Estimating aboveground woody biomass change in Kalahari woodland: combining field, radar and optical datasets

Vladimir R. Wingate 1,*, Stuart R. Phinn 2,3†, Nikolaus Kuhn 1† Peter Scarth3†

1Physical Geography and Environmental Change, University of Basel, Klingelbergstrasse 27 Basel 4056; Nikolaus.kuhn@unibas.ch

2Remote Sensing Research Centre, School of Earth and Environmental Sciences, The University of Queensland, St Lucia, QLD 4072, Australia; s.phinn@uq.edu.au

3Joint Remote Sensing Research Program, School of Earth and Environmental Sciences, University of Queensland, St Lucia, QLD 4072, Australia; p.scarth@uq.edu.au

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: Vladimir.wingate@unibas.ch;

Abstract

Maps that accurately quantify aboveground vegetation biomass (AGB) are essential for ecosystem monitoring and conservation. Throughout Namibia, four vegetation change processes are widespread, namely, deforestation, woodland degradation, the encroachment of the herbaceous and grassy layers by woody strata (woody thickening), and woodland regrowth. All of these vegetation change processes affect a range of key ecosystem services, yet their spatial and temporal dynamics and contributions to AGB change remain poorly understood. This study quantifies AGB associated with the different vegetation change processes over an eight-year period, for a region of Kalahari woodland savannah in northern Namibia. Using data from 101 forest inventory plots collected during two field campaigns (2014-2015), we model AGB as a function of the Advanced Land Observing Satellite (ALOS) Phased Array L-band Synthetic Aperture Radar (PALSAR and PALSAR-2) and dry season Landsat vegetation index composites, for two periods (2007 and 2015). Differences in AGB between 2007 and 2015 were assessed and validated using independent data, and changes in AGB for the main vegetation processes are quantified for the whole study area (75,501 km²). We find that woodland degradation and woody thickening contributed a change in AGB of -14.3 Tg
and 2.5 Tg over 14% and 3.5% of the study area, respectively. Deforestation and regrowth contributed a smaller portion of AGB change, i.e. -1.9 Tg and 0.2 Tg over 1.3% and 0.2% of the study area, respectively.

**Keywords:** Aboveground Biomass; Time-series; Random Forest; Landsat; Namibia; Savannah; ALOS PALSAR; Woodland; SAR; ICESat; LiDAR; Change detection

1 Introduction

1.1 Dryland vegetation

Dryland ecosystems encompass hyper-arid to sub-humid regions which include savannas and associated woodlands and forests (Pachauri et al. 2014). Dryland vegetation provides such essential ecosystem services as biodiversity conservation, forage, timber and water, critical for human livelihoods (Adeel et al. 2005). These ecosystems cover roughly 41% of the planet, but are heavily impacted by anthropogenic pressures, such as deforestation, land degradation and urbanization (Sörensen 2007; Scholes and Archer 1997; Bastin et al. 2017). Consequently, high rates of land-use and land cover change (LULCC) in drylands contribute significantly to global carbon (C) emissions (Ahlström et al. 2015; Liu et al. 2015).

Maps of vegetation carbon, or aboveground biomass (AGB), in these ecosystems are essential for modelling vegetation carbon dynamics, assessing the impact of different land-use and management strategies and monitoring vegetation change processes (Hall et al. 2011; Birdsey et al. 2013; Jantz, Goetz, and Laporte 2014; Carreiras, Melo, and Vasconcelos 2013). Despite their importance, AGB in drylands has received little attention compared to humid biomes, since drylands are thought to be less significant in terms of carbon storage (Liu et al. 2015). Yet, woodlands in sub-Saharan Africa were found to sequester a quantity of carbon comparable to that of the Congo basin humid forests (Ryan et al. 2016; Nasi et al. 2009). Moreover, a number of studies have clearly demonstrated their importance to global biogeochemical cycles, including net ecosystem carbon balance (Scheffer et al. 2001). For example, savannah biomes in Africa and Australia are a significant component of the carbon cycle (Liu et al. 2015), and semi-arid vegetation in the southern Hemisphere is an important carbon sink. In contrast, tropical humid forests are a major source of C emissions due to widespread deforestation (Poulter et al. 2014; de Jong et al. 2013). Due to their global significance, a number of recent studies have focused on mapping AGB change in drylands (Brandt et al. 2017; Naidoo et al. 2015; Mograbi et al. 2015; Wessels et al. 2013; Odipo et al. 2016) but also vegetation cover (Scanlon et al. 2002; Hansen et al. 2003).
Two processes affect AGB in drylands, (i) the persistent disappearance of herbaceous and woody vegetation, and (ii) the thickening or encroachment of the woody strata with a consequent loss of the herbaceous layer (Hudak and Wessman 1998; Mitchard and Flintrop 2013). Both these functional and structural changes affect ecosystem processes, services, biodiversity as well as the way the land is used economically (Scholes and Archer 1997; Briggs et al. 2005). Drivers of vegetation loss include deforestation, urbanization and land-use intensification (Bai et al. 2008). Woody encroachment, which occurs in Australia, America, Africa, the northern latitudes and at altitude (Jia, Epstein, and Walker 2003; Archer, Schimel, and Holland 1995; Hudak and Wessman 1998; Fensham, Fairfax, and Archer 2005; Ward 2005; Mitchard and Flintrop 2013; O'Connor, Puttick, and Hoffman 2014; Caviezel et al. 2014), is driven by rising atmospheric carbon dioxide (Bond and Midgley 2000; Donohue et al. 2013), shifts in fire activity (Bond and Midgley 2000; Bowman, Murphy, and Banfai 2010), over stocking, (Asner et al. 2004), loss of browsers (Ward 2005), long-term rainfall changes (Fensham, Fairfax, and Archer 2005) and the synergy of these (Ward 2005). Nevertheless, there are very few studies looking at the effect of rising CO$_2$ in drylands and there is an overall paucity of ground data (Stevens et al. 2017).

Notwithstanding, satellite and field-based studies have uncovered complex climate-vegetation interactions and subtle change processes, characterized by alternating positive and negative vegetation cover trends, present throughout drylands and in particular the African Sahel savannah (de Jong et al. 2013; Bai et al. 2008; Dardel et al. 2014; Nutini et al. 2013; Martínez et al. 2011). Some of these studies have established extensive “greening” and challenge the widely accepted view that land degradation is occurring, yet, in many instances what the observed greening trend represents in the field has yet to be fully explored (Herrmann and Tappan 2013). For example, processes associated with soil degradation, such as erosion, are exacerbated by the loss of herbaceous strata and replacement by hardy shrubs, with consequent species impoverishment; these process are often concealed or masked from satellite-based studies showing greening trends (Brandt 2014; Archer, Schimel, and Holland 1995) (Brandt et al. 2014; Herrmann and Tappan 2013; Gonzalez 2001).

1.2 Remote sensing woodlands and AGB

Tropical dry deciduous forests, hereafter referred to as woodlands, are an important component of savannah biomes globally and occur as forests with sparse, discontinuous canopies, composed of a mixture of trees, shrubs and grasses. They cover over 36% of Africa and thus constitute a large but low density AGB stock (Mayaux et al. 2004; Dewees et al. 2010).
However, African countries harbouring woodland often also have large and rapidly growing populations and consequently, high rates of LULCC, resulting in these ecosystems contributing substantially to Africa’s overall carbon emissions (FAO 2010; Mayaux et al. 2004).

Mapping woodland vegetation communities with satellite remote sensing is challenging, as they consist of mixed tree-grass strata and, in contrast to tropical humid forests, they undergo significant intra- and inter-annual changes in vegetation biophysical properties in each strata (Adams, Goudie, and Orme 1996; Lanly 1982; Grainger 1999). They are distinguished by marked phenological cycles manifesting as pronounced deciduousness and seasonality of the woody and herbaceous strata. Also contributing to the fluctuations of these cycles are variable rainfall and anthropogenic impacts such as fires and urbanization (Houghton and Hackler 2006; Ryan et al. 2012). Furthermore, African woodlands are generally characterized by small-scale (approximately 1 ha) and sometimes shifting cultivation, usually in combination with timber extraction and extensive grazing, so that the resulting mosaic landscape is highly variable over space and time (Ryan et al. 2012).

Remote sensing data are fundamental tools for measuring AGB in woodlands (Röder and Hill 2009), yet a critical drawback of freely available optical sensors is their saturation at closed canopy and moderate to high biomass densities (i.e. tropical forests) (Ahl et al. 2006; Zhang et al. 2003; Fensholt et al. 2012; Tucker et al. 2005; Jensen 1983). Similarly, L-band Synthetic Aperture Radar (SAR) is more sensitive to a greater range of AGB, approximately 100 t ha\(^{-1}\), as indicated in the majority of studies focused on boreal, temperate and tropical humid forests (Austin, Mackey, and Van Niel 2003; Dobson et al. 1992; Dobson et al. 1995; Fransson 1999; Hoekman and Quiriones 2000; Imhoff 1993; Le Toan et al. 1992; Lucas, Armston, et al. 2010a; Luckman et al. 1997; Saatchi et al. 2007). Combining data sources in a multi-sensor strategy can take advantage of each sensor’s relative sensitivities for biophysical modelling, and overcome some of the limitations associated with measurements using single classes of sensors. For example, to quantify AGB, this can be achieved by using the complementary data on forest vertical and horizontal structure, as well as greenness by fusing Light Detection and Ranging (LiDAR), radar and optical sensors (Montesano et al. 2013; Cartus et al. 2012b; Lu et al. 2016; Baccini et al. 2011).

### 1.3 Aims
This study maps and discusses AGB and its changes for part of the Namibian Kalahari woodland biome. This region provides the natural resources for the livelihoods of greater part of the Namibian population, yet its spatial and temporal dynamics and associated four
vegetation change processes (i.e. deforestation and woodland degradation, woody encroachment, and woodland regrowth) remain poorly quantified (Strohbach 2001; Mendelsohn and El Obeid 2002; Wingate et al. 2016; Mendelsohn and El Obeid 2005a; Tian et al. 2016). Hence, we use field measurements of AGB in combination with high spatial resolution multi-temporal maps L-band SAR and dry season Landsat vegetation index composites for two periods (2007 and 2015) to define the extent, intensity and severity of these change processes. In quantifying AGB changes, we aim to and address the following research questions,

(a) Can AGB be mapped and validated using independent data?
(b) Can AGB changes resulting from the four vegetation change processes be identified?
(c) Within which land-use categories are the most important AGB changes taking place and which vegetation change processes are they associated with?
(d) How much AGB has been lost and gained as a result of the different vegetation change processes?
2 Methodology

![Workflow schematic illustrating datasets, initial processing and methodological steps taken for the modelling, change detection, validation and quantitative analysis of this study.](image)

2.1 Study area

Namibia has three predominant land-uses with a number of subdivisions, namely, commercial (urban, government agriculture, agriculture and tourism on freehold land and other government and parastatal), communal (small and large scale agriculture on communal land and resettlement) and conservation (state protected), which occupy 45%, 40%, and 15% of the land area, respectively (Van der Merwe 1983). A long standing debate remains as to the sustainability of these land-uses; it is widely believed that commercial farming better manages resources, while communal farming is resulting in a typical “tragedy of the commons” situation (Hardin 1968; Strohbach 2001). Permanent vegetation loss and woody encroachment with the consequent loss of economically important herbaceous layer species, are the principal environmental concerns faced by both farming sectors and these processes are also believed to be impacting conservations areas (Tian et al. 2016; Ward 2005; Wingate et al. 2016; Ward and Ngairorue 2000). Changes in AGB resulting from deforestation or degradation are often limited
to approximately 2 t ha\(^{-1}\) due to lack of mechanical means, while woody encroachment occurs at regional scales (approximately 10,000 ha\(^{-1}\)) (Wingate et al. 2016; De Klerk 2004b).

The study area encompasses part of the Kalahari woodland biome in north eastern Namibia, forming part of the Miombo forests of southern Africa (Frost et al. 1996) (Figure 2). This includes the communally held regions of Oshikoto, Kavango and Ohangwena. The area’s main vegetation types include mopane savanna in the western-most region, mostly composed of *Colophospermum mopane*, and Kalahari woodland composed mainly of *Burkea Africana* and *Pterocarpus angolensis* comprising the remaining area. Tree species are semi-deciduous or deciduous and grow on the Kalahari sand sheet (Verlinden and Laamanen 2006; Verlinden and Dayot 2005; Campbell et al. 2007; Wang et al. 2007). The region extends over 75,501 km\(^2\) with a mean elevation of 1,100 m and annual precipitation ranging from 400 to 800 mm year\(^{-1}\). Communal land-use is characterised by labour-intensive and subsistence-based farming, with limited conventional input and use rather than ownership rights defining tenure, as well as a relatively high population density (Mendelsohn and El Obeid 2002). Agriculture is distinguished by agro-silvo-pastoralism with widespread small-scale pearl millet (*Pennisetum glaucum*) cultivation (Erkkilä 2001; National Planning 2012; Verlinden and Kruger 2007). Cultivation is depends on rainfall for plant growth and this results in the arable land being free of crop cover for the greater part of the year, with crops being harvested four to five months after planting between April and July (Mendelsohn 2006). In contrast, commercial farming centres mainly on ranching, tourism and conventional irrigated cropping. Land cleared of woodland for small-scale cultivation and urbanization has increased in extent. For example, cleared land increased from 6% to 12% since 1975, with new roads allowing access to arable land, although land abandonment and vegetation succession also occur (Wingate et al. 2016). Woodland loss is thought to be driven principally by widespread urbanization and small-scale pearl millet cultivation on approximately 2 ha plots, the expansion of these and small-scale timber extraction. In addition, areas which were once open woodland with a well-developed grass and herbaceous layer are being replaced with a denser woody layer. The predominant factors believed to be driving this vegetation change process are high stocking densities, loss of browsers and altered fire regimes (Mendelsohn and El Obeid 2005b; Wingate et al. 2016; Ward 2005; Tian et al. 2016; Strohbach 2001).
2.2 Field dataset

We measured woody AGB in Mg ha$^{-1}$ as this constitutes the main biomass component and omit estimates for dead trees, litter and grasses (Figure 9.4). Field data were collected at two sites, (i) adjacent to the town of Eenhana (Figure 2f) within an un-gazetted community forest, and (ii) within Kakekete (Kandjara) village (Figure 2c). Measurements of diameter at breast height (DBH) and species were collected for all trees with a circumference > 10 cm (diameter > 3.2 cm) in geo-located 30 × 30 m plots (900 m$^2$), and time-averaged GPS coordinates with an accuracy of <5 m were taken from the centre of each plot. Tree height was not measured, since the allometric biomass equation selected uses DBH as the main predictor. Here, we used a pantropical allometric model which includes wood density, trunk diameter, and the compound bioclimatic variable ($E$). Mean woody density for Zambia (0.69) was used, since measurements from this region were assumed to be the most closely related, both climatically and eco-floristically, to our study site and a mean $E$ value (0.87) calculated from the spatial area of our study site, as described in (Chave et al. 2014). Data were collected during two field campaigns (January-March) in 2014 and 2016, which resulted in a total of 101 sample plots. At the Eenhana field site, measurements were collected using a defined sampling protocol (Pearson,
Walker, and Brown 2005): high spatial resolution aerial imagery (0.5 m) was used to stratify the sampling area into two different apparent vegetation density classes (i.e. high and low woody cover) using supervised classification to account for landscape heterogeneity, then measurements from a preliminary set of field plots were taken. A minimum number of sites per stratified class were sampled in order to provide a statistically robust estimate for each strata and achieve an acceptable level of precision (i.e. error of 10% of the mean 95% confidence interval), with the aim of providing representative vegetation samples and appropriate statistical probability. The Kakekete field site was only stratified into density classes and plots then sampled along transects (due to limited time and access constraints). An additional “very low” biomass class was included in the final dataset, representing sparsely vegetated areas (i.e. more than half of the plot consisting of bare ground), where AGB was visually estimated based on previous field measurements gained from the forest inventory surveying, in an attempt at capturing coincident satellite data variability. At these sites, an average AGB value for shrubs (DBH = 2-5 cm; AGB = 0.02 t ha\(^{-1}\)), trees (DBH = 5-20 cm; AGB = 0.2 t ha\(^{-1}\)) and large trees (DBH = 20-100 cm; AGB = 4.1 t ha\(^{-1}\)) was calculated based on previous inventory measurements, and these estimates were multiplied by the number of trees or shrubs present. Hence, field data distribution is assumed to be representative of the satellite data. Subsequently, plots were assessed against Landsat scenes for disturbances (i.e. fire and deforestation), to ensure no occurrence before satellite data acquisition. Although each plot was found to not have been affected by deforestation during the study period (2007-2015), more gradual changes, for example, due to encroachment or degradation, cannot be qualitatively accounted for. AGB was then estimated for each measured species, summed and scaled to Mg ha\(^{-1}\) units (Lucas, Armston, Fairfax, Fensham, Accad, Carreiras, Kelley, Bunting, Clewley, Bray, et al. 2010).

2.3 Space-borne Radar

Synthetic Aperture Radar (SAR) facilitates the mapping of AGB due to its effectiveness at penetrating cloud and its sensitivity to woody canopy components (i.e. stems and trunks), thus providing a more direct link to the three dimensional forest structures than optical data. This is especially true in low biomass vegetation communities, despite signal saturation being identified for biomass values ranging from 40 to 180 Mg ha\(^{-1}\) (Cartus, Santoro, and Kellndorfer 2012; Le Toan et al. 1992; Dobson et al. 1992; Ranson and Sun 1994; Lucas et al. 2006; Lucas, Armston, et al. 2010b; Sandberg et al. 2011; Mitchard et al. 2009). An active SAR sensor emits a pulse of energy while simultaneously quantifying the return echoes, which is measured as a unitless variable and referred to as backscatter intensity or normalized radar cross section. Backscatter is quantified as a ratio of the energy returning from an area of ground, assuming it
scatters at the same intensity for any angle (isotropic). Hence, energy returning to the sensor is affected by the amount of incident energy reflected by the land surface as well as its directionality. It follows that energy pulses of distinct wavelengths interact differently with varying land surface components. L-Band, 23 cm wavelengths show a strong coefficient of determination with woody vegetation components, such as tree trunks, branches and canopy, with higher backscatter intensities (more return energy), being associated with higher woody biomass. In this study, the coefficient of determination is shown as ‘$R^2$’ and the Pearson’s correlation coefficient as ‘$r$’. Notwithstanding, backscatter intensities vary as a function of many confounding factors often unrelated to biomass, including soil and canopy structure and water content, which impacts scattering directionality and proportion, respectively. For example, the relationship between biomass and signal saturation varies across vegetation classes and moisture, phenological and seasonal conditions (Lucas et al. 2000; Mitchard et al. 2009; Mitchard et al. 2011b; Le Toan et al. 1992; Woodhouse 2005). Saturation levels for biomass retrieval with SAR sensors ranged from 40 to 180 t ha$^{-1}$, rendering it adequate for the biomass mapping of the Kalahari woodland biome (Le Toan et al. 1992; Lucas et al. 2006; Lucas, Armston, et al. 2010a; Mitchard et al. 2009).

2.4 ALOS PALSAR
The Advanced Land Observing Satellite’s Phased Array-type L-Band Synthetic Aperture Radar (ALOS PALSAR) is an L-Band, 23 cm wavelength SAR capturing cross-polarized data in horizontal send vertical receive (HV) and horizontal send and horizontal receive (HH), which have been used to map AGB at regional scales (Mitchard et al. 2011b; Lucas, Armston, Fairfax, Fensham, Accad, Carreiras, Kelley, Bunting, Clewley, Bray, et al. 2010; Mitchard et al. 2009; Morel et al. 2011). HH and HV polarizations are highly correlated yet the latter are more correlated to AGB. For this study we use the global 25 m resolution PALSAR/PALSAR-2 seamless mosaic for the years 2007 and 2015, distributed by the Japan Aerospace Exploration Agency (JAXA) (Shimada et al. 2014). Data are in gamma naught and both polarizations were used as model inputs, as they were found to reduce the error when modelling AGB in a regression equation (Rignot et al. 1994). Imagery speckle was not assessed as the ALSO PALSAR mosaic generation process already successfully addressed this using the 16-looks approach (Shimada et al. 2014). Several authors demonstrated a reduction in noise, speckle, positional errors associated with GPS localization and a strengthening of the SAR polarization – AGB relationship, with an averaging of the polarization data to a higher spatial resolution (Mitchard et al. 2011a; Carreiras, Vasconcelos, and Lucas 2012), hence to account for these sources of error all predictor layers were resampled to a 60 m resolution. The datasets were
resampled from the original using bilinear resampling, which calculates the value of each pixel by averaging the distance weighted values of the four pixels adjacent and is suitable for continuous data. Digital Numbers (DN) were converted to gamma naught values ($\gamma^0$) using the following equation:

$$
\gamma^0 = 10 \log_{10} (\text{DN})^2 + (\text{CF})
$$

were the calibration factor (CF) is equal to -83.0 (Shimada et al. 2009). DNs were first converted to backscattering intensity then averaged to a 60 m spatial resolution, before being converted to gamma naught values in decibel units (dB), in order to calculate the arithmetic rather than geometric averages (Lucas, Armston, Fairfax, Fensham, Accad, Carreiras, Kelley, Bunting, Clewley, Bray, et al. 2010; Mitchard et al. 2009). Small-scale deforestation may go undetected by moderate to coarse resolution (i.e. 250 m – 8 km) satellite analyses, but can effectively be detected with high resolution imagery, such as ALOS PALSAR (Ryan et al. 2012), which allows change be mapped at a spatial resolution sufficient to capture deforestation, but also a temporal scale sufficient to identify woody encroachment (Chidumayo 1997; Williams et al. 2008). A qualitative comparison of the ALOS PALSAR imagery with Landsat scenes in relation to permanent landscape features (i.e. road intersections) revealed minor geo-referencing discrepancies (i.e. approximately on the order of 1-5 pixels). Here, we omit co-registration due an important lack of permanent landscape features proportionately distributed through the study area (i.e. the study area is mainly savannah with very few roads and buildings identifiable at the imageries resolution), therefore we assume geo-positional errors.

### 2.5 Landsat vegetation index time-series

Trends in satellite-derived time-series are extensively used to monitor vegetation change in drylands (Fensholt et al. 2009), while the increasing availability of high resolution Landsat data make such analyses possible for these sensors (Kuenzer, Dech, and Wagner 2015). Satellite derived vegetation indices (VI) including the Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI), merge reflectance measurements from different parts of the electromagnetic spectrum, for instance the red: near-infrared ratio, to create new biophysical information on vegetation productivity. These include vegetation leaf area biomass and physiological mechanisms such as the fraction of absorbed photosynthetically active radiation (Baig et al. 2014; Myneni et al. 1995), and correspondingly, they have been used to estimate vegetation change processes (Higginbottom and Symeonakis 2014; Eisfelder, Kuenzer, and
Dech 2012). In addition, the estimation AGB in semi-arid regions based on low-resolution optical imagery and seasonal integrated VI metrics with herbaceous biomass measurements has been successfully been carried out (Tucker et al. 1985). Similarly, a high degree of correlation was also found between field measurements of AGB and NDVI in savannah biomes (Sannier, Taylor, and Plessis 2002; Eisfelder, Kuenzer, and Dech 2012).

2.6 Seasonal time-series

The creation of radiometrically consistent seasonal reflectance mosaics often requires considerable data, therefore we use Landsat 5 Thematic Mapper (TM), 7 Enhanced Thematic Mapper plus (ETM+), and 8 Operational Land Imager (OLI), Top of the Atmosphere Reflectance imagery (2007-2015), which provide high resolution (30 m) multi-spectral data (Roy et al. 2014; Kuenzer, Dech, and Wagner 2015). The Fmask algorithm was first applied to all scenes to remove cloud and cloud shadows (Zhu and Woodcock 2012; Zhu, Wang, and Woodcock 2015). All scenes were then composited into a time-series of seasonal images by taking the median pixel value of all available images for four seasons, namely, January-March, April-June, July-September, and October-December, (i.e. four images representing distinguishable seasons were created for each year resulting in an 8-year 2007-2015 seasonal time-series). To reduce data gaps, Landsat 5 and 7 datasets were merged. The time-series attempts to capture part of the seasonal phenological vegetation variation, while simultaneously it maximises the use of low quality images (i.e. by removing clouds and cloud shadows) and reduces image noise from extreme pixel values, such as may result from disturbances including fire, by taking the seasonal median value.

Seasons were derived based on phenology metric analysis of the Moderate Resolution Imaging Sprectroradiometer (MODIS) MCD15A3H product, consisting of a time-series of averaged leaf area index (LAI) values over the whole study area (2007-2015), and calculated using the methods proposed by (Forkel et al. 2015; White, Thornton, and Running 1997). For the study period, the mean start of season (SOS) and end of season (EOS) was a Julian day of the year (DOY) of 340 +/- 8.2 and 130 +/- 16 days, respectively, corresponding to 6 December and 10 May. Similarly, positions of seasonal maximum and minimum values occur at DOY 50 +/- 24 and 230 +/- 4.7 days, respectively, corresponding to 19 February and 18 August (Figure 9.5). The first season available was from January to April 2007, and quarterly composites were created up until December 2015, resulting in a total of 35 images. These time intervals were assumed to be representative of the rainfall and vegetation phenological, annual seasonal cycle,
in that they correspond roughly to four obvious seasons (Mendelsohn and El Obeid 2002). Finally, trend analyses were applied to the resulting dataset.

2.7 Dry Season composite imagery as model predictors
Seasonal composites allow the phenological differences between the woody and herbaceous strata to be distinguished: during the early dry season, the difference between the woody canopy and herbaceous layer are the most pronounced, as the trees generally remain in leaf while the herbaceous layer dries out (Archibal and Scholes 2007; Ryan et al. 2017; Verlinden and Laamanen 2006). Therefore, in addition to the 8-year time-series, EVI and short wave infrared (SWIR) composites based only on Landsat 5 (i.e. to bypass the Landsat 7 SLC-off issue) (Chander, Markham, and Helder 2009), coincident with the year of the ALOS PALSAR imagery (i.e. 2007 and 2015), were generated for the period between the EOS and SOS (i.e. the dry season), to accentuate the presence of the woody canopy. Since reflectance from the SWIR band shows a strong relationship with field-measured AGB (Avitabile et al. 2012; Baccini et al. 2012), this band composite was also included as a model predictor.

2.8 Univariate analysis
Biophysical modelling has often relied on mapping a given dependant variable as a function of single or multiple independent (or predictor) variables, using parametric linear regression models. Several assumptions are inherent to these models, namely, normal distribution of values and associated error, and homoscedasticity (Jensen 1983). As such, predictor layer data were extracted for coincident AGB field plots, and linear logarithmic models fitted to estimate the strength of the relationship.

2.9 Random forest algorithm
Machine learning algorithms including Random Forest (RF) are frequently applied to regression and classification problems with large datasets (Lu 2010), as they allow the inclusion of multiple variables with different predictive strengths (Moisen and Frescino 2002), for instance, correlated, categorical or continuous variables, while being computationally effective (Breiman 2001). The algorithm is increasingly used in ecological studies due to its capacity to rank multiple variables by their predictive power and not “over-fit” data (Cutler et al. 2007; Breiman 2001). Compared to parametric models, it does not presume particular statistical distributions of the data, nor assume any distinct relation (i.e. logarithmic), between dependent and independent variables, and in addition to being multivariate it requires no assumption of normality (Prasad, Iverson, and Liaw 2006). This modelling approach uses modified non-parametric ensemble techniques derived from the classification and regression tree (CART)
method (Breiman et al. 1984; Breiman 2001). Rather than selecting the “best split” from all variable options, it instead samples variables randomly for every node and subsequently selects the best from these (Breiman et al. 1984).

2.10 Random forest in remote sensing

RF is increasingly used by remote sensors to scale measurements of biophysical variables to satellite data (Foody, Boyd, and Cutler 2003; Baccini and Asner 2013), and has been used extensively for broad-scale biophysical mapping, due to its ability to capture non-linear relationships and effectiveness at fusing of multi-sensor datasets (Baccini et al. 2012; Simard et al. 2011; Avitabile et al. 2012; Baccini and Asner 2013). It has been used to map tropical AGB with different data sources (Avitabile et al. 2012; Baccini 2004; Baccini et al. 2008), for instance, by modelling forest inventory measurements as a function of ALOS PALSAR backscatter (Karlson et al. 2015; Mishra and Crews 2014; Baccini 2004; Baccini and Asner 2013; Carreiras, Melo, and Vasconcelos 2013). We therefore use this algorithm to model AGB by combining seasonal Landsat metrics, ALOS PALSAR backscatter and forest inventory measurements.

2.11 Model inputs and validation with training data

A predictive RF model of AGB as a continuous variable was calibrated using field-measured AGB (training data), both HH/HV polarizations and Landsat dry season EVI/SWIR composites, and measures of model accuracy were then computed (Freeman, Frescino, and Moisen 2009; Breiman et al. 1984). Here, a separate model for each period (2007 and 2015) was created using the training data. Model parameters included 500 trees, predictors at every node were calculated as the square root of all predictors, and to assess the relative importance of the predictor the model was calibrated using different combinations of these.

RF model validation diagnostics are usually performed with out of bag estimation (OOB), which we use in this study (Liaw and Wiener 2002). For every model tree grown, the algorithm makes predictions on the OOB data (data excluded from the bootstrap sample from which it validates the models dynamically with a fraction of observations withheld from the model fit). To measure the error for every prediction, the coefficient of variation (CV) was calculated (pixel-wise), by taking the standard deviation of a modelled observation divided by the corresponding average values.

Models were created from the training data and these were validated with an independent test set, namely, the OOB predictions on the training data, where the training data was randomly
We divide the data into training and test sets (20%). We plot variable importance, measured as (i) percentage change in mean standard error (MSE) with the random permutation of each predictor layer, and (ii) mean increase in node purity (NP) of the residual sum of squares, of the OOB sample or from all the splits in the forest. This results from the inclusion of every predictor and is measured using the gini coefficient. Pearson’s and Spearman’s correlation coefficients and Root Mean Squared Error (RMSE) are included for the observed and modelled values.

2.12 Vegetation change processes and change detection

The change detection method aims to gain an estimation of the change in AGB resulting from the main vegetation change processes (i.e. deforestation, woodland degradation, woody thickening and woodland regrowth). Over the time-scale of this study, woodland degradation is characterized by small and gradual changes in woody biomass, often resulting from the felling on individual trees, while deforestation is identified by large and abrupt woody biomass changes generally cause by land clearing and burning. Similarly, woodland regrowth is characterized by relatively large and rapid woody biomass gains, characteristic of fast growing, post-disturbance or successional woody species. In contrast, woody thickening is distinguished by small, gradual changes in woody biomass, which corresponds to the slow growth rate of the shrub species functional type associated with woody thickening, (e.g. *Acacia mellifera* subsp. *detinens*, *Dichrostachys cineria*, *Terminalia sericea*, *Terminalia prunioides*, *Acacia erubescens* and *Acacia reficiens*) (De Klerk 2004a). Therefore, in order to differentiate between the two “gain” process (i.e. woodland regrowth and woody thickening) we assume that only a difference in the magnitude of change between the two dates. The same assumption applies the two “loss” processes.

To detect these changes, we use the image ratioint method, which calculates the relative differences (giving emphasis to differences at the low end of the scale), and effectively discriminates noise from change in SAR imagery (Ryan et al. 2012). Since the result of ratioing is not linear or symmetrical about the zero, the natural logarithm transformation is applied using the following formula:

\[
(Image \ Ratio) = \ln\left(\frac{(AGB)_{2015}}{(AGB)_{2007}}\right)
\]
where AGB_{2015} and AGB_{2007} are the modelled datasets at both dates. The resulting image ratio change map has values ranging from 2.05 to -2.31 (205% to -231%). Here, AGB in 2015 is shown as a percentage of AGB in 2007; values greater than 0% signify AGB increases while values lower than 0% signify decreases.

As outlined above, deforestation implies relatively large AGB losses (>20 Mg ha\(^{-1}\)), while woodland degradation and woody encroachment imply low AGB gains, manifesting as AGB increases over several years (i.e. 1.4–2.0 Mg ha\(^{-1}\) year\(^{-1}\)) (Chidumayo 1997; Williams et al. 2008; Ryan et al. 2012). In order to distinguish AGB changes characteristic of the four vegetation change processes under investigation, thresholds of change were applied to the image ratio change map to create different classes. These aim to capture the magnitude of AGB fluctuations associated with the different vegetation change processes between the periods assessed. Here, values ranging from 50 – 100% were assumed to constitute woody thickening, and from 100 – 200%+ regrowth. Similarly, all values ranging from -50 – -100% were assumed to represent degradation, and values from -100 – -200%+ deforestation (Table 1). All values ranging from 50% to -50% were considered as no significant AGB change or noise. Here, we assumed that 95.5% of the AGB Mg ha\(^{-1}\) values are due to normal variation while 2.3% in each "tail" constitute significant change. To choose relevant thresholds, we compared an area of known deforestation (i.e. an 80 km road approximately 70 m wide, built between the two periods assessed and visited during field campaigns), to subjectively assess the validity of the chosen vegetation change classes. To illustrate this, a pixel with a value of 39.5 Mg AGB ha\(^{-1}\) in 2007 decreased to 12 Mg AGB ha\(^{-1}\) in 2015, a loss of the 27.5 Mg AGB ha\(^{-1}\) and equivalent to a log ratio of -1.19 or -119%, which places the pixel in the deforestation class. In contrast, a pixel with a value of 12.5 Mg AGB ha\(^{-1}\) in 2007 increased to 43 Mg AGB ha\(^{-1}\) in 2015, a gain of the 31 Mg AGB ha\(^{-1}\) and equivalent to a log ratio of 1.14 or 114%, which places the pixel in the regrowth class. A number of assumptions and limitations are implicit in this approach. Firstly, the vegetation change process in question is identified simply based on the magnitude of the observed change. For example, in the regrowth class, we do not assume a low starting AGB value, as would be expected from woodland regrowth after clearing; instead, we simply identify the process based on the fraction of AGB in found 2015 compared to 2007. Therefore, the change classes constitute proxies for different vegetation change processes.

Total AGB for both dates and comparison maps was calculated by aggregating each pixel to the resolution of a ha (using bilinear resampling), summing the each pixel value in Mg ha\(^{-1}\), and
converting to Tg. To quantify total AGB change for each vegetation change process, the various vegetation change classes were used to extract the difference between the 2015 and 2007 AGB maps. The sum of the values then represents the amount of AGB lost or gained per vegetation change class (Tg ha\(^{-1}\)). Finally, the amount of AGB changed per vegetation class was converted to a percentage, both of the total AGB found in the study area in 2007 and within each land-use category.

2.13 Validation

Independent data were used to validate modelled and changed AGB: (i) the change map was compared to a classified trend map derived from a time-series Landsat seasonal Normalized Difference Vegetation index (NDVI) composites using an error matrix; (ii) regression and correlation analysis were undertaken for LiDAR derived height metrics with modelled AGB at both dates, and (iii) results were compared to published biomass datasets.

2.14 Trend estimation

A validation dataset is generated to estimate of the accuracy of the change map. Here, we apply a trend estimation method to the Landsat seasonal NDVI composite time-series and map trends in vegetation greenness (Forkel et al. 2013). Trend calculation is dependent on the time-series length, its temporal and spatial resolution, data quality and method of analysis (Sulkava et al. 2007; Badreldin and Sanchez-Azofeifa 2015). The various approaches for calculating trends produce comparable results with regard to significant trends, although differences occur for weaker trends (de Jong et al. 2011). In this study we identify overall trends using seasonal trend model (STM); please refer to Forkel et al. (2013) for details. Calculated trends were classified as either significantly negative or positive (p=0.05), signifying vegetation loss or gain, respectively, and no significant trend suggesting no vegetation shifts. We compare this validation map to the change map, which is the map of applied thresholds aggregated into positive and negative change, using an error matrix and calculate the user’s, producer’s and overall accuracy. Please refer to Foody (2010) for a discussion on implementing accuracy assessments.

2.15 ICESat

The space-borne LiDAR sensor Geoscience Laster Altimeter System (ICESat) provides continuous observations of Earth (2003-2009) and is used to measure a range of environmental variables including vegetation structure (Lefsky et al. 2005; Simard et al. 2011). ICESat is a waveform sampling LiDAR sensor with a 70 m footprint spaced at 170 m intervals; due to its limited sampling area it is often used in conjunction with other sensors (Badreldin and Sanchez-
Azofeifa 2015; Cartus et al. 2012a). It emits short duration (5 ns) laser pulses and records the echo of those pulses as they reflect from the ground (Zwally et al. 2002). When the surface is vegetated, the return echoes (waveforms) are a function of the vertical distribution of vegetation and ground surfaces. For forests on flat ground, stand height is calculated as the difference between the elevation of the first returned energy minus the mean elevation of the ground return (Harding and Carabajal 2005). LiDAR waveform metrics are related to vertical forest structure, which in turn are correlated with AGB (Lefsky et al. 2002; Drake et al. 2003; Baghdadi et al. 2014; Baccini and Asner 2013; Simard et al. 2011), such that ICESat has been used to map continental scale vegetation vertical structure and AGB (Saatchi et al. 2011; Koch 2010; Scarth 2014).

The linear relationship described by the $R^2$ between LiDAR height metrics and AGB is used as an independent proxy method for validating the AGB models (Baccini et al. 2008; Lefsky et al. 2005). An ICESat metric proposed by (Scarth 2014) for 134,282 waveforms (2003-2009) was aggregated within 5 Mg ha$^{-1}$ modelled AGB classes, averaged and correlation statistics calculated. Aggregated height metrics were also converted to AGB; these were first multiplied by the mean number of stems per ha (as measured from both field campaigns), converted to DBH, and finally, to AGB Mg ha$^{-1}$ using the equations proposed by (Mugasha, Bollandsås, and Eid 2013) (Mugasha et al. 2013). The equation by (Mugasha, Bollandsås, and Eid 2013) was modified to allow for the modelling of low height values by changing “1.3” (i.e. DBH) to “0.4”, where:

$$\text{Height} = 1.3 + \exp(10.5116 - 10.6039 \times (\text{DBH})^{-0.0823})$$

(3)

The temporal mismatch between the ICESat returns and modelled AGB, as well as spatial autocorrelation, were not accounted for.

2.16 Published map comparison

A comparison with published maps derived from different data sources is essential to evaluate model accuracy. Comparison datasets include Baccini et al. (2008) (BAC) (Baccini et al. 2008), and Saatchi et al. (2011) (SAA) (Saatchi et al. 2011). Comparative statistics were calculated, while covariance, variance correlation matrices are also included (Table 4.2) (Snedecor and Cochran 1989). Finally, pixel values of all maps coincident with the field plots were plotted and variances compared. The temporal mismatch between published and modelled data was not accounted for.
3 Results and discussion

3.1 Field data

A total of 7,435 trees of 36 species were measured during the course of two field campaigns. AGB estimates for the 101 sample sites varied from 0 to 108.3 Mg ha\(^{-1}\), with a mean of 36.1 Mg ha\(^{-1}\), median of 38.7 Mg ha\(^{-1}\), standard deviation of 29.6 Mg ha\(^{-1}\) and standard error of the mean of 3.8 Mg ha\(^{-1}\). Training data are skewed towards lower AGB values, whereas those of modelled predictions are not (Figure 3a); such an uneven frequency distribution may have biased model results. Training data frequency distribution may have resulted from the (i) inclusion of visually estimated low AGB plots, and (i) sampling of low biomass areas (i.e. lacking large trees).

Observed and predicted data (Figure 3b), and show that the modelled AGB is able to reproduce the 101 field AGB measurements with a moderate degree of accuracy \((R^2 = 60\%\) and RMSE = 20 Mg ha\(^{-1}\) (Figure 3b). However, results from the Baccini et al., (2008) model with both and modelled AGB and field measured AGB, show a lower level of agreement \((R^2 = 40\%; \text{RMSE} = 16.1 \text{ Mg ha}^{-1} \text{ and } R^2 = 30\%; \text{RMSE} = 27 \text{ Mg ha}^{-1}, \text{respectively}) \text{ (Figure 3 c, d).}

Figure 3.1. Frequency distribution of training data showing skewed distribution (grey) and modelled output for 2015 (yellow) (a); regression plot of measured and modelled AGB for each field site (b); observed and predicted results compared to published results, respectively.

Tree cover, canopy structure and AGB in Miombo woodlands are inherently spatially and temporally variable; a response compounded by the often pronounced anthropogenic impact, yet, mean field-measured AGB (36.1 Mg ha\(^{-1}\)) is comparable to analogous field studies. For example, Tanzanian woodland AGB varied from 13 – 30 Mg ha\(^{-1}\); AGB in Mozambique varied between 31 – 45 Mg ha\(^{-1}\) (Carreiras, Melo, and Vasconcelos 2013; Ryan, Williams, and Grace 2011), while AGB ranged from 58 – 92 Mg ha\(^{-1}\) for Kenyan coastal dry forest (Glenday 2008).
3.2 Univariate analysis

The proportion of the variance explained by fitting the equations were low, with the strongest coefficients of determination found for the HV polarizations, in accordance with previous studies, explaining 74% of the variance for 2015 (Figure 4a) (Lucas et al. 2006; Mitchard et al. 2009). Weak relationships may be due to (i) heterogeneous vertical and horizontal woodland structure. Indeed, sample sites contained a wealth of structural form, vegetation functional types and growth stages, counting mixed aged tree species and transition zones from forest to woodland and grassland, resulting in a broad range of AGB values (0 to 108.3 Mg ha\(^{-1}\)) and hence DBH (Mendelsohn and el Obeid 2005a); (ii) sample plot size, which is known to impact the correlation between SAR polarization and field-measured AGB: larger plot sizes increase the strength of the relationship, hence where vegetation structural characteristics are heterogeneous, sample plots covering larger areas would be more effective (Carreiras, Vasconcelos, and Lucas 2012); lastly, (iii) the low biomass plots for which AGB was visually estimated may have resulted in inaccurate results.

![Figure 3.2](image)

Figure 3.2. Linear logarithmic models for HV/HH polarizations (a, b) and dry season EVI/SWIR composites (c, d) with field measured AGB (\(n = 101\)).

3.3 Model output validation

Observed and predicted AGB resulting from fitting the RF model to a subset of the training data, show that the 2007 and 2015 models explained 88% and 92% of the variance of the OOB data, respectively (Figure 5 a, d). Regression plots of the CV and modelled AGB are shown (Figure 5 c, f). The coefficient of variation ranges from 0.13 – 1.33% and 0.12 – 1.33% in 2007 and 2015, respectively.
Predictor variable importance plots show EVI is the most important predictor, followed by SWIR in 2007 (Figure 5b), while HV, HH and SWIR were the most important predictors in 2015 (Figure 5e). The weaker importance of the HV polarization in 2007 may indicate confounding factors influencing model outcome (i.e. variations in soil and canopy moisture), accentuated in low biomass sparse cover forests (Cartus et al. 2012a). For instance, Radar backscatter intensity (and subsequent AGB estimates), in particular from the HH polarization, may have been affected by, (i) the roughness and exposure of the soil, (ii) and the sparse and deciduous canopy cover and consequent strong influence of the woody vegetation components (Rignot et al. 1994; Carreiras, Melo, and Vasconcelos 2013). Sample sites were often located in areas close to farms and villages, with important amounts of bare ground exposed through the canopy and low biomass density, resulting from the influence of human activity (i.e. intensive grazing, fire, small-scale timber extraction). The important role of the SWIR band in predicting AGB for the 2007 model is in agreement with earlier studies (Baccini et al. 2008; Baccini 2004; Avitabile et al. 2012).
3.4 Change detection

Changes in AGB and respective areal extent are shown in Table 1, while Figure 6 illustrates the change detection results (image ratio) and the image ratio map with thresholds applied to distinguish the vegetation change processes under investigation (thresholds). Plots of AGB change (Tg) and areal extents (%) of the change for each change class are also included (Figure 6a,b). We find that most changes in AGB are associated with the “degradation” class (-14.3 Tg), followed by the “thickening” (2.5 Tg), “deforestation” (-1.9 Tg) and “regrowth” (0.2 Tg) classes. This pattern is also reflected by the areal extent of changes, where we see that woodland degradation affected the largest portion of the study area (14%), followed by woody thickening (3.5%), deforestation (1.3%) and regrowth (0.5%). The area of the “No change” class also contributed to the overall change in total AGB (i.e. -16.5 Tg and 81% of the total area), suggesting widespread but low intensity change.

<table>
<thead>
<tr>
<th>Thrids</th>
<th>Vegetation change</th>
<th>Area (km²)</th>
<th>Area (%)</th>
<th>AGB change (Tg)</th>
<th>AGB change since 2007 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100% to -200%</td>
<td>Deforestation</td>
<td>957.5</td>
<td>1.3</td>
<td>-1.9</td>
<td>-4.2</td>
</tr>
<tr>
<td>-50% to -100%</td>
<td>Degradation</td>
<td>10555.0</td>
<td>14.0</td>
<td>-14.3</td>
<td>-31.5</td>
</tr>
<tr>
<td>-50% to 50%</td>
<td>No change</td>
<td>61188.1</td>
<td>81.0</td>
<td>-16.5</td>
<td>-36.5</td>
</tr>
<tr>
<td>50% to 100%</td>
<td>Thickening</td>
<td>2648.6</td>
<td>3.5</td>
<td>2.5</td>
<td>5.5</td>
</tr>
<tr>
<td>100% to 200%</td>
<td>Regrowth</td>
<td>151.4</td>
<td>0.2</td>
<td>0.2</td>
<td>0.4</td>
</tr>
</tbody>
</table>
Figure 3.4. Map of the thresholds applied to the image ratio map (50%, 100% and 200% +) to identify the degree of AGB change and hence the vegetation change processes. Plots of AGB (Tg) change between each date, for each class (a), and its respective area (b) as a percentage for each vegetation change class. Changes in AGB for each vegetation change process as a percentage of the total AGB contained within each land-use in 2007 (y-axis) (c). Likewise, the changes in AGB for each vegetation change process, as a percentage of the total AGB of the study area, in 2007 (y-axis) (d). AGB change in relation to land-use

Figure 6 (c, d) summarises the AGB changes and associated areal extents of those changes, undergone by the different vegetation change processes, in relation to land-use. We find that in terms of percentage of the total AGB per land-use category, deforestation was most pronounced on land designated as “agriculture and tourism on freehold land” (7.8%), followed by “resettlement” (6.7%), “government agriculture” (5%), “state protected” (5.1%) and “urban” (4.3%). In terms of degradation, the most important AGB changes were associated with “government agriculture” (43%), “resettlement” (43%), “agriculture on and tourism on freehold land” (39%), and “small scale agriculture on communal land” (18%). Woody thickening was most prevalent on “urban” (16.4%), “large scale agriculture on communal land” (4.4%) and “small scale agriculture on communal land” (4.3%), while regrowth was most widespread on “urban” land (4.3%). However, for the whole study area, the highest percentage of AGB change associated with deforestation (1.7%) and degradation (11.4%) occurred on communal land.
These results are reflected by the areal extent of each vegetation change process for every land-use, shown both as a percentage of the respective land-use and study area. However, for the whole study area most deforestation and degradation occurred on communal land.

3.5 Validation

3.5.1.1 Map comparison

Differences in model date were not accounted for in the comparison and hence may have contributed to amplifying differences. We compare only the 2015 map as it is most coincident with the date of the field data collection. Mean model AGB for 2015 was 25.3 Mg ha\(^{-1}\), an estimate which is higher than for studies conducted by BAC and SAA. The maximum value (98.8 Mg ha\(^{-1}\)) is comparable to those of the published maps, but the minimum value (5.9 Mg ha\(^{-1}\)) is higher. Pearson’s correlation coefficients suggest a weak linear relationship between maps (i.e. 0.3 and 0.5 for BAC and SAC, respectively), while the RMSE shows comparable results between maps (Table 2). A moderate relationship exists between model predictions and those of BAC (Figure 3 c, d), \((R^2 = 40\%)\), while for those of SAA none is apparent.

Table 3.2. Descriptive and comparative statistics for model and published AGB predictions: Root Mean Squared Error (RMSE) and Pearson’s correlation coefficients are included for all three maps.

<table>
<thead>
<tr>
<th>AGB model</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
<th>RMSE</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1=Model 2015</td>
<td>5.90</td>
<td>68.80</td>
<td>25.30</td>
<td>12.10</td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2=Baccini et al. (2008)</td>
<td>0.00</td>
<td>97.00</td>
<td>24.91</td>
<td>8.18</td>
<td>1</td>
<td>1</td>
<td>12.40</td>
<td>21.10</td>
</tr>
<tr>
<td>3=Saatchi et al. (2011)</td>
<td>1.00</td>
<td>80.42</td>
<td>10.11</td>
<td>11.96</td>
<td>1</td>
<td>2</td>
<td>12.10</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>21.50</td>
<td>17.90</td>
</tr>
</tbody>
</table>

3.5.1.2 ICESat and model validation

ICESat height metrics and modelled AGB in 2015 revealed a strong relationship \((R^2 = 85\% \text{ and } 87\%, \text{ respectively})\). This relationship weakened for the higher biomass classes (i.e. standard error and deviation), suggesting decreased model accuracy with at higher AGB estimates (Figure 7). The weak \(R^2\) in the lowest biomass class (0-5 Mg ha\(^{-1}\)) suggests inaccurate values. This may be linked to the ICESat sensors known lower accuracies when sampling short vegetation. This comes as a result of the interplay between the sensors footprint diameter (90 m), the terrain slope and roughness of the terrestrial surface and the transmitted pulse width, which when acting together render it highly complex to derived short vegetation from the actual
ground return (Scarth 2014). Both correlations lend support to model predictions (i.e. vegetation height is strongly correlated with biomass) (Baccini et al. 2008; Lefsky et al. 2002), while providing an independent validation method. A similar relationship is evident when the height metric is converted to AGB (Figure 7b).

Figure 3.5. $R^2$ and Pearson’s coefficients ($r$) resulting from ICESat height metrics and modelled AGB 2015 (a), and ICESat height metrics converted to AGB (b). Red error bars show standard errors and black error bars show the standard deviation.

3.5.1.3 Change detection validation

Overall accuracy was 65% when comparing all available pixels of the STM reference map and threshold change map for part of the study area, however, when applying the mapped area proportions variable ($W_i$) (Olofsson et al. 2014) overall accuracies decreased to 22% (Table 9.6 and 9.7). User’s accuracies tended to be higher than producers accuracy. Results indicate that although there was a moderate parallel between model and validation datasets (overall accuracy), most of the equivalence was found in the “No change” class, followed by “Deforestation/Degradation”, with very little relationship between the “Thickening/Regrowth” classes (user’s accuracy).

3.5.2 Total AGB estimates

Estimates of total carbon stocks for the study area in 2007 where of 221 Tg and 191 Tg in 2015 (overall loss of 30 Tg). These results are proportionate to the comparison maps (i.e. 187.08 Tg for BAC and 75.82 Tg for SAA). Similarly, Ryan et al. (2012) (Ryan et al. 2012) found AGB for their >1000 km$^2$ study site in central Mozambique to be of 2.13 ± 0.12 Tg in 2007 (equivalent to approximately 159.75 Tg for the same area). Our results are slightly higher than
those of published maps, which may be the result of including small stem diameter classes. In effect, the above studies did not include <5 cm stems in their model. In order to test this, we ran the models with the stems <5 cm excluded and found that total AGB in 2007 was 211 Tg while in 2015 it was 199 Tg, revealing a net change of only 12 Tg. These results suggest the exclusion of the small stems from the model has caused an overall decrease in estimates of carbon density (Mg ha\(^{-1}\)) in both models and hence smaller estimates of the total change in AGB between the periods assessed, and has therefore resulted in a marked effect upon model carbon density estimates (Figure 8).

Figure 3.6. Descriptive statistics (interquartile range, upper quartile, lower quartile, max, min, sd and mean) and sum of the carbon density (Tg) for modelled AGB including stems with a diameter <5 cm in 2007 and 2015 (panels a and b, respectively), and excluding stems with a diameter <5 cm in 2007 and 2015 (panels c and d, respectively). Results show that the exclusion of <5 cm diameter stems impacts the overall estimates of carbon density (Mg ha\(^{-1}\)).

4 Conclusion

We modelled AGB for a part of the Kalahari woodland savannah biome in Namibia using ALOS PALSAR backscatter, dry season Landsat metrics and field measurements DBH. We then quantify the changes between two dates for the four main different change processes (i.e. deforestation, woodland degradation, woody thickening and woodland regrowth). Variance explained from fitting both RF models was high, with comparable RMSE values. Model validation using independent data was successfully carried out, with ICESat height metrics and model predictions revealing strong coefficients of determination, supporting model validity. However, change map validation using a reference dataset demonstrates only moderate overall classification accuracy with intermediate to low accuracies for change classes. Although our models give slightly higher biomass estimates compared to published maps, we suggest this small effect is likely due to our sampling, and inclusion in model calibration, of smaller stem diameters. To test this, we ran the model with <5 cm stem diameter classes excluded and find that overall estimates of carbon density are lower. We use the image ratio change detection method to measure the degree of AGB change between the two dates compared, which serves
as a proxy to identify the four main types vegetation changes expected. We then apply thresholds to distinguish AGB changes associated with the four different vegetation change processes. Results demonstrate that most change in AGB is associated with woodland degradation, followed by woody thickening. Deforestation and regrowth contribute a smaller portion of AGB change, a pattern which is also reflected by the areal extent of these. Change in AGB in relation to different land-uses categories were also assessed, revealing that land designated as “agriculture and tourism on freehold land” experienced the greatest amount of AGB losses associated with deforestation, while degradation was most prevalent on “resettlement” land. These results are in agreement with previous studies, which identify both extensive greening but at the same time small-scale deforestation and woodland degradation. The Kalahari woodlands are structurally complex, presenting a range of plant functional groups with contrasting stand ages and successional stages, and this variability is exacerbated by intensive anthropogenic influences. Accurately mapping AGB in such a dynamic biome is hindered by a number of issues, chiefly environmental variability and consequent vegetation phenological fluctuations which affect satellite data, together with field datasets which may not adequately describe the biophysical parameter being measured, lead to variable estimates.

5 References


Brandt, Martin. 2014. "Environmental Change Based on Earth Observation and Field Data--A Local Study in the Sahel Zone of Mali and Senegal."


Mendelsohn, John. 2006. *Farming systems in Namibia: Research & Information Services of Namibia*.


Ryan, Casey M, Mathew Williams, John Grace, Emily Woollen, and Caroline ER Lehmann. 2017. "Pre-rain green-up is ubiquitous across southern tropical Africa: implications for


