbot et al. (1997) because only species defined as timber species were included in the modelling data set.

The developed models offer options for forest inventory scenarios in which data on dbh only or both dbh and ht are available. None of the selected general total tree models had MPE values significantly different from zero (see Table 5). Not surprisingly, the general performance criteria (RMSE and MPE) of the models with both dbh and ht as independent variables were better than those of the models with dbh only. Still much of the variation in volume was explained by dbh alone and the improvement was relatively small when including ht. This result also conforms to previous studies (Abbot et al. 1997; Mauya et al. 2014; Mwakalukwa et al. 2014a).

If dbh and ht of all trees are measured in an inventory, the model including both these variables should, of course, be applied. Height measurements, however, are time consuming. Therefore height measurements in most cases are only done for sample trees, and subsequently height–diameter models are developed to predict heights that can be used as input for a volume model. When applying a height–diameter model, however, we include additional uncertainty related to the volume prediction that may reverse the improvements we get from applying a model with both dbh and ht as independent variables. It is therefore important that the prevalent tree forms are represented among the sample trees used to develop the height–diameter model. It is also important that we are conscious of the challenges in measuring height in closed-canopy forests, where differences in crown shapes, inaccurate sights of tree tops and steep terrain (Hunter et al. 2013; Larjavaara and Muller-Landau 2013) may lead to large height measurement errors. With appropriate sample trees and small measurement errors, the accuracy of the volume predictions will probably be improved by including ht as independent variable, in spite of the uncertainty added by using a height–diameter model.

When our selected general total tree volume models were tested on our data set over different sites, none of the MPE values were significantly different from zero, except for the model with both dbh and ht as independent variables in Tsamba, where volume was over-estimated (see Table 6). The MPE values, however, varied from −14.6% to 13.6%. It is, of course, not surprising that we obtain MPE values different from zero (significant or not) when we apply the general model on individual sites. The obvious conclusion is therefore to apply site-specific total tree models (Table 7) for local forest inventories taking place in the respective sites (Abbot et al. 1997; Mauya et al. 2014). The MPE values shown in Table 6 are also an indication of the kind of errors we can expect when we

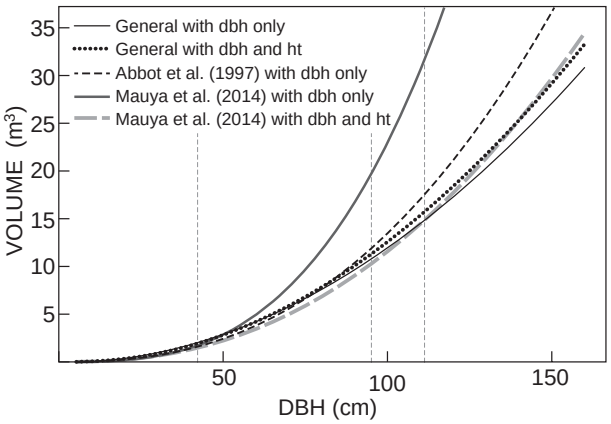


Figure 3: Display of total tree volume over diameter at breast height (dbh) for models developed in this study and previously. For the models with ht included as an independent variable, a height–diameter model developed from our sample trees was applied. Vertical dotted lines are the maximum dbh of the modelling data sets used by Abbot et al. (1997a), Mauya et al. (2014) and in this study (CS), respectively

Table 8: Previously developed total tree volume models tested on our data set. MPE = mean prediction error

Model	Observed	Predicted	MPE			SD	
	(m^3)	(m^3)	(m^3)	(%)	*p*-value	(m^3)	(%)
Abbot et al. (1997a) (dbh only)	1.797	2.221	−0.424	−23.6	0.1489	2.499	139.1
Mauya et al. (2014) (dbh only)	1.797	1.652	0.150	8.1	0.2054	0.981	54.6
Mwakalukwa et al. (2014b) (dbh only)	1.797	1.287	0.510*	28.4	<0.0001	1.000	55.6
Mauya et al. (2014) (dbh and ht)	1.797	1.481	0.316*	17.6	0.0033	0.893	49.7
Malimbwi et al. (1994) (dbh and ht)	1.797	1.126	0.671*	37.4	<0.0001	1.133	63.0
Chamshama et al. (2004) (dbh and ht)	1.797	0.919	0.879*	48.9	<0.0001	1.397	77.7
Mwakalukwa et al. (2014b) (dbh and ht)	1.797	1.482	0.316*	17.6	0.0023	0.859	47.8

* Significantly different from zero ($p < 0.05$)

apply our general model across Malawi (e.g. in general forest inventories such as a National Forestry Inventory). For individual sites there might be quite large deviations from the 'correct' volume while summarised over many sites the deviation will be close to zero.

As previously mentioned, tree section models for stems may be useful when planning commercial extraction of timber, whereas models for branches may be useful for quantification of volume for domestic fuelwood or production of charcoal. All selected tree section models produced MPE values not significantly different from zero (Table 5) indicating an appropriate performance. Generally, the model fit statistics are a bit poorer when the section models are compared with the total tree volume models. Similar results have been observed previously (Dadzie 2013; Mauya et al. 2014) and can probably be attributed to very diverse branching patterns among the trees. The stem volume, for example, is defined as the volume up to where the first large branch protrudes the stem. This volume varies a lot between trees, and thus affects variations in both stems and branches volumes used for modelling.

The results when testing the previously developed models on our data set (Table 8) highlight the dangers of applying models beyond their geographical ranges because a change in geographical site in most cases also mean changes in ecological, climatic and edaphic conditions. All the site-specific (modelling data set collected from one site) total tree volume models developed in Tanzania (Malimbwi et al. 1994; Chamshama et al. 2004; Mwakalukwa et al. 2014b) significantly under-estimated volume (MPE % values varied from 18% to 48%). For the general models (modelling data set collected from one site) developed in Tanzania (Mauya et al. 2014), the under-estimations were somewhat lower (MPE % varied from 8% to 18%), and they were only partly statistically significantly different from zero. One reason for getting higher MPE % values from site-specific models as compared with general models could be that the site-specific models are based on narrower data ranges (see Table 4), whereas the general models cover much larger data ranges, and are thus less vulnerable when applied outside this range.

Since the general total tree volume model (modelling data set collected from one site) by Abbot et al (1997) (Table 8) was developed based on selected timber species only, it was not surprising that total tree volume was over-estimated (although not statistically significant because of large variation between observations, i.e. standard deviation of 139%). This was confirmed by the significant drop in MPE % when this model was applied on our data set solely to the selected timber species. This also shows that the general models developed in this study perform much better than the general model developed by Abbot et al. (1997) if we aim for quantifying total volume of trees in general forest inventories over large areas. This does not mean, however, that the species-specific models developed by Abbot et al. (1997) should not be applied when relevant, i.e. when predicting volumes for the tree species present in the data used for the development of his models. These models were based on a very large data set, and as such they are very likely to produce accurate volumes for the respective tree species.

The number of sample trees in our data set was 74. This is relatively few trees as compared with those used for the general models developed by Mauya et al. (2014) (158 sample trees) and Abbot et al. (1997) (88 sample trees). Given that our data set covers wide ranges in ecological conditions and in tree size, we do not consider this to be a major problem. The fact, however, that our data set comprised 33 species out of 139 species observed in the inventories indicate that if more tree species were included it would probably reduce uncertainty when applying the models over large areas. The data set displayed in the Appendix therefore in future could be supplemented by more species and the models recalibrated.

Conclusions

We have developed general models for predicting total tree, merchantable stem and branches volume. The model performances and the evaluations of the models suggest that they can be used over a wide range of geographical and ecological conditions in Malawi with an appropriate accuracy in predictions. The appropriateness of our models was also supported by the fact that the mean prediction errors of these models were much lower than the mean prediction errors of some previously developed models tested on our data.

Acknowledgements — The main financier for the main author of this study is the Norwegian Government through the Norwegian Quota Scheme. We would also like to thank the Norwegian University of Life Sciences and World Wildlife Fund (WWF) under the Russel E Train Education for Nature Fellowship for funding the fieldwork. Thanks should also go to the Forest Research Institute of Malawi (FRIM) for their assistance in facilitating the grant for a research permit from the National Research Council of Malawi (NRCM). Special thanks should also go Mike Chirwa, Devey Lekapo, Happy Kachamba, Humphrey Chapama, Boyd Zulu, Kola Daitoni and Steve Hara for their hard-working spirit during data collection. I would also like to thank the District Forest Officers for Mzimba, Dowa, Dedza and Neno districts for their cooperation and guidance during data collection.

References

Abbot P, Lowore J, Werren M. 1997. Models for the estimation of single tree volume in four Miombo woodland types. *Forest Ecology and Management* 97: 25–37.

Abdallah JM, Monela GG. 2007. Overview of Miombo woodlands in Tanzania. *Working Papers of the Finnish Forest Research Institute* 50: 9–23.

Adekunle VAJ. 2007. Nonlinear regression models for timber volume estimation in natural forest ecosystems, South Nigeria. *Research Journal of Forestry* 1: 40–54.

Adekunle VAJ, Nair KN, Srivastava AK, Singh NK. 2013. Models and form factors for stand volume estimation in natural forest ecosystems: a case study of Katarniaghat wildlife sanctuary (KGWS), Bahraich District, India. *Journal of Forestry Research* 24: 217–226.

Akindele SO, LeMay VM. 2006. Development of tree volume equations for common timber species in the tropical rain forest area of Nigeria. *Forest Ecology and Management* 226: 41–48.

Chamshama S, Mugasha A, Zahabu E. 2004. Stand biomass and volume estimation for miombo woodlands at Kitulangalo, Morogoro, Tanzania. *Southern African Forestry Journal* 200: 59–70.

Chidumayo EN. 1988. Estimating fuelwood production and yield in

regrowth dry miombo woodland in Zambia. *Forest Ecology and Management* 24: 59–66.
Chidumayo EN, Gumbo DJ. 2010. *The dry forests and woodlands of Africa: managing for products and services.* London: Earthscan.
Clutter JL, Fortson JC, Piennar LV, Brister GH, Bailey RL. 1983. *Timber management: a quantitative approach*, vol. 6. New York: John Wiley and Sons.
Dadzie PK. 2013. Potential contribution of branchwood quantity, left after logging operations, towards reducing depletion rate and preserving Ghana's forest ecosystem. *American Journal of Agriculture and Forestry* 1: 32–39.
Dewees PB, Campbell BM, Katerere Y, Sitoe A, Cunningham AB, Angelsen A, Wunder S. 2011. *Managing the miombo woodlands of southern Africa: policies, incentives, and options for the rural poor.* Washington, DC: Program on Forests (PROFOR).
Giliba RA, Boon EK, Kayombo CJ, Musamba EB, Kashindye AM, Shayo PF. 2011. Species composition, richness and diversity in miombo woodlands of Bereku Forest Reserve, Tanzania. *Journal of Biodiversity* 2: 1–7.
Government of Malawi. 2010a. Malawi state of environment and outlook report. Lilongwe: Environmental Affairs Department, Ministry of Natural Resources, Energy and Environment.
Government of Malawi. 2010b. Forestry Act Amendment. *Malawi Gazette Supplement*, *Government Notice* no. 23.
Government of Malawi. 2012. Forest resource mapping project under the Japanese grant for the forest preservation programme to the Republic of Malawi: final report for implementation phase. Lilongwe: Department of Forestry Ministry of Environment and Climate Change Management.
Guendehou GHS, Lehtonen A, Moudachirou M, Mäkipää R, Sinsin B. 2012. Stem biomass and volume models of selected tropical tree species in West Africa. *Southern Forests* 74: 77–88.
Hardcastle PD. 1978. *A preliminary silvicultural classification of Malawi*. Zomba: Forest Research Institute of Malawi.
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: 326–330.
Henry M, Picard N, Trotta C, Manlay RJ, Valentini R, Bernoux M, Saint-André,L. 2011. Estimating tree biomass of sub-Saharan African forests: a review of available allometric equations. *Silva Fennica* 45: 477–569.
Hofstad O. 2005. Review of biomass and volume functions for individual trees and shrubs in southeast Africa. *Journal of Tropical Forest Science* 17: 151–162.
Hunter MO, Keller M, Victoria D, Morton DC. 2013. Tree height and tropical forest biomass estimation. *Biogeosciences* 10: 8385–8399.
James G, Witten D, Hastie T, Tibshirani R. 2013. *An introduction to statistical learning*. New York: Springer-Verlag.
Kalaba FK, Quinn CH, Dougill AL. 2012. Carbon storage, biodiversity and species composition of miombo woodlands in recovery trajectory after charcoal production and slash and burn agriculture in Zambia's Copperbelt. *SRI Papers* no. 40. Leeds: Sustainability Research Institute, Schools of Earth and Environment, University of Leeds.
Larjavaara M, Muller-Landau HC. 2013. Measuring tree height: a quantitative comparison of two common field methods in a moist tropical forest. *Methods in Ecology and Evolution* 4: 793–801.
Lowore JD, Abbot PG, Warren M. 1994. Stackwood volume estimations for miombo woodlands of Malawi. *Commonwealth Forestry Review* 73: 193–197.
Luhanga J. 2009. The timber trade in Malawi. *Resource Insight* no. 7. Johannesburg: Southern Africa Resource Watch.
Malimbwi RE, Solberg B, Luoga E. 1994. Estimation of biomass and volume in Miombo woodland at Kitulangalo Forest Reserve, Tanzania. *Journal of Tropical Forest Science* 7: 230–242.
Mate R. 2014. Biomass and volume estimates for valuable timber species in Mozambique. PhD thesis, Swedish University of Agricultural Sciences. Sweden.
Mauya EW, Mugasha WA, Zahabu E, Bollandsås OM, Eid T. 2014. Models for estimation of tree volume in the miombo woodlands of Tanzania. *Southern Forests* 76: 209–219.
Missanjo E, Kamanga-Thole G, Mtambo C, Chisinga, O. 2014. Evaluation of natural regeneration and tree species diversity in miombo woodlands in Malawi. *Journal of Biodiversity Management and Forestry* 3: 3.
Mugasha WA, Eid T, Bollandsås OM, Malimbwi RE, Chamshama SAO, Zahabu E, Katani JZ. 2013. Allometric models for prediction of above- and belowground biomass of trees in the miombo woodlands of Tanzania. *Forest Ecology and Management* 310: 87–101.
Munishi PK, Temu PR-A, Soka G. 2011. Plant communities and tree species associations in a miombo ecosystem in the Lake Rukwa basin, southern Tanzania: implications for conservation. *Journal of Ecology and the Natural Environment* 3: 63–71.
Mwakalukwa EE, Meilby H, Treue T. 2014a. Floristic composition, structure, and species associations of dry miombo woodland in Tanzania. *International Scholarly Research Notices: Biodiversity* 2014: Art. ID 153278, 15 pages.
Mwakalukwa EE, Meilby H, Treue T. 2014b. Volume and aboveground biomass models for dry Miombo woodland in Tanzania. *International Journal of Forestry Research* 2014: Art. ID 531256, 11 pages.
Mwase W, Bjørnstad Å, Bokosi JM, Kwapata MB, Stedje B. 2007. The role of land tenure in conservation of tree and shrub species diversity in miombo woodlands of southern Malawi. *New Forests* 33: 297–307.
Parresol BR. 1999. Assessing tree and stand biomass: a review with examples and critical comparisons. *Forest Science* 45: 573–593.
Picard N, Saint-André L, Henry M. 2012. Manual for building tree volume and biomass allometric equations: from field measurement to prediction. Rome: Food and Agriculture Organization of the United Nations; Montpellier: Centre de Coopération Internationale en Recherche Agronomique pour le Développement.
Pinard MA, Putz FE. 1996. Retaining forest biomass by reducing logging damage. *Biotropica* 28: 278–295.
Ryan CM, Williams M, Grace J. 2011. Above-and belowground carbon stocks in a miombo woodland landscape of Mozambique. *Biotropica* 43: 423–432.
SAS Institute. 2012. SAS 9.4 TS Level 1 M1. Cary: SAS Institute.
Sileshi G. 2014. A critical review of forest biomass estimation models, common mistakes and corrective measures. *Forest Ecology and Management* 329: 237–254.
van Laar A, Akça A. 2007. *Forest mensuration. Managing Forest Ecosystems* vol. 13. Dordrecht: Springer.
Zahabu E. 2008. Sinks and sources: a strategy to involve forest communities in Tanzania in global climate policy. PhD thesis, University of Twente, The Netherlands.

Received 5 April 2015, revised 24 July 2015, accepted 12 August 2015

Appendix: Data set used for model development. dbh = diameter at breast height, ht = total tree height

Site	Scientific name	Local name	dbh (cm)	ht (m)	Stem volume (m^3)	Branch volume (m^3)	Total tree volume (m^3)
Mtangatanga	*Brachystegia glaucescens*	Musani	6.0	4.4		0.015	0.015
	Erica banguelensis	Msankhanya	9.4	4.0		0.028	0.028
	Erica banguelensis	Msankhanya	13.0	6.0		0.056	0.056
	Uapaca kirkiana	Msuku	15.0	5.0		0.116	0.116
	Julbernadia paniculata	Mtondo	16.5	8.0	0.077	0.059	0.136
	Isoberlinia angolensis	Kabale	19.4	12.0	0.191	0.180	0.371
	Brachystegia boehmii	Mombo	20.2	8.0	0.184	0.134	0.318
	Uapaca kirkiana	Msuku	21.0	7.8	0.143	0.175	0.318
	Brachystegia utilis	Nzale	23.0	10.0	0.212	0.206	0.418
	Brachystegia taxifolia	Mchinji	23.6	10.0	0.158	0.218	0.376
	Julbernadia paniculata	Mtondo	27.3	10.0	0.214	0.441	0.655
	Isoberlinia angolensis	Kabale	28.6	11.0	0.218	0.541	0.759
	Brachystegia spiciformis	Chumbe	36.0	12.0	0.650	0.697	1.347
	Brachystegia longifolia	Sanga	36.0	10.8	0.673	0.852	1.525
	Brachystegia glaucescens	Musani	48.4	14.5	0.832	1.362	2.194
	Brachystegia glaucescens	Musani	50.0	16.0	1.138	1.879	3.017
	Brachystegia spiciformis	Chumbe	60.0	18.0	1.630	3.996	5.626
	Brachystegia glaucescens	Musani	68.0	10.5	0.645	3.424	4.069
	Brachystegia glaucescens	Musani	78.0	17.0	1.148	5.127	6.275
	Brachystegia glaucescens	Musani	111.2	18.0	3.817	7.437	11.254
Kongwe	*Julbernadia paniculata*	Mtondo	9.0	5.0		0.024	0.024
	Uapaca kirkiana	Msuku	10.2	5.0		0.034	0.034
	Brachystegia spiciformis	Chumbe	13.0	6.0		0.109	0.109
	Brachystegia manga	Mpapa/Bovo	17.3	10.0	0.080	0.112	0.192
	Julbernadia paniculata	Mtondo	21.0	11.0	0.152	0.166	0.318
	Brachystegia utilis	Nzale	24.5	9.0	0.225	0.294	0.519
	Monotes africanus	Mkalakate	27.2	10.0	0.145	0.399	0.544
	Brachystegia spiciformis	Chumbe	27.5	9.3	0.205	0.473	0.678
	Brachystegia boehmii	Mombo	30.8	14.0	0.266	0.446	0.712
	Uapaca kirkiana	Msuku	34.0	12.0	0.245	0.764	1.009
	Parinari curatellifolia	Muula	37.3	12.0	0.426	0.926	1.352
	Brachystegia manga	Mpapa/Bovo	38.1	17.0	0.754	1.575	2.329
	Cussonia arborea	Mbwabwa	38.4	7.0	0.153	0.633	0.786
	Brachystegia spiciformis	Chumbe	38.7	14.0	0.473	1.259	1.732
	Brachystegia manga	Mpapa/Bovo	52.0	22.0	1.035	2.461	3.496
	Brachystegia manga	Mpapa/Bovo	61.7	17.0	1.728	4.935	6.663
	Brachystegia spiciformis	Chumbe	71.2	18.0	2.209	4.417	6.626
	Erythrina abyssinica	Muwale	75.7	12.0	1.460	6.238	7.698
Mua-livulezi	*Markhamia obtusifolia*	Msewa	5.3	5.0		0.009	0.009
	Combretum apiculatum	Kakunguni	6.0	3.0		0.011	0.011
	Bauhinia petersiana	Mphandula	9.1	5.2		0.037	0.037
	Bauhinia thoningii	Msekese	13.0	5.0		0.088	0.088
	Diplorhynchus condylocarpon	Thombozi	16.2	10.4	0.077	0.058	0.135
	Anonna senegalensis	Mpoza	17.0	8.0	0.081	0.039	0.120
	Markhamia obtusifolia	Msewa	20.0	9.3	0.100	0.122	0.222
	Pterocarpus rotundifolius	M'balitsa	22.0	8.5	0.218	0.182	0.400
	Albizzia versicolor	Mtangatanga	23.2	10.4	0.150	0.195	0.345
	Bauhinia petersiana	Mphandula	27.4	7.9	0.157	0.333	0.490
	Diplorhynchus condylocarpon	Thombozi	31.8	13.0	0.351	0.381	0.732
	Vachellia galpinni	Mgundanjira	37.7	14.0	0.703	0.543	1.246
	Pterocarpus rotundifolius	M'balitsa	42.0	15.0	0.680	0.732	1.412
	Brachystegia spiciformis	Chumbe	47.1	15.5	1.018	1.795	2.813
	Brachystegia bussei	Mtwana	61.0	20.5	1.352	3.216	4.568
	Pseudolachnostaylis maprouneifolia	Msolo	61.0	15.6	1.674	3.853	5.527
	Brachystegia spiciformis	Chumbe	69.5	21.0	3.332	3.262	6.594
	Brachystegia bussei	Mtwana	81.7	22.0	3.871	6.571	10.442
Tsamba	*Uapaca kirkiana*	Msuku	8.4	6.5		0.025	0.025
	Brachystegia floribunda	Tsamba	9.5	9.0		0.037	0.037
	Pseudolachnostaylis maprouneifolia	Msolo	12.4	7.0		0.102	0.102
	Brachystegia floribunda	Tsamba	19.9	8.0	0.154	0.173	0.327
	Parinari excelsa	Mpembu	21.0	10.0	0.254	0.071	0.325

Appendix (cont.)

Site	Scientific name	Local name	dbh (cm)	ht (m)	Stem volume (m^3)	Branch volume (m^3)	Total tree volume (m^3)
Tsamba	*Julbernadia globiflora*	Kachumbe	21.0	10.0	0.199	0.114	0.313
	Brachystegia spiciformis	Chumbe	21.2	12.0	0.205	0.152	0.357
	Uapaca sansibarica	Msokolowe	22.2	10.0	0.180	0.220	0.400
	Pericorpsis angolensis	Muwanga	28.4	8.0	0.306	0.177	0.483
	Uapaca kirkiana	Msuku	30.3	11.0	0.326	0.674	1.001
	Pterocarpus angolensis	Mlombwa	31.3	17.0	0.370	0.603	0.973
	Brachystegia utilis	Nzale	32.0	14.0	0.307	0.528	0.835
	Uapaca sansibarica	Msokolowe	32.0	12.0	0.305	0.621	0.926
	Brachystegia floribunda	Tsamba	34.5	12.0	0.597	0.146	0.743
	Julbernadia globiflora	Kachumbe	37.5	18.0	0.734	0.732	1.466
	Faurea speciosa	Chisese	38.7	16.0	0.608	0.351	0.959
	Julbernadia globiflora	Mchenga	67.4	21.0	2.426	3.392	5.818
	Brachystegia spiciformis	Chumbe	75.0	25.0	4.446	4.581	9.027

Paper II

forests

MDPI

Article

Above- and Belowground Biomass Models for Trees in the Miombo Woodlands of Malawi

Daud J. Kachamba *, Tron Eid and Terje Gobakken

Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, 1432 Ås, Norway; tron.eid@nmbu.no (T.E.); terje.gobakken@nmbu.no (T.G.)

* Correspondence: daud.kachamba@nmbu.no or dkachamba@gmail.com; Tel.: +47-649-658-00 or +265-995-658-387

Academic Editors: Shibu Jose and Eric J. Jokela

Received: 25 October 2015; Accepted: 25 January 2016; Published: 5 February 2016

Abstract: In this study we present general (multiple tree species from several sites) above- and belowground biomass models for trees in the miombo woodlands of Malawi. Such models are currently lacking in the country. The modelling was based on 74 trees comprising 33 different species with diameters at breast height (dbh) and total tree height (ht) ranging from 5.3 to 2 cm and from 3.0 to 25.0 m, respectively. Trees were collected from four silvicultural zones covering a wide range of conditions. We tested different models including dbh, ht and wood specific gravity (ρ) as independent variables. We evaluated model performance using pseudo-R^2, root mean square error (RMSE), a covariance matrix for the parameter estimates, mean prediction error (MPE) and relative mean prediction error (MPE%). Computation of MPE% was based on leave-one-out cross-validation. Values of pseudo-R^2 and MPE% ranged 0.82–0.97 and 0.9%–2.8%, respectively. Model performance indicated that the models can be used over a wide range of geographical and ecological conditions in Malawi.

Keywords: root to shoot ratio; root sampling; tropical dry forest; error estimation; carbon

1. Introduction

Miombo woodlands, classified as dry forests, are dominated by woody plants, primarily trees, whose canopy cover more than 10% of the ground surface, occurring in a climate with a dry season of three months or more [1]. The woodlands are dominated by deciduous trees of the genera *Brachystegia*, *Julbernadia* and *Isoberlinia*, which cover an area of approximately 2.7 million km^2 spanning ten countries in eastern and central Africa including Malawi [1–5]. Miombo woodlands may be divided into dry and wet miombo. Dry miombo occur in areas receiving less than 1000 mm of rainfall annually in Zimbabwe, central Tanzania, and in the southern areas of Mozambique, Malawi and Zambia. Wet miombo occur in areas receiving more than 1000 mm of annual rainfall in eastern Angola, northern Zambia, south western Tanzania and central Malawi [1,6]. In Malawi, miombo woodlands constitute 92.4% of the country's total forested area, and are mainly located in forest and game reserves established for water catchment as well as for soil and biodiversity conservation [2,7].

Miombo woodlands play a critical role in the livelihoods of Malawian communities because they provide social, economic, and environmental benefits, such as firewood, timber, medicinal plants, food, and catchment protection, among others [8]. Increased population growth, currently estimated at an annual growth rate of 2.8% [9], has led to higher demand for firewood, charcoal, and timber. Thus, the woodlands are being deforested at an annual rate of approximately 0.9%, which is among the highest rates in Africa [10].

To sustain the provision of these services, there is an urgent need to implement sustainable forest management measures including estimation of growing stock, productivity, forest biomass

and yield [11,12]. Estimation of forest biomass is the first step towards calculation of carbon stocks. Due to the natural capacity of trees to sequester carbon dioxide, miombo woodlands are considered an important element in global climate change mitigation programs such as the reducing emissions from deforestation and forest degradation mechanism (REDD+), which provides a framework where developing countries may be financially rewarded for reducing carbon emissions.

The Malawi government recently established a baseline for forest biomass and carbon stock estimates for targeted forest reserves in miombo woodlands of Malawi [7]. However, the estimates are unreliable because of the nature of the allometric models that were used. For example, the aboveground forest biomass estimates were based on a pan-tropical biomass model developed by Chave *et al.* [13]. This model was developed using data from trees in tropical America and Asia but did not include Africa or the miombo woodlands. In addition, belowground forest biomass estimates were based on an allometric model developed by Cairns *et al.* [14]. Unlike Chave *et al.* [13], this dataset included some trees from Africa, *i.e.*, Democratic Republic of the Congo, Ghana and the Ivory Coast. However, the trees were from moist evergreen tropical forests, whose structure is different from miombo woodlands.

By 2011 there were approximately 370 allometric models for predicting tree biomass in sub-Saharan Africa [15]. The majority of these models were developed for tropical rainforests in western Africa. Among the models in south-eastern Africa, only a few were developed for miombo woodlands. These models consisted of: (a) models developed for specific tree species based on a dataset from one site (Mwakalukwa *et al.* [16]); (b) models developed for specific tree species based on a dataset from several sites (Mate *et al.* [17]); (c) models developed for multiple tree species based on a dataset from one site (Chidumayo [18], Chamshama *et al.* [19], Malimbwi *et al.* [20], Ryan *et al.* [4], Mwakalukwa *et al.* [16]) and (d) models developed for multiple tree species based on a dataset from several sites (Mugasha *et al.* [21]).

Miombo woodlands are characterized by high tree species diversity, and the reported number of species from assessments at different spatial scales ranges from 80 to 300 [22–27]. Due to such large number of tree species, the applicability of species-specific models is limited. Furthermore, applicability of single-site models over different ecological zones is also limited due to their narrow geographical range. A scenario with general models, combining multiple species collected over several sites, would therefore be the best alternative, for example, in cases where national forest inventories are to be carried out. No such models exist for miombo woodlands in Malawi.

Most of the previously mentioned studies focused on aboveground biomass. However, estimation of belowground biomass in miombo woodlands is also vital. Belowground tree biomass, as a basis for model development, can be determined using complete excavation of roots, soil core sampling for fine and medium roots, and root sampling (complete excavation of a few sampled roots of a tree). Estimating belowground tree biomass can be done by using the root to shoot ratio (RS-ratio), *i.e.*, the ratio between belowground and aboveground dry weights (see e.g., [28,29]), or through allometric models. Belowground biomass models for miombo woodlands in neighbouring countries were developed by Mugasha *et al.* [21], Chidumayo [18] and Ryan *et al.* [4].

The Inter-governmental Panel on Climate Change (IPCC) (see [30]) requires biomass and carbon reporting under the REDD+ mechanism to be accompanied by appropriate measures of uncertainty. Uncertainties are likely to occur in the following steps in biomass quantification: (i) when applying the sampling design (number and size of plots); (ii) during tree measurements and (iii) when applying the biomass model (e.g., [31,32]). The model-related uncertainty in this context stems from sources such as: (a) model misspecifications; (b) uncertainties in values of independent variables; (c) residual variability and (d) uncertainty in the model parameter estimates (see e.g., [33,34]). Among these, uncertainty in model parameter estimates has a great influence [35]. However, very few studies report uncertainty in the model parameter estimates, *i.e.*, the covariance structure of the parameter estimates of developed models, and this makes it impossible to analyse the totality of the uncertainty related to estimated forest biomass (see [35]).

The objective of this study is to develop general (multiple tree species from several sites) above- and belowground biomass models applicable across the entire distribution of miombo woodlands in Malawi. The models are also accompanied with information on their covariance structure to enable quantification of model-related uncertainties in biomass and carbon estimation. Furthermore, we provide basic statistics on RS-ratios and compare the performance of our models with existing models from miombo woodlands.

2. Materials and Methods

2.1. Site Description

The sample trees for model development were selected from four forest reserves, namely Mtangatanga, Kongwe, Mua-livulezi and Tsamba (Figure 1, Table 1). The selection of sites was based on geographical location, management regime, silvicultural classification, and climatic conditions to capture a wide range of factors influencing tree growth [1].

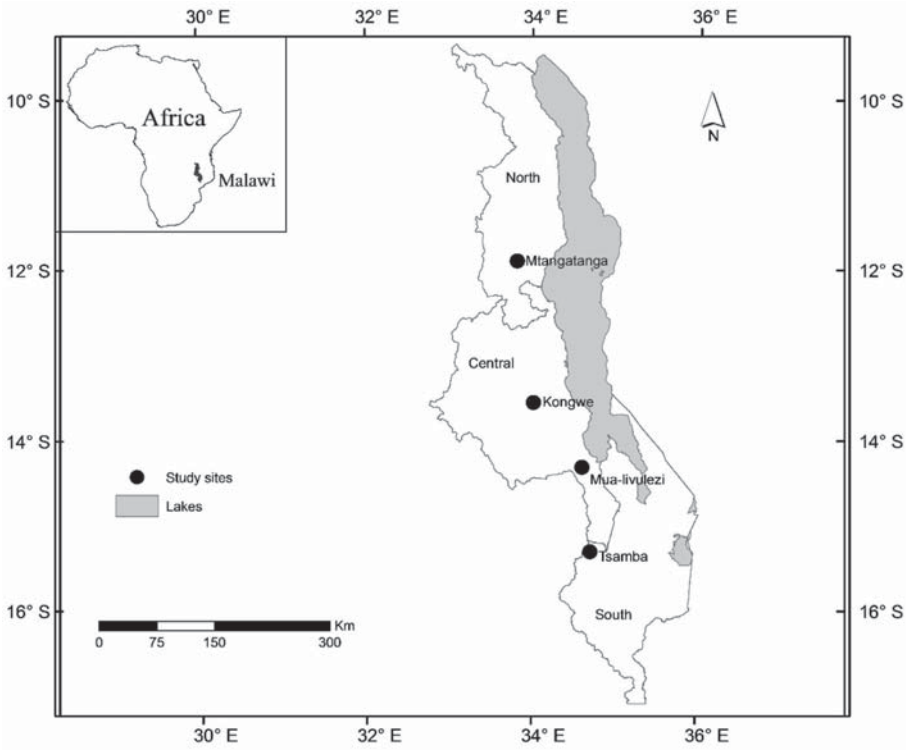


Figure 1. Map of Malawi showing the location of the study sites.

Table 1. Geographical location, management regime, silvicultural classification and climatic conditions of study sites.

	Mtangatanga	Kongwe	Mua-livulezi	Tsamba
Region	Northern	Central	Central	Southern
District	Mzimba	Dowa	Dedza	Neno
Location	11°56′ S 33°42′ E	13°35′ S 33°55′ E	14°21′ S 34°37′ E	15°21′ S 34°36′ E
Area (ha)	8443	1813	12147	3240
Management regime	Co-management	Government	Co-management	Government
Altitude (m)	1500–1700	1000–1500	400–900	700–1500

Table 1. *Cont.*

	Mtangatanga	Kongwe	Mua-livulezi	Tsamba
Dominant soil type	Humic ferrallitic	Ferruginous	Lithosols	Ferrallitic
Silvicultural classification	Moist *Brachystegia*	Moist *Brachystegia*	Dry *Brachystegia*	Moist *Brachystegia*
Mean minimum annual temp (°C)	6	6	13	8
Mean maximum annual temp (°C)	29	29	32	28
Total annual rainfall range (mm)	960–1050	960–1050	840–960	1200–1600
Rain period	December–April	November–April	November–April	November–April
Dry months	May–November	May–October	May–October	May–October

Data sources: Rainfall and temperature data (1975–2005) from Ministry of Natural Resources, Energy and Mining, Department of Climate Change and Meteorological Services in Malawi. Soil and silvicultural classification according to Hardcastle [36].

2.2. Selection of Sample Trees

We conducted systematic sample plot inventories covering each site to collect information on ranges in tree size and species distribution to guide the selection of sample trees (e.g., [21]). We used circular plots with radius 11.28 m (400 m^2). On each plot, we identified all tree species and measured their diameter at breast height (dbh) for all trees with dbh >4 cm. In addition, we sampled three trees within each plot, (one with the smallest, one with a medium and one with the largest dbh), and measured their total height (ht). The inventories covered a total of 221 plots with 70, 30, 71 and 50 plots for Mtangatanga, Kongwe, Mua-livulezi and Tsamba, respectively. The maximum recorded dbh values based on all sample plots in Mtangatanga, Kongwe, Mua-livulezi and Tsamba were 61 cm, 73 cm, 70 cm and 56 cm, respectively, while the number of species identified for the respective sites were 66, 45, 77 and 65. In total, for all the study sites, we identified 139 species. The most frequent species for Mtangatanga, Kongwe, Mua-livulezi and Tsamba were *Uapaca kirkiana* Müll. Arg., *Brachystegia spiciformis* Benth., *Diplorhynchus condylocarpon* (Müll. Arg.) Pichon and *Uapaca kirkiana* Müll. Arg, respectively.

A total of 74 trees were selected based on the observed dbh and tree species frequency within the sites. We ensured that the trees were selected from all dbh classes observed in the sample plot inventories. In addition, we selected a total of eight trees with larger dbh than those observed in the sample plot inventories to reduce uncertainty when predicting biomass of very large trees. We also selected at least one tree among the eight most frequent species observed in each site. The remaining sample trees were selected randomly among all species. In total, 33 tree species were selected, comprising 10, 10, 12 and 10 different tree species in Mtangatanga, Kongwe, Mua-livulezi and Tsamba, respectively.

Before felling (at a stump height of 30 cm), we recorded scientific and local names, and measured dbh, stump diameter (at 30 cm above ground) and ht (Tables 2 and A1). We used either a calliper or a diameter tape, depending on tree sizes, to measure dbh and stump diameter, while a Suunto hypsometer was used for all ht measurements. These trees have previously been used to develop general volume models for miombo woodlands in Malawi [37].

Table 2. Mean, minimum, maximum and standard deviation (STD) of diameter at breast height (dbh) and total tree height (ht) for sample trees at each site.

Site	No. of Trees	dbh (cm)				ht (m)			
		Mean	Min.	Max.	STD	Mean	Min.	Max.	STD
Mtangatanga	20	35.5	6.0	111.2	26.7	10.7	4.0	18.0	4.3
Kongwe	18	34.9	9.0	75.7	19.6	11.7	5.0	22.0	4.7
Tsamba	18	30.2	8.4	75.0	17.5	12.6	6.5	25.0	5.1
Mua-livulezi	18	32.8	5.3	81.7	23.0	11.6	3.0	22.0	5.8
All	74	33.4	5.3	111.2	21.8	11.6	3.0	25.0	4.9

2.3. Destructive Sampling

We separated each sample tree into the following aboveground components: merchantable stem (from the stump at 30 cm above ground to the point where the first branches start), branches (all parts of the tree above the defined merchantable stem and up to a minimum diameter of 2.5 cm) and twigs (all branches with a diameter less than 2.5 cm). For small trees not considered suitable for timber production (dbh < 15 cm, in total 14 trees), merchantable stem biomass were allocated to branches (e.g., [37,38]). During the work on destructive sampling most of the trees had already started to shed leaves. We therefore excluded leaves from twigs, and leaves were not included in the models.

To facilitate measurements, stems and branches were crosscut into manageable pieces of approximately 1–2 m in length and then weighed for fresh weight using a mechanical hanging spring balance (0–200 kg). Twigs from each tree were separately bundled and weighed for fresh weight. Three small sub-samples, varying in weight between 0.1 and 1.0 kg, from each of the components (merchantable stem, branches and twigs) were taken from each sample tree and weighed with an electronic balance for fresh weight and finally brought to laboratory for drying. The sub-samples were taken from the biggest, medium and smallest diameter parts of each tree component.

For determination of belowground biomass of the sample trees, our strategy involved root sampling at two levels, namely main roots (roots branching directly from the root crown) and side roots (roots branching from the main roots). The first step in excavation involved clearing the topsoil around the tree base to expose the points at which the roots were branching. We then selected three main roots, *i.e.*, the main roots with the largest, medium and smallest diameters, and recorded their diameters at the points where they joined the root crown. The diameters of all main roots not excavated were recorded at the point where they were joined the root crown. From each of the selected main roots, we selected up to three side roots, *i.e.*, the side roots with the largest, medium and smallest diameters. For each of the selected side roots, we recorded the diameter where they joined the main root. For the remaining side roots, we also recorded the diameters at the branching point from the main root. The selected side and main roots were then fully excavated up to the minimum diameter of 1 cm and then weighed. In cases where the full roots could not be excavated due to obstacles such as rocks, the diameter of the last bit of the root was recorded and we treated the remaining unexcavated part as a side root. An effort was made to ensure that all the main, side and taproots were fully excavated up to the last 1 cm. In total, 38 out of the 41 trees, had taproots. Out of the 38 trees, we were not able to fully excavate the taproots for 16 trees. In such cases, the diameter at the breaking point of the unexcavated taproot was recorded and treated as a side root. On average, trees were dug down to 2.5 m depth. Lastly, we recorded the fresh weight of the root crown for each tree.

For all sample trees, three small sub-samples, varying in weight between 0.1 and 1.0 kg, were taken from each main and side root, and one was taken from the root crown. We obtained the fresh weight of the sub-samples using an electronic balance and brought them to the laboratory for drying.

2.4. Laboratory Analyses and Determination of Biomass Dry Weight

All sub-samples, from both above- and belowground, for each tree were dried in an oven in a laboratory at a temperature of 80 °C until a constant weight was achieved (constant weight was observed in 2–3 days). We then recorded dry weights of the individual sub-samples. Subsequently, we used the sub-sample dry and fresh weights to determine the tree- and section specific dry to fresh weight ratios (DF-ratios) (see Table A2).

We then calculated the dry weight of each section as a product of tree- and section specific DF-ratios and the fresh weights of the respective trees and tree sections. Subsequently, we computed the total aboveground dry weight by summing the dry weights of the merchantable stem, branches and twigs of each tree (Figure 2, Table A1).

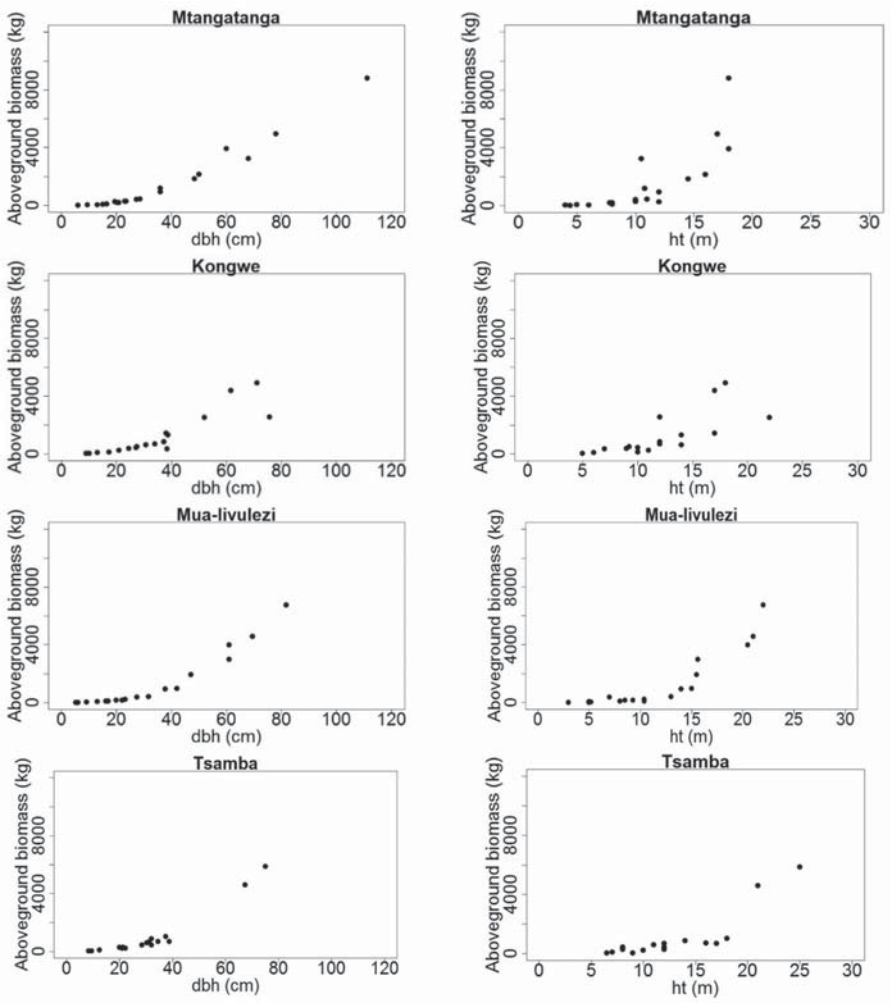


Figure 2. Total aboveground tree biomass (kg dry weight) distribution over dbh (cm) and ht (m) for Mtangatanga, Kongwe, Mua-livulezi and Tsamba and forest reserves.

To determine the total belowground dry weights of the excavated parts of the trees we first converted all the fresh weights from the different sections to dry weight biomass by multiplying the tree- and section specific DF-ratios and their respective fresh weights. We then developed a general (combining data from all sites) side root model by regressing the dry weight biomass of the fully excavated side roots and their diameters (cm). We assumed the relationship between side root biomass and root diameter (similarly for main roots, see below) to exhibit a power-law relationship described as:

$$B = a \times d^{b} \quad (1)$$

where B = dry weight biomass of a side root or main root (kg); d = diameter (cm) of a side or main root at the point it is joining the main root or the root crown, respectively; a and b are parameter estimates. The following side root model was developed:

$$B = 0.198102 \times d^{1.656968},\ \text{Pseudo} - R^2 = 0.67,\ \text{MPE\%} = -2.0 \quad (2)$$

where

$$\text{Pseudo} - R^2 = 1 - \left(\frac{\text{SSR}}{\text{CSST}}\right) \quad (3)$$

SSR is the sum of squared residuals and CSST is the corrected total sum of squares. The mean prediction error (MPE), and the relative mean prediction error (MPE%) is calculated as

$$\text{MPE} = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)}{n}, \ \text{MPE\%} = \frac{\text{MPE}}{\overline{y}} \times 100 \quad (4)$$

where y_i is the observed biomass of tree i, $\hat{y}_i$ is the predicted biomass of tree i and $\overline{y}$ is the mean observed biomass. Both MPE and MPE% are based on leave-one-out-cross validation. The MPE% value for the side root model was not significantly different from zero indicating appropriate model performance.

The side root model was used to predict the dry weight biomass of all the side roots that were not excavated for the main sample root. The total dry weight of all side roots for each main sample root was then determined by summing dry weights of the excavated side roots and predicted dry weights of unexcavated side roots. Finally the complete dry weight of the sample main root was determined by summing the total dry weights of all side roots and the excavated parts of the main root. The following main root model was then developed and applied to predict the dry weights of main roots not excavated;

$$\text{B} = 0.063132 \times \text{d}^{2.174388}, \ \text{Pseudo} - R^2 = 0.79, \ \text{MPE\%} = -0.4 \quad (5)$$

The MPE% value for the main root model was not significantly different from zero indicating appropriate model performance. To determine the dry weight of unexcavated parts of the taproots (16 trees), we applied the general side root model.

Total belowground dry weight biomass for each tree was finally determined by adding the dry weights of all excavated and unexcavated main roots, dry weight of the taproot, and the dry weight of the root crown (Figure 3, Table A1).

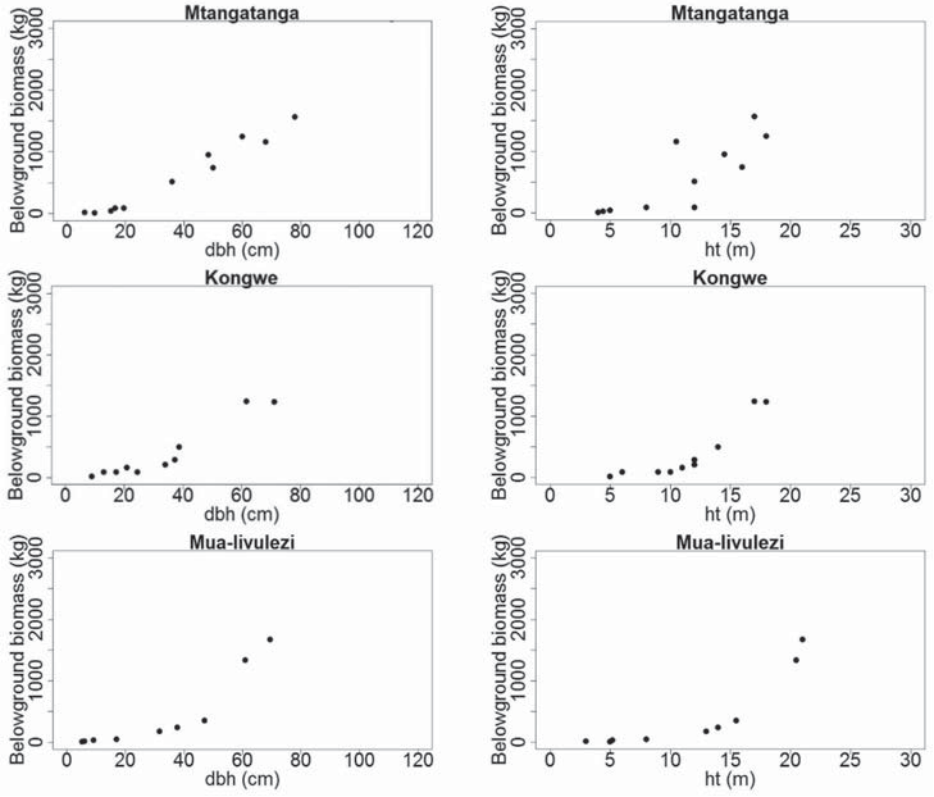


Figure 3. *Cont.*

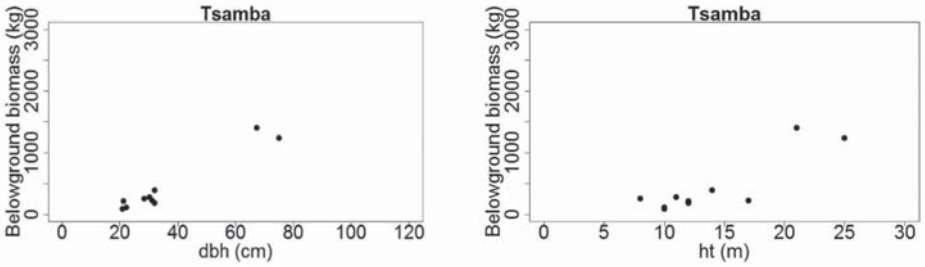


Figure 3. Total belowground tree biomass (kg dry weight) distribution over dbh (cm) and ht (m) for Mtangatanga, Kongwe, Mua-livulezi and Tsamba and forest reserves.

2.5. Model Development and Evaluation

Before fitting the models, we assessed the basic diagnostic plots of biomass over dbh and ht. As expected, the plots indicated non-linear patterns in the relationships between biomass and dbh and ht (Figures 2 and 3). Since wood specific gravity (ρ) is considered as an important factor for explaining variation in biomass (e.g., [39]), we included this variable in the models. We therefore tested the following models:

$$\text{Model 1}: \ B = a \times dbh^b \tag{6}$$

$$\text{Model 2}: \ B = a \times dbh^b \times ht^c \tag{7}$$

$$\text{Model 3}: \ B = a \times dbh^b \times \rho^c \tag{8}$$

$$\text{Model 4}: \ B = a \times dbh^b \times ht^c \times \rho^d \tag{9}$$

where B is biomass (kg), dbh is diameter at breast height (cm), ht is total tree height (m) and ρ is the species-specific mean wood specific gravity (g/cm^3) extracted from the global wood density database [40,41] and a, b, c and d are parameter estimates.

Since the data was collected from different study sites in different geographical regions and silvicultural zones across Malawi, we anticipated that individual tree attributes would be different depending on site. We therefore initially fitted mixed effects models, with site as a random effect, to the side root, main root and total belowground and aboveground biomass datasets using PROC NLMIXED of SAS 9.4 [42]. This procedure fits nonlinear mixed models, that is, models in which both fixed and random effects enter nonlinearly. PROC NLMIXED fits nonlinear mixed models by maximizing an approximation to the likelihood integrated over the random effects using the maximum likelihood estimation method.

These mixed effects models were compared with weighted nonlinear regression models fitted using PROC MODEL in SAS 9.4 [42]. This procedure fits models in which the relationships among the variables comprise a system of one or more nonlinear equations using the full information maximum likelihood (FIML) estimation method.

For each model derived from the two procedures, *i.e.*, mixed effects and weighted regression, we computed Akaike Information Criterion (AIC) values [43]. AIC measures the model goodness-of-fit whilst correcting for model complexity. We used the AIC values to compare mixed effects models with the weighted regression nonlinear models. The results showed that, in all cases, weighted regression models produced lower AIC values relative to mixed effects models. We thus decided to develop our final models based on weighted regression.

Model efficiency and performance were assessed based on results from a leave-one-out cross validation procedure [44]. This splits the dataset of n observations into two parts, namely, a validation dataset and a training dataset. The validation dataset comprises a single observation (x_1, y_1) and the training dataset comprises the remaining $\{(x_2, y_2), \ldots .., (x_n, y_n)\}$ observations. The model is fitted on the n-1 observations in the training dataset and a prediction $\hat{y}$ is made for a single observation in

the validation dataset, using its value x_1. Since (x_1, y_1) was not used in the fitting process, the square error (SE) = $(y_1 - \hat{y})^2$ provides an estimate of the test error. This procedure is repeated n times, thus producing n test errors, $SE_1 \ldots \ldots .SE_n$. The leave-one-out-cross validation estimate for the test error is the mean of these n test error estimates (MSE).

The cross validation results were then used to calculate the root mean square error (RMSE) as follows;

$$\text{RMSE} = \sqrt{\text{MSE}} \tag{10}$$

$$\text{RMSE } (\%) = \frac{\text{RMSE}}{\overline{y}} \times 100 \tag{11}$$

where $\overline{y}$ is the mean observed biomass and RMSE (%) is the relative Root Mean Square Error.

Model comparison was based on AIC values. Models with insignificant parameter estimates were not considered irrespective of AIC values. For all the models, we presented pseudo-R^2, RMSE, RMSE (%), covariance matrix for the parameter estimates, and the MPE and MPE% values based on leave-one-out cross validation. Student *t*-tests were conducted to determine whether the MPE values were significantly different from zero.

In addition, we tested a number of previously developed biomass models (Table 3) on our data. This included models developed for miombo woodlands in neighbouring countries, *i.e.*, Ryan *et al.* [4] in Mozambique, Mugasha *et al.* [21] in Tanzania and Chidumayo [18] in Zambia, and the pan-tropical model developed by Chave *et al.* [39]. MPE values were computed, and student *t*-tests were applied to determine whether the MPE values were significantly different from zero.

For a graphical display of the behaviour of models with ht as independent variable, *i.e.*, Mugasha *et al.* [21] and Chave *et al.* [39] (see Table 3), we applied a height-diameter model developed from our sample trees:

$$\text{ht} = 1.3 + \exp 3.787685 - 6.62809 \times \text{dbh}^{-0.45222} \tag{12}$$

Furthermore, when applying Chave *et al.* [39], we extracted ρ values from the global wood density database [40,41] and subsequently calculated a mean ρ value, which was then used for the graphical display of this model.

Table 3. Number of sites, sample trees and dbh ranges (cm) of previously developed models tested on our data.

Tree Section	Author	Model	No. of Sites	No. of Trees	dbh Range (cm)	Species
Above-Ground	Mugasha *et al.* [21]	$B = 0.1027 \times \text{dbh}^{2.4798}$	4	167	1.1–110	60
	Mugasha *et al.* [21]	$B = 0.0763 \times \text{dbh}^{2.2046} \times \text{ht}^{0.4918}$	4	167	1.1–110	60
	Ryan *et al.* [4] [a]	$C = -3.629 + 2.601 \times \log(\text{dbh})$	1	29	5–73	6
	Chidumayo [18]	$B = -2.5265 + 2.5553 \times \log(\text{dbh})$	1	113	2–39	19
	Chave *et al.* [39] [b]	$B = 0.0673 \times \left(\rho \times \text{dbh}^2 \times \text{ht}\right)^{0.976}$	58	4004	5–180	Unclear
Below-Ground	Mugasha *et al.* [21]	$B = 0.2113 \times \text{dbh}^{1.9838}$	4	80	3.3–95	60
	Mugasha *et al.* [21]	$B = 0.1766 \times \text{dbh}^{1.7844} \times \text{ht}^{0.3434}$	4	80	3.3–95	60
	Ryan *et al.* [4] [a]	$C = -3.370 + 2.262 \times \log(\text{dbh})$	1	23	5–72	6
	Chidumayo [18]	$B = -1.9439 + 2.1712 \times \log(\text{dbh})$	1	12	4–35	19

[a] Carbon converted to biomass as $B = \left(\frac{C}{47}\right) \times 100$; [b] ρ (g/cm^3) is the species-specific mean wood specific gravity (g/cm^3) extracted from the global wood density database [40,41].

3. Results

The mean RS-ratio of the 41 trees sampled both above- and belowground was 0.47 (Table 4). No significant differences in RS-ratios were found between sites (p = 0.8684, F = 0.2400). The RS-ratio decreased nonlinearly with increasing dbh (Figure 4).

Table 4. Mean, standard deviation (STD) and range of root to shoot ratios (RS-ratio) over sites.

Site	No. of Trees	Mean	Min	Max	STD
Mtangatanga	12	0.49	0.32	1.15	0.25
Kongwe	10	0.44	0.22	0.91	0.22
Mua-livulezi	9	0.51	0.27	0.92	0.27
Tsamba	10	0.44	0.21	0.78	0.16
All	41	0.47	0.18	1.15	0.22

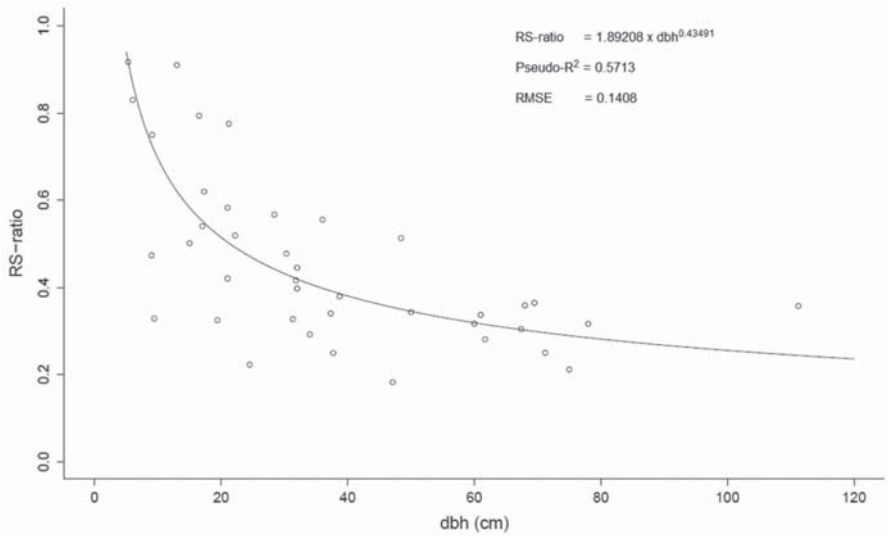


Figure 4. Root to shoot ratio (RS-ratio) *vs.* diameter at breast height (dbh). The dots represent observations for individual trees and the line represents the fitted nonlinear model.

For aboveground biomass, all models, except Model 4, had significant parameter estimates and appropriate performance criteria, *i.e.*, none of the models had MPE% values significantly different from zero ($p > 0.05$) (Table 5). Among these models, Model 2 with dbh and ht as independent variables had the smallest AIC value. The pseudo-R^2 values for all models ranged from 0.93 to 0.97. For belowground biomass, Model 1 was the only one where all parameter estimates were significant. Covariance matrices for all models with significant parameter estimates in Table 5 are shown in Table A3.

Among the models with significant parameter estimates for twigs, branches and merchantable stem biomass (Table 6), Models 1, 2 and 2, respectively, provided the smallest AIC values. The pseudo-R^2 values for the twigs, branches and merchantable stem models with significant parameter estimates were 0.82, 0.91–0.92 and 0.77–0.88, respectively.

We further evaluated the above- and belowground biomass models over sites and dbh classes (Table 7). None of the tested models produced MPE values significantly different from zero ($p > 0.05$) overall or for any site. However, a significant MPE was observed for dbh class 0–20 cm for Model 3. For the aboveground biomass models, MPE% values ranged from 0.4% to 15.1% while for the belowground biomass model, the MPE% values ranged from 2.1% to 3.9%.

Finally, we tested previously developed models (Table 3) on our dataset (Table 8). The MPE% when applying the aboveground biomass models developed by Mugasha *et al.* [21], Ryan *et al.* [4],

Chidumayo [18] and Chave *et al.* [39] ranged from 2.8 to 30.8 (under prediction). The above- and belowground biomass models developed by Chidumayo [18] generally produced the lowest MPE% values, *i.e.*, 2.8% and −4.7%, respectively. Figures 5 and 6 display above- and belowground biomass over dbh for some of the models developed in the current study and some from the previous studies.

Table 5. Model parameters and performance criteria of above- and belowground biomass models.

Component	Model No.	No. of Trees	Model	Pseudo-R^2	RMSE		MPE		AIC
					(kg)	(%)	(kg)	(%)	
Aboveground	1	74	$0.21691 \times dbh^{2.318391}$	0.93	751.2	60.6	−22.2	−1.8	981.89
	2	**74**	$\mathbf{0.103685 \times dbh^{1.921719} \times ht^{0.844561}}$	**0.97**	**426.6**	**34.4**	**−19.8**	**−1.6**	**954.27**
	3	74	$0.290457 \times dbh^{2.283998} \times \rho^{0.443619}$	0.94	923.1	74.5	−53.2	−4.3	977.05
	4	74	$0.129899 \times dbh^{1.90203} \times ht^{0.828647} \times \rho^{0.296271NS}$	0.97	542.1	43.7	−37.1	−3.0	952.35
Belowground	1	41	$\mathbf{0.284615 \times dbh^{1.992658}}$	**0.94**	**161.7**	**30.7**	**−4.8**	**−0.9**	**481.23**
	2	41	$0.224132 \times dbh^{1.899061} \times ht^{0.222554NS}$	0.94	169.7	32.2	−7.4	−1.4	481.73
	3	41	$0.415451 \times dbh^{1.933905} \times \rho^{0.465663NS}$	0.94	175.4	33.3	−5.0	−1.0	481.74
	4	41	$0.350488 \times dbh^{1.732452} \times ht^{0.395127NS} \times \rho^{0.836154NS}$	0.94	170.2	32.3	−6.5	−1.2	479.66

^{NS} Parameter estimate not significant ($p > 0.05$). Note: biomass in kg, dbh in cm, ht in m, and ρ in g/cm^3, Bold: best model according to AIC (Akaike Information Criterion (AIC)).

Table 6. Model and performance criteria for twigs, branches and merchantable stem biomass models.

Component	Model No.	No. of Trees	Model	Pseudo-R^2	RMSE		MPE		AIC
					(kg)	(%)	(kg)	(%)	
Twigs	**1**	**72**	$\mathbf{0.07239 \times dbh^{1.858897}}$	**0.82**	**39.3**	**62.5**	**−0.8**	**−1.2**	**602.84**
	2	72	$0.070224 \times dbh^{1.598204} \times ht^{0.384139NS}$	0.84	37.8	60.1	−1.1	−1.7	634.09
	3	72	$0.130116 \times dbh^{1.764995} \times \rho^{0.687581NS}$	0.83	44.3	70.4	−1.4	−2.2	616.35
	4	72	$0.109969^{NS} \times dbh^{1.446488} \times ht^{0.549901} \times \rho^{0.951968}$	0.83	40.2	63.9	−1.5	−2.4	631.63
Branches	1	74	$0.137316 \times dbh^{2.328104}$	0.91	659.8	80.9	−25.4	−3.1	973.58
	2	**74**	$\mathbf{0.051157 \times dbh^{2.161115} \times ht^{0.598879}}$	**0.92**	**565.2**	**69.3**	**−22.6**	**−2.8**	**933.50**
	3	74	$0.123375 \times dbh^{2.379626} \times \rho^{0.30643NS}$	0.91	789.0	96.8	−47.5	−5.8	946.79
	4	74	$0.059792 \times dbh^{2.150762} \times ht^{0.582627} \times \rho^{0.20053NS}$	0.92	692.2	84.9	−39.3	−4.8	934.74
Merchantable	1	60	$0.145576 \times dbh^{2.116265}$	0.77	299.9	66.9	−1.2	−0.3	773.89
Stem	**2**	**60**	$\mathbf{0.030811 \times dbh^{1.572422} \times ht^{1.345696}}$	**0.88**	**249.9**	**55.8**	**5.3**	**1.2**	**737.16**
	3	60	$0.221213 \times dbh^{2.059249} \times \rho^{0.544483}$	0.79	330.2	73.7	−13.9	−3.1	772.48
	4	60	$0.039194 \times dbh^{1.554911} \times ht^{1.329832} \times \rho^{0.36722NS}$	0.88	256.2	57.2	2.6	0.6	737.75

^{NS} Parameter estimate not significant ($p > 0.05$). Note: biomass in kg, dbh in cm, ht in m, and ρ in g/cm^3. Bold: best model according to AIC.

Table 7. Mean prediction errors (MPE) of the models over study sites and dbh classes.

Component	Model No.	Variable		No. of Trees	Observed	Predicted	MPE	
					(kg)	(kg)	(kg)	(%)
Aboveground	1	Site	Mtangatanga	20	1465.5	1606.2	−140.8	−9.6
			Kongwe	18	1195.4	1200.6	−5.2	−0.4
			Mua-livulezi	18	1316.7	1231.5	85.3	6.5
			Tsamba	18	956.3	882.4	73.8	7.7
		dbh class (cm)	0–20	21	79.7	93.4	−13.6	−17.1
			21–40	35	545.6	584.1	−38.5	−7.1
			41–60	6	2225.4	1911.6	313.9	14.1
			>60	12	4801.5	4826.0	−24.5	−0.5
		All		74	1239.7	1240.3	−0.6	−0.1

Table 7. *Cont.*

Component	Model No.	Variable		No. of Trees	Observed (kg)	Predicted (kg)	MPE (kg)	MPE (%)
Aboveground	2	Site	Mtangatanga	20	1465.5	1424.1	41.4	2.8
			Kongwe	18	1195.4	1146.8	48.6	4.1
			Mua-livulezi	18	1316.7	1337.8	−21.1	−1.6
			Tsamba	18	956.3	1032.2	−75.9	−7.9
		dbh class (cm)	0–20	21	79.7	86.3	−6.5	−8.2
			21–40	35	545.6	578.5	−32.9	−6.0
			41–60	6	2225.4	2111.9	112.5	5.1
			>60	12	4801.5	4754.1	47.4	1.0
		All		74	1239.7	1240.3	−0.6	−0.1
	3	Site	Mtangatanga	20	1465.5	1686.3	−220.8	−15.1
			Kongwe	18	1195.4	1133.3	62.1	5.2
			Mua-livulezi	18	1316.7	1205.4	111.4	8.5
			Tsamba	18	956.3	889.3	67.0	7.0
		dbh class (cm)	0–20	21	79.7	96.3	−16.5	−20.7 *
			21–40	35	545.6	585.6	−40.0	−7.3
			41–60	6	2225.4	1936.4	289.1	13.0
			>60	12	4801.5	4807.8	−6.3	−0.1
		All		74	1239.7	1240.9	−1.2	−0.1
Belowground	1	Site	Mtangatanga	12	795.9	777.0	18.9	2.4
			Kongwe	10	386.3	401.4	−15.1	−3.9
			Mua-livulezi	9	427.7	418.9	8.8	2.1
			Tsamba	10	435.3	450.9	−15.5	−3.6
		dbh class (cm)	0–20	12	42.1	46.1	−4.0	−9.5
			21–40	16	240.3	259.3	−19.0	−7.9
			41–60	4	821.3	736.9	84.4	10.3
			>60	9	1553.4	1551.9	1.5	0.1
		All		41	527.2	527.2	−0.0	−0.0

* MPE is significantly different from zero ($p < 0.05$).

Table 8. Mean prediction error (MPE) of previously developed models.

Component	Model	Independent Variable(s)	No. of Trees	Observed (kg)	Predicted (kg)	MPE (kg)	MPE (%)
Aboveground	Mugasha *et al.* [21]	dbh	74	1239.7	1135.7	104.0	8.4
	Mugasha *et al.* [21]	dbh, ht	74	1239.7	1076.7	163.0	13.2 **
	Ryan *et al.* [4]	dbh	74	1239.7	1068.8	170.9	13.8 *
	Chidumayo [18]	dbh	74	1239.7	1205.6	34.1	2.8
	Chave *et al.* [39]	dbh, ρ, ht	74	1239.7	953.7	286.1	23.1 ***
Belowground	Mugasha *et al.* [21]	dbh	41	527.2	377.5	149.7	28.4 ***
	Mugasha *et al.* [21]	dbh, ht	41	527.2	364.8	162.4	30.8 ***
	Ryan *et al.* [4]	dbh	41	527.2	426.9	100.3	19.0 ***
	Chidumayo [18]	dbh	41	527.2	551.9	−24.7	−4.7

* MPE is significantly different from zero at ($p < 0.05$); ** MPE is significantly different from zero at ($p < 0.01$) and *** MPE is significantly different from zero at ($p < 0.001$).

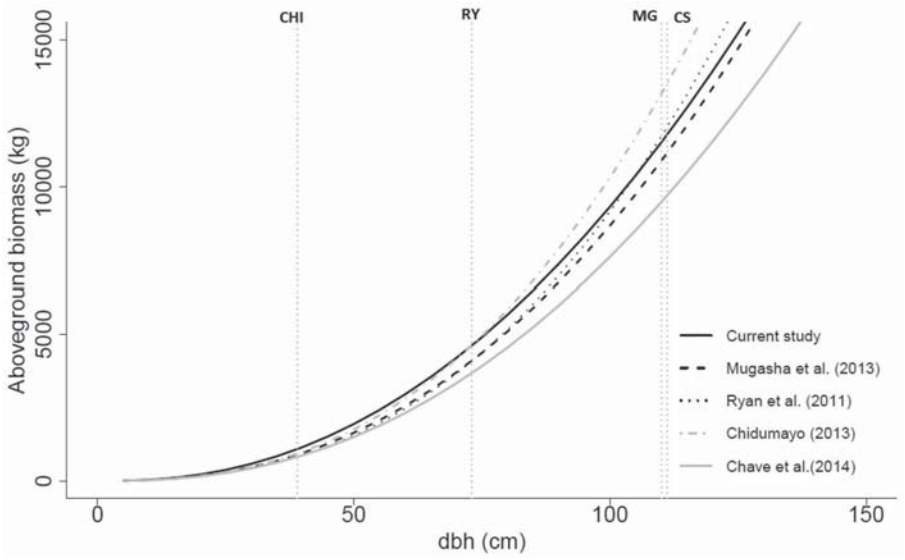


Figure 5. Aboveground biomass (dry weight) over dbh based on the general models developed in this study (with dbh and ht as independent variables), by Mugasha *et al.* [21] (with dbh and ht as independent variables), by Ryan *et al.* [4] (with dbh as only independent variable), by Chidumayo [18] (with dbh only as independent variable) and by Chave *et al.* [39] (with dbh, ρ, and ht as independent variables). CHI, RY, MG and CS are the maximum dbh values for the data used in the models developed by Chidumayo [18], Ryan *et al.* [4], Mugasha *et al.* [21] and current study, respectively.

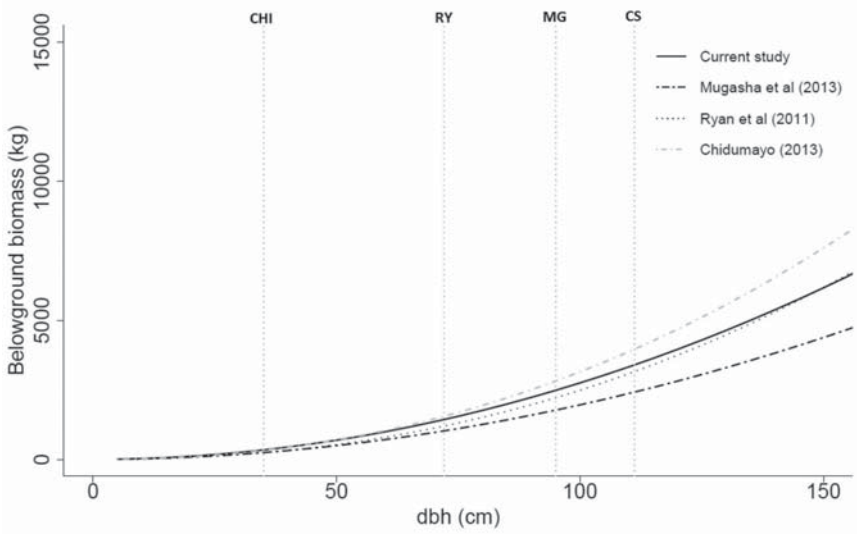


Figure 6. Belowground biomass (dry weight) over dbh based on the general models developed in this study, by Mugasha *et al.* [21], by Ryan *et al.* [4] and by Chidumayo [18]. All models had dbh as the only independent variable. CHI, RY, MG and CS are the maximum dbh values for the data used in the models developed by Chidumayo [18], Ryan *et al.* [4], Mugasha *et al.* [21] and current study, respectively.

4. Discussion

Capturing a wide range of natural variability in factors influencing tree growth, such as soil types, temperature, and rainfall, is important for developing robust biomass models [21,39]. The modelling dataset for this study was collected from sites located in all the three regions of Malawi, *i.e.*, north,

central and south (Figure 1). Application of sound sampling procedures when selecting sample trees is also critical in development of models as it may help in reducing the probability of biases. In this study, we selected trees based on information from forest inventories for each site done prior to tree selection. We also included a number of very large trees to avoid extrapolation beyond the data ranges as much as possible. Our sample trees comprised 33 out of the 139 tree species identified during the forest inventories for the four study sites. Compared to most previous studies our dataset included a large number of tree species. The modelling dataset of Mugasha *et al.* [21] comprised 60 tree species, while those of Chidumayo [18] and Ryan *et al.* [4] had 19 and 6 species, respectively. Although the proportion of the number of tree species in our dataset is relatively low compared to total number of tree species reported in miombo woodlands [22–27], the information from the prior forest inventories ensured that the most common species were represented, in addition to species selected randomly among the remaining less frequent species.

Leaves and fine roots were excluded from our biomass sampling. Leaves were excluded because most of the trees had started to shed leaves when we carried out destructive sampling. This is a common challenge for biomass studies in seasonally dry forests, as acknowledged by Chave *et al.* [39]. A recent study on miombo woodlands in Mozambique by Mate *et al.* [17] found that leaves comprised only 3% of the total aboveground biomass during the peak leaf season. Such a number would probably be a good estimate of how much aboveground biomass is missing in our data.

Root parts with diameter < 1 cm were also excluded from the biomass data, mainly to reduce workload. Similar root parts (diameter < 2 cm) were also excluded in the Chidumayo [18] and Ryan *et al.* [4] models. Chidumayo [18] analysed cumulative root biomass *vs.* diameter and found that the trend levelled off considerably around a diameter of 2 cm. Thus, the underestimation of the belowground biomass in our data is small, especially since we used a diameter threshold of 1 cm.

RS ratio is an important alternative for estimating belowground tree biomass in cases where allometric models are not available (see [14,45]). The mean RS ratio observed here (0.47, Table 4) is higher than the value reported by Mugasha *et al.* [21] and Ryan *et al.* [4] for miombo woodlands (0.40 and 0.42 respectively). However, Chidumayo [18] reported a mean value of 0.54 which is comparatively higher. Since the RS ratio decreases with increasing dbh (Figure 4), the mean value depends on the size distribution of the trees. The high mean RS-ratio found by Chidumayo [18] is likely to result from the relatively smaller tree sizes in his dataset (see Table 3) rather than a higher proportion of root biomass in trees from miombo woodlands in Zambia. Mean RS ratios are frequently used to estimate belowground biomass (e.g., [46]). However, by using a fixed mean RS ratio for a relationship that most probably is non-linear (Figure 4), a bias will be introduced (see e.g., [47]). Therefore, application of mean RS-ratios to estimate belowground should be done with caution.

Among the aboveground biomass models with significant parameter estimates, Model 2 (with dbh and ht as independent variables) had the smallest AIC value (Table 5). Models with both dbh and ht as independent variables had a better fit than those with only dbh (e.g., [21,39]). Inclusion of ρ as independent variable (Model 3), in place of ht (Model 2), did not improve aboveground biomass prediction. This could be attributed to the fact that the ρ values were not obtained directly from the sampled trees, but from the global wood density database [40,41]. The model with the smallest AIC value (Model 2) is similar (pseudo-R^2 = 0.97) to the models developed by Mugasha *et al.* [21] (pseudo-R^2 = 0.95), Chidumayo [18] (R^2 = 0.98) and Ryan *et al.* [4] (adjusted-R^2 = 0.93).

Although Model 2 is generally considered the best aboveground biomass model, it should be noted that Models 1 and 3 can still be applied during forest inventories in cases where ht is lacking or considered as inaccurate. Application of Model 2 requires measuring both dbh and ht during a forest inventory. In such cases, ht for individual trees is usually predicted based on dbh-ht models developed from sample trees collected from the study site because measuring ht in all trees is too time consuming. However, measuring tree height is also prone to errors especially in closed-canopy forests, due to differences in crown shapes and the difficulty of observing the top of the tree crown [48,49].

Among the belowground biomass models, the only viable model, *i.e.*, with significant parameter estimates, was the one with dbh as an independent variable (Model 1). The fit of this model is similar (pseudo-R^2 = 0.94) to that of the models developed by Mugasha *et al.* [21] (pseudo-R^2 = 0.92), Chidumayo [18] (R^2 = 0.95) and Ryan *et al.* [4] (adjusted-R^2 = 0.94).

Proper implementation of the REDD+ mechanism in participating countries, including Malawi, requires biomass estimates to be accompanied with an estimate of uncertainty (see [30]). Uncertainty of biomass estimates is usually computed from error estimates of model parameters for the employed biomass models [34]. In Table A3, we have therefore presented the covariance matrices of model parameters for all the valid models in Table 5 to enable potential users to estimate uncertainty in biomass and carbon quantities during national forest inventories and monitoring, reporting and verification systems under the REDD+ mechanism (see [30,34]).

Tree component biomass models, *i.e.*, models for twigs, branches and merchantable stem, may be useful when planning commercial extraction of timber or quantification of biomass for domestic fuelwood or charcoal production (see [37]). All tree component models with significant parameter estimates produced MPE% values not significantly different from zero (see Table 6). This is an indication of appropriate performance.

The evaluation of the developed above- and belowground models on our own data showed that no models produced MPE% significantly different from zero for any site (Table 7), thus indicating that the models can be applied over a wide range of conditions. The trend was the same over dbh classes except for the smallest dbh class under Model 3. It should be noted that the magnitude of MPE% seen over sites in Table 7 is the kind of error that should be expected across sites if we were to apply our models across Malawi, e.g., in a national forest inventory.

The previously developed models from neighbouring countries resulted in large prediction errors significantly different from zero when applied to our data (Table 8). Exceptions were observed in the aboveground biomass models of Mugasha *et al.* [21] and the above- and belowground models of Chidumayo [18]. For the recently developed pan tropical aboveground biomass model of Chave *et al.* [39], the prediction error was also large (23.1% underestimation). The generally large prediction errors when appliying the previously developed models are, of course, not surprising since they are applied outside their respective data ranges. These results, however, also demonstrate the importance of developing local models.

5. Conclusions

We developed general above- and belowground, stem, branch and twig biomass models for the miombo woodlands of Malawi. Our models can be used over a wide range of geographical and ecological conditions in Malawi. The generally large prediction errors seen when applying previously developed models from neighbouring countries to our data demonstrated the importance of developing local models.

Acknowledgments: The first author of the study was mainly funded by the Norwegian Government through the Norwegian Quota Scheme. Special thanks should also go to the Norwegian University of Life Sciences and World Wildlife Fund (WWF) under the Russel E. Train Education for Nature Fellowship for funding the fieldwork. We would also like to thank the Forest Research Institute of Malawi for facilitating the grant of a research permit from the National Research Council of Malawi. We thank Wilson Mugasha, Norwegian University of Life Sciences, for his guidance during initial stages of fieldwork. Our gratitude is also extended to Devey Lekapo, Happy Kachamba, Humphrey Chapama, Boyd Zulu, Kola Daitoni and Steve Hara for working tirelessly during data collection. Finally, we also like to say thank you to District Forest Officers for Mzimba, Dowa, Dedza and Neno districts for their cooperation and guidance during data collection.

Author Contributions: Daud Jones Kachamba was responsible for developing the research design, data collection and analyses and manuscript development. Tron Eid guided the development of the research design and analyses, and participated in manuscript development. Terje Gobakken participated in manuscript development.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

Table A1. Dataset used for model development.

Site	Scientific Name	Local Name	dbh (cm)	ht (m)	Total Above-Ground Biomass (kg)	Total Below-Ground Biomass (kg)	Use
Mtangatanga	*Brachystegia glaucescens* Burtt Davy & Hutch.	Musani	6.0	4.4	15.5	17.8	a,b
	Erica benguelensis (Welw. ex Engl.) E.G.H. Oliv.	Msankhanya	9.4	4.0	18.7	6.2	a,b
	Erica benguelensis (Welw. ex Engl.) E.G.H. Oliv.	Msankhanya	13.0	6.0	31.6	N/A	a
	Uapaca kirkiana Müll. Arg.	Msuku	15.0	5.0	73.4	36.8	a,b
	Julbernadia paniculata (Benth.) Troupin	Mtondo	16.5	8.0	106.0	84.2	a,b
	Isoberlinia angolensis (Benth.) Hoyle & Brenan	Kabale	19.4	12.0	266.3	86.5	a,b
	Brachystegia boehmii Taub.	Mombo	20.2	8.0	194.0	N/A	a
	Uapaca kirkiana Müll. Arg.	Msuku	21.0	7.8	175.2	N/A	a
	Brachystegia utilis Burtt Davy & Hutch.	Nzale	23.0	10.0	278.8	N/A	a
	Brachystegia taxifolia Burtt Davy & Hutch.	Mchinji	23.6	10.0	271.8	N/A	a
	Julbernadia paniculata (Benth.) Troupin	Mtondo	27.3	10.0	412.2	N/A	a
	Isoberlinia angolensis (Benth.) Hoyle & Brenan	Kabale	28.6	11.0	445.0	N/A	a
	Brachystegia spiciformis Benth.	Chumbe	36.0	12.0	921.1	N/A	a
	Brachystegia longifolia Benth.	Sanga	36.0	10.8	1173.0	511.1	a,b
	Brachystegia glaucescens Burtt Davy & Hutch.	Musani	48.4	14.5	1847.1	947.1	a,b
	Brachystegia glaucescens Burtt Davy & Hutch.	Musani	50.0	16.0	2160.7	741.5	a,b
	Brachystegia spiciformis Benth.	Chumbe	60.0	18.0	3935.0	1247.0	a,b
	Brachystegia glaucescens Burtt Davy & Hutch.	Musani	68.0	10.5	3233.2	1160.7	a,b
	Brachystegia glaucescens Burtt Davy & Hutch.	Musani	78.0	17.0	4952.1	1567.2	a,b
	Brachystegia glaucescens Burtt Davy & Hutch.	Musani	111.2	18.0	8798.5	3144.9	a,b

Table A1. *Cont.*

Site	Scientific Name	Local Name	dbh (cm)	ht (m)	Total Above-Ground Biomass (kg)	Total Below-Ground Biomass (kg)	Use
Kongwe	*Julbernadia paniculata* (Benth.) Troupin	Mtondo	9.0	5.0	23.4	11.1	a,b
	Uapaca kirkiana Müll. Arg.	Msuku	10.2	5.0	26.2	N/A	a
	Brachystegia spiciformis Benth.	Chumbe	13.0	6.0	94.9	86.4	a,b
	Brachystegia manga De Wild.	Mpapa/Bovo	17.3	10.0	138.4	85.7	a,b
	Julbernadia paniculata (Benth.) Troupin	Mtondo	21.0	11.0	264.9	154.3	a,b
	Brachystegia utilis Burtt Davy & Hutch.	Nzale	24.5	9.0	368.6	82.1	a,b
	Monotes africanus A. DC.	Mkalakate	27.2	10.0	437.2	N/A	a
	Brachystegia spiciformis Benth.	Chumbe	27.5	9.3	499.9	N/A	a
	Brachystegia boehmii Taub.	Mombo	30.8	14.0	636.5	N/A	a
	Uapaca kirkiana Müll. Arg.	Msuku	34.0	12.0	689.6	201.4	a,b
	Parinari curatellifolia Planch. ex Benth.	Muula	37.3	12.0	839.7	285.7	a,b
	Brachystegia manga De Wild.	Mpapa/Bovo	38.1	17.0	1445.1	N/A	a
	Cussonia arborea Hochst. ex A. Rich.	Mbwabwa	38.4	7.0	346.6	N/A	a
	Brachystegia spiciformis Benth.	Chumbe	38.7	14.0	1298.4	492.7	a,b
	Brachystegia manga De Wild.	Mpapa/Bovo	52.0	22.0	2534.0	N/A	a
	Brachystegia manga De Wild.	Mpapa/Bovo	61.7	17.0	4400.3	1235.3	a,b
	Brachystegia spiciformis Benth.	Chumbe	71.2	18.0	4914.1	1228.4	a,b
	Erythrina abyssinica Lam. ex DC.	Muwale	75.7	12.0	2559.1	N/A	a
Mua-livulezi	*Markhamia obtusifolia* (Baker) Sprague	Msewa	5.3	5.0	6.3	5.8	a,b
	Combretum apiculatum Sond.	Kakunguni	6.0	3.0	15.0	12.4	a,b
	Bauhinia petersiana Bolle	Mphandula	9.1	5.2	36.5	27.4	a,b
	Bauhinia thonningii Schum.	Msekese	13.0	5.0	61.6	N/A	a
	Diplorhynchus condylocarpon (Müll. Arg.) Pichon	Thombozi	16.2	10.4	84.7	N/A	a
	Annona senegalensis Pers.	Mpoza	17.0	8.0	83.9	45.3	a,b
	Markhamia obtusifolia (Baker) Sprague	Msewa	20.0	9.3	142.8	N/A	a
	Pterocarpus rotundifolius (Sond.) Druce	M'balitsa	22.0	8.5	170.1	N/A	a
	Albizia versicolor Welw. ex Oliv.	Mtangatanga	23.2	10.4	214.6	N/A	a
	Bauhinia petersiana Bolle	Mphandula	27.4	7.9	366.6	N/A	a
	Diplorhynchus condylocarpon (Müll. Arg.) Pichon	Thombozi	31.8	13.0	407.3	169.5	a,b
	Acacia galpinni Burtt Davy	Mgundanjira	37.7	14.0	932.3	232.6	a,b
	Pterocarpus rotundifolius (Sond.) Druce	M'balitsa	42.0	15.0	959.6	N/A	a
	Brachystegia spiciformis Benth.	Chumbe	47.1	15.5	1916.1	349.5	a,b
	Brachystegia bussei Harms	Mtwana	61.0	20.5	3970.5	1337.1	a,b
	Pseudolachnostaylis maprouneifolia Pax	Msolo	61.0	15.6	2994.0	N/A	a
	Brachystegia spiciformis Benth.	Chumbe	69.5	21.0	4580.6	1669.5	a,b
	Brachystegia bussei Harms	Mtwana	81.7	22.0	6758.4	N/A	a

Table A1. *Cont.*

Site	Scientific Name	Local Name	dbh (cm)	ht (m)	Total Above-Ground Biomass (kg)	Total Below-Ground Biomass (kg)	Use
Tsamba	*Uapaca kirkiana* Müll. Arg.	Msuku	8.4	6.5	18.0	N/A	a
	Brachystegia floribunda Benth.	Tsamba	9.5	9.0	37.5	N/A	a
	Pseudolachnostaylis maprouneifolia Pax	Msolo	12.4	7.0	105.9	N/A	a
	Brachystegia floribunda Benth.	Tsamba	19.9	8.0	287.6	N/A	a
	Parinari excelsa Sabine	Mpembu	21.0	10.0	205.1	86.2	a,b
	Julbernadia globiflora (Benth.) Troupin	Kachumbe	21.0	10.0	223.0	N/A	a
	Brachystegia spiciformis Benth.	Chumbe	21.2	12.0	271.7	210.7	a,b
	Uapaca sansibarica Pax	Msokolowe	22.2	10.0	212.3	110.2	a,b
	Pericorpsis angolensis (Baker) Meeuwen	Muwanga	28.4	8.0	440.2	249.4	a,b
	Uapaca kirkiana Müll. Arg.	Msuku	30.3	11.0	581.9	277.7	a,b
	Pterocarpus angolensis DC.	Mlombwa	31.3	17.0	667.7	218.5	a,b
	Brachystegia utilis Burtt Davy & Hutch.	Nzale	32.0	14.0	867.6	385.8	a,b
	Uapaca sansibarica Pax	Msokolowe	32.0	12.0	445.9	177.3	a,b
	Brachystegia floribunda Benth.	Tsamba	34.5	12.0	683.0	N/A	a
	Julbernadia globiflora (Benth.) Troupin	Kachumbe	37.5	18.0	1014.0	N/A	a
	Faurea speciosa (A. Rich.) Pic.Serm.	Chisese	38.7	16.0	693.7	N/A	a
	Julbernadia globiflora (Benth.) Troupin	Mchenga	67.4	21.0	4591.2	1397.5	a,b
	Brachystegia spiciformis Benth.	Chumbe	75.0	25.0	5866.2	1240.0	a,b

Note: a = aboveground only; a, b = both above-and belowground; N/A = not applicable.

Table A2. DF-ratio for merchantable stem, branches, twigs, roots, and root crown in the four study sites.

Site	Aboveground									Belowground						
	No. of Trees	Merchantable Stem			Branches			Twigs			No. of Trees	Roots			Root crown		
		Mean	Min	Max	Mean	Min	Max	Mean	Min	Max		Mean	Min	Max	Mean	Min	Max
Mtangatanga	20	0.61	0.48	0.77	0.57	0.44	0.73	0.51	0.30	0.66	12	0.58	0.46	0.69	0.59	0.53	0.67
Kongwe	18	0.56	0.36	0.71	0.54	0.35	0.65	0.52	0.37	0.65	10	0.58	0.48	0.72	0.57	0.39	0.67
Mua-livulezi	18	0.60	0.54	0.69	0.61	0.54	0.67	0.56	0.44	0.66	9	0.60	0.49	0.70	0.59	0.46	0.68
Tsamba	18	0.59	0.45	0.70	0.58	0.45	0.67	0.53	0.38	0.65	10	0.52	0.45	0.63	0.57	0.51	0.61
All	74	0.59	0.36	0.77	0.58	0.35	0.73	0.53	0.30	0.66	41	0.57	0.45	0.73	0.58	0.39	0.68

Note: The number of sample trees for Kongwe for the twigs section is 16 instead of 18 because two trees did not have twigs.

Table A3. Covariance matrices of all valid above- and belowground models.

Component	Model No.	Model	Variable	Covariance Matrix		
Belowground	1	$0.284615 \times dbh^{1.992658}$		intercept	dbh	
			intercept	0.01184		
			dbh	−0.01018	0.00882	
Aboveground	1	$0.21691 \times dbh^{2.318391}$		intercept	dbh	
			intercept	0.00251		
			dbh	−0.00263	0.00278	
	2	$0.103685 \times dbh^{1.921719} \times ht^{0.844561}$		intercept	dbh	ht
			intercept	0.00058		
			dbh	−0.00045	0.00166	
			ht	−0.00129	−0.00094	0.00580
	3	$0.290457 \times dbh^{2.283998} \times \rho^{0.443619}$		intercept	dbh	ρ
			intercept	0.00592		
			dbh	−0.00422	0.00305	
			ρ	0.00557	−0.00346	0.01364

References

1. Chidumayo, E.N.; Gumbo, D.J. *The Dry Forests and Woodlands of Africa: Managing for Products and Services*; Earthscan: London, UK, 2010.
2. Abbot, P.; Lowore, J.; Werren, M. Models for the estimation of single tree volume in four miombo woodland types. *For. Ecol. Manag.* **1997**, *97*, 25–37. [CrossRef]
3. Munishi, P.K.T.; Temu, R.P.C.; Soka, G. Miombo ecosystem in the Lake Rukwa basin, southern Tanzania: Implications for conservation. *J. Ecol. Nat. Environ.* **2010**, *3*, 63–71.
4. Ryan, C.M.; Williams, M.; Grace, J. Above- and belowground carbon stocks in a miombo woodland landscape of Mozambique. *Biotropica* **2011**, *43*, 423–432. [CrossRef]
5. Walker, S.M.; Desanker, P.V. The impact of land use on soil carbon in miombo woodlands of Malawi. *For. Ecol. Manag.* **2004**, *203*, 345–360. [CrossRef]
6. Abdallah, J.M.; Monela, G.G. Overview of miombo woodlands in Tanzania. In *Management of Indigenous Tree Species for Ecosystem Restoration and Wood Production in Semi-Arid Miombo Woodlands in Eastern Africa*; Metla, Ed.; Working Papers of the Finnish Forest Research Institute: Metla, Finland, 2007; Volume 50, pp. 9–23.
7. Government of Malawi. *Forest Resource Mapping Project under the Japanese Grant for the Forest Preservation Programme to the Republic of Malawi*; Ministry of Environment and Climate Change Management: Lilongwe, Malawi, 2012.
8. Government of Malawi. *Malawi State of Environment and Outlook Report*; Ministry of Natural Resources, Energy and Environment: Lilongwe, Malawi, 2010.
9. Government of Malawi. *Malawi Population Data Sheet*; Malawi National Statistics office: Zomba, Malawi, 2012.
10. FAO (Food and Agriculture Organisation of the United Nations). *Global Forest Resources Assessment 2015*; FAO: Rome, Italy, 2015.
11. Akindele, S.O.; LeMay, V.M. Development of tree volume equations for common timber species in the tropical rain forest area of Nigeria. *For. Ecol. Manag.* **2006**, *226*, 41–48. [CrossRef]
12. West, P.W. *Tree and Forest Measurement*, 3rd ed.; Springer International Publishing: Cham, Switzerland, 2015.
13. Chave, J.; Andalo, C.; Brown, S.; Cairns, M.A.; Chambers, J.Q.; Eamus, D.; Fölster, H.; Fromard, F.; Higuchi, N.; Kira, T.; *et al*. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **2005**, *145*, 87–99. [CrossRef] [PubMed]
14. Cairns, M.A.; Brown, S.; Helmer, E.H.; Baumgardner, G.A. Root biomass allocation in the world's upland forests. *Oecologia* **1997**, *111*, 1–11. [CrossRef]
15. Henry, M.; Picard, N.; Trotta, C.; Manlay, R.; Valentini, R.; Bernoux, M.; Saint-André, L. Estimating tree biomass of sub-Saharan African forests: A review of available allometric equations. *Silva Fenn.* **2011**, *45*, 477–569. [CrossRef]
16. Mwakalukwa, E.E.; Meilby, H.; Treue, T. Volume and aboveground biomass models for dry miombo woodland in Tanzania. *Int. J. For. Res.* **2014**, *2014*. [CrossRef]
17. Mate, R.; Johansson, T.; Sitoe, A. Biomass equations for tropical forest tree species in Mozambique. *Forests* **2014**, *5*, 535–556. [CrossRef]
18. Chidumayo, E.N. Estimating tree biomass and changes in root biomass following clear-cutting of *Brachystegia-Julbernardia* (miombo) woodland in central Zambia. *Environ. Conserv.* **2014**, *41*, 54–63. [CrossRef]
19. Chamshama, S.A.O.; Mugasha, A.G.; Zahabu, E. Stand biomass and volume estimation for miombo woodlands at Kitulangalo, Morogoro, Tanzania. *South. Afr. For. J.* **2004**, *200*, 59–70. [CrossRef]
20. Malimbwi, R.E.; Solberg, B.; Luoga, E. Estimation of biomass and volume in miombo woodland at Kitulangalo forest reserve, Tanzania. *J. Trop. For. Sci.* **1994**, *7*, 230–242.
21. Mugasha, W.A.; Eid, T.; Bollandsås, O.M.; Malimbwi, R.E.; Chamshama, S.A.O.; Zahabu, E.; Katani, J.Z. Allometric models for prediction of above- and belowground biomass of trees in the miombo woodlands of Tanzania. *For. Ecol. Manag.* **2013**, *310*, 87–101. [CrossRef]
22. Giliba, R.A.; Boon, E.K.; Kayombo, C.J.; Musamba, E.B.; Kashindye, A.M.; Shayo, P.F. Species composition, richness and diversity in miombo woodlands of Bereku forest reserve, Tanzania. *J. Biodivers.* **2011**, *2*, 1–7.
23. Kalaba, F.K.; Quinn, C.H.; Dougill, A.J.; Vinya, R. Floristic composition, species diversity and carbon storage in charcoal and agriculture fallows and management implications in Miombo woodlands of Zambia. *For. Ecol. Manag.* **2013**, *304*, 99–109. [CrossRef]

24. Mwakalukwa, E.E.; Meilby, H.; Treue, T. Floristic composition, structure, and species associations of dry miombo woodland in Tanzania. *ISRN Biodivers.* **2014**, *2014*, 153278. [CrossRef]
25. Dewees, P.A.; Campbell, B.M.; Katerere, Y.; Sitoe, A.; Cunningham, A.B.; Angelsen, A.; Wunder, S. *Managing the Miombo Woodlands of Southern Africa: Policies, Incentives and Options for the Rural Poor*; World Bank, Program on Forests (PROFOR): Washington, DC, USA, 2010; Volume 2, pp. 57–73.
26. Mugasha, W.A.; Bollandsås, O.M.; Eid, T. Relationships between diameter and height of trees in natural tropical forest in Tanzania. *South. For. J. For. Sci.* **2013**, *75*, 221–237. [CrossRef]
27. Missanjo, E.; Kamanga-Thole, G.; Mtambo, C.; Chisinga, O. Evaluation of natural regeneration and tree species diversity in miombo woodlands in Malawi. *J. Biodivers. Manag. For.* **2014**, *3*. [CrossRef]
28. Brown, S. Measuring carbon in forests: current status and future challenges. *Environ. Pollut.* **2002**, *116*, 363–372. [CrossRef]
29. Njana, M.A.; Eid, T.; Zahabu, E.; Malimbwi, R.E. Procedures for quantification of belowground biomass of three mangrove tree species. *Wetl. Ecol. Manag.* **2015**, *23*, 749–764. [CrossRef]
30. Intergovernmental panel on climate change (IPCC). *Revised Supplementary Methods and Good Practice Guidance Arising from the Kyoto Protocol*; IPCC: Geneva, Switzerland, 2014.
31. Chave, J.; Condit, R.; Aguilar, S.; Hernandez, A.; Lao, S.; Perez, R. Error propagation and scaling for tropical forest biomass estimates. *Philos. Trans. Royal Soc. Lond. B Biol. Sci.* **2004**, *359*, 409–420. [CrossRef] [PubMed]
32. Clark, D.B.; Kellner, J.R. Tropical forest biomass estimation and the fallacy of misplaced concreteness. *J. Veg. Sci.* **2012**, *23*, 1191–1196.
33. Magnussen, S.; Carillo Negrete, O. Model errors in tree biomass estimates computed with an approximation to a missing covariance matrix. *Carbon Balance Manag.* **2015**, *10*. [CrossRef] [PubMed]
34. McRoberts, R.E.; Westfall, J.A. Effects of uncertainty in model predictions of individual tree volume on large area volume estimates. *For. Sci.* **2013**, *60*, 34–42. [CrossRef]
35. Breidenbach, J.; Anton-Fernandez, C.; Petersson, H.; McRoberts, R.E.; Astrup, R. Quantifying the model-related variability of biomass stock and change estimates in the Norwegian National Forest Inventory. *For. Sci.* **2014**, *60*, 25–33. [CrossRef]
36. Hardcastle, P.D. *A Preliminary Silvicultural Classification of Malawi*; Forestry Research Institute of Malawi: Zomba, Malawi, 1978.
37. Kachamba, D.J.; Eid, T. Total tree, merchantable stem and branch volume models for miombo woodlands of Malawi. *South. For.* **2016**. [CrossRef]
38. Mauya, E.W.; Mugasha, W.A.; Zahabu, E.; Bollandsås, O.M.; Eid, T. Models for estimation of tree volume in the miombo woodlands of Tanzania. *South. For.* **2014**, *76*, 209–219. [CrossRef]
39. Chave, J.; Réjou-Méchain, M.; Búrquez, A.; Chidumayo, E.N.; Colgan, M.S.; Delitti, W.B.C.; Duque, A.; Eid, T.; Fearnside, P.M.; Goodman, R.C.; *et al*. Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Chang.Biol.* **2014**, *20*, 3177–3190. [CrossRef] [PubMed]
40. Chave, J.; Coomes, D.A.; Jansen, S.; Lewis, S.L.; Swenson, N.G.; Zanne, A.E. Towards a worldwide wood economics spectrum. *Ecol. Lett.* **2009**, *12*, 351–366. [CrossRef] [PubMed]
41. Zanne, A.E.; Lopez-Gonzalez, G.; Coomes, D.A.; Ilic, J.; Jansen, S.; Lewis, S.L.; Miller, R.B.; Swenson, N.G.; Wiemann, M.C.; Chave, J. Global Wood Density Database. Available online: http://dx.doi.org/10.5061/dryad.234 (accessed on 15 October 2015).
42. SAS Institute Inc. *SAS 9.4 TS Level 1 M1*; SAS Institute Inc.: Cary, NC, USA, 2012.
43. Akaike, H. A new look at the statistical model identification. *Inst. Electr. Electron. Eng. IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
44. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning with Application in R*; Springer: New York, NY, USA, 2013; Volume 103.
45. Mokany, K.; Raison, R.J.; Prokushkin, A.S. Critical analysis of root : Shoot ratios in terrestrial biomes. *Glob. Chang. Biol.* **2006**, *12*, 84–96. [CrossRef]
46. Munishi, P.K.T.; Shear, T.H. Carbon storage in afromontane rainforests of the eastern arc mountains of Tanzania: Their net contribution to atmospheric carbon. *J. Trop. For. Sci.* **2004**, *16*, 78–93.
47. Gertner, G. Prediction Bias and Response Surface Curvature. *For. Sci.* **1991**, *37*, 755–765.
48. Larjavaara, M.; Muller-Landau, H.C. Measuring tree height: A quantitative comparison of two common field methods in a moist tropical forest. *Methods Ecol. Evol.* **2013**, *4*, 793–801. [CrossRef]

49. Hunter, M.O.; Keller, M.; Victoria, D.; Morton, D.C. Tree height and tropical forest biomass estimation. *Biogeosciences* **2013**, *10*, 8385–8399. [CrossRef]

Paper III

remote sensing

MDPI

Article

Biomass prediction using an unmanned aerial vehicle in a tropical woodland

Daud Jones Kachamba[1,2,*], Hans Ole Ørka[1], Terje Gobakken[1], Tron Eid[1] and Weston Mwase[2]

1 Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432, Ås, Norway; e-mails: hans.ole.orka@nmbu.no (H.O.Ø.); terje.gobakken@nmbu.no (T.G.); tron.eid@nmbu.no (T.E.);

2 Department of Forestry, Lilongwe University of Agriculture & Natural Resources, P.O Box 219, Lilongwe, Malawi; e-mails: westmwase@yahoo.co.uk (W.M.)

* Correspondence: daud.kachamba@nmbu.no or dkachamba@gmail.com; Tel.: +47-939-836-91 or +265-995-468-387.

Academic Editor: name

Received: date; Accepted: date; Published: date

Abstract: Application of unmanned aerial vehicles (UAVs) in forest biomass estimation have shown great potential in reducing costs, time and improving the estimates. However, UAVs have never been tested in miombo woodlands. UAV-based biomass estimation rely on the availability of reliable digital terrain models (DTMs). The main objective of this study was to evaluate application of UAVs in biomass prediction and to compare impacts of DTMs generated based on different methods and parameter settings. Biomass was modelled using data acquired on 107 sample plots from a forest reserve in miombo woodlands of Malawi. The results indicated that there are no significant differences (p = 0.985) between tested DTMs except for that based on shuttle radar topography mission (SRTM). A model developed using unsupervised ground filtering based on a grid search approach, had the smallest root mean square error (RMSE) of 46.7% of a mean biomass value of 38.99 Mg ha^{-1}. Amongst the independent variables, maximum canopy height (*Hmax*) was the most frequently selected. In addition, all models included spectral variables incorporating the three colour bands, red, green and blue. The study has demonstrated that UAV acquired image data can be used in biomass estimation in miombo woodlands using automatically generated DTMs.

Keywords: Biomass; digital terrain model; miombo woodlands; unmanned aerial vehicle

PACS: J0101

1. Introduction

The Reducing Emissions from Deforestation and Forest Degradation, plus forest conservation, sustainable management of forest, and enhancement of carbon stocks (REDD+) mechanism has given a financial incentive to developing countries for their efforts in reducing deforestation and forest degradation. According to requirements of the United Nations Framework Convention on Climate Change (UNFCCC), participating countries are supposed to report verified national level carbon estimates to benefit from the mechanism. Therefore, it is expected that each country implementing REDD+ should have a carbon monitoring system capable of collecting reliable data at designated time intervals using consistent methodologies.

Many countries rely on data captured through national forest inventories (NFIs) to estimate country level biomass and carbon stock estimates. The first step in a NFI involves developing a national land cover map displaying different forest strata across the country. Next, sample plot inventories are conducted on permanent sample plots distributed across the country. Allometric models are then applied to estimate average biomass and carbon stocks for sample plots lying within a given stratum. National level carbon stock is then estimated by applying the average biomass and carbon density values across the map with the same forest strata [1]. However, many developing countries, including Malawi, do not have comprehensive NFIs. Malawi is currently in the preparatory phase of implementing the REDD+ mechanism and an NFI is one of the planned activities [2,3].

In Malawi, miombo woodlands constitute 92.4% of the country's total forest area [4]. These woodlands are dominated by *Brachystegia-Julbernardia-Isoberlinia* species and they occur between 5° and 25° south of the equator in the upland plateau ecoregion of eastern and southern Africa within an altitude ranging from 600 – 4200 m. The region has a mean annual rainfall of between 800 – 1400 mm [5]. In Malawi, a relatively small part of these woodlands are in large continuous blocks, instead they are scattered over the landscape in 112 forest reserves as "islands" in cultivated land. The government of Malawi is targeting these forest reserves as potential REDD+ project areas. These forest reserves have variable sizes ranging between 42 and 114780 ha [6] and approximately 50% of these reserves may be characterized as small- to medium-sized (i.e. up to 2240 ha).

Ground-based forest inventories usually form a key component in NFIs. However, comprehensive ground-based inventories are associated with high labour and operational costs hence restrictive to most developing countries [7-9]. This has prompted researchers to search for alternative reliable and cost effective biomass estimation methodologies. A promising approach aimed at reducing labour and operational costs as well as improving the reliability of estimated biomass in NFIs involve combining data from ground-based forest inventories and remote sensing [8,10-13].

For forestry applications, remotely sensed data is mainly sourced from three main systems, namely, airborne laser scanning (ALS), radio detection and ranging (RADAR)(e.g. synthetic aperture radar (SAR)) and optical images (e.g. satellite and aerial) [14,15]. Remote sensing has been widely applied in forestry for several decades in most countries although with various degrees of success due to differences in data types, forest canopy cover, geographical and environmental conditions and methods used [15,16]. For example, estimation of forest biomass based on data from optical systems is usually challenged by clouds, shadows, saturation problems in forests with high biomass density, intra-crown spectral variance, low spectral variability and its two-dimensional (2D) nature [16,17]. RADAR systems are capable of improving biomass estimations due their ability to capture high quality three-dimensional (3D) data in all weather and light conditions. However, biomass estimation using such data is also affected by saturation problems in complex mature forest stands and also have difficulty in distinguishing vegetation types [17]. Data from ALS systems have shown great potential for forest biomass estimations in different forest types including boreal [18], temperate [19-21] and tropical forests [22,23] due to ALS's ability to overcome saturation problems associated with data collected from the other sensors when biomass density is ≥100Mg/ha [13,14,17,24]. However, wide application of ALS data for large-scale forest biomass estimation has been limited due to high data acquisition costs.

On the other hand, application of unmanned aerial vehicles (UAVs) enable acquisition of high quality 3D remote sensing data for estimating forest biomass on small- to medium-sized forests with relatively low costs [25-29]. Relatively little technical expertise is required for operating and acquiring images using UAVs, thus costs associated with hiring airborne image acquisition platforms are reduced. Furthermore, the availability of user-friendly structure from motion (SfM) and stereo-matching software that uses a photogrammetric approach to obtain 3D data reduces the need for hiring image processing services [26,27,30].

Results from recent research on small- to medium-sized boreal forests have demonstrated the potential of using UAVs for estimating forest biomass [27,31]. Application of this technology for REDD+ implementation in Malawi could be an attractive option since a substantial proportion of the potential project areas are small- to medium-sized forest reserves scattered across the country. However, the application of a UAV system for biomass estimation in tropical forests such as miombo woodlands of Malawi first needs to be investigated because the forest structure in miombo woodlands is different from the boreal forests. Miombo woodlands are dry tropical forest woodlands dominated by relatively short trees with highly variable canopy structures [6,32,33]. Boreal forests may have relatively tall trees with an almost uniform size and shape, few species and a relatively simple canopy structure [34,35].

Several previous studies, including Mauya, *et al.* [36], Gregoire, *et al.* [37], Yang and Prince [38] and Fuller, *et al.* [39], among others, have attempted to estimate forest attributes in miombo woodlands using a combination of remote sensing and ground-based inventory data. Amongst the previous studies, Mauya, et al. [36] and Gregoire, et al. [37] are the only studies that utilized 3D ALS data. To the best of our knowledge, there are currently no studies in which 3D remotely sensed data captured by UAVs have been applied for biomass estimation in dry tropical forests such as miombo woodlands.

Successful estimation of forest characteristics from 3D remotely sensed data is conditioned on the ability of the sensor to capture data also from the ground. Thus, the accuracy and precision of the final estimates of forest characteristics are a function of the quality of the digital terrain model (DTM). ALS sensors are generally well suited to collect point clouds describing the ground for generating a DTM even in dense forest areas. The photogrammetric approach can yield very dense point clouds in forested environments, but these data are mostly located in the top of the canopy.

When applying the photogrammetric approach, DTMs may be generated through a two stage process. The first stage involves separating ground from vegetation points (ground filtering). In the second stage, the data from the ground points are interpolated to estimate ground elevation at places where there are no data to get a continuous ground surface [40].

Most studies utilizing the UAVs technology utilize DTMs derived from ALS data because they are regarded as being most accurate and reliable [27,41]. However, due to the high costs associated with acquiring ALS data, it is imperative for researchers utilizing UAVs in developing countries to search for alternative and cost efficient DTM generating approaches. Furthermore, several algorithms for DTM generation are available including Terrascan [42] and Fusion [43]. Most of these algorithms can be parameterized in different ways. Thus, the main objective of this study was to evaluate the application of photogrammetric point cloud data generated from UAV acquired images in biomass estimation (aboveground only) for the miombo woodlands. DTMs generated from the photogrammetric point cloud based on different methods and parameter settings were also compared.

2. Materials and Methods

2.1. Study area

The study was conducted in Muyobe community forest reserve (11° 35′ S, 33° 65′ E, 1169 – 1413 m above sea level), located in Mpherembe traditional authority in Mzimba district in the northern region of Malawi (Figure 1). This forest reserve was established in 2000 to save it from rampant deforestation happening in and around the area due to tobacco farming. The forest reserve is now managed through a committee whose members are representatives from all villages surrounding the forest. The forest reserve is 486 ha, which is common for most small- to medium-sized forest reserves in Malawi. Ferrosols are the dominant soil type in the study area [44]. For the period 1975 – 2005, the mean annual rainfall was 889±146 mm and the mean annual daily minimum and maximum temperatures were 15±1.6° C and 26±0.6° C, respectively (the nearest weather station in Mzimba about 69 km south of the study area). The rainy season lasts from December to April. The study area comprises miombo woodlands with *Julbernadia globiflora, Diplorhychus condylocarpon* and *Combretum zeyheri* as dominant species.

2.2. Data collection

2.2.1. Sampling design and ground reference data collection

Ground reference biomass was based on data from a forest inventory conducted on 107 systematically distributed circular sample plots (radius = 17.84 m, 0.1 ha each) for 15 days from 25th April to 10th May, 2015. Precise registration of the positions of centres for sample plots is very important in remote sensing-assisted forest inventories. In this study, positions of the plot centres were measured with a differential Global Navigation Satellite Systems (dGNSS) unit. The dGNSS unit is comprised of two Topcon legacy- E +40 dual frequency receivers. One of the receivers was used as base station unit and the other as a rover field unit. The receivers observe pseudo-range and carrier phase of both the Global Positioning System (GPS) and the Global Navigation Satellite System (GLONASS). During the study, the baseline between the base station and the rover units was approximately 25 kilometres. The position of the base station was determined using Precise Point Positioning (PPP) with GPS and GLONASS data collected continuously for 24 hours as suggested by Kouba [45]. The rover field unit was placed was placed at the centre of each sample plot on a 2.98 m rod for an average of 33±20 minutes using a one-second logging rate. The recorded plot centre coordinates were post-processed using the RTKLIB software [46] and the results revealed that the maximum deviations for northing, easting and height were 1.16 cm, 3.02 cm and 3.06 cm, respectively.

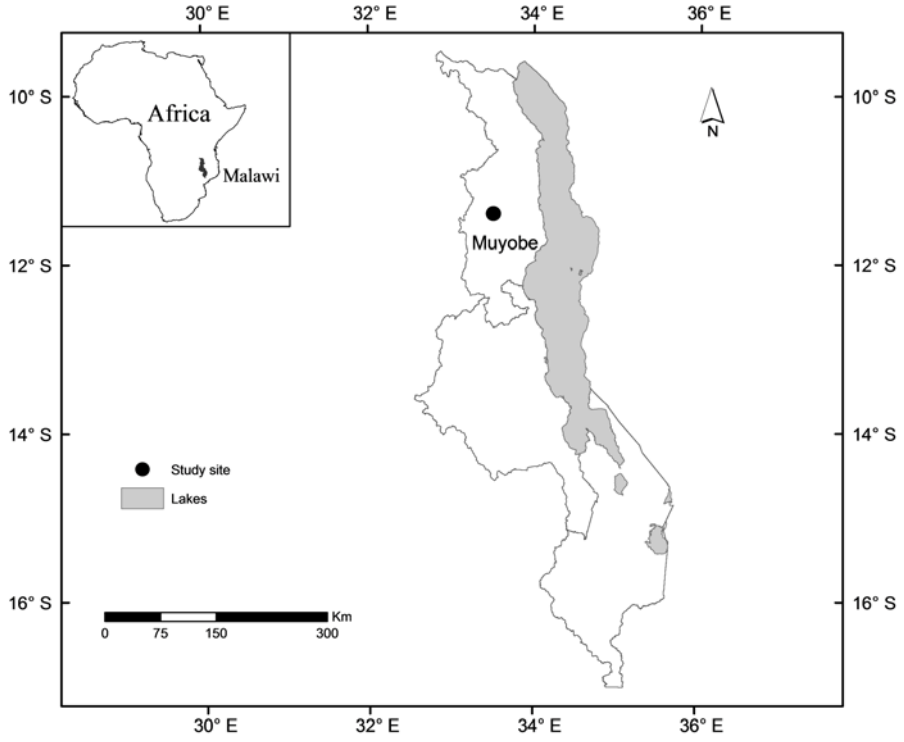


Figure 1. Map of Malawi showing the location of the study site.

On each plot, the following tree variables were recorded: diameter at breast height (*dbh*) (using a calliper or a diameter tape), and scientific name of all trees ≥ 5 cm. Furthermore, total tree height (*ht*) of up to 10 randomly selected sample trees within each plot were measured using a Vertex hypsometer.

For each tree in respective sample plots, we calculated biomass by using a model developed by Kachamba, *et al.* [47] with *dbh* and *ht* as independent variables:

$$Biomass = 0.103685 \times dbh^{1.921719} \times ht^{0.844561} \quad (1)$$

Before calculating biomass, we predicted the *ht* of trees whose heights were not measured in the respective sample plots, using a height-diameter model developed from the sample trees from all sample plots:

$$ht = 1.3 + \exp(3.75876 - 6.01583 \times dbh^{-0.42991}) \quad (2)$$

The performance criteria for the *ht* model were as follows: pseudo-R^2 of 0.65, root mean square error (RMSE) of 1.9 m and a mean prediction error (MPE) of 0.1 m. The pseudo-R^2, RMSE and MPE were calculated as follows:

$$\text{pseudo-}\ R^2 = 1 - \left(\frac{SSR}{CSST}\right) \quad (3)$$

where *SSR* is the sum of squared residuals and *CSST* is the corrected total sum of squares.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(yi - \hat{y}_i)^2}{n}} \quad (4)$$

$$\text{MPE} = \frac{\sum_{i=1}^{n}(yi-\hat{y}_i)}{n} \quad (5)$$

y_i is the ground reference biomass for plot i and $\hat{y}_i$ is the predicted biomass for plot i, respectively and n is the number of sample plots.

Biomass (Mg ha^{-1}), basal area (m^2 ha^{-1}) and number of stems (ha^{-1}) of the respective sample plots were calculated by summing up the individual tree biomass and basal area values and the number of stems, within a given plot and scaling them to per hectare values by plot area. The mean height of trees for each plot was calculated as Lorey's mean height, i.e. mean height weighted by basal area of individual trees. The ground reference values are presented in Table 1.

Table 1. Ground reference values for the sample plots.

Characteristic	Ground reference values				
	Range	**Mean**	**Std**[1]	**Cv**[2]	**Stderr**[3]
Biomass (Mg ha^{-1})	0 – 125.59	38.99	29.49	75.62	2.85
Basal area (m^2 ha^{-1})	0 – 16.10	5.32	3.78	71.06	0.37
Number of stems (ha^{-1})	0 – 830	337	178	53	17
Lorey's mean height (m)	3.76 – 14.58	8.81	2.41	27.31	0.23

[1] Std = Standard deviation, [2] Cv = Coefficient of variation, [3] Stderr = Standard error.

2.2.2. UAV imagery collection

The UAV images were acquired during four days from 23rd to 26th April 2015. At this time of the year, the rain season had just ended and trees still had leaves on them. Due to time constraints, the images had to be acquired over the entire day, i.e. morning, noon and evening. Therefore, differences in shadow effects were expected in the acquired images. The images were acquired using a SenseFly eBee fixed-wing UAV [48] equipped with a Canon IXUS127 HS Digital camera. The dimensions and weight of camera with battery and memory card were 93.2 x 57.0 x 20.0 mm and 135g, respectively. The camera produces 16.1 megapixel images in the red, green and blue spectral bands. The camera automatically set with a shutter speed of 1/2000 sec. The eBee is made from flexible foam and the weight is 537 g without camera. The eBee is also equipped with an inertial measurement unit as well as an on-board Global Navigation Satellite Systems GNSS to control the flight and to provide rough positioning [48].

Prior to taking images, positions of ground control points (GCPs) as well as landing and take-off points, e.g. on open areas with no trees within the forest and agricultural fields near the forest, were identified and measured. The GCPs were made of a set of 1 × 1 m cross-shaped timber planks painted white and some black and white 50 × 50 cm checkerboards. The position of the centre of each GCP was measured using the same procedure as used when locating plot centres for the sample plot inventory described above. The data were collected for an average of 13±6 minutes for each GCP with a 1-second logging rate. The recorded coordinates for each GCP were post-processed similarly as the sample plots. The results from the RTKLIB software revealed that maximum deviations for northing, easting and height were 2.24 cm, 4.50 cm and 4.46 cm, respectively.

Acquisition of images was controlled from a laptop computer with a mission control software called eMotion 2 version 2.4 [48]. All the flights were planned in the mission control software prior to flying. For navigation purposes, we used a georeferenced base map from Microsoft Bing maps covering the study area. For this study we applied percentage end and side image overlaps of 80% and 90% respectively, as well as a flight height above the ground of 325 m. Finally, images covering the entire forest were taken in all the designated strips. A summary of flight characteristics for each flight day is presented in Table 2.

Table 2. Summary of flight characteristics for each flight block of imagery collection.

Date	Number of flights	Number of images	Flight time (minutes)	Wind speed (m sec^{-1})	Cloud cover (%)
23.04.2015	7	1241	518	6.0 – 9.5	10 – 80
24.04.2015	6	1301	489	6.0 – 9.0	20 – 80
25.04.2015	6	1118	502	5.0 – 8.5	10 – 100
26.04.2015	1	273	22	3.0 – 4.0	50

2.3. Image processing

Agisoft Photoscan Professional version 1.1 [49] was used to generate a 3D point cloud from the acquired images. This software uses both SfM and stereo-matching algorithms for image alignment and multi-view stereo reconstruction. The process for generating a 3D point cloud involved: a) image alignment using the inbuilt SfM algorithms, b) mesh building, c) guided marker positioning and optimization of camera alignment (georeferencing of created scene) where the guided marker positioning involved importing coordinates for the GCPs and manually refining the estimated positions of the GCPs so as to improve camera orientation and position estimates (since the GNSS onboard the SenseFly eBee provide rough positioning) and d) building the dense point cloud. Parameter settings for the processes in (a) – (d) are displayed in Table 3. All parameters were chosen based on empirical experience in Puliti, et al. [27]. Finally, we added spectral information from the point cloud, i.e. red, green and blue image bands to the point cloud. According to Wallace, et al. [29], spectral information from a UAV point cloud can present additional useful information for estimating other non-structural properties of the canopy.

Table 3. Processing steps with corresponding parameter settings in Agisoft Photoscan Professional software for generation of 3D point cloud from UAV imagery.

Task	Parameters
a) Image alignment	Accuracy: high Pair selection: reference Key points: 40000 Tie points: 1000
b) Mesh building	Surface type: height field Source data: dense cloud Face count: High
c) Guided marker positioning	Manual positioning of markers on the 14 GCPs for all the photos where a GCP was visible
d) Building dense point cloud	Quality: medium Depth filtering: mild

2.4. DTM generation methods

2.4.1. Supervised ground filtering based on visual classification.

The first step involved producing an orthophoto for the entire area to guide the visual identification of areas with and without vegetation. Points, within and around all the sample plots, were then visually assigned to either ground or non-ground classes depending on whether they fell on an area with or without vegetation. The ground-class points were then used to create DTM as a triangular irregular network (TIN) surface. This DTM was denoted as DTM 01 for further analysis. The DTM was finally used to calculate the height relative to the ground for all points by subtracting respective TIN values from each point.

2.4.2. Supervised ground filtering based on logistic regression

In order to filter the respective point clouds, a supervised ground filtering approach with a standard binary logistic regression classifier was applied [50]. To train the model, visual classification was carried out on 132 circular areas of size 314 m² (radius 10 m) located with a systematic offset of 110 m in north and east from the sample plots. A total of 495457 points were classified with approximately 23% classed as ground points. The following logistic regression model was then fitted:

$$ln\left(\frac{P_j}{1-P_j}\right) = \beta_0 + \beta_1 R_j + \beta_2 G_j + \beta_3 B_j + \varepsilon_j \quad (6)$$

where P_j is the probability of point j being a ground point. R, G and B are the red, green and blue spectral information in each point, β_0, β_1 and β_2 are regression coefficients and ε_{ij} is the random error component.

All identified ground-class points were then used to generate a DTM and calculate the height relative to the ground for all points using the approach described in 2.4.1.This DTM was denoted as DTM 02 for further analysis.

2.4.3. Supervised ground filtering based on quantile regression

Quantile regression enable fitting of regression curves to other parts of the distribution of the response variable than the mean [51,52]. An assumption was made that the points representing the 0.01 quantile of the z-values of each sample plot could be considered ground heights. The relationship between the z-values against the y and x values of each point (j) at the sample plot (i) was then modelled as follows:

$$z_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 y_{ij} + \varepsilon_{ij} \quad (7)$$

where β_0, β_1 and β_2 are regression coefficients and ε_{ij} is the random error component. The fitted values ($\hat{z}_{ij}$) of the regression model were then used to calculate the height relative to the ground for all points in a given plot. This ground classification was denoted as DTM 03 for further analysis.

2.4.4. Unsupervised ground filtering based on shuttle radar topography mission (SRTM).

In remote areas of the earth as well as in developing countries, the best source of terrain heights are usually based on the SRTM [53]. In this study, the centre points in the SRTM pixels and the corresponding height value were used to create a TIN surface. The DTM was denoted as DTM 04 for further analysis. The DTM was finally used to calculate the height relative to the ground for all points in a given plot using the approach described in 2.4.1.

2.4.5. Unsupervised ground filtering based on the progressive TIN algorithm

A standard method for classifying ground points in point cloud data is based on the principles of the progressive TIN algorithm developed by Axelsson [54]. A variant of this algorithm is implemented in Agisoft Photoscan software [49]. The algorithm divides the point cloud into cells of a certain size. In this study, a cell size of 50 m was applied. In each cell, the lowest point is detected and is triangulated to produce a first approximate terrain model. Next, parameters describing the angle between a point and the DTM surface and the maximum distance between a point and the DTM surface are set. In this study a grid search approach was applied to test different values for maximum distance and angle parameters. Based on the preliminary tests, the following angle-distance parameter combinations were chosen for further testing: 3-1, 3-3, 3-6, 6-1, 6-3, 6-6, 9-1, 9-3 and 9-6. For each of these combinations, a DTM was generated as a TIN surface of the points classified as ground. For further analysis, the DTMs were denoted as DTM 05, DTM 06, … and DTM 13,

respectively. The DTMs were finally used to calculate the height relative to the ground for all points using the approach described in 2.4.1.

2.6. Variable extraction

For each of the generated DTMs, variables describing plot-level canopy height and canopy density were then extracted as described by Næsset [55]. Variables describing canopy height included maximum and mean values (*Hmax, Hmean*), standard deviation (*Hsd*), coefficient of variation (*Hcv*), kurtosis (*Hkurt*), skewness (*Hskewness*) and percentiles at 10% intervals (H_{10}, H_{20},…, H_{90}). A height threshold of 0.5 m was applied in order to separate trees from low vegetation. Furthermore, canopy density variables were derived by dividing the height between a 95% percentile height and the 0.5 m threshold into 10 equally tall vertical layers and calculating the proportion of points above each layer to the total number of points. These variables were denoted as follows: D_0, D_1,…, D_9. In addition, spectral variables derived from the RGB (red-green-blue) spectral bands were included. The spectral variables were computed as the maximum (*Smax*), mean (*Smean*), standard deviation (*Ssd*), coefficient of variation (*Scv*), kurtosis (*Skurt*), skewness (*Sskewness*) and 9 percentiles (S_{10}, S_{20}, …,S_{90}) for each of the three bands. For example, the *Smax* variable was denoted as follows: *Smax.red*, *Smax.green* and *Smax.blue*. The remaining spectral variables were also denoted similarly. In total, 64 variables describing canopy height, canopy density and canopy spectral properties were extracted.

2.7. Model development and evaluation

Multiple linear regression models relating reference biomass (as dependent variable) and the generated variables (as independent variables) were initially fit on untransformed variables. Since multicollinearity normally occurs between remotely sensed variables [56], we applied the best subset variable selection procedure using the *leaps* package [57] in R statistical software [58]. Selection of potential independent variables was restricted to a combination of up to five independent variables with minimum Bayesian information criterion (BIC) and variance inflation factors (VIF) values as selection criteria. This procedure was repeated for logarithmic and square root transformed dependent variable. An empirical ratio estimator for bias correction proposed by Snowdon [59] was employed when converting the logarithmic and square root predictions to an arithmetic scale. The proportional bias was estimated from the ratio of the mean of the observed values to the mean of the back-transformed predicted values. The estimates were finally corrected by multiplying them with the estimated ratio.

Preliminary results indicated that models developed using both untransformed and logarithmically transformed dependent variable produced unsatisfactory results in comparison to those developed using square root transformed dependent variable. Thus, results from models developed using square root transformed dependent variable (equation 8) were considered for further analysis.

$$\sqrt{y_j} = \beta_0 + \beta_1 x_{jk} + \cdots + \beta_k x_{jk} + \varepsilon_j \quad (8)$$

where y_j is the ground reference biomass of the j^{th} sample plot, x_{j1}…,x_{jk} are the k independent variables, β_0,…, β_k are the parameter estimates, n is the number of sample plots and j is the sample plot level residual, j = 1,……..,n; $\varepsilon_j \sim N(0, \sigma_\varepsilon^2)$.

For each model, reported values included relative root mean square error (RMSE%), relative mean prediction error (MPE%) and squared Pearson's linear correlation coefficient (r^2). RMSE% and MPE% were calculated as follows:

$$\text{RMSE\%} = \frac{\text{RMSE}}{\bar{y}} \times 100 \quad (9)$$

$$\text{MPE\%} = \frac{\text{MPE}}{\bar{y}} \times 100 \quad (10)$$

$\bar{y}$ is the mean ground reference biomass for all sample plots.

Both RMSE% and MPE% values were calculated using the leave-one-out-cross-validation (LOOCV) procedure [60]. However, comparison of the models was based on RMSE% values. RMSE is a reliable measure for model performance as it accounts for both variance and bias of the predicted value [61].

The MPE% values for each model were tested to check if they were significantly different from zero using a two-sided student's t-tests at 95% confidence level. Similar tests were also applied to test the significance of the differences in height deviations between GPS readings and the different DTMs.

3. Results

3.1. Comparison of the DTM generation methods

The initial results showed that there were statistically significant differences in height deviations from the GPS readings between the different DTMs ($p < 0.001$). However, when the values from SRTM (DTM 04) were excluded, the results indicated that there were no significant differences in height deviations from GPS readings ($p = 0.997$) amongst the remaining DTMs. Figure 2 displays the distribution of mean height differences of plot centre values for GPS and corresponding DTMs.

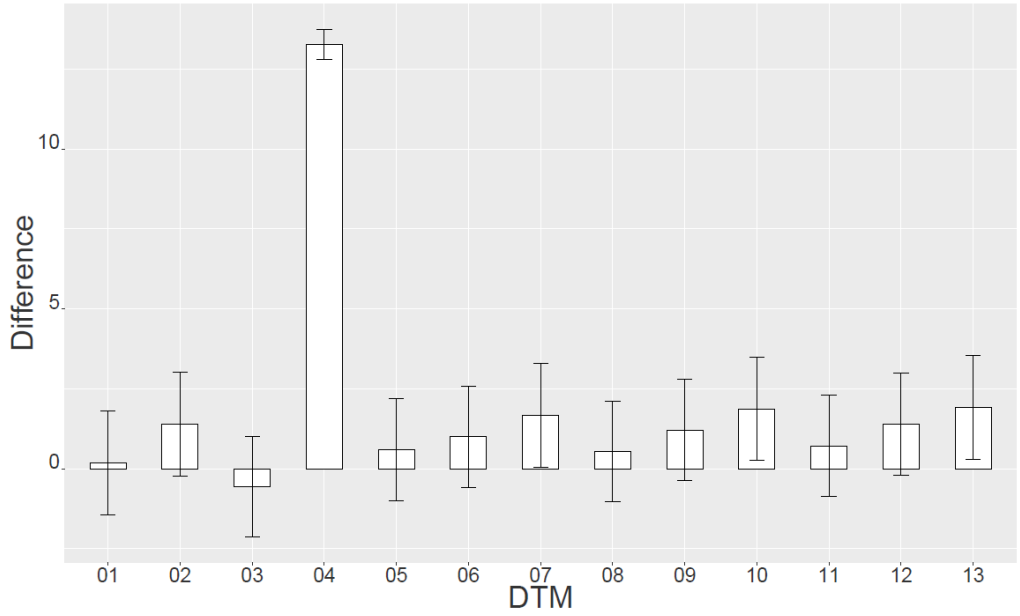


Figure 2. Mean height differences (m) between measured GPS (true values) and predicted heights for the different DTM generation methods with standard errors. 01: DTM using supervised ground filtering based on visual classification; 02: DTM using supervised ground filtering based on logistic regression; 03: DTM using supervised ground filtering based on quantile regression; 04: DTM using unsupervised ground filtering based on shuttle radar topography mission (SRTM); 05-13: DTMs using unsupervised ground filtering based on a grid search approach for optimal parameter settings in Agisoft Photoscan Professional software (See section 2.4. for details).

3.2. Regression analysis

All the models produced appropriate model performance criteria, i.e. none of the models had MPE% values that were significantly different from zero ($p > 0.05$) (Table 4). The models produced RMSE values in the range of 46.7 – 81.7% of mean biomass of 38.99 Mg ha^{-1} from ground reference biomass data. The r^2 values for all the models ranged from 0.12 to 0.67. Model 07, developed using unsupervised ground filtering based on a grid search approach, had the smallest RMSE% value amongst the models. Figure 3 displays the relationship between ground reference and predicted biomass for models 01 -13. There is a general trend of under predictions for sample plots with higher biomass amongst all the models.

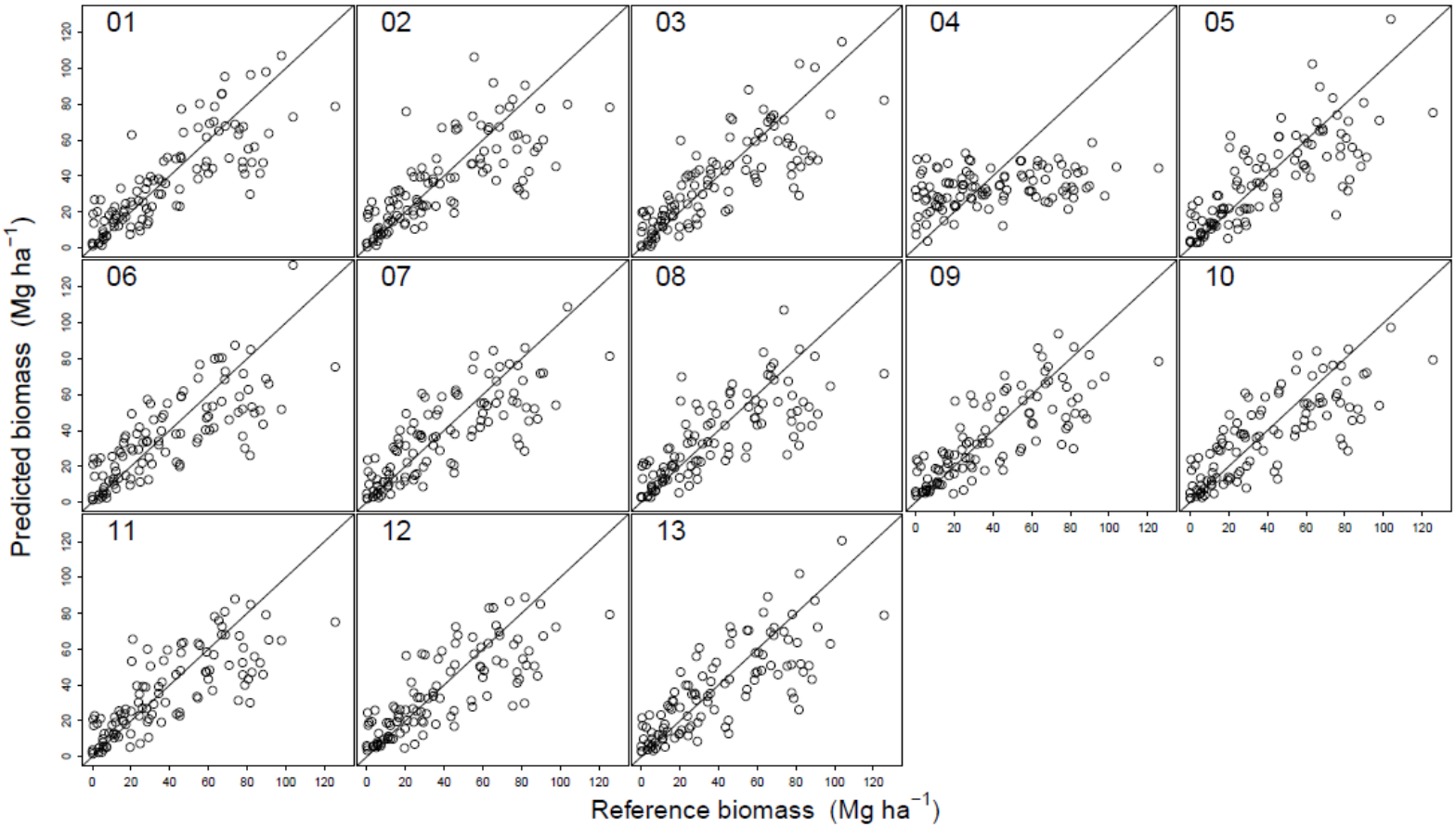


Figure 3. Ground reference versus predicted biomass for different DTMs. 01 – 13. 01: DTM using supervised ground filtering based on visual classification; 02: DTM using supervised ground filtering based on logistic regression; 03: DTM using supervised ground filtering based on quantile regression; 04: DTM using unsupervised ground filtering based on shuttle radar topography mission (SRTM); 05-13: DTMs using unsupervised ground filtering based on a grid search approach for optimal parameter settings in Agisoft Photoscan Professional software (See section 2.4. for details).

Amongst the developed models, approximately 20%, i.e., 13 out of 64 of the candidate independent variables, were selected. In total nine of the variables belonged to canopy height while the remaining four belonged to canopy density. Amongst the selected canopy height variables, *Hmax*, was the most frequently selected variable amongst the models.

In addition, all models (see Table 4) included variables incorporating spectral variables with the three colour bands, red, green and blue. Amongst the models, the red colour band is the most frequently selected.

Table 4. Performance of models over different digital terrain models (DTMs) and ground classification approaches.

DTM[1]	Independent variables[2]	r^2	Predicted Biomass	RMSE		MPE		*p*-value
			Mg ha^{-1}	Mg ha^{-1}	%	Mg ha^{-1}	%	
01	*Hmax, D9, Ssd.blue*	0.67	39.64	18.36	46.8	-0.41	-1.1	0.82
02	*Hmax, D2, Ssd.green*	0.58	39.26	21.44	55.0	-0.27	-0.7	0.90
03	*Hmax, D5, Ssd.blue*	0.65	39.21	18.73	48.0	-0.22	-0.6	0.91
04	*S50.red*	0.12	38.95	31.87	81.7	0.04	0.1	0.99
05	*Hsd, D0, Ssd.blue*	0.61	39.44	20.38	52.2	-0.45	-1.2	0.82
06	*Hmax, S70.red, S90.green*	0.61	39.15	19.76	50.7	-0.16	-0.4	0.93
07	*Hmax, S70.red, S90.green*	0.64	39.03	18.21	46.7	-0.03	-0.1	0.99
08	*Hmax, D0, Ssd.blue*	0.59	38.60	22.73	58.3	0.39	1.0	0.86
09	*Hmax, D0, Scv.red*	0.63	38.90	20.40	52.3	0.09	0.2	0.96
10	*Hmax, S70.red, S90.green*	0.62	39.02	19.54	50.1	-0.03	-0.1	0.99
11	*Hmax, D0, Scv.green*	0.62	39.03	20.36	52.2	-0.03	-0.1	0.99
12	*Hmax, D0, Scv.red*	0.62	38.90	19.68	50.5	0.09	0.2	0.96
13	*Hmax, D0, S70.red, S90.green*	0.63	39.69	20.19	51.8	-0.71	-1.8	0.72

[1] 01: DTM using supervised ground filtering based on visual classification; 02: DTM using supervised ground filtering based on logistic regression; 03: DTM using supervised ground filtering based on quantile regression; 04: DTM using unsupervised ground filtering based on shuttle radar topography mission (SRTM); 05-13: DTMs using unsupervised ground filtering based on a grid search approach for optimal parameter settings in Agisoft Photoscan Professional software (See section 2.4. for details).

[2] *S50.red, S70.red* and *S90.green*: spectral variables for the 50%, 70% and 90 percentile for the red, red and green colour bands, respectively; *D0, D2, D5* and *D9*: canopy densities corresponding to the proportions of points above fraction number 0, 2, 5 and 9, respectively; *Hmax, Hsd, Scv.red, Scv.green, Ssd.green, Ssd.blue*: maximum canopy height, canopy height standard deviation, coefficient of variation for the spectral variables with red and green colour bands respectively; standard deviation for the spectral variables with green and blue colour bands, r^2 = squared Pearson's correlation coefficient, RMSE= Root mean square error, MPE = Mean prediction error.

4. Discussion

Successful implementation of the REDD+ initiative is dependent on the availability of reliable biomass estimating methods. Application of UAVs in monitoring forest ecosystems and in biomass estimation is gaining increased attention [28,62,63]. This study aimed at evaluating the application of 3D data generated from UAV acquired images in forest biomass estimation for the miombo woodlands. This could be of particular interest for Malawi since most of the forests in the country comprise of miombo woodlands scattered over the landscape as small- to medium sized reserves, a situation that suits the application of the UAV both when it comes to technical and costs associated with its execution.

Reliable biomass estimates from remotely sensed 3D data are heavily reliant on the availability of a good DTM. This study first tested different methods of generating DTMs. The comparisons of plot centre height estimates from different DTMs showed that estimates from the DTM generated using SRTM data are unreliable as compared to the DTMs derived from the other methods (Figure 2). This is indicating that when a DTM based on SRTM data is used in biomass estimation, the estimates can hardly be trusted. Despite the fact that there were no statistically significant differences amongst the DTMs developed from the different methods, i.e. excluding SRTM, the DTM model producing the smallest RMSE should always be applied when estimating biomass.

Biomass estimates from the 13 different DTM generation approaches shown that the DTM developed from unsupervised ground filtering based on a grid search approach (model 07) performed slightly better compared with the other models (Table 4). This performance demonstrated that with some effort, it is possible to get a good combination of angle-distance parameter settings in the AgiSoft Photoscan software [49]. Furthermore, despite performing slightly less than model 07, model 01 developed from a DTM based on supervised ground filtering using visual classification, is equally good. However, the process for generating a DTM using supervised ground filtering based on visual classification is quite arduous. Therefore, since unsupervised ground filtering is easier to implement and performs equally (or even better) to visual classification, future studies should consider application of this approach. On the other hand, the relatively poor performance of the DTMs developed from unsupervised ground filtering based on shuttle radar topography mission (SRTM) could be attributed to the inherent random errors in heights associated with SRTM data [64-66].

The findings from our study demonstrated that data generated by the UAV system have potential of being successfully used in estimating forest biomass in dry tropical forests such as miombo woodlands. The fact that no MPE values were significantly different from zero (Table 4) is indicating that the models can be appropriately used in biomass estimation. Similar studies for other dry tropical forest sites are, however, recommended to validate the results from the current study because of the wide range of forest conditions seen in dry tropical forests such as miombo woodlands.

The observed RMSE% value for the best model from our study (46.7%) is similar to that reported in a study conducted in miombo woodlands of Tanzania (46.8%) by Mauya, et al. [36]. Direct comparison of these results should however be done with caution because ALS data was used in that study. In addition, there were differences in sample size, plot size and forest conditions for the respective studies. Further studies on the assessment of the effect of sample size, plot size, and forest conditions are therefore required to validate the results from the current study.

On the other hand, in a study by Puliti, et al. [27] , where data acquired from UAV was applied in boreal forests, an RMSE% value of 14.95% was observed when estimating forest stand volume. This might be attributed to the differences in forest structures between miombo woodlands and boreal forests. It should also be noted that Puliti, et al. [27] utilized ALS data for DTM determination, which are superior in describing forest ground surface compared to optical sensors such as those applied in the current study [67]. It is worth noting that the observed RMSE% in Puliti, et al. [27] is comparable to that observed in a study by Gobakken, *et al.* [68] conducted in similar forest conditions. However, Gobakken, et al. [68] used exclusively ALS data. This demonstrates the efficiency of UAV data in forest inventories.

During image acquisition, wind speeds up to 9.5 m sec^{-1} were encountered. According to Dandois and Ellis [26], images captured when wind speeds are around or over 9.5 m sec^{-1} generally make some trees sway as well as affect the UAV motion. This results in incomplete image overlaps, thus affecting the quality of the generated point cloud and DTM, for some parts of a study area. We thus anticipated that this scenario would have some influence on the results. We also anticipated that the presence of shadow effects for some of the captured images could have had a profound effect on the results. This is the case because automatic detection of reliable image features in areas of shadow is reportedly difficult due to large brightness range between shadows and sunlit areas such as top of the canopy and the properties and settings of the camera resulting in insufficient image contrast in areas of shadow [17,29]. Furthermore, we also anticipated that the leaf-on nature of the forest during data collection may have led to inaccurate interpolation of ground points due to uncertain point locations and low point density in areas of high canopy cover [29]. It is worth noting that despite the fact the data for the current study were collected under varying wind speed, light and terrain conditions, the results are similar with results from using ALS data [36]

The results from the current study demonstrate the ability of the UAV to capture reliable imagery under varying conditions in miombo woodlands. However, to fully comprehend the potential of utilising UAV systems in miombo woodlands, future studies should aim at acquiring the images at times of a day when shadow effects are minimized, e.g. at noon as well as when wind speeds are low. However, it is often costly to wait for optimal weather conditions in remote sites due to associated logistical costs. Thus, the current study, together with the study by Puliti, et al. [27] provide examples that it is still possible to obtain high accuracies under varying weather conditions. Acquiring images under leaf-off conditions might improve the ground classification, but might decrease the possibility to reconstruct trees in the point cloud and thus decrease biomass estimates. Likely, acquiring imagery under both leaf-off and leaf-on conditions might provide the best results. Further, it should also be noted that the study was conducted on a relatively small area and these results need to validated over different study sites with variable sizes.

The importance of spectral information is highlighted by the selection of variables incorporating the red, green and blue colour bands in all the generated models. According to Wallace, et al. [29], spectral components of UAV point clouds tend to present additional useful information for estimating other non- structural properties of the canopy.

During the study a fixed plot size was used. However, the size of the field plots has been identified as one of the sources of model uncertainty in remote sensing based biomass estimation [36,69-71]. Future studies should therefore aim at examining the effects of field plot size and even number of plots, on biomass estimates. Furthermore, during the study the images were acquired at a constant flight height of approximately 325 m above the ground. Future studies utilizing UAVs in Malawi could also aim at testing the effect of flight height on biomass estimates. Such a study would be possible because the directorate of civil aviation in Malawi does not currently impose any restrictions on flight height unlike in other countries where UAV flight heights are restricted by law [16,27,29,72].

5. Conclusions

Remote sensing data were captured using a SenseFly eBee fixed-wing UAV system on a small- to medium sized potential REDD+ project site in miombo woodlands. The prediction errors observed are similar to results from previous studies using ALS data in miombo woodlands. The DTM developed using unsupervised ground filtering based on a grid search approach, had the smallest RMSE value thus indicating that DTMs developed through this method can produce reliable results in miombo woodlands. Additional studies, however, are recommended to validate these results under other conditions using different flight settings, plot sizes and plot numbers.

Acknowledgments: The main author is indebted to the Norwegian Government for funding the study through the Norwegian Quota scholarship scheme. Special thanks should also be extended to the Capacity Building for Managing Climate Change (CABMACC) programme in Malawi for funding the fieldwork. We would also like to thank the Malawi Government and the communities around the Muyobe forest reserve for granting us a permit to use the UAV for data collection. We also say thank you to Steven Mphamba, Herbert Jenya, Martin Nyoni and Kola Daitoni for their support during data collection.

Author Contributions: Daud Jones Kachamba, the main author, was involved in designing the research, collected and analyzed the data and developed the manuscript. Hans Ole Ørka was involved in designing, collecting and analyzing UAV data as well as developing the manuscript. Terje Gobakken was involved in designing the research, analyzing the data and manuscript development. Tron Eid was involved in designing the research, data collection and manuscript development. Weston Mwase was involved in data collection and manuscript development.

Conflicts of Interest: The authors declare no conflicts of interest

References

1. Gibbs, H.K.; Brown, S.; Niles, J.O.; Foley, J.A. Monitoring and estimating tropical forest carbon stocks: Making REDD a reality. *Environmental Research Letters* **2007**, *2*, 045023.
2. Government of Malawi. Government of Malawi REDD+ action plan 2014-2019. Department of Forestry, Ministry of Environment and Climate Change Management: Lilongwe, Malawi, 2015.
3. Tomppo, E.; Malimbwi, R.; Katila, M.; Mäkisara, K.; Henttonen, H.M.; Chamuya, N.; Zahabu, E.; Otieno, J. A sampling design for a large area forest inventory: Case Tanzania. *Canadian Journal of Forest Research* **2014**, *44*, 931-948.
4. Government of Malawi. Forest resource mapping project under the Japanese grant for the forest preservation programme to the Republic of Malawi. Department of Forestry, Ministry of Environment and Climate Change Management: Lilongwe, Malawi, 2012.
5. Chidumayo, E.N. *Miombo ecology and management: An introduction*. IT Publications in Association with the Stockholm Environment Institute: London, 1997; p 166.
6. Government of Malawi. Malawi state of environment and outlook report. Ministry of Natural Resources, Energy and Environment: Lilongwe, Malawi, 2010.
7. Angelsen, A. *Moving ahead with REDD: Issues, options and implications*. Center for International Forestry Research (CIFOR): Bogor, Indonesia, 2008.
8. DeFries, R.; Achard, F.; Brown, S.; Herold, M.; Murdiyarso, D.; Schlamadinger, B.; de Souza Jr, C. Earth observations for estimating greenhouse gas emissions from deforestation in developing countries. *Environmental Science & Policy* **2007**, *10*, 385-394.
9. Sinha, S.; Jeganathan, C.; Sharma, L.; Nathawat, M. A review of radar remote sensing for biomass estimation. *International Journal of Environmental Science and Technology* **2015**, *12*, 1779-1792.
10. McRoberts, R.E.; Næsset, E.; Gobakken, T.; Bollandsås, O.M. Indirect and direct estimation of forest biomass change using forest inventory and airborne laser scanning data. *Remote Sensing of Environment* **2015**, *164*, 36-42.
11. Masek, J.G.; Hayes, D.J.; Joseph, H.M.; Healey, S.P.; Turner, D.P. The role of remote sensing in process-scaling studies of managed forest ecosystems. *Forest Ecology and Management* **2015**, *355*, 109-123.
12. Barrett, F.; McRoberts, R.E.; Tomppo, E.; Cienciala, E.; Waser, L.T. A questionnaire-based review of the operational use of remotely sensed data by national forest inventories. *Remote Sensing of Environment* **2016**, *174*, 279-289.
13. Næsset, E.; Gobakken, T.; Bollandsås, O.M.; Gregoire, T.G.; Nelson, R.; Ståhl, G. Comparison of precision of biomass estimates in regional field sample surveys and airborne lidar-assisted surveys in Hedmark county, Norway. *Remote Sensing of Environment* **2013**, *130*, 108-120.
14. Lu, D.; Chen, Q.; Wang, G.; Liu, L.; Li, G.; Moran, E. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth* **2014**, 1-43.
15. Kumar, L.; Sinha, P.; Taylor, S.; Alqurashi, A. Review of the use of remote sensing for biomass estimation to support renewable energy generation. *Journal of Applied Remote Sensing* **2015**, *9*, 29.

16. Dandois, J.; Olano, M.; Ellis, E. Optimal altitude, overlap, and weather conditions for computer vision UAV estimates of forest structure. *Remote Sensing* **2015**, *7*, 13895.
17. Kumar, L.; Sinha, P.; Taylor, S.; Alqurashi, A.F. Review of the use of remote sensing for biomass estimation to support renewable energy generation. *Journal of Applied Remote Sensing* **2015**, *9*.
18. Næsset, E.; Gobakken, T. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment* **2008**, *112*, 3079-3090.
19. Skowronski, N.S.; Clark, K.L.; Gallagher, M.; Birdsey, R.A.; Hom, J.L. Airborne laser scanner-assisted estimation of aboveground biomass change in a temperate oak–pine forest. *Remote Sensing of Environment* **2014**, *151*, 166-174.
20. Patenaude, G.; Hill, R.A.; Milne, R.; Gaveau, D.L.A.; Briggs, B.B.J.; Dawson, T.P. Quantifying forest above ground carbon content using lidar remote sensing. *Remote Sensing of Environment* **2004**, *93*, 368-380.
21. Gonzalez, P.; Asner, G.P.; Battles, J.J.; Lefsky, M.A.; Waring, K.M.; Palace, M. Forest carbon densities and uncertainties from lidar, quickbird, and field measurements in california. *Remote Sensing of Environment* **2010**, *114*, 1561-1575.
22. Mauya, E.; Hansen, E.; Gobakken, T.; Bollandsås, O.; Malimbwi, R.; Næsset, E. Effects of field plot size on prediction accuracy of aboveground biomass in airborne laser scanning-assisted inventories in tropical rain forests of Tanzania. *Carbon Balance and Management* **2015**, *10*, 1-14.
23. Drake, J.B.; Dubayah, R.O.; Clark, D.B.; Knox, R.G.; Blair, J.B.; Hofton, M.A.; Chazdon, R.L.; Weishampel, J.F.; Prince, S. Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment* **2002**, *79*, 305-319.
24. Vauhkonen, J.; Maltamo, M.; McRoberts, R.E.; Næsset, E. Introduction to forestry applications of airborne laser scanning. In *Forestry applications of airborne laser scanning-concepts and case studies*, Maltamo, M.; Næsset, E.; Vauhkonen, J., Eds. Springer: Dordrecht, 2014; pp 1-6.
25. Whitehead, K.; Hugenholtz, C.H. Remote sensing of the environment with small unmanned aircraft systems (uass), part 1: A review of progress and challenges. *Journal of Unmanned Vehicle Systems* **2014**, *02*, 69-85.
26. Dandois, J.P.; Ellis, E.C. High spatial resolution three-dimensional mapping of vegetation spectral dynamics using computer vision. *Remote Sensing of Environment* **2013**, *136*, 259-276.
27. Puliti, S.; Ørka, H.; Gobakken, T.; Næsset, E. Inventory of small forest areas using an unmanned aerial system. *Remote Sensing* **2015**, *7*, 9632.
28. Tang, L.; Shao, G. Drone remote sensing for forestry research and practices. *Journal of Forestry Research* **2015**, *26*, 791-797.
29. Wallace, L.; Lucieer, A.; Malenovský, Z.; Turner, D.; Vopěnka, P. Assessment of forest structure using two UAV techniques: A comparison of airborne laser scanning and structure from motion (sfm) point clouds. *Forests* **2016**, *7*, 62.
30. Siebert, S.; Teizer, J. Mobile 3d mapping for surveying earthwork projects using an unmanned aerial vehicle (UAV) system. *Automation in Construction* **2014**, *41*, 1-14.

31. Getzin, S.; Wiegand, K.; Schöning, I. Assessing biodiversity in forests using very high-resolution images and unmanned aerial vehicles. *Methods in ecology and evolution* **2012**, *3*, 397-404.
32. Chidumayo, E.N.; Gumbo, D.J. *The dry forests and woodlands of africa: Managing for products and services*. Earthscan: London, UK, 2010.
33. Chidumayo, E.N. Forest degradation and recovery in a miombo woodland landscape in Zambia: 22 years of observations on permanent sample plots. *Forest Ecology and Management* **2013**, *291*, 154-161.
34. Næsset, E.; Økland, T. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment* **2002**, *79*, 105-115.
35. Ravindranath, N.H.; Ostwald, M. *Carbon inventory methods - handbook for greenhouse gas inventory, carbon mitigation and roundwood production projects.* Springer: Heidelberg, 2008; Vol. 29.
36. Mauya, E.; Ene, L.; Bollandsås, O.; Gobakken, T.; Næsset, E.; Malimbwi, R.; Zahabu, E. Modelling aboveground forest biomass using airborne laser scanner data in the miombo woodlands of Tanzania. *Carbon Balance and Management* **2015**, *10*, 28.
37. Gregoire, T.G.; Næsset, E.; McRoberts, R.E.; Ståhl, G.; Andersen, H.-E.; Gobakken, T.; Ene, L.; Nelson, R. Statistical rigor in lidar-assisted estimation of aboveground forest biomass. *Remote Sensing of Environment* **2016**, *173*, 98-108.
38. Yang, J.; Prince, S.D. Remote sensing of savanna vegetation changes in eastern Zambia 1972-1989. *International Journal of Remote Sensing* **2000**, *21*, 301-322.
39. Fuller, D.O.; Prince, S.D.; Astle, W.L. The influence of canopy strata on remotely sensed observations of savanna-woodlands. *International Journal of Remote Sensing* **1997**, *18*, 2985-3009.
40. Maguya, A.; Junttila, V.; Kauranne, T. Algorithm for extracting digital terrain models under forest canopy from airborne lidar data. *Remote Sensing* **2014**, *6*, 6524.
41. Niethammer, U.; James, M.R.; Rothmund, S.; Travelletti, J.; Joswig, M. UAV-based remote sensing of the super-sauze landslide: Evaluation and results. *Engineering Geology* **2012**, *128*, 2-11.
42. Soininen, A. Terrascan user' s guide. Terrasolid: 2004.
43. McGaughey, R. Fusion/ldv: Software for lidar data analysis and visualization. United States Department of Agriculture (USDA), Forest Service: Pacific Northwest Research Station: Portland, OR, USA, 2015; Vol. Version 3.50.
44. Hardcastle, P.D. *A preliminary silvicultural classification of Malawi*. Forestry Research Institute of Malawi, : Zomba (Malawi), 1978.
45. Kouba, J. A guide to using international gnss service (IGS) products. Geodetic Survey Division, Natural Resources Canada: 615 Booth Street, Ottawa, Ontario, K1A 0E9, 2015.
46. Takasu, T. In *Rtklib: Open source program package for rtk-gps*, Free and open source software for Geospatial Sydney, 2009; Open Source Geospatial Foundation (OSGeo): Sydney.
47. Kachamba, D.J.; Eid, T.; Gobakken, T. Above- and belowground biomass models for trees in the miombo woodlands of Malawi. *Forests* **2016**, *7*, 38.
48. Sensefly. Ebee sensefly: Extended user manual. Sensefly, Ltd: 2015.

49. AgiSoft. Agisoft photoscan user manual: Professional edition. AgiSoft LLC: 2015; Vol. Version 1.1.
50. Stal, C.; Briese, C.; De Maeyer, P.; Dorninger, P.; Nuttens, T.; Pfeifer, N.; De Wulf, A. Classification of airborne laser scanning point clouds based on binomial logistic regression analysis. *International Journal of Remote Sensing* **2014**, *35*, 3219-3236.
51. Cade, B.S.; Noon, B.R. A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment* **2003**, *1*, 412-420.
52. Fahrmeir, L.; Kneib, T.; Lang, S.; Marx, B. *Regression*. Springer-Verlag Berlin Heidelberg: Verlag Berlin Heidelberg, 2013; p XIV, 698.
53. van Zyl, J.J. The shuttle radar topography mission (srtm): A breakthrough in remote sensing of topography. *Acta Astronautica* **2001**, *48*, 559-565.
54. Axelsson, P. Processing of laser scanner data-algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing* **1999**, *54*, 138 - 147.
55. Næsset, E. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research* **2004**, *19*, 164-179.
56. Meng, J.; Li, S.; Wang, W.; Liu, Q.; Xie, S.; Ma, W. Mapping forest health using spectral and textural information extracted from spot-5 satellite images. *Remote Sensing* **2016**, *8*, 719.
57. Lumley, T. Leaps: Regression subset selection. In *R package version 2.9*, 2009.
58. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, 2016.
59. Snowdon, P. A ratio estimator for bias correction in logarithmic regressions. *Canadian Journal of Forest Research* **1991**, *21*, 720-724.
60. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An introduction to statistical learning with application in R*. Springer New York: 2013; Vol. 103.
61. Gregoire, T.G.; Valentine, H.T. *Sampling strategies for natural resources and the Environment*. Taylor & Francis: 2007.
62. Zhang, J.; Hu, J.; Lian, J.; Fan, Z.; Ouyang, X.; Ye, W. Seeing the forest from drones: Testing the potential of lightweight drones as a tool for long-term forest monitoring. *Biological Conservation* **2016**, *198*, 60-69.
63. Chianucci, F.; Disperati, L.; Guzzi, D.; Bianchini, D.; Nardino, V.; Lastri, C.; Rindinella, A.; Corona, P. Estimation of canopy attributes in beech forests using true colour digital images from a small fixed-wing UAV. *International Journal of Applied Earth Observation and Geoinformation* **2016**, *47*, 60-68.
64. Rodríguez, E.; Morris, C.; Belz, J. A global assessment of the srtm performance. *Photogrammetric Engineering & Remote Sensing* **2006**, *72*, 249–260.
65. Karwel, A.; Ewiak, I. In *Silk road for information from imagery*, Estimation of the accuracy of the SRTM terrain model on the area of Poland, Beijing, China, 2008; Chen, J.; Jiang, J.; Genderen, J.v., Eds. International Society for Photogrammetry and Remote Sensing: Beijing, China, pp 169-172.
66. Hofton, M.; Dubayah, R.; Blair, J.; Rabine, D. Validation of srtm elevations over vegetated and non-vegetated terrain using medium footprint lidar. *Photogrammetric Engineering & Remote Sensing* **2006**, *72*, 279-285.

67. Baltsavias, E.P. A comparison between photogrammetry and laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* **1999**, *54*, 83-94.
68. Gobakken, T.; Bollandsås, O.M.; Næsset, E. Comparing biophysical forest characteristics estimated from photogrammetric matching of aerial images and airborne laser scanning data. *Scandinavian Journal of Forest Research* **2015**, *30*, 73-86.
69. Frazer, G.; Magnussen, S.; Wulder, M.; Niemann, K. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of lidar-derived estimates of forest stand biomass. *Remote Sensing of Environment* **2011**, *115*, 636 - 649.
70. Keller, M.; Palace, M.; Hurtt, G. Biomass estimation in the tapajos national forest, brazil: Examination of sampling and allometric uncertainties. *Forest Ecology and Management* **2001**, *154*, 371-382.
71. Hansen, E.; Gobakken, T.; Bollandsas, O.; Zahabu, E.; Naesset, E. Modeling aboveground biomass in dense tropical submontane rainforest using airborne laser scanner data. *Remote Sensing* **2015**, *7*, 788 - 807.
72. Zahawi, R.A.; Dandois, J.P.; Holl, K.D.; Nadwodny, D.; Reid, J.L.; Ellis, E.C. Using lightweight unmanned aerial vehicles to monitor tropical forest recovery. *Biological Conservation* **2015**, *186*, 287-295.

Paper IV

Influence of plot size and sample size on efficiency of biomass estimates in inventories of dry tropical forests assisted by photogrammetric data from an unmanned aerial vehicle

Daud Jones Kachamba, Hans Ole Ørka, Terje Gobakken, Tron Eid, Erik Næsset

Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432, Ås, Norway

Abstract

Applications of unmanned aerial vehicle (UAV) to assist in forest inventories have provided promising results in biomass estimation for different forest types. Recent studies demonstrating use of different types of remotely sensed data to assist in biomass estimation have shown that accuracy and precision of estimates are influenced by the sample size as well as the size of field sample plots used to obtain reference values for biomass. The objective of this case study was to assess the influence of sample size and sample plot size on efficiency of UAV-assisted biomass estimates in the dry tropical miombo woodlands of Malawi. The results of a design-based field sample inventory assisted by three-dimensional point clouds obtained from aerial imagery acquired with a UAV showed that the root mean square errors as well as the standard error estimates of mean biomass decreased as sample sizes and sample plot sizes increased. Furthermore, relative efficiency values over different sample sizes and sample plot sizes were above 1.0 in a design-based and model-assisted inferential framework, indicating that UAV-assisted inventories were more efficient than purely field-based inventories. The results on relative costs for UAV-assisted and pure field-based sample plot inventories revealed that there is a trade-off between inventory costs and required precision. For example, in our study if a standard error of less than approximately 3 Mg ha^{-1} was targeted, then a UAV-assisted forest inventory should be applied to ensure more cost effective and precise estimates. Future studies should therefore focus on finding optimum plot and sample sizes for particular applications, like for example in projects under the Reducing Emissions from Deforestation and Forest Degradation, plus forest conservation, sustainable management of forest and enhancement of carbon stocks (REDD+) mechanism with different geographical scales.

Key words*: miombo woodlands, plot size, forest inventory, remote sensing, image matching*

1.0 Introduction

Carbon emissions in tropical forests is currently estimated at 1 Pg C yr^{-1} (Houghton et al. 2015). The mechanism for Reducing Emissions from Deforestation and Forest Degradation, plus forest conservation, sustainable management of forest and enhancement of carbon stocks (REDD+) is one of the key measures for reducing carbon emissions in tropical forests (Gizachew & Duguma 2016; Goetz et al. 2015; Houghton et al. 2015; UNFCCC 2014). As of 2015, there were 33 countries in the tropics in preparatory phase of implementing REDD+ (Nhantumbo & Camargo 2015). The preparatory phase involve establishing administrative structures, determining reference levels for carbon stocks and development of credible monitoring, reporting and verification (MRV) systems, among others. Malawi, one of the African countries currently in the preparatory phase of the implementation of REDD+, is targeting 112 forest reserves scattered across the country as potential REDD+ project areas. These forest reserves have variable sizes ranging between 42 and 114780 ha (Government of Malawi 2010) and approximately 50% of these reserves may be characterized as small- to medium-sized (i.e., up to 2240 ha). Thus, in Malawi forest carbon estimates for REDD+ reporting should be based on forest inventory methods aiming at providing reliable biomass estimates at this geographical scale (forest reserve level) in a consistent and statistically rigorous manner (Næsset et al. 2011).

Conventionally, credible estimation of biomass stocks in Malawi's forest reserves would require field-based sample plot inventories in each of these forest reserves. The inventories are usually conducted using different probability sampling designs. Application of probability sampling when allocating sample plots enables the use of the design-based inferential framework to biomass estimation and inference (Kangas & Maltamo 2006; Næsset et al. 2011). In the design-based inferential framework approach to sampling, values of a variable of interest of the population are viewed as fixed quantities and the selection probabilities introduced with the design are used in determining the expectations, variances, biases and other properties of estimators (Thompson 2012a).

Field-based forest inventories with plots allocated as a probability sample are usually associated with high operational and logistical costs. Recent studies indicate that forest inventories involving combination of data from field-based probability samples and auxiliary information from remote sensing platforms, i.e. design-based and model-assisted inferential framework, are being preferred because they tend to reduce costs while improving precision of the estimates (Gobakken et al. 2012; McRoberts & Tomppo 2007; Su et al. 2016). These

types of forest inventories have successfully been used for estimation of forest biomass in different forest types, including boreal (Margolis et al. 2015; Næsset et al. 2011), temperate (Lefsky et al. 2001; Su et al. 2016) and tropical (Mauya et al. 2015; Ota et al. 2015).

In forestry, remotely sensed data are sourced from three main systems, namely, optical (e.g. satellite and aerial), radio detection and ranging (RADAR) (e.g. synthetic aperture radar (SAR)) and airborne laser scanning (ALS) (Lu et al. 2014). For biomass estimation in particular, data captured from satellite borne optical and RADAR systems usually have challenges with saturation in forests with high biomass density (Kumar et al. 2015). Furthermore, data based on optical systems are also challenged by clouds, shadows, intra-crown spectral variance, low spectral variability and its two-dimensional (2D) nature (Dandois et al. 2015; Kumar et al. 2015). On the other hand, three-dimensional (3D) data from ALS systems seem to overcome the problems associated with optical and RADAR data (Kumar et al. 2015; Lu et al. 2014; Næsset et al. 2013; Vauhkonen et al. 2014). ALS data have shown great potential for forest biomass estimation in different forest types, including boreal (Næsset & Gobakken 2008), temperate (Gonzalez et al. 2010; Patenaude et al. 2004; Skowronski et al. 2014) and tropical forests (Drake et al. 2002; Mauya et al. 2015). However, due to high data acquisition costs, wide application of ALS data for large-scale forest biomass estimation has been limited.

Aerial imagery from optical systems have a long history in forest inventory and monitoring (Campbell & Wynne 2011; Kangas & Maltamo 2006). During the last two decades improvements in aerial imaging, including digital photographs and computer processing capacity, have made it possible to automatically produce 3D data from overlapping aerial images. Such photogrammetric point clouds are similar to the ones derived from ALS data except that only the exterior of the canopy is reconstructed whereas ALS data provide information on both the interior structure and ground under the forest canopies. Forest biomass estimation using photogrammetric point clouds acquired from manned aerial images has been carried out in different forest types, e.g. boreal (Gobakken et al. 2015), temperate (Chen et al. 2012) and tropical (Ota et al. 2015). However, manned aerial images can also be difficult to acquire for many small and scattered forest reserves such as those in Malawi due to the high associated logistical costs.

Unmanned aerial vehicles (UAVs) could be a cost-effective alternative to data acquisition over small and scattered forest reserves. UAVs are also capable of capturing very high-

resolution images from which 3D point clouds can be derived using photogrammetric principles (Paneque-Gálvez et al. 2014; Tang & Shao 2015). Recent research on application of UAVs over small-to-medium sized forests for biomass and volume estimation in dry tropical forests (Kachamba et al. 2016b), boreal forests (Puliti et al. 2015) and temperate forests (Getzin et al. 2012) have demonstrated their potential. Unlike manned aerial platforms, application of UAVs have proved to produce reliable forest biomass estimates at the small-to-medium forest scale with relatively small logistical costs and thus opens an opportunity for local adaption under the REDD+ mechanism (DeFries et al. 2007; Gibbs et al. 2007; Kachamba et al. 2016b; Næsset et al. 2016; Puliti et al. 2015). It is therefore prudent for aspiring REDD+ implementation in developing countries, such as Malawi, to assess the application of UAVs in forest inventories.

Determination of field plot size is an important design decision when planning field-based probability sample inventories. In estimation based on field-based probability sample data combined with auxiliary data from remote sensing, an appropriate geographical correspondence between plots on the ground and the remotely sensed data is paramount. An increased sample plot size can reduce the effects of errors arising from co-registration problems (Frazer et al. 2011). Larger plots will also tend to reduce the plot boundary effects (McRoberts et al. 2014). Several authors have studied the effects of sample plot sizes on biomass estimates and other forest attributes in inventories assisted by remotely sensed data in tropical wet forests (Asner et al. 2009; Hansen et al. 2015; Keller et al. 2001; Mascaro et al. 2011; Mauya et al. 2015; Saatchi et al. 2011), temperate forests (Frazer et al. 2011; Levick et al. 2016) and a boreal forest (Gobakken & Næsset 2008; Næsset et al. 2015), among others.

Apart from sample plot size, sample size, i.e., the number of sample plots employed during an inventory, will also have a large effect on the efficiency of biomass estimates and the associated total inventory costs (Eid et al. 2004; Gobakken & Næsset 2008; Strunk et al. 2012). To the best of our knowledge, no studies on the influence of sample plot sizes and sample sizes on efficiency of biomass estimates have been conducted in UAV-assisted sample inventories, i.e. using design-based and model assisted inferential framework, in miombo woodlands.

The main objective of this case study was therefore to assess the efficiency of UAV-assisted biomass estimation in miombo woodlands of Malawi based on different sample sizes and sample plot sizes, which subsequently may inform design decisions for MRV of fragmented forest reserves in Malawi.

2.0 Materials and methods

This section describes the study area, sampling design, collection and processing of both field reference biomass and remotely sensed data, model development and evaluation and biomass estimation methods under the model-assisted inferential framework. We also describe biomass estimation using purely field-based data as well as the analysis for assessing the efficiency of both field-based and UAV-assisted estimates.

2.1 Study area

The study area, Muyobe community forest reserve, is located in Mpherembe traditional authority in Mzimba district in the northern region of Malawi (11° 35’S, 33° 65’E, 1169 – 1413 m above sea level) (Figure 1). The forest reserve is 486 ha in size, which is common for most small- to medium-sized forests in Malawi. The dominant soil type in the area is Ferrosols (Hardcastle 1978). For the nearest weather station in Mzimba, located 69 km south of the study area, the mean annual rainfall was 889 ± 146 mm and the mean annual daily minimum and maximum temperatures were 15 ± 1.6°C and 26 ± 0.6°C, respectively, for the period 1975 to 2005. The forest is covered by miombo woodlands dominated by *Julbernadia globiflora, Diplorhychus condylocarpon* and *Combretum zeyheri*. For more details on the study area, see Kachamba et al. (2016b).

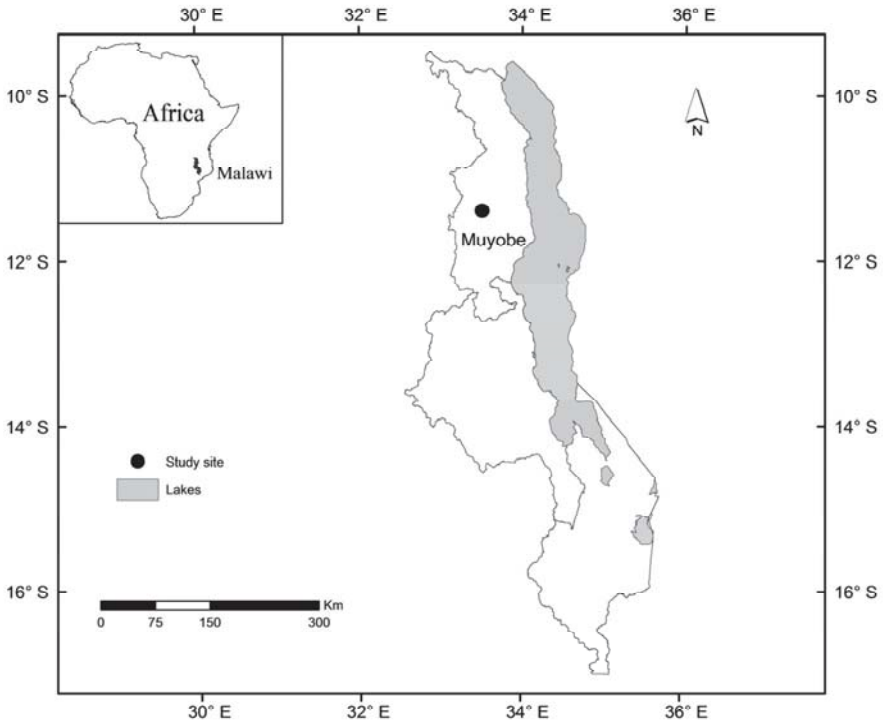


Figure 1. Map of Malawi showing the location of the study site.

2.2 Sampling design and field data collection

Field reference biomass was based on a systematically distributed probability sample collected from the entire forest reserve. The sample size, i.e. the number of plots, was guided by the budget that could support approximately 100 sample plots. The systematic sample was distributed on a grid of 220 m by 220 m. The grid spacing was calculated as follows:

$$\sqrt{\frac{A}{n}} \quad (1)$$

where A = size of the forest reserve (m^2) and n = is the initial number of sample plots. The grid axes were oriented to the north–south and east–west directions with the starting point randomly selected. Based on the grid, 107 sample plots were located within the boundary of the forest reserve.

Field reference biomass was based on data from a forest inventory conducted on the 107 sample plots which were circular (radius = 17.84 m, 0.1 ha each) for 15 days from the 25th April to 10th May, 2015. Differential Global Navigation Satellite System (dGNSS) was used

to ensure precise registration of the positions of centres for field plots. Two Topcon legacy- E +40 dual frequency receivers were used, one operating as a base station and the other as a rover field unit. These receivers observe pseudorange and carrier phase of global positioning system (GPS) and Global Navigation Satellite System (GLONASS). During the study, the baseline between the base station and rover units was approximately 25 km. The position of the base station was determined using Precise Point Positioning (PPP) with GPS and GLONASS data collected continuously for 24 hours as suggested by Kouba (2015) before the commencement of the forest inventory. The rover field unit was placed at the centre of each sample plot on a 2.98 m rod for an average of 33±20 minutes using a one-second logging rate. The recorded sample plot centre coordinates were post-processed using RTKLIB software (Takasu 2009) and the results revealed that the maximum deviations for northing, easting and height were 1.16 cm, 3.02 cm and 3.06 cm, respectively.

The following information was recorded on each sample plot: diameter at breast height (*dbh*) (using a caliper or a diameter tape), scientific name and total horizontal distances to sample plot centres of all trees with *dbh* ≥5 cm. The total horizontal distance from a plot centre to each tree was calculated as the sum of the horizontal distance to the front of each tree and half of the tree's *dbh*. These distances were subsequently used to subset the sample plot data into different sizes of 250, 500, 750 and 1000 m^2, respectively, for further analysis. Furthermore, total tree heights (*ht*) of up to 10 randomly selected sample trees within each plot were also measured. Horizontal distances to sample plot centres and *ht* measurements were made using a Haglöf vertex hypsometer.

In order to assess the effect of sample size on precision of biomass estimates we considered three different systematic samples of different sizes, i.e., the full sample of 107 plots, one sample with half the size (54 plots) in which every second plot was excluded and finally one sample of one third of the full size (36 plots) in which every third plot was retained. In total 12 datasets (i.e. four sample plot sizes × three sets of sample sizes) were created and used for the analyses.

Prior to calculating field reference biomass for each sample plot, we used a height-diameter model developed by Kachamba et al. (2016b) to predict the *ht* of trees whose *ht* was not measured. Aboveground biomass for the individual trees in each sample plot was then calculated using a model developed by Kachamba et al. (2016a) with *dbh* and *ht* as independent variables. Field reference biomass for the sample plots was subsequently

calculated by summing up the individual tree aboveground biomass values within a given distance defined by the radius of the four sample plot sizes. These values were then scaled to per hectare values for the different sample plot sizes. These scaled values were denoted AGB and used as reference values and treated as if they were free from errors, although such plot-wise field values will be subject to both measurement errors and allometric errors. Table 1 shows statistical summary of AGB for the 12 different datasets.

Table 1. Statistical summary of aboveground biomass (AGB) on sample plots based on different sample plot sizes and sample sizes.

Plot size (m^2)	**Statistic**	**AGB (Mg ha^{-1})**		
		(n = 107)	(n = 54)	(n = 36)
250	Range	0 – 150.77	0 – 136.48	0 – 150.77
	Mean	36.86	36.23	36.37
	Std	34.07	33.68	37.27
	Cv (%)	92.43	92.95	102.45
	Stderr	3.29	4.58	6.21
500	Range	0 – 132.53	0 – 132.53	0 – 105.67
	Mean	37.38	39.87	34.21
	Std	30.61	33.57	28.06
	Cv (%)	81.89	84.19	82.00
	Stderr	2.96	4.57	4.68
750	Range	0 – 127.53	0.18 – 106.80	0 – 88.67
	Mean	38.12	39.50	32.63
	Std	28.91	30.35	24.91
	Cv (%)	75.82	76.83	76.35
	Stderr	2.79	4.13	4.15
1000	Range	0 – 125.59	0.14 – 103.79	0 – 86.72
	Mean	38.99	39.59	33.12
	Std	29.49	30.08	26.33
	Cv (%)	75.62	75.96	79.49
	Stderr	2.85	4.09	4.39

Std = Standard deviation, Cv = Coefficient of variation, Stderr = Standard error and n = number of sample plots.

2.3 Remote sensing data collection and processing

2.3.1 UAV imagery collection

Image acquisition took place from 23rd to 26th April 2015. At this time of the year, the rain season had just ended and the trees still had leaves on them. A SenseFly eBee fixed-wing UAV (Sensefly 2015) was used. The UAV was made from flexible foam weighing 537 g without camera. The UAV was equipped with a Canon IXUS127 HS digital RGB (red-green-blue) camera. The dimensions and weight of camera with battery and memory card were 93.2 × 57.0 × 20.0 mm and 135 g, respectively. The camera produces 16.1 megapixel images in the red, green and blue spectral bands. The UAV is also equipped with an inertial measurement unit as well as an on-board Global Navigation Satellite Systems (GNSS) to control the flight and to provide rough positioning (Sensefly 2015).

Prior to taking images, positions of ground control points (GCPs) were identified and measured. The GCPs markers were made of a set of 1 × 1 m cross-shaped timber planks painted white and some black and white 50 × 50 cm checkerboards. The position of the centre of each GCP was fixed using the same procedure used in locating sample plot centres for the sample plot inventory described above. The data were collected for an average of 13±6 minutes for each GCP with a 1-second logging rate. The recorded coordinates for each GCP were post-processed similarly as the sample plots. The results revealed that maximum deviations for northing, easting and height were 2.24, 4.50 and 4.46 cm respectively.

Acquisition of images was controlled from a laptop computer with mission control software eMotion 2 version 2.4 (Sensefly 2015). All the flights were planned in the mission control software prior to flying. For navigation purposes we used a georeferenced base map from Microsoft Bing maps covering the study area. For this study, we applied percentage end and side image overlaps of 80 and 90%, respectively, as well as a fixed flight height above the ground of 325 m. In total 20 flights were carried out to cover the forest. A summary of flight characteristics for each flight day is presented in Table 2. The procedure described here was also applied in Kachamba et al. (2016b).

Table 2. Summary of flight characteristics for each flight block of imagery collection.

Date	Number of flights	Number of images	Flight time (minutes)	Wind speed (m sec^{-1})	Cloud cover (%)
23.04.2015	7	1241	518	6.0 – 9.5	10 – 80
24.04.2015	6	1301	489	6.0 – 9.0	20 – 80
25.04.2015	6	1118	502	5.0 – 8.5	10 – 100
26.04.2015	1	273	22	3.0 – 4.0	50

2.3.2 Image processing

Agisoft Photoscan Professional version 1.1 (AgiSoft 2015) was used to generate a 3D point cloud from the acquired images. This software uses both structure for motion (SfM) and stereo-matching algorithms for image alignment and multi-view stereo reconstruction. Generation of the point cloud involved the process displayed in Table 3. Finally, we added spectral information from the point cloud, i.e., red, green and blue image bands to the point cloud. According to Wallace et al. (2016), spectral information from a UAV point cloud can present additional useful information for estimating other non-structural properties of the canopy.

Table 3. Processing steps with corresponding parameter settings in Agisoft Photoscan Professional software for generation of 3D point clouds from UAV imagery.

Task	Parameters
a) Image alignment	Accuracy: high Pair selection: reference Key points: 40000 Tie points: 1000
b) Mesh building	Surface type: height field Source data: dense cloud Face count: high
c) Guided marker positioning	Manual positioning of markers on the 14 GCPs for all the photos where a GCP was visible
d) Building dense point cloud	Quality: medium Depth filtering: mild

2.3.3 Point cloud normalization

The study by Kachamba et al. (2016b), conducted in the same study area, revealed that a digital terrain model (DTM) developed from unsupervised ground filtering of the photogrammetric point cloud performed well. The ground filtering was conducted using a version of the progressive triangular irregular network (TIN) algorithm (Axelsson 1999) implemented in Agisoft Photoscan software (AgiSoft 2015). The algorithm divides the point

cloud into cells of a certain size. In this study, a cell size of 50 m was applied. In each cell, the lowest point is detected and is triangulated to produce a first approximate terrain model. Next, a new point is added to the ground class, providing that it satisfies two conditions: it lies within a certain distance from the TIN and that the angle between the TIN and the line to connect this new point with a point from a ground class is less than a certain angle. This step is repeated while there still are points to be checked. The ground filtering was based on the angle-distance parameter combinations of 3° and 6 m, respectively, as recommended by Kachamba et al. (2016b). In this study, we applied this DTM to calculate the height relative to the ground for all points by subtracting respective TIN values from each point.

2.3.4 Variable extraction

Variables describing canopy height were derived from the normalized point cloud and they included maximum and mean values (*Hmax, Hmean*), standard deviation (*Hsd*), coefficients of variation (*Hcv*), kurtosis (*Hkurt*), skewness (*Hskewness*) and percentiles at 10% intervals (H_{10}, H_{20}, …, H_{90}). A height threshold of 0.5 m was applied in order to separate tree points from low vegetation and ground. Apart from the variables describing canopy height, variables describing canopy density were derived by dividing the height between a 95^{th} percentile of height and the 0.5 m threshold into 10 equally tall vertical layers and calculating the proportion of points above each layer to the total number of points. These variables were denoted as follows: D_0, D_1, …, D_9. In addition, spectral variables derived from the RGB spectral bands were included. The spectral variables were computed as the maximum (*Smax*), mean (*Smean*), standard deviation (*Ssd*), coefficient of variation (*Scv*), kurtosis (*Skurt*), skewness (*Sskewness*) and nine percentiles (*S10, S20, …, S90*) for each of the three bands. For example, the *Smax* variable was denoted as follows: *Smax.red*, *Smax.green* and *Smax.blue*. The remaining spectral variables were also denoted similarly. In total, 64 variables describing canopy height, canopy density and canopy spectral properties were extracted. Some of the extracted height and density variables had missing values for some sample plots for one or several of the plot sizes. In such cases we replaced the missing values with zero in the datasets.

2.4 Model development and evaluation

Kachamba et al. (2016b) found that square root-transformed dependent variables performed better than untransformed and logarithmically transformed variables when modelling AGB with variables derived from the UAV imagery. We therefore fit multiple linear regression models relating square root-transformed AGB (as dependent variable) and the calculated

variables (as independent variables) for each of the 12 samples constituting different combinations of sample plot sizes and sample sizes. Since multicollinearity normally occurs between remotely sensed variables (Meng et al. 2016), we applied the best subset variable selection procedure using the *leaps* package (Lumley 2009) in R statistical software (R Core Team 2016). The selection of potential independent variables was restricted to a combination of up to four variables (i.e., based on preliminary tests) with minimum Bayesian information criterion (BIC) and variance inflation factor (VIF) values as selection criteria. An empirical ratio estimator for bias correction proposed by Snowdon (1991) was employed when converting predictions to arithmetic scale since square root transformation introduce a bias during back-transformation. The empirical ratio estimator was estimated from the ratio of the mean of the observed values to the mean of the back-transformed predicted values. The estimates were finally corrected by multiplying them with the estimated ratio.

For each model, reported values included root mean square error (RMSE), relative root mean square error (RMSE%), mean prediction error (MPE), relative mean prediction error (MPE%) and squared Pearson's correlation coefficient (r^2). However, comparison of the models was based on RMSE values. RMSE is a reliable measure for model performance as it accounts for both variance and bias of the predicted value (Gregoire & Valentine 2007). RMSE, MPE, RMSE% and MPE% were calculated as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (2)$$

$$\text{MPE} = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{n} \quad (3)$$

$$\text{RMSE\%} = \frac{\text{RMSE}}{\bar{y}} \times 100 \quad (4)$$

$$\text{MPE\%} = \frac{\text{MPE}}{\bar{y}} \times 100 \quad (5)$$

where y_i is the AGB field value for plot *i* and $\hat{y}_i$ is the predicted AGB for sample plot *i*, respectively, *n* is the number of plots and $\bar{y}$ is mean AGB for all plots.

Both RMSE and MPE values were calculated using leave-one-out cross validation. The MPE value of each model was tested to check for statistical significance using a two-sided student's t-tests with α=0.05.

2.5 Biomass estimation methods

2.5.1. Field-based sample estimation

We applied a simple random sampling estimator to estimate AGB and standard errors of AGB estimates based on the sample plots (Gregoire & Valentine 2007). Application of systematic sampling design brings positive bias on the estimates (Thompson 2012b). However, according to Gregoire and Valentine (2007), this bias is negligible and unknown as it is a property of the estimator and not a particular sample (Gregoire & Valentine 2007). Since the focus of the analysis in the current study was on the assessment of the influence of sample plot sizes and sample sizes on the precision of the estimates, we assumed that the calculated variance estimates based on the simple random sampling design were adequate for the task at hand. We thus estimated mean AGB per hectare for the study area as follows:

$$\hat{B}_{field} = \frac{\sum_{i=1}^{n} y_i}{n} \quad (6)$$

where y_i is AGB for plot *i* and *n* is the total number of sample plots.

The variance for the estimator in (6) was estimated as follows:

$$\hat{V}(\hat{B}_{field}) = \frac{\sum_{i=1}^{n} (y_i - \hat{B}_{field})^2}{n(n-1)} \quad (7)$$

where $\hat{V}(\hat{B})_{field}$ is the variance estimator for estimated mean AGB.

2.5.2 UAV-assisted estimation

First, the study area was tessellated into grid cells using regular grids with sizes equivalent to respective sample plot sizes, i.e., 250 m^2, 500 m^2, 750 m^2 and 1000 m^2 for the purpose of estimating AGB for the entire area. For each grid cell, variables were derived from the image data as described from the sample plots above. For each of the plot sizes and sample sizes, the previously fitted regression model (Section 2.6) was applied to predict AGB values for each grid cell. Then mean AGB for the study area was estimated based on the model-assisted regression estimator described in Särndal et al. (1992) (Page 231) as follows:

$$\hat{B}_{uav} = \frac{\sum_{i=1}^{N} \hat{y}_i}{N} + \frac{\sum(e_i)}{n} \quad (8)$$

where $\hat{y}_i$ is predicted AGB for the i^{th} grid cell, N is the total number of grid cells for the study area and $e_i = y_i - \hat{y}_i$ is the model prediction residual for plot i.

The variance for the model-assisted regression estimator was estimated as follows (Särndal et al. 1992) (Page 234):

$$\hat{V}(\hat{B}_{uav}) = \frac{\sum_{i=1}^{n}(e_i - \bar{e})^2}{n(n-1)} \quad (9)$$

where e_i is the model prediction residual for plot i and $\bar{e} = \frac{\sum_{i=1}^{n} e_i}{n}$ is the mean residual for all plots.

2.5.3 Efficiency of UAV-assisted estimations

The relative efficiency (RE) of UAV-assisted over the field-based inventories was calculated as follows:

$$\text{RE} = \frac{\hat{V}(\hat{B}_{field})}{\hat{V}(\hat{B}_{uav})} \quad (10)$$

where RE is the relative efficiency of UAV-assisted over field-based inventories, $\hat{V}(\hat{B}_{field})$ is the variance of the field-based biomass estimate (eq. 7) and $\hat{V}(\hat{B}_{uav})$ is the variance of the UAV-assisted biomass estimate (eq. 9).

An RE value greater than 1.0 indicates higher efficiency of UAV-assisted estimates than field-based estimates for a given plot size (see Cochran 1977). For each of these datasets, biomass estimates, standard error of the estimate (SE) and RE values were calculated for both field-based and UAV-assisted inventories for all the 12 datasets. The SE values were calculated as the square root of the variance of the biomass estimates based on field-based and UAV-assisted inventories, respectively.

2.6 Cost efficiency analysis

During field work, we randomly selected 16 sample plots and for each plot recorded three categories of time consumption, i.e. fixed time (time spent when recording sample plot attributes such as plot number, date, etc.), variable time (time spent on measuring trees) and walking time (time spent during walking from one plot to another). The average recorded

time consumption was 7.5, 25.0 and 7.0 minutes for each of the categories, respectively. We then set the relative cost of a sample plot inventory of 107 sample plots (1000 m^2 each) in a 220 m by 220 m grid to 100% based on the recorded information. We then used the cost information from the current inventory (4 persons working for 15 days with a daily salary of USD 25.13 each) to calculate the variable costs for each plot scaled according to plot size and walking distance.

The costs of UAV data acquisition were fixed for all sample plot sizes and sample sizes because the need for auxiliary remotely sensed information would be the same regardless of plot size and sample size. The cost was computed based on the experience from the current study. The costs included pre-flight preparations and the actual flying where a two-man crew was required. Each person worked five days with a salary similar to the field crew. Post-processing of the acquired images required four days.

3.0 Results

3.1 Model performance based on different sample plot sizes and sample sizes

For assessment of the effect of sample plot sizes and sample sizes on model performance, we fit regression models using the data from the different sample plot sizes and sample sizes based on a leave-one-out- cross validation procedure. None of the models had MPE values that were significantly different from zero ($p > 0.05$) (Table 4). The models produced RMSE% values (on arithmetic scale) in the range of 46.6 – 142.5% of the mean AGB from the respective sample plot sizes and sample sizes. The model for sample plot size 1000 m^2 with a sample size of 107 produced the smallest RMSE% value amongst the models. On the other hand, the model based on a sample plot size of 250 m^2 with a sample size of 36 produced the highest RMSE% value. The RMSE% values decreased as the sample plot sizes increased for the respective sample sizes. Furthermore, for a given sample plot size, RMSE% values increases as the number of sample plots decreases (see Table 4). However, the pattern for MPE% values with increase in sample plot size and sample size is irregular (see Table 4). The r^2 values for the final models ranged from 0.31 to 0.71. The model for sample plot size 1000 m^2 with a sample size of 36 produced the highest r^2 value amongst all the 12 developed models. On the other hand, the model based on a sample plot size of 250 m^2 with a sample size of 107 produced the smallest r^2 value (see Table 4). Figure 2 displays the relationship between field reference and predicted AGB for the different plot sizes (n = 107). The results indicate a pattern of under prediction of AGB for higher AGB values for all plot sizes.

The models contained a minimum of one and a maximum of four independent variables. All the models contained variables describing canopy height. Further, all models, except two, contained variables incorporating red, green and blue spectral bands and variables describing canopy density were only incorporated in the model for sample plots of 750 m^2, with sample size 36.

Table 4. Model performance over different sample plot sizes and sample sizes.

Plot size (m^2)	Sample size (*n*)	r^2	Independent variables	Predicted AGB	RMSE		MPE		
				Mg ha^{-1}	Mg ha^{-1}	%	Mg ha^{-1}	%	p-value
250	107	0.31	Hmax, S20.red	37.02	29.61	80.3	-0.17	-0.5	0.95
250	54	0.32	Hcv, Hmax.red	36.51	36.09	99.6	-0.27	-0.8	0.96
250	36	0.30	Hmax, H60.red	41.84	51.84	142.5	-5.47	-15.0	0.53
500	107	0.44	Hmax, S10.red	37.70	23.82	63.7	-0.33	-0.9	0.89
500	54	0.39	Hmax	41.92	29.02	72.8	-2.04	-5.1	0.61
500	36	0.43	Hmax, H50.red	35.35	31.26	91.4	-1.14	-3.3	0.83
750	107	0.56	Hmax, Hcv, D0, Ssd.blue	37.86	21.99	57.7	0.26	0.7	0.90
750	54	0.46	Hmax	40.17	23.86	60.4	-0.67	-1.7	0.84
750	36	0.62	Hcv, D9	34.46	19.38	59.4	-1.83	-5.6	0.58
1000	107	0.64	Hmax, S70.red, S90.green	39.18	18.18	46.6	-0.19	-0.5	0.91
1000	54	0.69	Hmax, H70.red, H90.blue	40.02	19.50	49.2	-0.43	-1.1	0.87
1000	36	0.71	Hmax, H50	33.18	19.05	57.5	-0.06	-0.2	0.99

*Predicted AGB, RMSE and MPE based on leave-one-out cross validation.

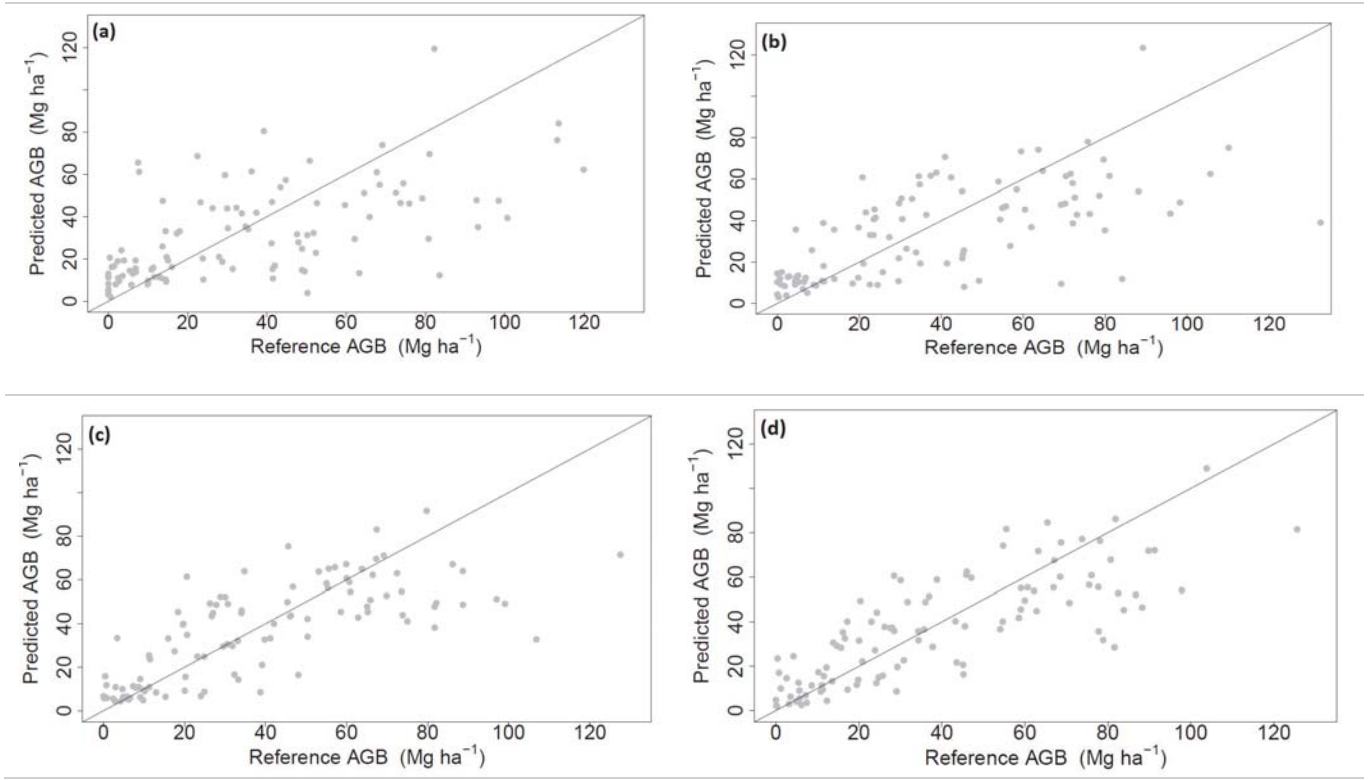


Figure 2. Field reference AGB versus predicted AGB for models from different sample plot sizes (n = 107). a) 250 m^2, b) 500 m^2, c) 750 m^2 and d) 1000 m^2.

3.2 Assessment of the efficiency of UAV-assisted AGB estimation

The standard error (SE) values for field-based AGB estimates were larger compared to those of UAV-assisted for all sample plot sizes and sample sizes (Table 5). For the respective sample plot sizes, SE values for both field-based and UAV-assisted AGB estimates generally increased as the sample sizes decreased. The RE values showed an increasing trend with increasing sample plot sizes. However, the pattern of the RE values was irregular as the sample sizes increased for the respective plot sizes (Figure 4).

Table 5. Estimated mean AGB and associated standard error (SE) estimates based on field-based and UAV-assisted estimation for different sample plot sizes and sample sizes.

Plot size	**Sample size**	**Field-based (Mg ha^{-1})**		**UAV-assisted (Mg ha^{-1})**	
(m^2)	(n)	$\hat{B}_{field}$	SE	$\hat{B}_{uav}$	SE
250	107	36.86	3.29	44.12	2.75
250	54	36.23	4.58	43.63	3.79
250	36	36.37	6.21	49.69	5.20
500	107	37.38	2.96	42.49	2.22
500	54	39.87	4.57	45.60	3.59
500	36	34.21	4.68	42.42	3.52
750	107	38.12	2.79	42.16	1.86
750	54	39.50	4.13	43.39	3.07
750	36	32.63	4.15	43.81	2.56
1000	107	38.99	2.85	43.30	1.72
1000	54	39.59	4.09	43.11	2.30
1000	36	33.12	4.39	40.96	2.36

$\hat{B}_{field}$ = Estimated mean AGB from field-based sample, $\hat{B}_{uav}$ = Estimated mean AGB from UAV-assisted estimation, SE = Estimated standard error of mean AGB.

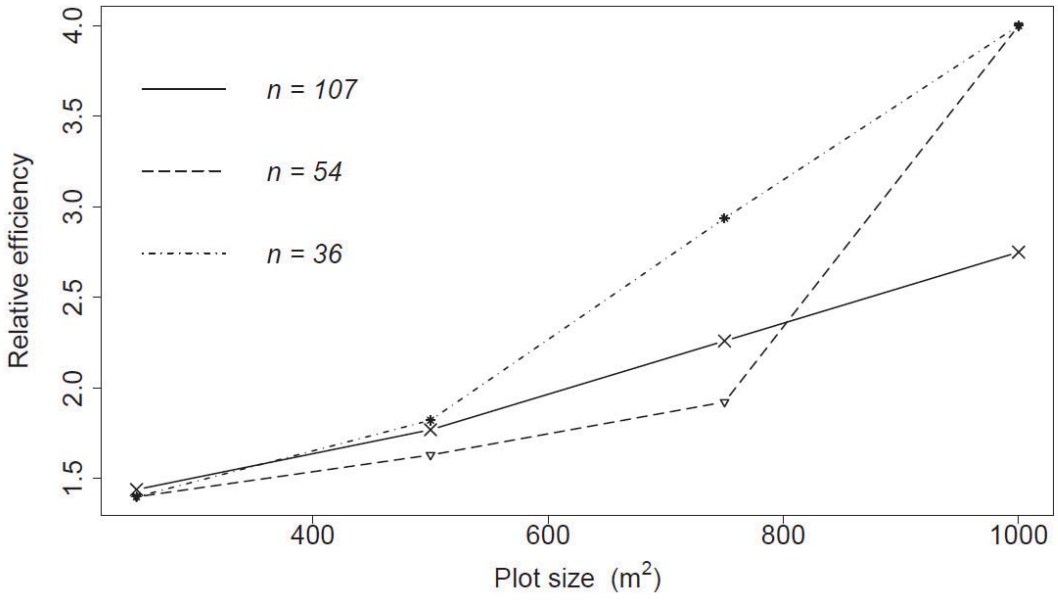


Figure 4. Relative efficiency of UAV-assisted inventory for different sample plot sizes and sample sizes.

3.3 Cost efficiency analysis

The smallest SE was obtained using 1000 m^2 sample plot size and sample size of 107 in a UAV-assisted inventory. However, relative cost of this inventory was the highest amongst the tested sample plot sizes. Furthermore, for a sample plot size of 1000 m^2 the reductions in SE were 0.71 Mg ha^{-1} for UAV-assisted and 1.54 Mg ha^{-1} for pure field-based inventory when increasing the number of sample plots from 36 to 107 (see Table 5). However, the relative increase in costs were approximately 57% and 54% for UAV-assisted and pure field-based inventory, respectively (see Fig 5).

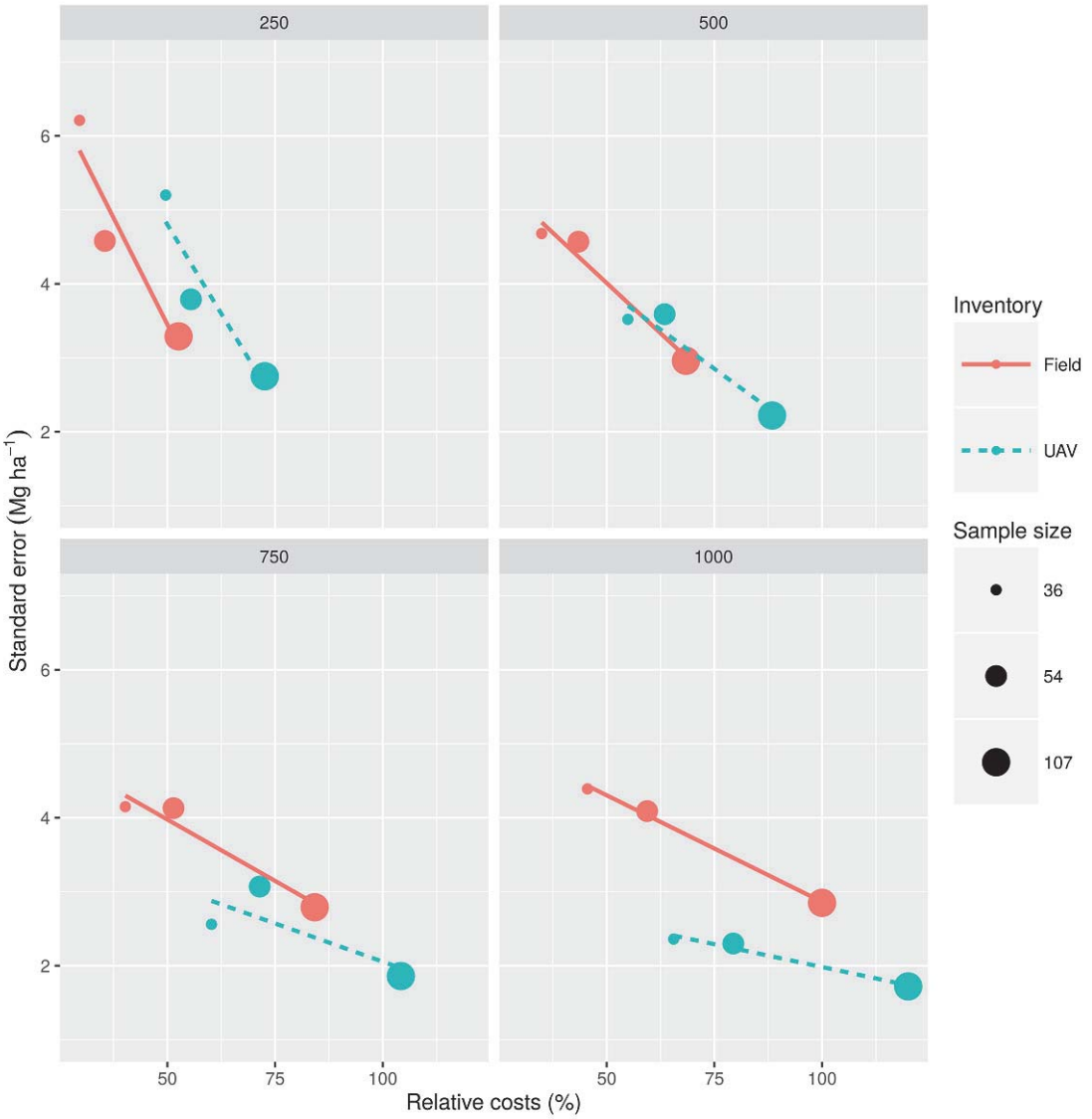


Figure 5. Standard error (SE) estimates and relative costs (RC) for field-based inventories (red) and UAV-assisted inventories (green) with different sample plot sizes (point size) and sample sizes (text label). Linear regression lines of the form *SE = b0+b1*RC* are added for the field-based and UAV-assisted inventory, respectively, to indicate the trends in the results.

4.0 Discussion

Integration of efficient forest inventory techniques in biomass estimation is a key to the successful implementation of the REDD+ mechanism. This study aimed at estimating the efficiency of a UAV-assisted inventory for estimation of biomass based on different sample plot sizes and sample sizes in a forest reserve in miombo woodlands of Malawi which is typical for the size and properties of such reserves in the country. Thus, the results may be of general interest and relevance beyond the scope of this case study. The study has demonstrated that incorporating UAV-derived photogrammetric data in a forest inventory can improve biomass estimates beyond what can be achieved by purely field-based sample inventories. The relatively small SE values for the UAV-based estimates indicate that inclusion of remotely sensed data from UAV imagery can improve the precision of the biomass estimates. Thus, the application of UAV-assisted inventories for REDD+ implementation in Malawi could potentially result in improved biomass estimates compared to pure field-based inventories.

Just like in any remote sensing assisted forest inventory, the precision of biomass estimates based on a UAV-assisted inventory is highly influenced by the choice of sample plot sizes and sample sizes (Frazer et al. 2011). This is demonstrated by several aspects in the results. First, the trends in the RE values show that the magnitude of the efficiency of UAV-assisted inventories is influenced by the sample plot size employed. For instance, the average RE values for the 250, 500, 750 and 1000 m^2 sample plot sizes were 1.45, 1.72, 2.23 and 3.12, respectively, indicating an increase in the magnitude of the efficiency of UAV-assisted inventory with increasing sample plot sizes.

It should however be noted that when reducing the sample sizes, i.e., from 107 to 36, for the different sample plot sizes, the trends for the RE values were irregular. This could be attributed to small sample sizes, e.g. selecting the 36 plots in a different way might give different results. Furthermore, changes in the mean biomass estimates for the respective datasets due to the systematic exclusion/inclusion of plots with high biomass values when selecting plots. This demonstrates the impact the application of systematic sampling design can have on estimated biomass in forests with irregular biomass distribution such as miombo woodlands where forest canopy cover is highly variable.

As in our study, systematic designs are often applied in forest inventories because they are more efficient. However, when the sampling intensity is small for any sampling design, the

risk of not capturing the whole range of the variable of interest increases. Subsequently, poorer predictive models are developed thus leading to extrapolation when predicting forest attributes in forests with attributes that are not represented in the field data. In order to develop a good predictive model, more efficient sampling designs should be adopted by spreading the sample in the space of the variables derived from the remotely sensed data (Grafström et al. 2014; Hawbaker et al. 2009). This issue should therefore be taken into account when planning future UAV forest inventories.

The influence of sample plot and sample sizes was further demonstrated by the trends in RMSE values. The results have shown that the RMSE values decreased by approximately 39%, 46% and 63% between the smallest (250 m^2) to the largest (1000 m^2) sample plot sizes for sample sizes 107, 54 and 36 respectively. The same trend was observed by Frazer et al. (2011) and Mauya et al. (2015) (see Table 6). The magnitude of the decrease in RMSE values reported are quite varied due to differences in the range of sample plot sizes and sample sizes.

Table 6. Effect of sample plot sizes and sample size on the root mean square errors (RMSE) of biomass estimates for different studies.

Authors	*n*	Sample plot size (m^2)	RMSE ($Mg\ ha^{-1}$)	Change (%)	Remotely sensed data
Frazer et al. (2011)	60	314 – 1964	157.60 – 97.80	37.94	ALS
Mauya et al. (2015)	30	200 – 3000	262.47 – 93.73	64.29	ALS
Current study	107	250 – 1000	29.61 – 18.18	38.60	UAV

The improvement of biomass estimates with increasing sample plot sizes shows that large sample plot sizes favour UAV-assisted inventories. This could be attributed to reduction in plot boundary effects as the sample plot sizes increases as suggested by Goetz and Dubayah (2011). Thus, for small sample plots canopies of trees with wide crowns such as those in miombo woodlands (Frost 1996) tend to be partially included and thus under predicting sample plot biomass. On the other hand, as the sample plot sizes increase, this effect tends to decrease substantially since these variations are averaged out at larger sample plot sizes (Saatchi et al. 2011).

The fact that the results in the current study indicate that larger plots and larger sample sizes favour UAV-assisted forest inventories does not imply that larger plots and larger sample sizes should always be applied during the UAV-assisted inventory because of the associated costs. The results of the cost efficiency analysis presented in Figure 5 indicate that there is a

trade-off between costs and required precision. On one hand, acquiring UAV data and field reference data from many large plots is expensive but produces more precise results. On the other, acquiring the data from many small plots is less expensive but produces less precise results. Based on the trends observed in Figure 5, in this study if SE less than approximately 3 Mg ha^{-1} was targeted during a forest inventory, then a UAV-assisted forest inventory should be applied to ensure cost efficient and precise estimates. This demonstrates the need for carrying out a cost analysis during UAV-assisted inventory in order to determine the optimal sample plot size and sample size to apply.

Finally, it should be noted that careful planning is needed for application of UAV-assisted inventories under the REDD+ mechanism in Malawi to be accomplished. For example, if the inventory is intended for smaller forest reserves, wall-to-wall coverage using a UAV is possible. On the other hand, in cases where inventories are conducted in larger forest reserves, the UAV can be applied as a sampling tool because wall-to-wall operations may be found economically and logistically infeasible. Furthermore, this study was conducted on a single site and thus represents a forest inventory scenario at specific location. Although this case study has provided evidence of great efficiency of UAV-assisted inventory, similar studies should be conducted in other reserves across the country in order to be able to generalize and provide guidance for future operational inventories.

5.0 Conclusions

Successful incorporation of UAV-based remote sensing technologies for biomass estimation in REDD+ projects in Malawi requires studying the influence of key factors, such as sample plot size and sample size, which affect the precision of the estimates. The study has demonstrated that UAV-assisted inventories tend to produce more precise estimates compared to those utilising exclusively field-based methods. Furthermore, larger sample plot sizes and sample sizes tend to favour UAV-assisted inventories. We therefore recommend that UAV-assisted inventories be incorporated in the implementation of REDD+ in Malawi to improve biomass estimates. However, more studies are needed to check the influence of sample plot sizes and sample sizes for different study sites where also costs are considered.

6.0 References

AgiSoft. (2015). *Agisoft PhotoScan User Manual: Professional Edition*, Version 1.1: AgiSoft LLC.

Asner, G. P., Flint Hughes, R., Varga, T. A., Knapp, D. E. & Kennedy-Bowdoin, T. (2009). Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. *Ecosystems*, 12 (2): 261-278.

Axelsson, P. (1999). Processing of laser scanner data-algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54 (2-3): 138 - 147.

Campbell, J. B. & Wynne, R. H. (2011). *Introduction to Remote Sensing*. 5 ed. New York: Guilford press. 667 pp.

Chen, Q., Vaglio Laurin, G., Battles, J. J. & Saah, D. (2012). Integration of airborne lidar and vegetation types derived from aerial photography for mapping aboveground live biomass. *Remote Sensing of Environment*, 121: 108-117.

Cochran, W. (1977). *Sampling techniques*. Third ed. Probability and mathematical statistics. Toronto, Canada: John Wiley and Sons, Inc. 442 pp.

Dandois, J., Olano, M. & Ellis, E. (2015). Optimal altitude, overlap, and weather conditions for computer vision UAV estimates of forest structure. *Remote Sensing*, 7 (10): 13895.

DeFries, R., Achard, F., Brown, S., Herold, M., Murdiyarso, D., Schlamadinger, B. & de Souza Jr, C. (2007). Earth observations for estimating greenhouse gas emissions from deforestation in developing countries. *Environmental Science & Policy*, 10 (4): 385-394.

Drake, J. B., Dubayah, R. O., Clark, D. B., Knox, R. G., Blair, J. B., Hofton, M. A., Chazdon, R. L., Weishampel, J. F. & Prince, S. (2002). Estimation of tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of Environment*, 79 (2–3): 305-319.

Eid, T., Gobakken, T. & Næsset, E. (2004). Comparing stand inventories for large areas based on photo-interpretation and laser scanning by means of cost-plus-loss analyses. *Scandinavian Journal of Forest Research*, 19 (6): 512-523.

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. *Remote Sensing of Environment*, 115 (2): 636-649.

Frost, P. (1996). The ecology of miombo woodlands. In Campbell, B. (ed.) *The miombo in transition: woodlands and welfare in Africa*, pp. 11-57. Bogor, Indonesia: Centre for International Forestry Research.

Getzin, S., Wiegand, K. & Schöning, I. (2012). Assessing biodiversity in forests using very high-resolution images and unmanned aerial vehicles. *Methods in Ecology and Evolution*, 3 (2): 397-404.

Gibbs, H. K., Brown, S., Niles, J. O. & Foley, J. A. (2007). Monitoring and estimating tropical forest carbon stocks: making REDD a reality. *Environmental Research Letters*, 2 (2): 13.

Gizachew, B. & Duguma, L. A. (2016). Forest carbon monitoring and reporting for REDD+: What future for Africa? *Environmental Management*: 1-9.

Gobakken, T. & Næsset, E. (2008). Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research*, 38 (5): 1095-1109.

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. *Remote Sensing of Environment*, 123: 443-456.

Gobakken, T., Bollandsås, O. M. & Næsset, E. (2015). Comparing biophysical forest characteristics estimated from photogrammetric matching of aerial images and airborne laser scanning data. *Scandinavian Journal of Forest Research*, 30 (1): 73-86.

Goetz, S. & Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Management*, 2 (3): 231-244.

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+. *Environmental Research Letters*, 10 (12): 123001.

Gonzalez, P., Asner, G. P., Battles, J. J., Lefsky, M. A., Waring, K. M. & Palace, M. (2010). Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. *Remote Sensing of Environment*, 114 (7): 1561-1575.

Government of Malawi. (2010). *Malawi state of environment and outlook report*. Lilongwe, Malawi: Ministry of natural resources, energy and environment.

Grafström, A., Saarela, S. & Ene, L. T. (2014). Efficient sampling strategies for forest inventories by spreading the sample in auxiliary space. *Canadian Journal of Forest Research*, 44 (10): 1156-1164.

Gregoire, T. G. & Valentine, H. T. (2007). *Sampling Strategies for Natural Resources and the Environment*: Taylor & Francis.

Hansen, E., Gobakken, T., Solberg, S., Kangas, A., Ene, L., Mauya, E. & Næsset, E. (2015). Relative Efficiency of ALS and InSAR for Biomass Estimation in a Tanzanian Rainforest. *Remote Sensing*, 7 (8): 9865.

Hardcastle, P. D. (1978). *A preliminary silvicultural classification of Malawi*. Forestry Research Record (Malawi), no. 57. Zomba (Malawi): Forestry research institute of Malawi, .

Hawbaker, T. J., Keuler, N. S., Lesak, A. A., Gobakken, T., Contrucci, K. & Radeloff, V. C. (2009). Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *Journal of Geophysical Research: Biogeosciences*, 114 (G2): n/a-n/a.

Houghton, R. A., Byers, B. & Nassikas, A. A. (2015). A role for tropical forests in stabilizing atmospheric CO^2. *Nature climate change*, 5 (12): 1022-1023.

Kachamba, D. J., Eid, T. & Gobakken, T. (2016a). Above- and belowground biomass models for trees in the miombo woodlands of Malawi. *Forests*, 7 (2): 38.

Kachamba, D. J., Ørka, H. O., Gobakken, T., Eid, T. & Mwase, W. (2016b). Biomass estimation using unmanned aerial vehicle in a tropical woodland in Malawi. *Under review*.

Kangas, A. & Maltamo, M. (2006). *Forest inventory: methodology and applications*. Managing Forest Ecosystems, vol. 10. Dordrecht, Netherlands: Springer. 368 pp.

Keller, M., Palace, M. & Hurtt, G. (2001). Biomass estimation in the Tapajos National Forest, Brazil: Examination of sampling and allometric uncertainties. *Forest Ecology and Management*, 154 (3): 371-382.

Kouba, J. (2015). *A guide to using international GNSS service (IGS) products*. 615 Booth Street, Ottawa, Ontario, K1A 0E9: Geodetic Survey Division, Natural Resources Canada.

Kumar, L., Sinha, P., Taylor, S. & Alqurashi, A. F. (2015). Review of the use of remote sensing for biomass estimation to support renewable energy generation. *Journal of Applied Remote Sensing*, 9 (1): 097696-097696.

Lefsky, M. A., Cohen, W. B. & Spies, T. A. (2001). An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Canadian Journal of Forest Research*, 31 (1): 78-87.

Levick, S. R., Hessenmöller, D. & Schulze, E. D. (2016). Scaling wood volume estimates from inventory plots to landscapes with airborne LiDAR in temperate deciduous forest. *Carbon Balance and Management*, 11 (1): 7.

Lu, D., Chen, Q., Wang, G., Liu, L., Li, G. & Moran, E. (2014). A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth*: 1-43.

Lumley, T. (2009). *Leaps: regression subset selection*. R package version 2.9.

Margolis, H. A., Nelson, R. F., Montesano, P. M., Beaudoin, A., Sun, G., Andersen, H.-E. & Wulder, M. A. (2015). 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*, 45 (7): 838-855.

Mascaro, J., Detto, M., Asner, G. P. & Muller-Landau, H. C. (2011). Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment*, 115 (12): 3770-3774.

Mauya, E., Hansen, E., Gobakken, T., Bollandsås, O., Malimbwi, R. & Næsset, E. (2015). Effects of field plot size on prediction accuracy of aboveground biomass in airborne laser scanning-assisted inventories in tropical rain forests of Tanzania. *Carbon Balance and Management*, 10 (1): 1-14.

McRoberts, R., Andersen, H.-E. & Naesset, E. (2014). Using airborne laser scanning data to support forest sample surveys. Forestry applications of airborne laser scanning. In *Forestry applications of airborne laser scanning--concepts and case studies*, pp. 269 - 292. Dordrecht: Springer.

McRoberts, R. E. & Tomppo, E. O. (2007). Remote sensing support for national forest inventories. *Remote Sensing of Environment*, 110 (4): 412-419.

Meng, J., Li, S., Wang, W., Liu, Q., Xie, S. & Ma, W. (2016). Mapping Forest Health Using Spectral and Textural Information Extracted from SPOT-5 Satellite Images. *Remote Sensing*, 8 (9): 719.

Nhantumbo, I. & Camargo, M. (2015). *REDD+ for profit or for good? Review of private sector and NGO experience in REDD+ projects*. Natural Resource Issues. 80-86 Grays Inn Road, London WC1X 8NH, United Kingdom: International institute for environment and development (IIED).

Næsset, E. & Gobakken, T. (2008). Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment*, 112 (6): 3079-3090.

Næsset, E., Gobakken, T., Solberg, S., Gregoire, T. G., Nelson, R., Ståhl, G. & Weydahl, D. (2011). Model-assisted regional forest biomass estimation using LiDAR and InSAR as auxiliary data: A case study from a boreal forest area. *Remote Sensing of Environment*, 115 (12): 3599–3614.

Næsset, E., Gobakken, T., Bollandsås, O. M., Gregoire, T. G., Nelson, R. & 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.

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., Ørka, H. O., Solberg, S., Bollandsås, O. M., Hansen, E. H., Mauya, E., Zahabu, E., Malimbwi, R., Chamuya, N., Olsson, H., et al. (2016). Mapping and estimating forest area and aboveground biomass in miombo woodlands in Tanzania using data from airborne laser scanning, TanDEM-X, RapidEye, and global forest maps: acomparison of estimated precision. *Remote Sensing of Environment*, 175: 282-300.

Ota, T., Ogawa, M., Shimizu, K., Kajisa, T., Mizoue, N., Yoshida, S., Takao, G., Hirata, Y., Furuya, N., Sano, T., et al. (2015). Aboveground biomass estimation using structure from motion approach with aerial photographs in a seasonal tropical forest. *Forests*, 6 (11): 3882.

Paneque-Gálvez, J., McCall, M., Napoletano, B., Wich, S. & Koh, L. (2014). Small drones for community-based forest monitoring: An assessment of their feasibility and potential in tropical areas. *Forests*, 5 (6): 1481.

Patenaude, G., Hill, R. A., Milne, R., Gaveau, D. L. A., Briggs, B. B. J. & Dawson, T. P. (2004). Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sensing of Environment*, 93 (3): 368-380.

Puliti, S., Ørka, H., Gobakken, T. & Næsset, E. (2015). Inventory of small forest areas using an unmanned aerial system. *Remote Sensing*, 7 (8).

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Saatchi, S., Marlier, M., Chazdon, R. L., Clark, D. B. & Russell, A. E. (2011). Impact of spatial variability of tropical forest structure on radar estimation of aboveground biomass. *Remote Sensing of Environment*, 115 (11): 2836-2849.

Sensefly. (2015). *eBee sensefly: Extended user manual*: Sensefly, Ltd.

Skowronski, N. S., Clark, K. L., Gallagher, M., Birdsey, R. A. & Hom, J. L. (2014). Airborne laser scanner-assisted estimation of aboveground biomass change in a temperate oak–pine forest. *Remote Sensing of Environment*, 151: 166-174.

Snowdon, P. (1991). A ratio estimator for bias correction in logarithmic regressions. *Canadian Journal of Forest Research*, 21 (5): 720-724.

Strunk, J., Temesgen, H., Andersen, H.-E., Flewelling, J. P. & Madsen, L. (2012). Effects of lidar pulse density and sample size on a model-assisted approach to estimate forest inventory variables. *Canadian Journal of Remote Sensing*, 38 (5): 644-654.

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. *Remote Sensing of Environment*, 173: 187-199.

Särndal, C. E., Swensson, B. & Wretman, J. (1992). *Model assisted survey sampling.* New york: Springer-Verlag, inc.

Takasu, T. (2009). *RTKLIB: Open source program package for RTK-GPS*. Free and open source software for Geospatial Sydney: Open Source Geospatial Foundation (OSGeo).

Tang, L. & Shao, G. (2015). Drone remote sensing for forestry research and practices. *Journal of Forestry Research*, 26 (4): 791-797.

Thompson, S. K. (2012a). Design and Model. In *Sampling*, pp. 131-137: John Wiley & Sons, Inc.

Thompson, S. K. (2012b). Systematic and Strip Adaptive Cluster Sampling. In *Sampling*, pp. 339-352: John Wiley & Sons, Inc.

UNFCCC. (2014). *Key decisions relevant for reducing emissions from deforestation and forest degradation in developing countries (REDD+): Decision booklet REDD+*. Bonn, Germany: UNFCCC secretariat. p. 44.

Vauhkonen, J., Maltamo, M., McRoberts, R. E. & Næsset, E. (2014). Introduction to forestry applications of airborne laser scanning. In Maltamo, M., Næsset, E. & Vauhkonen, J. (eds) *Forestry applications of airborne laser scanning—concepts and case studies*, pp. 1-6. Dordrecht: Springer.

Wallace, L., Lucieer, A., Malenovský, Z., Turner, D. & Vopěnka, P. (2016). Assessment of forest structure using two UAV techniques: A comparison of airborne laser scanning and structure from motion (SfM) point clouds. *Forests*, 7 (3): 62.